HMS 2009

Puerto de la Cruz, Spain
September 23-25

Contents

WELCOME TO HMS 2009

The International Conference on Harbor, Maritime & Multimodal Logistics Modeling and Simulation (HMS Conference) concentrates on applications of simulation and computer technologies to logistics, supply chain management, multimodal transportation, maritime environment and industrial logistics.

The HMS Conference is already a classical event: it was held successfully world-wide, usually in the surroundings of major international ports (HMS1999 Genoa, HMS2000 Portofino, HMS2001 Marseille, HMS2002 Bergeggi, HMS2003 Riga, HMS2004 Rio de Janeiro, HMS2006 Barcelona); however HMS foundations are even older, originally a "Harbor & Maritime Transport Workshop" was included as part of Simulation in Industry Conference, Genoa (1996) and in 1998 a "Modeling and Simulation within a Maritime Environment Workshop" was organized in Riga; in 2007 the HMS2007 conference moved back to Italy as part of I3M2007, after Barcelona (2006) and Marseille (2005).

In 2008 the Workshop was held in Campora SG Calabria, Italy co-located with I3M2008: the event was a great success in terms of both quality and attendance.

This year the HMS Conference is held in Tenerife, Canary Islands, Spain as a joint event within I3M2009, (co-locating HMS2009, EMSS2009, and MAS2009). The set of contributions included into the HMS2009 program looks very promising in ensuring a high international standard for this event, thus continuing the sequence of successful conferences of the HMS series.

The HMS2009 application context is related to a sector that today is strongly affected by the crisis; as example during last 18 months the Daily Chart for Cape Dry Bulk Ship moved up to 200'000 USD/day to low values such as 3'000 USD/day, to be back on 40'000 USD/Day; similar situations are happening in many different areas such as containers, heavy haul, terminal activities, rails, industrial logistics, etc.

Concurrently, huge investments are still on-going in re-organizing infrastructures and organization for being ready for the new start-up of the world economy. In this situation and considering the impact of risks and opportunities, it is evident that simulation, in its broad meaning, represents the strategic advantage to properly support decisions.

In the HMS Conference, a special attention is always paid to selecting papers focusing on these sharp subjects, but emphasising innovation in methodologies and applications; the selection process, based on multiple viewers, is traditionally very effective and in fact, even this year as in the past, the authors of best papers will be provided an opportunity to further extend their researches for being published in two Special Issues of International Journals connected to our multiconference.

So HMS2009 provides a very important opportunity for sharing innovation in Harbor, Maritime & Multimodal Logistics Modeling and Simulation.

Proceedings of the International Conference on Harbor, Maritime and Multimodal Logistics Modeling & Simulation, HMS 2009
ISBN 978-84-692-5416-5

In the HMS Conference, a special attention is always paid to selecting papers focusing on these sharp subjects, but emphasizing innovation in methodologies and applications; the selection process, based on multiple reviewers, is traditionally very effective and in fact, even this year as in the past, the authors of best papers will be provided an opportunity to further extend their researches for being published in two Special Issues of International Journals connected to our multiconference.

We are glad to thank all the supporters: International Program Committee, Local Organizers, Track/Session Chairs and Authors and we welcome you into HMS2009, wishing you will enjoy the wonderful Canary Island resources while attending this Conference.

HMS Proceedings Editors,

Agostino G. Bruzzone, MISS-DIPTEM, University of Genoa, Italy
Gerson Cunha, MISS Rio de Janeiro Center. Federal Univ. of Rio de Janeiro, Brasil
Rafael Martínez, Universidad de Valencia, Spain
Yuri Merkuryev, MISS Latvian Center. Riga Technical University, Latvia

I3M Organizers

Academic Sponsors

Sponsors

REAL-TIME SIMULATORS DEVELOPMENTS: CURRENT AND NEW TRENDS

Rafael J. Martínez Durá

Instituto de Robótica, Universidad de Valencia
C/ Catedrático José Beltrán, 2
46980, Paterna, SPAIN

Rafael.Martinez@uv.es

ABSTRACT
This article briefly describes some technologies applied to current real-time simulation training systems. It focuses on nowadays and future trends both in hardware and software architectures. The most important open source and commercially available physics engines are briefly described along with some scene graph management libraries. Additionally, cluster-based simulator issues are also discussed, including the main problems that arise from using distributed GPU programming. An analysis of current state of the art regarding web browser-based simulators is also done. Finally new trends and cutting-edge research that is done in the real-time simulation field are discussed.

Keywords: real-time simulation, browser-based simulation, GPU clusters, shaders

1. INTRODUCTION

Simulation issues always deal with a set of techniques that allow reproducing the behaviour of both physical processes, such nuclear reactions, store stocks fluctuations or vehicles, and human behaviours, such as evacuation systems or collaborative works. Nowadays, due mainly to the availability of high performance computer systems, it is possible to simulate in real time the behaviour of devices, machinery, vehicles, control rooms, allowing the user to be trained to obtain the necessary skills to handle the real system with high confidence.

A training simulator is built based upon a set of functional blocks, being the most important the image generation subsystem, the dynamical models that compute the behaviour of the simulated objects, the input/output subsystem that connect the user interfaces, the instructor subsystem, the user management subsystem and the evaluation subsystem. All of them are closely coupled, in a way that the actions over any simulator controller are computed by the dynamical models and produce a reaction that is observed immediately by the simulator operator, indicating some kind of interactivity.

This capability presented in any real-time simulator is achieved whenever the computation time that the systems dedicates to calculate the evolution of the simulated environment is smaller than the time that is being simulated. As a consequence, every simulation model, or the image rendering phases, should not be as complex as desired, having to assume some kind of simplifications (reduce the integration step, simplify the dynamical model, reduce the number of polygons in the scene, use levels of detail in the objects, etc.).

Training by means of simulation has several advantages against the use of the real machine, although in any case the simulator can substitute the real one, due mainly to the fact that the simulator will always simplify the working environment. The main characteristics that make the simulators an indispensable element in the training process are:

- Risk avoidance
- Training costs savings
- Ability to reproduce extreme situations
- Ability to train with faulty machinery
- Objective evaluation of user skills

When designing a training simulator, special care must be taken with the details of the simulated environment in order to keep it close to reality, since this will greatly influence in the decrease of the time involved in the training process. The term *presence* (Sheridan 1992 and Slater, Khanna, Mortensen and Insu Yu 2009) refers to the user impression of "being there", in the virtual world that the computer simulates. In order to achieve a high presence, every user sense must be stimulated as much as possible and also as real as possible. Therefore, when integrating a simulator, it is highly recommended to use immersive visualization technologies, moving platforms, real dimension cabins and controllers, 3D sound surround systems and mainly graphical scenes and dynamical models highly accurate.

Along this paper a review of the technology related with real-time training simulator design will be done. Section 2 deals with new trends in computers and I/O hardware that will be used in future simulators. Following, in Section 3, the most important simulation engines are described and also the new trends that exist in multi GPU programming. Next section describes the state in the art in browser-based simulators and finally some open problems and current research lines regarding simulation are explained.

2. HARDWARE ISSUES

When building a simulator, it can be done using many hardware alternatives. The decision will depend on the presence and accuracy that the simulator must provide. Following, a description of the future trends in the computational systems commonly used in the implementation of training simulators is done.

2.1. Computer Architectures

The computer used to execute a simulator always depends on the number of different views that the simulator employs. There are simulators that have one or two screens, like those used in crane simulators, three screens, used in vehicle simulators and even five in general purpose simulators implemented with CAVES (C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon and J. C.Hart 1992). Every view has to be computed by a graphic card, so the simulator computer system at least must provide the capability to include all of them inside, and additionally, it must have enough performance to be able for rendering the images at high refresh rates.

For these two main reasons, PC-based simulators and virtual reality systems were traditionally built with computer clusters, since a graphic card could be plugged in every node and there are also several CPUs that allow the implementation of parallel algorithms and complex dynamical and visual models. In such systems, designers had to cope with the problems derived from the synchronization of simulation data, scene graphs and user inputs.

Nowadays, the evolution of PCs performance and the ability to plug several graphic cards in the same computer, make clusters being superseded by low-cost single computers, avoiding the synchronization issues present in clusters.

2.1.1. PC based configurations examples

Following some advice is given for choosing the hardware to build a simulator, based, as viewed before, in the number of views it has to provide.

When building a simple simulator, a single graphic card with several heads can be used (the Matrox M9140 LP has four inside) for connecting several displays. If high quality images are needed, as motherboards with up to four PCI-e 16X slots are already available, it is possible to have several graphic cards plugged in the same PC. They can be configured to act as a single card with high resolution and antialiasing capabilities (SLI from NVidia and CrossFire from ATI) or in an independent way providing several outputs. Another similar solution is provided by NVidia. Its Quadro Plex Model II is capable of connecting eight displays (it uses four NVidia Quadro FX 4500 X2) and the NVIDIA Quadro Plex Model IV, is capable of connecting four synchronized displays (it uses two NVidia Quadro FX 5600 with GSync boards).

2.1.2. Game Consoles based Systems

Another important line to take into account is all the hardware related with video games. Video games are similar to simulators in many aspects, so the computer architectures specially developed for them also have enough power to drive a simulator, but with very low costs. The only restriction is that video game consoles are thought to have just a single screen, so they are not good for multi display systems, however clusters of them can be easily built in the way that traditionally has been done with the PC-based simulators. The most important game consoles that are currently available are:

Sony PlayStation 3 (http://www.playstation.com): Its parallel Cell processor and the NVidia RSX architecture (based on a GeForce 7800) provide high performance to this videoconsole, allowing the development of highly parallel applications. Although an official software development kit exists, it is also possible to install Linux Yellow Dog and other distributions for the PS3. They include a set of compilers and libraries for programming the Cell Processor. The main drawback they show is that their kernels restrict the access both the GPU and some peripherals, avoiding the development of high performance applications (See Barttlet 2007).

Microsoft XBOX 360: It is based on a 3 core Xenon CPU with an ATI Xenos graphic card that supports unified shaders. Microsoft has developed a free API (Application Program Interface) that allows the development of applications for this architecture. In order to develop professional games a subscription to the Xbox Registered Developer Program must be done (http://www.xbox.com/dev).

Computer architectures for real-time simulation are thus evolving in two main directions. On the one hand, the use of Personal Supercomputers offers an alternative to clusters, reducing cost and system complexity. On the other hand, the use of clusters is also being reviewed by the introduction of Video Game Consoles which provide a step ahead in parallelism and reduce costs.

2.2. Input Devices

Professional simulators trend to use the same devices employed in game simulators and other video game applications. Main reasons are that its interfaces are very well defined and supported by the operating system and that these devices have very low costs.

One example is the WiiRemote. It is one of the most common computer input devices in the world. It also happens to be one of the most sophisticated. It contains a 1024x768 infrared camera with built-in hardware blob tracking of up to 4 points at 100Hz, a 3-axis accelerometer also operating at 100Hz and an expansion port for even more capability. Its Bluetooth interface allows connecting the device to any personal computer. Many simulation related applications are including the wireless WiiRemote, since it can provide positional information that can be used as a tracker or a wand. See http://www.wiimoteproject.com

Future simulators will also include personal trackers. The high reality you get with this devices increment the presence of the simulator at very low costs. They are based on infrared cameras, like the ones included in the wiimote.

One example to take into account is the Microsoft Natal Project (http://www.xbox.com/live/projectnatal), that uses a video camera that tracks where your body is and what you're doing with it. It also uses a monochrome camera (it works with infrared) that reads depth — how far away your body and its component parts are — and a highly specialized microphone that can pick up voice commands. Along with all this hardware, it's got a ton of software that tells the Xbox how to find your body's various joints (it tracks 48 of them) and how to keep track of multiple players at the same time.

Others examples are the NaturalPoint low cost trackers (http://www.naturalpoint.com/trackir). One of its products, TrakcIR5, makes 6 DOF head tracking by means of detecting some marks on the head of the user. They provide a free software development kit that can be used to include tracking in your own simulators.

Also, new low-cost data acquisition boards are currently emerging, allowing the connection of any analogue o digital device to the computer through the usb interface in a straightforward way. They are HID compatible, so another advantage is that they are programmed very easily, as this protocol is supported by any operating system since it is the one employed to connect the standard game devices. In the same way, you can use your home-built simulator device to play with every commercial game simulator.

2.3. Motion Systems

In order to avoid sickness, motion systems must be included in simulators with immersive displays. This will be another future trend, since their installation is quite easy, programming the motion cueing algorithms it is already not necessary and their cost is not very high.

Electro-hydraulic technology has been used for over 30 years in probably 20,000 6DOF simulators around the world, since its response curves denote they can work at high frequencies (over 15Hz), however the shift to all electric systems is the ground breaking event that will change the world of simulation.

There will be small increments in software improvements, including special effects, buffets and the use of white noise generators to complement the increased fidelity of the visual displays. User interface continues to be improved. Automatic testing is being integrated into the core software so that acceptance testing can be run daily and compared to yesterday/last week/last year's data to show a trend in degraded performance indicating a need for maintenance, or no degradation which would indicate the system is being properly maintained. Moog Inc. is one of the main companies that supply full 6 DoF motion platforms. See http://www.moog.com/products/motion-systems.

Apart from this, in the near future, independent actuators will be used to build your own motions seats. They are mainly used in games, although for medium-size simulators it is very easy to build your own 2-3 DoF motion cabin. SimCraft has recently announced a motion simulation development kit, that allow to built your chassis from one of SimCraft's various free plan options, order a bolt-together kit, or design your own within the SimCraft architecture specification. See http://www.simcraft.com/star.html.

2.4. Display Systems

When dealing with high immersive simulators, today emerging technologies trend to substitute the projection systems that use rear-projection screens and mirrors with embedded in-cabin display systems. Stereoscopic images will be a must, since simulation for training demands a high degree of depth perception. Passive stereo technologies, although more comfortable, need to duplicate the number of visuals generated, so its uses will still remain be very restricted.

New solutions will use 3D LCD monitors with large 46-inch screens and Full HD 1920x1080 resolutions. Nowadays cutting-edge display technologies are based in new single chip solid state illuminated DLP projectors, with WUXGA resolutions (1920x1200). The main drawback of these systems is that brightness is only up to about 700 ANSI lumens, however they are capable of processing and displaying infrared content for simultaneous display of both visible light and the infrared spectrum, and the solid state illumination allow a nearly virtually maintenance-free system.

In the very near future, customizable high resolution panels will be built, by means of replicating low profile tileable single displays. Also some research is being done in auto-stereo systems, avoiding the necessity of wearing glasses, but this technology is still to come.

3. SOFTWARE TRENDS

The main problem that arises when designing a training simulator is that its software is compound by many different modules: the user data bases management system, the graphical user interfaces for the exercises and the instructor, the report generators, and the most important, the scene graph management system and the dynamical models for calculating the behaviour of the active objects. Everyone needs to be programmed using different tools.

The design and implementation of such software elements strongly depends on the computer systems they are thought to be executed. The introduction of new processor architectures, such as GPUs and PPUs, with many cores capable of execute instructions in parallel, has made that new programming paradigms have to be used for them. Furthermore, new algorithms and techniques to increase the simulation realism and performance should be applied to the existing methods. In order to understand what new problems arise when designing the software for current simulation systems, a brief description of the advances in computer architectures follows.

3.1. GPU and PPU Programming

Current trends in computer architectures use replicated parallel cores to increment performance. This is done both at CPU level (for example Intel is currently developing a processor with 80 cores inside, see S. Vangal et all 2007) and also at the graphic processing units (GPU) level (for example see NVidia Testla architecture in E. Lindholm, J. Nickolls, S. Oberman, J. Montrym 2008). The GPU evolved from a fixed function graphics pipeline to a programmable parallel processor with computing power exceeding that of multicore CPUs. The main impact of these current architectures was that now it is possible to execute programs at GPU levels through what are called *shaders*. They are simple programs that describe the traits of either a vertex or a pixel. Following the Direct3D version 9 it is possible to write vertex and pixel shaders in a more abstract, readable and reusable fashion, using a high level C-Style like language called High Level Shader Language (see St-Laurent 2005). Later OpenGL 2.0 introduced a common high level shading language for vertex and pixel programs called OpenGL Shading Language (see Rost 2006). New GPUs incorporate a new architecture called "unified shading core" (Blythe 2006) allowing that both vertex and pixel processing can be handled by one single programmable unit instead of having separate programmable units for both the vertex and pixel pipeline and also added to the pipeline another programmable stage called the geometry shader.

In parallel, NVIDIA developed CUDA (Compute Unified Device Architecture), where the GPU is seen as a massively parallel set of multiprocessors, capable of executing a high number of threads in parallel. In the same way, AMD's response to GPU programming for high-performance data parallel tasks was the AMD Stream Computing Model. Both programming models are based on an extension of the C programming language. Furthermore, NVIDIA and ATI recently have added to their consumer level graphics cards a new concept: the Physics Processing Units (PPUs). These units implement the most common physics simulation techniques, accelerating their calculation. They are also programmed by means of CUDA or other high level libraries. Additional details can be seen in NVidia 2008.

This hardware evolution has stated a revolution in the programming paradigms of the new simulation algorithms, since from then, GPUs can be used for general purpose computations, and therefore have a specific weigh much more important than before. Now, they are responsible of taking care of light, shadows, particle systems, and complex dynamical models more and more. In Borgo and Brodlie (2009) a beginner's introduction to GPUs, from both hardware and software point of view, is done.

3.2. New programming paradigms

Nowadays PC clusters are still being used for implementing multi-view training simulators. In the past, software developers must cope mainly with the problems derived from the multichannel synchronization (gen-lock & frame-lock issues) and the data distribution between the I/O devices (sensors, tracking), the dynamical model and the different nodes that were replicated in the visual scene graphs. Within this scheme, only a centralized node that executes a single dynamical model must cope with the dynamics properties of the scene, sending at visual loop rates the position of the different objects that move around the virtual word and the position of the observer. The rest of the nodes execute an instance of the scene graph, modifying it according to the data received from the dynamical models. Then every graphic card renders locally the images that were displayed on the screen.

Current architectures trend to build many cores for increasing program concurrency. This is not a problem since compilers can manage processor affinities very well and also operating systems provide the necessary services to cope with multi-thread applications. Additionally, programmable GPUs allow, also without any problem, the implementation of a dynamical scene graph whose nodes where directly controlled by shaders entirely executed in them. If everything runs in a single personal supercomputer with several graphics boards inside, still we don't have any problem with the renderization of our multiple views.

But however, if we get a cluster of GPUs connected, with everyone executing its own shaders, current software technologies are not ready to cope with the distributed pipelines configuration necessary to execute the shader in parallel, so new paradigms are still pending to come.

Some developments are already being done in what respect to general purpose computing in the GPU. For example OpenCL (Open Computing Language), described in http://www.khronos.org/opencl, aims are the design of an open standard for parallel programming of heterogeneous computational resources at processor level. More than just a programming language it includes an API, libraries and runtime system for software development.

The framework aspires at enabling portable and efficient access to general purpose parallel programming across CPUs, GPUs, Cell and ManyCores architectures for both HPC and commodity applications. The main key is to allow applications to use a host and one or more OpenCL devices as a single heterogeneous parallel computer system. Experienced programmers are supported throughout the process of developing general purpose algorithm without the necessity of mapping the algorithm onto architecture/platform specific features like 3D graphics API such as OpenGL or DirectX.

Also, other studies have been recently made, trying to uniform both the GPU and CPU calculations and the communications between them. The most relevant are the Zippy project (Fan, Qiu, and Kaufman 2008) and the CUDASA project (Strengert, Mller, Dachsbacher and Ertl 2008) that present a non-uniform memory access scheme between GPUS, the work done by Moweschell and Owens (2008), that implements a multi-GPU distributed shared memory architecture and DCGN (Stuart and Owens 2009), a message passing interface usable on systems with data-parallel processors. All of them ease the development and integration of parallel visualization, graphics, and computation modules on a heterogeneous cluster, but still lack of generalities to share a dynamic scene graph between the nodes of a GPU cluster.

3.3. Software for simulator developments

The core of a simulator is, at the end, a software application which computes the evolution of the simulated environment. There are a lot of libraries that allow developing such kind of applications; however it is difficult to choose which one better fits to our needs. Following a brief description is done of those that are the most active in the field of real-time simulation, divided in two: physic engines and scene graph management engines.

3.3.1. Physics Engines

Traditional physics engines are mainly based on solid-rigid simulations coupled with links. Their main drawbacks are that when developing the models, many simplifications have to be done in order to guarantee real-time and stability. Simulation aims, in this case, are limited to have a realistic behaviour of a few scene objects, the most important.

Nowadays, as viewed before, GPUs are in charge of executing the dynamical models, allowing the simulation of physically based visual effects, complex systems and the calculation of real-time collision detections. Nowadays everything is simulated, including plants movements, trees, atmospheric effects, particle systems, fire, fluids and deformable objects.

Programmers must choose one of the many physical engines available to develop the dynamical models of the scene objects. It is worth mentioning that if different engines are used for physics and graphics, special care must be taken when matching the dimensions and positions of the objects in the scene graph with those equivalents that the physical engine handles for calculating collisions. The whole information should be extracted from the 3D models and from the scene graph in order to avoid discrepancies between the calculated behaviour of the objects and what you see.

The most important physics engines that today are used in real-time simulations are the following:

ODE (http://www.ode.org): ODE is an open source library for simulating rigid body dynamics. It has advanced joint types and integrated collision detection with friction capabilities.

Bullet physics (http://www.bulletphysics.com): It is a professional open source collision detection, rigid body and soft body dynamics library. It is also integrated in MAYA and Blender3D.

Newton Dynamics (http://newtondynamics.com): It is an integrated solution for real time simulation of physics environments. The API provides scene management, collision detection and dynamic behaviour of objects.

Vortex (http://www.vxsim.com): It simulates the behaviour of vehicles, robotics, and heavy equipment in real-time synthetic environments for operator training and testing. It is integrated in OSG and VEGA.

PhysX (http://nvidia.com/object/physx_new.html): It delivers real-time, hyper-realistic physical and environmental gaming effects: explosions, reactive debris, realistic water, and lifelike character motion. Everything is computed in the NVidia GPU.

Havok FX (http://www.havok.com): It is a physic engine that runs entirely on the GPU and provides failure-free physic simulation using proprietary techniques for ensuring robustness, collision detection, dynamics and constraint solving. It provides integrated vehicle solutions and other tools available for simulating clothes, skeletons physics and rigid body destruction.

3.3.2. Scene graph management

Regarding scene graph management, there are many sdks dedicated to that. This topic is highly influenced by game technologies, where commercially available graphics engines use shaders executed in the GPU to increase performance. Mostly game engines also incorporate other features like sound, networking, artificial intelligence, collision, physics, etc.

Following are described the most important scene graph managers, game engines and cluster related applications that make possible the integration of a high performance real-time simulator.

OpenSceneGraph (http://openscenegraph.org): It is an open source high performance 3D graphics toolkit, used by application developers in fields such as visual simulation, games, virtual reality, scientific visualization and modelling. Written entirely in Standard C++ and OpenGL it runs on all Windows platforms, OSX, GNU/Linux, IRIX, Solaris, HP-Ux, AIX and FreeBSD operating systems.

OGRE (http://www.ogre3d.org): The Object-Oriented Graphics Rendering Engine is a scene-oriented, flexible 3D engine written in C++ designed to make it easier and more intuitive for developers to produce applications utilising hardware-accelerated 3D graphics. The class library abstracts all the details of using the underlying system libraries like Direct3D and OpenGL and provides an interface based on world objects and other intuitive classes.

Irrlicht Engine (http://irrlicht.sourceforge.net): It is a cross-platform high performance real-time 3D engine written in C++. It is a powerful high level API for creating complete 3D and 2D applications like games or scientific visualizations. It integrates all the state-of-the-art features for visual representation like dynamic shadows, particle systems, character animation, indoor and outdoor technology, and collision detection.

Delta3D (http://www.delta3d.org): It is a widely used and well-supported open source game and simulation engine. Delta3D is a fully-featured game engine appropriate for a wide variety of uses including training, education, visualization, and entertainment. Delta3D is unique because it offers features specifically suited to the Modelling and Simulation community such as High Level Architecture (HLA), After Action Review (AAR), large scale terrain support, and SCORM Learning Management System (LMS) integration.

Unreal Engine 3 (http://unrealtechnology.com): It is under the hood of the most visually intensive computer and video games on the market. It is available under license for PC, PlayStation3, and Wii. Its main features are: multi-threaded rendering, 64-bit high dynamic range rendering pipeline with gamma correction, dynamic composition and compilation of shaders, post-processing effects (ambient occlusion, motion blur, bloom, depth of field, tone mapping), artist-defined materials, dynamic fluid surfaces, soft body physics, deformable geometries, texture streaming system for maintaining constant memory usage, particle physics and skeletal animation.

Cry Engine 3 (http://www.crytek.com): It gives developers full control over their multi-platform creations in real-time. It features many improved efficiency tools to enable the fastest development of game environments and game-play available on PC, PlayStation® 3 and Xbox 360™. Its main characteristics are: road and river tools, vehicle creation, multi-core support, and multithreaded physics, deferred lighting, facial animation editor, dynamic pathfinding, rope physics, parametric skeletal animation and soft particle systems.

OpenSG (http://opensg.vrsource.org): It is a scene graph system to create real-time graphics programs, e.g. for virtual reality applications. It is developed following Open Source principles, LGPL licensed, and it can be used freely. It runs on Microsoft Windows, Linux, Solaris and Mac OS X and is based on OpenGL. Its main features are advanced multithreading and clustering support (with sort-first and sort-last rendering, amongst other techniques), although it is perfectly usable in a single-threaded single-system application as well.

VR Juggler (http://www.vrjuggler.org): It is a platform for virtual reality application development. This component allows a user to run an application on almost any VR system. VR Juggler acts as "glue" between all the other Juggler components. VR Juggler is scalable from simple desktop systems like PCs to complex multi-screen systems running on high-end work stations and super computers. Its development environment supports many VR configurations including desktop VR, HMD, CAVE™-like devices, and Powerwall™-like devices.

Equalizer (http://www.equalizergraphics.com): It is the standard middleware to create parallel OpenGL-based applications. It enables applications to benefit from multiple graphics cards, processors and computers to scale rendering performance, visual quality and display size. An Equalizer-based application runs unmodified on any visualization system, from a simple workstation to large scale graphics clusters, multi-GPU workstations and Virtual Reality installations.

4. BROWSER-BASED SIMULATION

Regarding games and simulation, one of the trends that emerged recently was the implementation of applications that are performed through a browser-based interface or even on hand-held mobile devices. The motivations were the necessity for extending the use of the simulators in order that they arrive to the maximum people. By means of the installation of an activex component, every simulator component (graphical models, dynamical models, textures and user interfaces) is downloaded automatically and the most important, is executed locally, without the necessity of installing anything, just the mentioned ocx. Additional advantages are that the software maintenance is done in a way completely transparent to the user (whenever you execute a new instance of the program) and the software protection can be performed using encrypted keys or restricting the clients by its internet address.

By the other side, the disadvantages that this methodology shows are that it is still very new, so there are still many software development kits quite immature, showing many instabilities and also poor performance when rendering the images. Furthermore, it is necessary to receive through the network connection the whole data that will be displayed, including textures and 3D models that can weigh hundreds of megabytes, often requiring a high bandwidth connection. Another aspect to take into account is that the server is a single point of failure, and can be also stressed if it receives many connections simultaneously. Security is another issue, since a failure when programming an interface can lead to intrusions in the computer system.

This kind of applications is used to implement cheap simulators that use a single computer with a single display, so the presence they provide is very low. They are similar to computer based training applications (CBT), but provide a real-time 3D interactive environment.

Anyhow, the browser-based game industry is currently very active and their technologies would be translated in a short period of time to the simulation market.

4.1. Browser-based simulation technologies

Following a brief description of the most important libraries and engines used in browser-based distributed applications is shown.

4.1.1. Java

Mostly browser-based applications use Java, from Sun Microsystems (http://java.com), as it is possible to integrate any Java application in the navigator through applets. However, as the Java virtual machine consumes a lot of cpu and many applications depend closely from the navigator they are being executed, they have relative small use in simulation applications. Anyway if it is only required a 2D interface, it is worth having a look to JavaFX (http://javafx.com), an open source application that allow making very impressive interfaces.

In order to eliminate the explorer dependencies, Sun introduced in 2001 a new technology called Java Web Start (JWS), allowing the execution of any Java application outside the browser. This is done specifying the compiled Java binary archives (Java ARchives) to be executed in the virtual machine inside a XML file called JNLP (Java Network Launch Protocol). Using this technology and whatever library that links Java applications with OpenGL it is possible to build a 3D simulator accessible by web. The client only downloads the JNLP file that is automatically executed, allowing the downloading of the simulator and if necessary also the Java engine required in a transparent way to the user.

In order to build the Java application that links with OpenGL, there are many libraries that can be used. However, the most important are:

JOGL (https://jogl.dev.java.net): The Java OGL project hosts the development version of the Java™ Binding for the OpenGL® API. It is designed to provide hardware-supported 3D graphics to applications written in Java, and integrates with the AWT and Swing widget sets allowing the implementation of GUIs.

LWJGL (http://www.lwjgl.org): The Lightweight Java Game Library is a solution aimed to enable commercial quality games to be written in Java. LWJGL provides developers access to high performance cross platform libraries such as OpenGL (Open Graphics Library) and OpenAL (Open Audio Library) allowing for state of the art 3D games and 3D sound. Additionally LWJGL provides access to controllers such as Gamepads, Steering wheel and Joysticks.

java3d (https://java3d.dev.java.net): The Java 3D API enables the creation of three-dimensional graphics applications and Internet-based 3D applets. It provides high-level constructs for creating and manipulation of scene graphs, 3D geometry and building the structures used in rendering that geometry. It is similar to the both described before, but it uses high level structures.

jME (http://www.jmonkeyengine.com): jMonkey Engine is a high performance scene graph based graphics API. Using an abstraction layer, it allows any rendering system to be plugged in. Currently, both LWJGL and JOGL are supported. jME also supports many high level effects, such as imposters (render to texture), environmental mapping, lens flare, tinting, particle systems, etc.

4.1.2. Adobe Flash

Flash (http://www.adobe.com) is commonly used to create animation, advertisements and various web page components, to integrate video into web pages, and more recently, to develop rich Internet applications. Flash can manipulate vector and raster graphics, and supports bidirectional streaming of audio and video. It contains a scripting language called ActionScript allowing the design of graphic user interfaces. In order to facilitate their design, the Adobe Flex SDK can be used. It comes with a set of user interface components including buttons, list boxes, trees, data grids, several text controls, charts, graphs and various layout containers.

Following, the flash based libraries and engines most commonly used for 3D graphics rendering through the navigator are briefly described:

PV3D (http://www.papervision3d.org): It is a technology that uses Flash and ActionScript allowing the renderization of 3D graphics through the web browser. It is very efficient, but still lacks the ability of using hardware accelerated rendering techniques (only hardware polygon rendering it is allowed). PapervisionX will be the next version of Papervision3D, built from the ground up based on Flash10's new 3D api, that will take full advantage of the 3D features of Flash Player 10.

Sandy3D (http://www.flashsandy.org): Sandy is an object oriented ActionScript 3D library, for programming 3D scenes for the Flash Player. The engine's capabilities are related to the performance of the virtual machine, which does not provide any native 3D nor hardware acceleration (although the Flash 10 player already introduces some 3D features). Sandy3D currently provides transparent materials, video textures, several shadow techniques and Collada, 3DS and ASE external formats are supported. However it lacks support for new materials with bump mapping and the use of multiple cameras.

Away3D (http://away3d.com): It is another real-time 3D engine for flash. It uses the hardware acceleration techniques included in Flash10 and supports object culling (using frustum calculations), interchangeable camera lenses to allow for different types of projection, collada bones animation, shadows, fog and optimized corrective z-sorting. It is being actively used by now and very well documented.

4.1.3. Adobe Shockwave

It allows publishing Adobe Director applications on the Internet and viewed in a web browser by anyone who has the Shockwave plug-in installed. The Shockwave file format is .dcr, and it is created by the Director authoring application. Features not replicated by Flash include a much faster rendering engine, including hardware-accelerated 3D, and support for various network protocols, including Internet Relay Chat. Due to its hardware rendering acceleration capabilities (OpenGL and directX), it is often used in online applications which require a very rich graphical environment. Online Learning tools which simulate real-world physics or involve significant graphing, charting, or calculation sometimes use Shockwave. It is worth seeing the many applications developed with Adobe Shockwave in http://www.shockwave3d.com

4.1.4. Other alternatives

Apart from the graphic engines related with Flash or Java reviewed before, there are still some other alternatives for building 3D browser applications that should be taken into consideration.

O3D (http://code.google.com/apis/o3d): It is an open-source web API for creating rich, interactive 3D applications in the browser. It is open source and it is developed by Google. It uses OpenGL or DirectX and depends strongly from JavaScript.

OSG4Web: (http://www.virtualrome.itabc.cnr.it): It provides a framework for in-browser OpenGL-based application wrapping. The framework allows the development of OpenGL and OpenSceneGraph based applications that open windows within the browsers, allowing JavaScript bidirectional interaction with surrounding page elements.

Unity3D (http://unity3d.com) Unity is a multiplatform game development tool, designed from the start to ease creation. Unity has a highly optimized graphics pipeline for both DirectX and OpenGL. It supports animated meshes, particle systems, advanced lighting, shadows, shaders, physics, sound and networking. The Unity Web Player enables you to view blazing 3D content created with Unity directly in your browser, and updates automatically as necessary.

Torque-3d (http://www.garagegames.com): This commercial application 3D game engine features multi-player network code, state of the art skeletal animation, seamless indoor/outdoor rendering engines, drag-and-drop GUI creation, a C-like scripting language and a built in world editor. A great thing with The Torque Game Engine is that you will receive all C++ source code to the engine, this makes it a lot easier to add any extra additions you might need for your game. Torque 3D supports all major browsers and operating systems, including IE7, FF3, OS X and Chrome. Games perform at 100% native speed, with no performance cost, completely in your browser.

5. SIMULATION CUTTING-EDGE RESEARCH

From the late 60's, training simulators have been used in many military and civil applications. Simulation technologies, therefore, are not new to the research community, however, as has been viewed in previous sections, new hardware advances and new user requirements have made that new programming paradigms be used when developing a simulator, both in the hardware and the software side.

Following are described some of the main problems that emerge when integrating a real-time training simulator and that still need to be improved. Also some important cutting-edge research lines are included.

- Simulation for training requires a way to know if the simulator reproduces accurately the real environment, and also if it provides a good instructional design making it suitable for training purposes. This is done by means of certification and homologation tasks. These are very well established in military applications but not yet in the civil world.
- Improvements in 3D sound are still to come. Auralization and real spatial sound systems (appropriate for caves) must be improved in current simulation systems. (See V. Pulkki 2002 and R. Furse 2009 and T. Lentz, D. Schröder, M. Vorländer and I. Assenmacher 2007)
- Risk prevention will be one of the main applications of future simulators. Avatars are already very well driven by realistic physical engines, but still we miss a collaborative behaviour to develop some risky tasks.
- New numerical methods for simulating multi-body systems will come, including multi-rate integrators (M. Arnold 2006), new distributed models (J Wang, Z. Ma and G. Hulbert 2003) and the use of level of details in physics (S.Redon, N.Galoppo, and M.Lin 2005).

6. CONCLUSIONS

Along this paper the most important technologies, both hardware and software, that are involved in the design of a real-time training simulator have been reviewed. Although the basic simulator building blocks are independent from the hardware used, the design of the software architecture will be determined by the number of displays that the simulator will provide to the user.

Serious games technologies and embedded display systems will lead the future high-end training simulation systems. Personal supercomputers are replacing mostly medium sized cluster-based simulators, however low cost web browser-based simulators will spread around, thanks to the new graphics engines able to render graphics using the GPU.

Physics engines already use parallel algorithms executed in a single or a cluster of GPUs, however new developments will have to come in order to solve the distribution of dynamic scene graphs driven by GPU computations into a set of several displays.

REFERENCES

M. Arnold. Multi-Rate Time Integration for Large Scale Multibody System Models. *IUTAM Symposium on Multiscale Problems in Multibody System Contacts*. p. 1-10. 2006

J. Bartlett. Programming high-performance applications on the Cell BE processor. Jan, 2007. https://www.ibm.com/developerworks/power/library/pa-linuxps3-1/

D. Blythe. The Direct3D 10 system. *ACM Transactions on Graphics*, vol. 25, nº 3, pp. 724-734, August 2006.

R. Borgo, K. Brodlie. State of the Art Report on GPU Visualization. School of Computing – University of Leeds. VizNET Report 2009.

C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon and J. C.Hart. The Cave Audio Visual Experience Automatic Virtual Environment, Communication of the ACM, vol. 35, no. 6, pp. 64-72, 1992.

Z. Fan, F. Qiu, and A. E. Kaufman. Zippy: A framework for computation and visualization on a GPU cluster. *Computer Graphics Forum*, 27(2), June 2008.

R. Furse. Building an OpenAL Implementation using Ambisonics. *AES 35th Int. Conference*, 2009, London, UK.

T. Lentz, D. Schröder, M. Vorländer and I. Assenmacher, Virtual Reality System with Integrated Sound Field Simulation and Reproduction, *EURASIP Journal on Advances in Signal Processing*, 2007.

E. Lindholm, J. Nickolls, S. Oberman, J. Montrym. NVIDIA Tesla: A Unified Graphics and Computing Architecture. *Micro, IEEE*, Vol. 28, No. 2. (2008), pp. 39-55.

A. Moerschell and J. Owens. Distributed texture memory in a multi-GPU environment. *Computer Graphics Forum*, 27(1):130–151, March 2008.

NVidia, 2008. Compute Unified Device Architecture Programming Guide Version 2.0, http://www.nvidia.com/object/cuda develop.html

V. Pulkki "Compensating displacement of amplitude-panned virtual sources." *Audio Engineering Society 22th Int. Conf. on Virtual, Synthetic and Entertainment Audio*. pp. 186-195. 2002 Espoo, Finland.

S. Redon, N. Galoppo, and M. Lin. 2005. Adaptive dynamics of articulated bodies. *In ACM SIGGRAPH 2005*. p. 936 - 945. 2005.

R. Rost, 2006 OpenGL Shading Language 2nd edn. Reading, MA: Addison-Wesley.

T. Sheridan (1992). Musings on telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments*, 1, 120–126. 1992

M. Slater, P. Khanna, J. Mortensen and Y. Insu. Visual Realism Enhances Realistic Response in an Immersive Virtual Environment. *Computer Graphics and Applications*, IEEE Volume 29, Issue 3, May-June 2009. Page(s):76 – 84.

M. Strengert, C. Mller, C. Dachsbacher, and T. Ertl. CUDASA: Compute Unified Device and Systems Architecture. In *Eurographics Symposium on Parallel Graphics and Visualization (EGPGV08)*, pages 49–56, 2008.

J. Stuart, J. Owens. Message Passing on Data-Parallel Architectures. *Proceedings of the 23rd IEEE International Parallel and Distributed Processing Symposium*, pp.1-12, May 2009.

S. St-Laurent, 2005 The Complete Effect and HLSL Guide. Redmond, WA: Paradoxal Press

S. Vangal, J. Howard, G. Ruhl, S. Dighe, H. Wilson, J. Tschanz, D. Finan, P. Iyer, A. Singh, A. Singh, T. Jacob, A10, S. Jain, A11, S. Venkataraman, A12, Y. Hoskote, A13, N. Borkar. “An 80-Tile 1.28TFLOPS Network-on-Chip in 65nm CMOS”. *Digest of Technical Papers. IEEE International In Solid-State Circuits Conference*, 2007. pp. 98-589.

J Wang, Z. Ma and G. Hulbert. A Gluing Algorithm for Distributed Simulation of Multibody Systems. *Nonlinear Dynamics*, 34 (1-2). p. 159-188. 2003

AUTHORS BIOGRAPHY

Rafael J. Martínez is an assistant professor at the University of Valencia in the Computer Architecture knowledge area. He received a PhD on Computer Engineering from the University of Valencia in 1997. He teaches Operating Systems and Fault Tolerant Computing. He belongs to the Robotics Institute since eighteen years ago and is the head of the Simulation & Modelling Laboratory (LSyM). His main interest research topics are computer graphics and simulation for training. He has experience both in management, research activities and consulting. He has coordinated and participated in more than 20 research projects, mostly of them related with simulation activities and technology transfer to enterprises. He is the main inventor of two patents related with harbour crane simulators and has published more than 40 papers in several congresses and research periodicals.

A NOVEL HYBRID APPROACH TO SEMANTIC INTEROPERABILITY FOR LOGISTIC APPLICATIONS

Alessandro Pierpaolo Beneduce[(a)], Mosè Gallo[(a)], Vincenzo Moscato[(a)], Antonio Picariello[(a)], Lucio Sansone[(a)]

[(a)]University of Naples "Federico II"

[(a)]a.bendecuce@studenti.unina.it, [(a)]mose.gallo@unina.it, [(a)]vmoscato@unina.it, [(a)]picus@unina.it, [(a)]sansone@unina.it

ABSTRACT

As manufacturing and commerce become ever more global, companies are dependent increasingly upon the efficient and effective sharing of information with their partners, wherever they may be. In this paper, we propose an architecture that, exploiting the Semantic Web technologies, has the objective of allowing semantic interoperability among software agents for logistic applications that preserves, not only the semantic of transmitted messages, but also the subjectivity of agent's world vision in the communication. Such an architecture takes advantage by a particular communication model called Semantic Triangle in which communication agents share the referents (real world objects) and not the references (mental image or impression of a real object from the sender agents point of view) ,thus ensuring an effective semantic interoperability in the information exchange process.

Keywords: semantic web, semantic interoperability

1. INTRODUCTION

In the Internet era, large-scale computer networks and the pervasive World Wide Web infrastructure have largely solved the problem of providing ubiquitous access to any kind of information. However, while information can now be easily retrieved and accessed, the problem of processing and interpreting its meaning by automatic approaches has not yet been solved adequately and it remains an important research topic.

In such a context. as manufacturing and commerce become ever more global, companies are dependent increasingly upon the efficient and effective sharing of information with their partners, wherever they may be. Leading manufacturers perform this sharing with computers, which must therefore have the required software to encode and decode the associated electronic transmissions. Because no single company can dictate that all its partners use the same software, standards for how the information is represented become critical for error-free transmission and translation. The term *Semantic Interoperability* is frequently used to refer to this error-free transmission and translation.

We could image a *virtual supermarket* logistic scenario in which a customer (e.g. web or 3D application) asks one or more providers for a generic product with specific features. It is necessary for provider information systems (e.g. accessible by web services or REST services) understand in the correct way customer request and respond by a list of products that satisfy user needs, thus ensuring an effective semantic interoperability in the information exchange process.

Generally speaking, Semantic Interoperability is referred to the ability of two or more computer systems of exchanging information and having the meaning of sent information automatically and correctly interpreted by the receiving system. Semantic interoperability can be distinguished from other forms of interoperability (such as functional interoperability between computational units that requires the interface compatibility of connected units) because it requires that the information exchange preserves the semantic of the original message.

As a consequence, a transferred message must include, in its expressive form, all the information required by the receiving system to interpret its meaning correctly whatever the algorithms used by the receiving systems (which may be unknown to the sending system).

Presently, the important challenge of semantic interoperability among systems cannot be completely met, since current document management applications have limited capabilities for structuring and interpreting documents.

With the advent of the *Semantic Web* (Berners-Lee 2001), approaches exploiting metadata or based on semantic annotations from shared ontologies (Gruber, 1993) provide the mostly used solution to the problem of extracting some kind of knowledge from documents. The result is a number of web resources with machine interpretable mark-up that can be easily managed by software agents.

The introduction of a shared ontology solves the problem of a unique interpretation of the meaning, but it still does not keep the "subjectivity" of data sources in the communication process. Moreover, the use of shared ontologies is affected by another fundamental problem, i.e. their *maintenance*: maintaining an ontology requires trusted centralized entities that periodically up-to-date concepts from the ontology.

Such a problem has limited the use of ontologies to restricted and well-defined application domains. A further problem with software infrastructures based on shared ontologies is the scalability one that is a well-known open issue of the Semantic Web.

This paper presents a novel communication approach and a possible implementation of it that, exploiting the Semantic Web technologies, allows semantic interoperability among software agents in the Web, preserving not only the semantic but also the subjectivity of the agent's world vision in the communication. The approach is inspired to the *Semantic Triangle* communication model presented by Odgen and Richards (Odgen and Richards 1989), where communication agents share the *referents* (real world objects) and not the *references* (mental image or impression about a real object from the sender agent point of view), thus ensuring an effective semantic interoperability in the information exchange process.

For evaluating strengths and weaknesses of the proposed communication model, an instantiation of the communication process based on *semantic machines* deployed on a distributed software architecture has been developed and a case study for a logistic application has been reported

2. THE SEMANTIC TRIANGLE MODEL

Odgen and Richards proposed a model of communication between agents that characterizes each message by three distinct entities, i.e., mental thought (*reference*), sign (*symbol*), and real object (*referent*).

A reference and a symbol are linked by a symbolization relationship, reference and referent are linked by a reference relationship, while symbol and referent are not linked by any direct relationship. The relationship between symbol and referent is mediated by the subjective mind of the person who encodes or decodes the message, and thus this relationship is variable and subjective. These relationships were modeled by Odgen and Richards by the so-called *Semantic Triangle Model*.

For better understanding the introduced concepts at the base of our work, the current section will be concluded with a simple example explaining the fundamentals of communication in the semantic triangle model.

Let us suppose that Bob asks his friend John (that has a electronic devices shop) if he still has the mouse that yesterday he sold to Maria, by composing the message: "Do you have a mouse?". Because the "mouse" word (symbol) has two meanings related to two different classes of real objects (referents) - i.e. computer device or animal -, John may not understand the exact sense of the word.

Moreover, even if John understands the right sense of "mouse", he has to remember that Bob does not refer to a generic mouse, but to the particular one that yesterday he sold to Maria. Thus, the correct communication outcome depends on the image that the "mouse" word creates in John's mind (reference). In order to preserve the semantic, Bob has to attach useful information (reference) to the message - e.g. by composing the message "Do you have the mouse that yesterday you sold to Maria?" - that permits John to understand what is the real object (referent) that Bob asked him, although John has a different personal concept of "mouse". In fact, the reference "that yesterday you sold to Maria" can be seen as a sort of pointer to a local but sharable concept of Bob's memory.

3. APPROACHES FOR MANAGING THE SEMANTIC INTEROPERABILITY

Selvage et al. outline in (Selvage 2006) different architectural patterns for achieving semantic interoperability in distributed environments and place at the opposite extremes of their list the *point-to-point semantic integration pattern* and the *Semantic Web* pattern.

In the point-to-point semantic interoperability pattern, the communication is based on messages which directly embed a complete description of their semantic. This aspect implies that the message minimum semantic unit includes both symbols and descriptions of referents, which instances could be maintained in sender local knowledge sources (e.g. ontologies).

This approach represents a *subjective* communication approach where the subjective vision of the sender knowledge is preserved, but information redundancy and incoherency problems may arise.

Vice-versa in the Semantic Web, based on the *Semantic Web Layer Cake* (Berners-Lee 2001)model, communication relies on the use of ontologies that are conceptualizations of specific domains in order to formalize semantic of data. The Semantic Web links and relates elements of a message to a common ontology, using the *Resource Description Framework* and the *Web Ontology Language* that allow data to be shared and reused on the Web. A symbol will be directly linked to its referent by means of an ontology which provides the correct semantic of real world objects, thus solving the problem of message ambiguity.

This kind of communication approach can be classified as an *objectivistic* one, where the knowledge (that is formalized by the ontology) is independent on the agents involved in the communication. This model obviously simplifies the communication problem and the implementation of systems based on such an approach, but some new problems: *ontology acceptance*, *ontology building* and *maintenance* and *ontology expressiveness* (Devedzic 2001, Harris 2004, Gangemi 2005).

In order to cope with weaknesses of both communication approaches, it is possible to propose a sort of *hybrid model* that exploits the interoperability mechanisms of both the subjective and objective model. In particular, in the hybrid model each sender agent has its own subjective knowledge that may be either mapped into shared objective knowledge sources (such as an ontology), or directly included in a coded message in order to preserve its personal interpretation of transmitted concept, coherently with Odgen and Richards model.

4. AN IMPLEMENTATION OF THE HYBRID MODEL

Internet and the Semantic Web offer the necessary infrastructure for implementing the hybrid communication model presented so far, and obtaining a semantic interoperability among software agents in the Web (Hendler 2001). In this section, the software requirements of a possible implementation of the communication model will be presented. In particular, the characteristics of transmitted messages and the communication process involving them in the hybrid communication model will be presented.

4.1. The Message Conceptual Model

Messages exchanged in the hybrid communication model are considered as the aggregation (*information objects*) of *digital assets* (e.g., images, textual documents, audio, etc...) containing a set of *information concepts* that constitute the message semantic content that an agent is interested to transmit; these concepts refer to the agent knowledge that can be mapped either into a local ontology, or a universal ontology that is a-priori accepted by all the agents involved in the communications.

The information concepts can be classified in two distinct types: *entities* and *facts*. An entity is a noun, verb or other part of a speech that can be retrieved on any language dictionary. A fact is an expression corresponding to peoples, places, events or any other thing that could not be retrieved in a language dictionary. Facts can further be classified into *Encyclopedic* and *Non-Encyclopedic* facts.

Encyclopedic facts have a general relevance such that they could be contained in a general encyclopedia or in a domain encyclopedia. Referents to encyclopedic facts can be called *Universal Fact Referents*, since they have to be universally shared and accepted by any communication agent (or at least, they have to be shared by communication agents belonging to a specific domain).

Non-Encyclopedic facts are relevant in the internal world of the agent but could be not universally relevant or there could not be a universally accepted semantic. Referents for non-encyclopedic facts are named *Local Fact Referents* and are maintained in the Semantic Machine of the agent.

To make explicit the binding of information concepts with real word objects referents, we propose to use *semantic tags* for providing a symbolic representation of information concepts, and references for binding together semantic tags and related referents

4.2. The Communication Process

The proposed communication process is supposed to be decomposed into four sequential activities:

1. *Information encoding* - a sender agent composes a message in a particular format that is understandable by other semantic machines (in this step the binding between information concepts and referents is performed using semantic tags and the reference mechanism;
2. *Information transmission* -sender agent publishes the message (e.g. by web pages) on the web, or transmits the message to a receiver agent (e.g. by e-mail);
3. *Information acquisition* - a receiver agent performs the message acquisition (download);
4. *Information decoding* - receiver agent binds the received information with the related referent by means of sender references.

With respect to the example discussed in section 2, the communication could be outlined as follows:

1. Bob uses an e-mail client (semantic machine) to compose the message “Do you have a mouse?” that is represented in a particular format (e.g. XML) and in which the personal concept of a “mouse” is bound to the related referent by the tag “mouse” and a reference (e.g. an hypertextual link pointing to a web page deployed in a web server containing the description of the mouse);
2. the e-mail is sent to the pop server of receiver;
3. the email is downloaded by John;
4. the semantic machine of John decodes the message and binds the concept of “mouse’ to the corresponding referent using Bob’s reference; after the communication, John may decide to add it to his local ontology for ordering a new mouse of the same kind to his provider.

The Message Sending and Receiving activities are specified by the UML Activity Diagrams reported in Figures 1 and 2. In particular, Sending a message includes the *Tag Extraction* and *Reference Binding* sequential activities (implementing the message encoding), besides the *Encoded Message Transmission* one.

The *Tag Extraction* activity performs the retrieval of information concepts from a message to be transmitted and their symbolic representation by semantic tags. The set of these semantic tags can be indicated directly by the message author or it can be retrieved automatically by tag/information extraction algorithms.

For each tag, the *Reference Binding* activity is performed in order to retrieve from the local ontology a reference to the information concept represented by the semantic tag. More precisely, for each tag the *Local Ontology Mapping* activity is executed, in order to search in the local ontology for possible references associated with the tag.

If more than one possible reference is found, then a *Local Disambiguation* activity is carried out, where a single reference correctly representing the tag has to be selected. After the eventual disambiguation, two cases may happen:

1. the reference correctly representing the tag has been found in the local ontology (*Mapping Hit*) and it is possible to bind the message tag with this reference (*Bind Tag with Reference* activity);
2. the tag has not a corresponding reference in the local ontology (*Mapping Miss*), thus, a *Referent Search* activity is entered, where a referent reporting the semantic of the information concept must be found in an available Referent Source, and therefore a corresponding reference (*Create Reference*) has to be built.

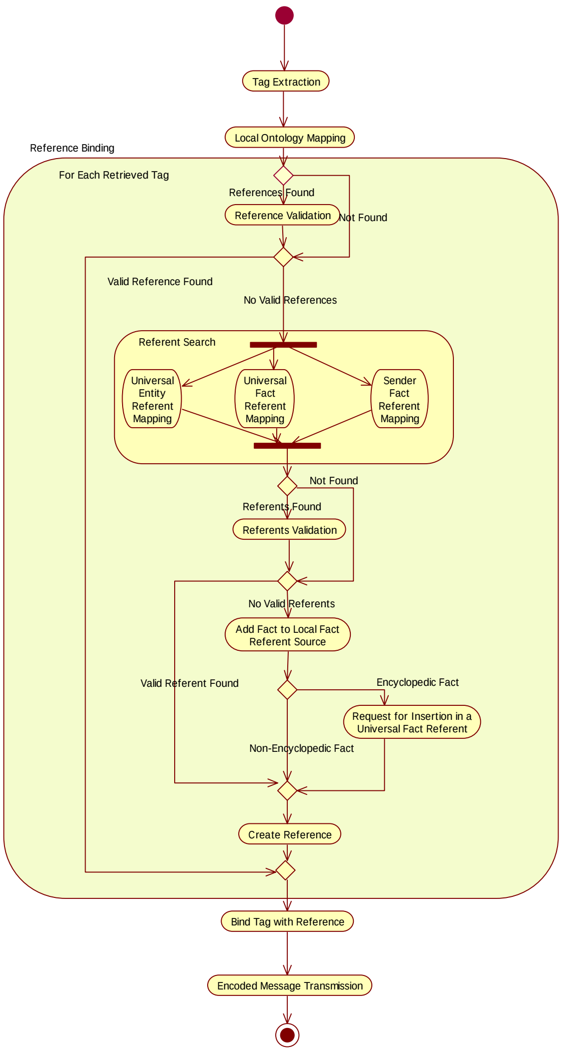


Figure 1: The Message Sending activity diagram

In the Referent Search activity, different sources can be queried depending on the type of information concept to be represented and on the available Referent Sources.

As an example, Entity Referents can be retrieved in Universal Entity Referent Sources, such as *Wordnet* or other on-line dictionaries. Encyclopedic relevant facts can be retrieved in Universal Fact Referent Sources, such as *Wikipedia*. Encyclopedic relevant facts related to a specific domain can be retrieved in sources containing domain ontologies. Non-Encyclopedic relevant facts, i.e. facts that are relevant only in the sender context, can be retrieved just in a Fact Referent Source published by the sender itself.

Facts for which the sender declares a semantic that is different, or more specific, from the one proposed by a universal fact source, can also be retrieved in the Local Fact Referent source of the sender.

In our architecture, each universal referent is a directly addressable resource. It can be accessed either by a HTTP GET service request, or by a web service that performs the wrapping of the related referent source [6]. For accessing local referents (also coded in XML/RDF) a possible solution is offered by the REST technology. In this case, a referent can be directly stored into the Local Fact Referent Source of the semantic machine via a HTTP PUT or POST request, and directly accessed via a HTTP GET request.

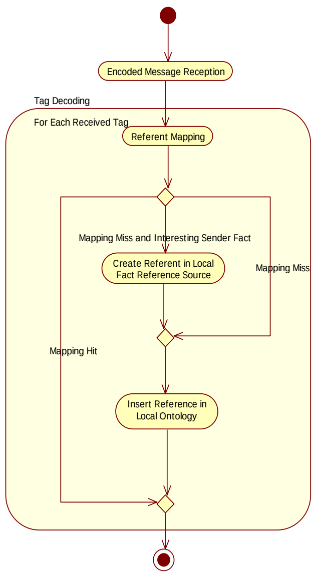


Figure 2: The Message Receiving activity diagram

At the end of the Referent Search activity, it is possible that more than one possible referent has been found. In this case, a *Validation* activity is needed, in order to state if any of the referents provides a satisfying semantic for the information concept, or to select the more suitable one, if more than one referent was found. If a satisfying referent has been found, then the *Create Reference* and *Bind Tag* activities are performed, and a new reference to the chosen referent is added to the local ontology and will be bound with the tag.

Elsewhere, a new referent has to be inserted in the Local Fact Referent source. Eventually, a *Request for*

Insertion activity for adding the new created referent to the Universal Fact Referent Source can be performed. This possibility is granted by some sources, such as Wikipedia, but the referent will be inserted only if it will be accepted by the referent source managers, and only in an asynchronous way.

At the end of the Reference Binding activity, the message minimum semantic unit will be composed by a set of couples *<tag, reference>*.

In our architecture it is considered as a web resource characterized by an URI with naming and addressing functionalities, and will be encoded by XML/RDF languages. Thus, using its URI, each information object can be accessed by HTTP protocol via a GET request, and the result of such a request will be an XML/RDF representation that separates the physical view from the resource view of contents, thus reflecting the semantic web layer cake schematization, and providing interoperability with applications of the Semantic Web. As to the receiving activity, its purpose is to map each received couple *<tag, reference>* into any concept from the receiver agent local ontology, in order to reconstruct the correct semantic (Referent Mapping Activity). If a reference is found in the local ontology (Mapping Hit), the information concept is known to the receiver and it must not be decoded. Otherwise, if no reference is found, then a new reference has to be created and inserted in the local ontology. Moreover, if a received concept points out from the sender fact referent source, then the receiver agent can choose to accept the referent and to add it to its Local Referent Fact source.

4.3. The Semantic Machine

The described communication process can be implemented in distributed environment including several *Semantic Agents*, each one equipped with a *Semantic Machine* that executes message sending and receiving activities. The realization of the Semantic Machine will require a number of design decisions, and will depend on several technological choices. A reference software architecture of the Semantic Machine is outlined by the UML Deployment Diagram reported in Figure 3, and a description of its five logical components is provided in the following.

Encoder is the Semantic Machine component responsible for implementing the information encoding and transmission activities. It asks the *Ontology Manager* component for retrieving/inserting references in the local ontology and asks the *Referent Manager* component for referents in the available Entity or Fact sources, and for inserting referents into the Local Fact source.

Decoder is the component responsible for the information acquisition and decoding activities. It queries the Ontology Manager to retrieve/insert references in the local ontology and asks the Referent Manager for inserting new referents into the local Fact source.

Tag Extractor is the component responsible for the automatic or human-assisted extraction of tags representing the relevant semantic concepts of the information object.. The extracted tag list is returned to the Encoder component for the elaboration.

Ontology Manager is the component responsible for managing the local ontology, offering reference searching and reference inserting services. The implementation of the Ontology Manager will depend on the technologies and languages adopted for the realization of the Local Ontology. Possible languages can be RDF or OWL , while apposite API provided by open-source tools (e.g. JENA) can be exploited to create, manage and query (via SPARQL) the set of references.

Referent Manager is the component responsible for managing the referent sources, offering referent searching and referent inserting services. Of course, the implementation of such a Referent Manager will depend on the technologies used to encode the referent sources.

Using the Internet infrastructure, the ***Universal Referent Sources*** can be provided by web sites or any other data source exposed on the Web. Some sources (e.g. *eBay* or *Wordnet*) offer APIs for providing direct access to the referents, while other sources do not provide them, and thus wrapping techniques (Canfora 20008) are needed to access referents. The ***Local Fact Referent source*** is deployed in the semantic agent and it can be developed by using local database or XML documents.

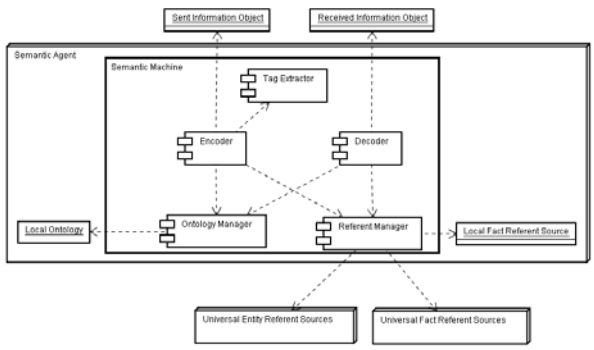


Figure 3: UML Deployment Diagram of the Semantic Machine

5. A CASE STUDY

In this section, the implementation of a semantic machine that supports the execution of the hybrid communication process in a specific context will be presented. We use as case study of the proposed semantic interoperability model a virtual supermarket scenario, where costumers (sending agents) during a purchase of a movie DVD can transmit on the web a message related to the desired product, while providers (receiver agents) give an interpretation to messages and respond (sender agents) with the description of a product that try to match user needs. Costumers (receiver agents), eventually, can decide to buy a given product.

For concepts related to the movie domain, it is a commonly accepted opinion that the *imdb.com* database (accessible via the *www.imdb.com* website) is the larger and reliable source of information available on the Web. Thus, in the case study imdb.com has been selected as a referent source for facts such as movies, directors, actors, cinema events and so on. For all the other facts, i.e. facts that are encyclopedically relevant but not directly related to the movie domain, the English version of the *Wikipedia* web-site has been selected as referent source. Since most of the musical concepts are also retrievable on Wikipedia, a disambiguation activity was needed in some cases.

As to the entities, the *Wordnet* repository has been considered as the unique Referent Source for entities in English language. Finally, for all the facts that are not encyclopedically relevant (i.e. for which no valid Referents can be retrieved in the considered Referent Sources) or for which the referent source is considered not reliable or not detailed enough, a Local Fact Referent Source deployed in the semantic agent was considered.

5.1. Working Example Execution

We assume that a costumer semantic agent wants to send to vendor semantic agents the digital asset consisting of the following text fragment:

"*I would like to see an action/adventure movie. I like gothic atmosphere with vampire and werewolves. I love the Dracula novel and I like actors as Keanu Reeves, Monica Bellucci and Winona Ryder and Anthony Hopkins and directors as Francis Ford Coppola* ".

The editor provides the following set of tags related to the digital asset: 'action movie' - 'adventure movie' - 'vampire' - 'werewolf' - 'gothic' - 'Dracula' - 'novel' - 'Keanu Reeves' - 'Monica Bellucci' - 'Winona Ryder' - 'Anthony Hopkins' – 'Francis Ford Coppola'.

During the encoding process, corresponding referents for each tag have been searched in the known referent sources, obtaining the following results:

1. added to the local fact source: nothing ;
2. retrieved on wordnet and wikipedia: 'novel', 'Dracula', 'vampire', 'werewolf';
3. retrieved on wikipedia: 'action movie', 'adventure movie';
4. retrieved on wikipedia and imdb.com: 'Keanu Reeves', 'Monica Bellucci', 'Winona Ryder', 'Anthony Hopkins' , 'Francis Ford Coppola'.

The result of the encoding process was a XML/RDF document reporting the list of tags with the corresponding URIs of chosen referents that vendors have to decode to respond with their products.

6. CONCLUSIONS

In this paper, we proposed an architecture that allows semantic interoperability among software agents for logistic applications. Future works will be devoted to implement wrapping modules for other referents sources for and realize a complete experimentation in order to obtain significant results

REFERENCES

Berners-Lee, T., et al., 2001. The Semantic Web, *Scientific American*, 284:34-43.

Canfora, G., Fasolino, A.R., Frattolillo, G., Tramontana, P., 2008. A wrapping approach for migrating legacy system interactive functionalities to service oriented architectures, *Journal of Systems and Software*, 81(4):463-480.

Devedzic, V., 2002. Understanding ontological development, *Communications of the ACM*, 45:136-144.

Gangemi, A., 2005. Ontology design patterns for semantic web content, *In ISWC 2005 (LNCS 3729)*, pag. 262-276.

Gruber, T., 1993. A translation approach to portable ontology specifications, *Knowledge Acquisition*, 5(2):199-220.

Harris, J., 2004. Judging the likely success of an ontology, *Technical report*

Hendler, J., 2001. Agents and the semantic web, *IEEE Intelligent Systems*, pages 30-37.

Ogden C.K., Richards, I.A., 1989. The meaning of meaning, *Ed. Harcourt Brace Jovanovich*.

Selvage., M., et al., 2006. Achieve semantic interoperability in a SOA - patterns and best practices, *Technical IBM report*.

AUTHORS BIOGRAPHY

Alessandro Pierpaolo Beneduce received his Bachelor degree in Logistic Management Engineering from University of Naples "Federico II". Actually, he is collaborating with the Computer Science Department for the development of 3D logistic and e-commerce applications.

Mosè Gallo received his Master Degree in Mechanics Engineering and the Ph.D. degree in Logistics from University of Naples "Federico II". His research activity concerns logistic modeling and simulation.

Vincenzo Moscato is san assistant professor of Database at University of Naples "Federico II". Actually, his main research interests are Multimedia Database and Semantic Interoperability.

Antonio Picariello is san associate professor of Database at University of Naples "Federico II". His research interest lie in the fields of Multimedia Database and Knowledge Management.

Lucio Sansone is a full professor of Information Systems at University of Naples "Federico II". His research interest lie in the fields of Multimedia Information Systems and Business Process Reengineering.

A SIMULATION DECISION SUPPORT MODEL FOR CHECK-IN DESKS IN AN AIRPORT

Murino T.[(a)], Romano E.[(b)], Santillo L. C.[(c)]

[(a)]Researcher (DIMP University of Naples "Federico II")
[(b)]PhD (DIMP University of Naples "Federico II")
[(c)]Full professor in Operations Management (DIMP University of Naples "Federico II")

[(a)]murino@unina.it, [(b)]elromano@unina.it, [(c)]santillo@unina.it

ABSTRACT

The goal of this study is to produce a simulation decision support model for check-in desks in the airport. In particular, this study develops a simulation model can be used to build a structure that helps predict dalay and to produce a logical and rational management of check-in and security checkpoint inside the terminal airport.
The model was tested in the medium size italian airport (Naples Capodichino Airport).

Keywords: airport terminal operation, passenger flow, simulation, decision support system, discrete event simulation.

1. INTRODUCTION

Although probably from the outside it cannot be realized, actually it is not easy to understand and to make people to realize the airport operations, since formed by a set of complex infrastructure and facilities and equipments.
Each airport is to be:
- An ideal point of interchange among different means of inland transport (such as road and rail system) and air system;
- A services' centre, because it must provide the necessary services for ticketing, boarding, documentation, control of passengers and goods;
- A point of connection between arrivals and departures, since it must ensure an easy transit from a system of continuous arrivals (which is the users flow) to a system of discrete departures (which are the scheduled flights).

The overall experience of a passenger at an airport can be demanding and time consuming. Delays occur with parking, checking in, security screening, and boarding. The less time the costumer spends in the systems, the higher the satisfaction. However, at the same time, the airport is obliged to hold standards that the passengers must meet. These standards include proper identification, limited luggage weight, and safety procedures at the security checkpoint.

The airport terminal for passengers represents the connecting interface with the air-side area, the place where the take-off and landing of aircraft take place, enabling the transfer of the aircrafts from the land transport system of land transport (rail or truck) to the aircraft one.

Most of the airports, especially those of large size, shows complex traffic and congestion problems, both as regards the runways and parking areas (airside), both the airport where the public has access (landside).

The characteristics that affect the airport landside capacity refer to the following structures:

- number of check-in desks;
- number of gates;
- baggage handling system;
- Police filters;
- departing, transit and arriving passengers flow systems.

All terminal operations (ticketing, check-in, passport control, security control) are obviously characterized besides a service time and a queue waiting time too. This last time cannot be too high, but so far we need to restrain it in order to ensure all passengers having regular ticket the observance of embarkation time thus avoiding the flight missing. In most the reduced times are for travelers synonymous of high quality, key element in developing a good business image. For both reasons, therefore, is crucial for an airport during the years the infrastructure adaptation due to the increasing demand passengers transportation.

2. BOARDING OPERATIONS

The main planned phases before the flight for departing passengers essentially are three:

1) The check-in operations at check-in desks, where the passenger delivers the baggage and receives the boarding pass;

2) The transit at the security controls in order to proceed to the boarding halls with their hand baggage;

3) The boarding operations on the aircraft: the passenger must be show at the boarding gate of his flight with a boarding pass and a valid identity document.

Generally, passengers must have completed the check-in operations:

- For domestic flights at least 35 minutes before departure;

- For international flights to Europe at least 45 minutes before the flight;

- For intercontinental flights at least 60 minutes before the flight.

Following the check-in the passenger, with boarding passes, valid identity document and hand luggage (weight and dimensions allowed by the company), passes through the expatriation controls and the metal detector where security operators perform a manual control on the person and on the hand luggage ascertaining the absence of improper weapons, real or other dangerous objects.

3. STATE OF ART

As mentioned earlier in the transport sector coexist two different concepts, that of the passenger who expect a customized and quick service and the one of that is suitable to a standard organization. As matter of fact the airline companies are highly geared to offer prompt and appealing service: booking via Internet, online payment, e-ticket services, check-in online. Instead the passenger badly tolerate the contrast between the simplicity that shines from the Internet and the hard reality that he must face on his arrival at the airport: several controls, slow registration procedures, queues to board at the gates.

In this section we show some of these works. The Nagoya University has conducted a simulation using the software Arena on departing flow passengers from the International Kansai airport in Japan, in order to reduce the number of passengers, because of long waiting times in peak periods and because of unavoidable delays, they lose their flights. Preliminary analysis on passengers waiting times showed that the total time spent by passengers in the airport: the 48% is spent moving from place to place within the terminal, the 25% is waiting and only the 4% is doing formalities such as process acceptance, embarkation, and so forth.
In addition, it has been found that the time spent at the check-in desks in waiting queue is more than 80% of total time before embarkation. This output highlights that the check-in should be considered as the main bottleneck.

The simulation was conducted considering the condition of congestion based on flight operations of a regular working day, usually adjusted on a number of 100 flights and boarding operations of 70%. In particular, it refers to the airline's "A Company" which operates about 25% of all flights of the Japanese airport. The results of the simulation suggests that the number of passengers loosing their flights can be drastically reduced by the addition of a staff supporting the standard working group, and by the use of check-in desks different for passengers class, such as tourists, business and first class. This is a solution already adopted by many carriers both Italian and International.

The departing passengers flow at the Buffalo International Airport (Niagara) has been studied by researchers from the Department of Industrial Engineering and Systems at the University of Buffalo, Particular attention has been shown to the check-in process, considering that the waiting system times besides varying depending on week day and the time of the day, are function of different check-in available to passengers, the number of bags and the chosen airline to fly.

Interesting is also the work proposed by the School of Mathematics and Statistics Carleton University (Ontario, Canada), where a linear programming model minimizing the total work hours at the check-in ensuring a satisfactory customer service level has been developed. The output of this alternative method shows a significant performance improvement since it provides a shorter queues length, reduced waiting times and an increase in satisfied customers percentage.

4. DATA COLLECTION

This study developed a simulation model able of managing in a logical and rational mode all check-in desks and security control inside the airport.

The importance of this model lies in the ability to identify, at any time of the year and depending on the airport traffic volume, the required number of check-in required and at the same time adequate to perform the check-in operations in full compliance and regularity of the scheduled flights ensuring a satisfactory service level for departing travelers. The close correlation between the check-in operations and the security controls allows to obtain in the same way the number of the security control accesses to be made operational during the referred period.

Building the model is equivalent to virtually retrace the journey made by departing passengers studying the issues and highlighting all the available alternatives. Essential to this development is the visit at the terminal, useful to understand in general the passengers approach and both in order to have all information needed for the continuation of this work. To fully clarify the issue we must know the number of check-in and the available security control checkpoint: these number are 56 and 12.

Obviously at the changing flight destination varies the check-in time at desks which is usually an interval of 90 minutes for international flights and approximately one hour for domestic one. While as regards the security passengers control of any airline company can be served at any desk regardless flight destination, related to the check-in process is essential to examine how the airport operator assigns the desks to airline companies or handler. The airlines that operate more flights simultaneously adopt the solution of the common check-in (as the case of AIRONE company in our study), in which passengers on same company flights may be served on any desk regardless destination. Also the management queue can be different: in Naples airport, e.g., Airone manages several desks through a single queue, while other companies like Air France arranges passengers to many queues as much available desks.

This clarification is crucial for an easier understanding of the logical model described in the next paragraph.

Today's airport terminals offer to tourists real business galleries, thus finished the check-in phase passengers are increasingly tempted to delay the next operation of security control since they prefer last shopping before departure or at least visiting other shops, always consistent with the remaining time for boarding.

4.1. Naples Airport Terminal

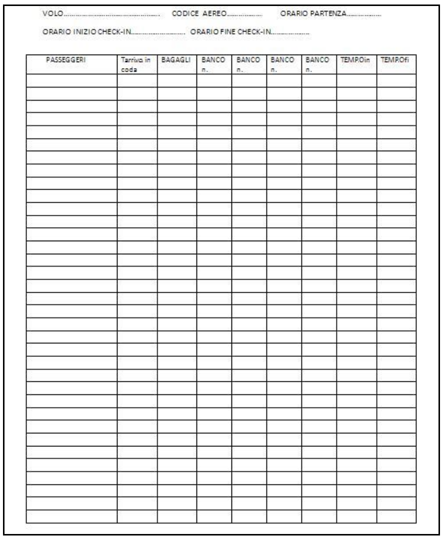
The landside of the Naples International airport consists of two terminals, Terminal 1 (T1) and Terminal 2 (T2). As the Terminal 2 is operating only in spring-summer season, and almost exclusively for charter flights, focus our study on passenger traffic on the only Terminal 1. The path related to passengers departing from Terminal 1 consists of two floors: on the ground floor we have the check-in desks, and on the first floor, accessible by escalators, the security control checkpoint and some shops and food services are located. On the other side of the ground floor besides the path related to arriving passengers there are the ticket office, the two currency exchange offices, several car rentals, and so forth.

Figure 1: Map of Terminal 1 at the Naples airport

4.2. Data Analisys

The most problematic phase of this study has been with no doubt that on data acquisition. Sampling took place in the passenger terminal 1 at Naples International airport. After a thorough inspection of Capodichino airport it came out that passenger traffic peaks during the whole week were on Monday morning and Friday afternoon. Therefore the focus data collection phase was on three week days: Monday, Wednesday and Friday. The survey was repeated for several weeks by 2 people with different recognition tasks: the first one monitored near the check-in the passengers arrival at the counter to get the determination of so-called 'curves of presentation' and the service time of the check-in formalities, instead the second one has carefully considered the queue near the desks by defining the queue length and service time.

The spreadsheets show on the header some information such as basic information flight (the airline and destination), time of departure, time of opening and closing its check-in and finally we added the flight code. This information is important because from the flight code it is possible to know the type of the aircraft used for that flight and then the total number of aircraft available seats. From this number it is possible the value of a important factor called load factor or filling factor, defined as the ratio between the number of seats occupied by passengers and the total number of available seats on the aircraft. In most of the studies this problem is simplified approximating the filling factor at a constant value and equal to 0.65.



VOLO.................... CODICE AEREO.............. ORARIO PARTENZA..............

ORARIO INIZIO CHECK-IN.................... ORARIO FINE CHECK-IN..............

PASSEGGERI	Tarrivo in coda	BAGAGLI	BANCO n.	BANCO n.	BANCO n.	BANCO n.	TEMP.Oin	TEMP.Ofi

Figure 2: Spreadsheet data collection

As for the security control at checkpoint (located on the first floor of the terminal), we used both historical data in our possession and new collected data. The survey in any case was simpler than the one made at check-in desks, because there is not flights distinction or other it is enough to consider the number of active locations and the time of queue clearing.

In order to carry out a correct analysis of collected data it was necessary to make an appropriate data stratification for highlighting some key aspects. This was a crucial step in order to get to the distributions to be included in the simulation model to be validated. The stratification concerns different features that are showed through aerograms summarizing the contents.

- ☐ Airline. Mainly we focused on check-in desks number dedicated at Airone, the airline company currently operate the largest number of flights per day at Naples International airport (61% vs 39% for the others). In particular, the Airone company exploits its available desks through a mode called 'Common Check-in generalised', with which all counters are indiscriminately opened to any flight operated by the company itself, so at the same desk are recorded passengers to different destinations. For most other airlines instead check-in is 'dedicated' to a specific flight.

- Flight destinations. The distinction between domestic (65%) and international flights (35%) is crucial in our study for the different approach that passengers show towards the second flight category compared to the next (eg arriving at the terminal generally more advance time) and for the different opening check-in desk times. The opening check-in desk times for the major airlines operating in Capodichino airport are:
- AIRONE: for domestic flights from 1.5 h to 25 ' before departure and for international flights from 2h 30' before the flight;

  BRITISH AIRWAYS: from 2h 30 ' before the flight;

  AirFrance: from 2h 30 ' before the flight;

  MERIDIANA: from 2h 30 ' before the flight;

  LUFTHANSA: 2.5 h at 35 ' before the flight.
- Type of passengers. Last but not least important distinction is that which concerns the two main categories of passengers: tourist and business passengers, 80% and 20% respectively

5. SIMULATION MODEL

The simulation is a very useful tool to predict the constant changes that result in a highly dynamic as the airport, and in particular, it refers to the discrete event simulation (DES), in which the state variables vary only at of discrete events (special moment), determined in turn by activity and delay, and not continuously, adapting perfectly to the systems characterized by complex processes, combined with limited infrastructure capacity.

It's possible to think in terms of modular system, developing model through different submodels that make cleaner the understanding and allow for easier management of the whole.

Three used submodels to represent whole system are:

- Generation passengers;
- Check-in;
- Security control.

In particular we considered it appropriate to divide the acceptance process depending on the type of servicing check-in: common desks and dedicated desks.

The passenger traffic flow at the airport is a discrete stochastic process. The simulation of discrete event is often used to model systems characterized by complex processes, combined with infrastructure at limited capacity. The airports are therefore ideal places to work with simulations having these features. The software used for the case study of our study is Rockwell Arena. Retracing the considerations made in the logical model, the simulation has been developed in submodels allowing to obtain all advantages inherent to a modular system representation. Having discussed the model widely from a logical point of view, to avoid unnecessary repetitions, now we restrict the study in highlighting only some key aspects.

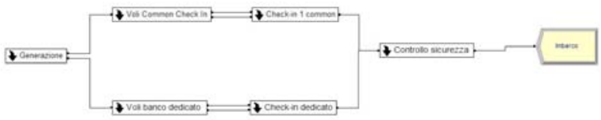


Figure 3: Model representation

As already said, the available resources (check-in desks and security control checkpoint) are considered as variables with a value equal to 1, in the case of free state, and 0 if engaged. At this point the queues are managed through hold of scan for condition type, in which block the phrase allowing users to leave the queue and to continue the path is inserted.

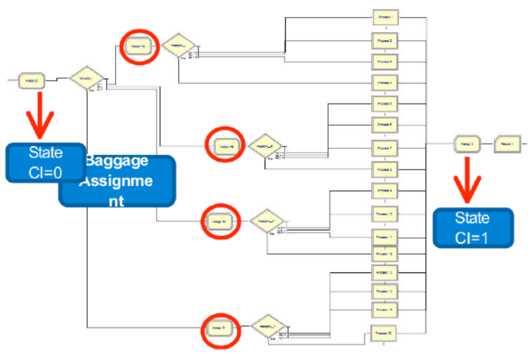


Figure 4: Common Check – in management

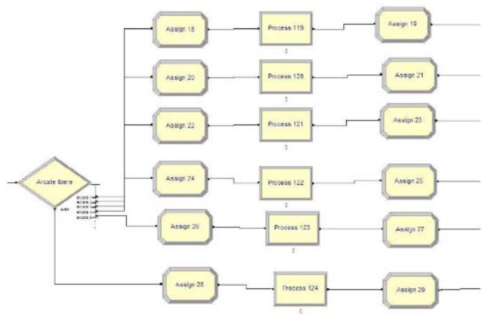


Figure 5: Security Checkpoint management

The first step has been to verify the correct model functionality towards implemented logics.

Verified the correct operation the simulation of the current situation is carried out. Initially are open all check-in desks (4 for the commons and 4 for the dedicated) and all six security control checkpoint inserted into the model and we advance the simulation with a number of 100 replications.

Waiting Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Coda A.Queue	0.3106	0,04	0.2516	0.4293	0.00001208	0.9017
coda Bd.Queue	0.02447783	0,01	0.01058678	0.04985998	0.00079936	0.08864483
coda4.Queue	0.0902	0,03	0.03910184	0.1457	0.00118771	0.3265
coda5.Queue	0.0903	0,03	0.03289742	0.1325	0.00005032	0.4291
coda6.Queue	0.08695318	0,03	0.04967483	0.1562	0.00017434	0.4000
Codabus.Queue	0.01748204	0,00	0.01181466	0.02965474	0.00001276	0.1225
CodaT.Queue	0.04333903	0,01	0.02537405	0.08383999	0.00004206	0.1613

ISBN 978-84-692-5416-5

Figure 6: Hourly queue waiting time with all available resources

Number Waiting	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Coda A.Queue	21.7389	3,53	15.9816	32.5216	0.00	199.00
coda Bd.Queue	0.02260145	0,01	0.00132335	0.06232497	0.00	3.0000
coda4.Queue	0.6582	0,22	0.2151	1.0447	0.00	14.0000
coda5.Queue	0.6593	0,22	0.1933	1.0351	0.00	15.0000
coda6.Queue	0.6678	0,21	0.2856	1.1128	0.00	15.0000
Codabus.Queue	0.04970299	0,02	0.02615481	0.08802838	0.00	5.0000
CodaT.Queue	1.6089	0,48	1.0276	3.2907	0.00	48.0000

Figure 7: Number of users in the queue with all available resources

As can be seen from Figure 7, while the number of users in the queue at check-in desk is quite low, which implies that too many desks are open, the values of the queue at the security control checkpoint is quite high therefore this scenario is not satisfactory to the customer. Then we add 2 more security control checkpoint in the model, initializing it with value 1 (free state) and we launch again the simulation.

Waiting Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Coda A.Queue	0.1442	0,03	0.08273419	0.1898	0.00001713	0.4012
coda Bd.Queue	0.02710774	0,01	0.00849423	0.07885684	0.00482387	0.1219
coda4.Queue	0.1394	0,03	0.08541390	0.1838	0.00013671	0.4811
coda5.Queue	0.1325	0,02	0.0948	0.1881	0.00001231	0.5248
coda6.Queue	0.1316	0,02	0.08474739	0.1666	0.00018533	0.4373
Codabus.Queue	0.01880890	0,00	0.01524115	0.02770415	0.00014332	0.0915
CodaT.Queue	0.04046939	0,01	0.02104750	0.05911443	0.00000303	0.1521

Figure 8: Hourly queue waiting time with 8 available security control checkpoint

Number Waiting	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Coda A.Queue	8.6206	1,91	4.5711	13.5705	0.00	130.00
coda Bd.Queue	0.01828298	0,02	0.00254316	0.08871395	0.00	3.0000
coda4.Queue	1.1504	0,21	0.8008	1.6249	0.00	17.0000
coda5.Queue	1.0865	0,22	0.6043	1.5284	0.00	19.0000
coda6.Queue	1.0870	0,20	0.6886	1.4993	0.00	19.0000
Codabus.Queue	0.06060394	0,01	0.04453302	0.0970	0.00	4.0000
CodaT.Queue	1.5202	0,39	0.7064	2.2537	0.00	48.0000

Figure 9: Number of users in queue with 8 available security control checkpoint

In this new context, we remark that the queue waiting time at the checkpoint is noticeably reduced and in confirmation of what previously developed we found for the check-in desks the values similar to the ones previously found.

6. CONCLUSION

This paper documents our study and development in creating a knowledge-based simulation system to predict check-in counter resource requirements at airports. Results generated by this knowledge-based simulation system are then used by a constraint-based scheduling system that generates a schedule for daily check-in counter allocation.

The results obtained through Arena provide queue average values and waiting times as well as the related minimum and maximum peaks in relation to the number of resources left open in the model.

We obtained, in this step analysis, a quantitative model that represent the airport passenger flow before boarding. The immediate use of this structure is to study the check-in desks numbers and the related security checkpoint numbers in different time interval.

Next phase is to optimize the model, and then to figure out the check-in desks and security control checkpoint best combination able to minimise costs.

REFERENCES

Joustra, P., and Van Dijk, N., 2001. Simulation of Check-in at Airports. *Proceedings of the 2001 Winter Simulation Conference*, 1023.

Verbraeck, A., and Valentin, E., 2002 Simulation Building Blocks for Airport Terminal Modeling," *Proceedings of the 2002 Winter Simulation Conference*, 1199-1200.

Appelt S., Batta R., Lin Li, Drury C., 2007. Simulation of passenger check-in at a medium sized airport. *Proceedings of the 2007 Winter Simulation Conference*, 1252-1260.

Chun Wai Hon, Wai Tak Mak R., Intelligent Resource Simulation for an Airport Check-In Counter Allocation System. IEEE Transaction on Systems, Man, adnd Cybernetics – part C: applications and reviews, vol. 29, no.3, august 1999.

Evans, W.A., 1994. Approaches to intelligent information retrieval. *Information Processing and Management*, 7 (2), 147–168.

Sargent R., 2003. Verification and validation of simulations models, *Proceedings of the 2003 Winter Simulation Conference (WSC 2003).*

Park Y., Ahn S., 2003. Optimal assignment for check-in counters based on passenger arrival behavior at an airport, *Korea Transport Institute*, 2003.

Rauch R., Kljajić M., 2006. Discrete Event Passenger Flow Simulation Model for an Airport Terminal Capacity Analysis.

Yanbing Ju, Aihua Wang, Haiying Che, 2007. Simulation and Optimization for the Airport Passenger Flow. *2007 IEEE*

ABOUT OPTIMAL REMANUFACTURING POLICIES AND SECONDARY MARKETS SUPPLYING

M. Gallo[(a)], L. Guerra[(a)], G. Guizzi[(a)]

(a) Department of Materials Engineering and Operations Management University of Naples "Federico II" - ITALY

'mose.gallo@unina.it' 'luigi.guerra@unina.it' 'g.guizzi@unina.it'

ABSTRACT
The synergic action of different actors involved in environmental protection, is pushing more and more companies to adapt and, in some cases, to revolutionize their strategies, plans and their business goals in an "*environmentally conscious*" way. It is crucial, therefore, to develop and adopt suitable production techniques and EOL management policies for products.
Product recovery involves concepts like reuse, remanufacturing and recycling. In many cases, moreover, reuse and remanufacturing could be simultaneously implemented if *secondary market* supplying is a profitable option. When hybrid remanufacturing/manufacturing systems (HRMSs) are implemented and when secondary markets are supplied with *high quality* returns, some new issues have to be faced and these systems have to be deeply analyzed to better manage them. With this aim three different policies to supply a secondary market will be compared (two *PUSH* policies and one PULL policy), which are based upon different stock level control.

Keywords: tactical analysis, remanufacturing, secondary markets, simulation

1. INTRODUCTION

The increasing rate of products' technological innovation is pushing towards new profit models, based on an integrated management of product life cycle. Innovative policies intended to the recovery of *end of life products*, in fact, not only improve the efficiency in natural resources consumption, but also open new business opportunities for producers (Gungor and Gupta 1999). Among the different recovery options, remanufacturing is of particular importance and is worthwhile of interest (Srivastava and Srivastava 2006).

The aim of this work is to investigate the multifaceted field of remanufacturing and to identify those factors make it a sustainable business from a financial point of view.

The operating cost of a remanufacturing system, in fact, is strictly linked both to strategic decisions (logistics network configuration, secondary markets opportunities, prices of new and remanufactured products, design of new generation products) and to tactical and operational decisions (Daniel, Teunter, and van Wassenhove 2003).

Products at the end of their life cycle can be recovered in many ways and with different levels of efficiency.

The option of materials recycling is at the lowest level of recovery efficiency, this process allows to retrieve the raw material but not the added value of the product.

Higher added-value recovery options are: reconditioning, remanufacturing or cannibalization, where are retrieved, respectively, products, modules or components. Recovery options like repair or reuse don't involve massive restoring activities (Thierry, Salomon, Van Nunen, and Van Wassenhove 1995).

The simultaneous presence of returns with a high residual value and demand for such products on secondary markets, puts the management in front of the dilemma to allocate these units to the secondary market rather than to remanufacture them and supply the primary market.

In an our previous analysis (Gallo, Guerra, and Guizzi 2009) we have assessed the convenience of supplying a secondary market considering some factors related to reverse logistics system (*r*), to the product (*quality mix*) and to the market (*price*). Being such factors external to the system, the analysis has to be considered from a "strategic" point of view. If supplying a secondary market is a profitable option, the production system has to be reorganized in order to best manage this new demand, determining the most suitable operating rules to be used. To this aim three stock level control strategies for secondary market supplying will be compared: the first two are based upon a PUSH logic and the last upon a PULL logic.

This paper is organized as follows. In the problem setting section the stock level control strategies and the logical model used are presented. In Section 3 some issues about the comparison of the different policies are discussed and the results are presented. Section 4 summarizes our findings and draw the conclusions.

2. PROBLEM SETTING

Many studies confirm that the increased uncertainty and variability in a remanufacturing system makes problematic the use of traditional tools for production planning and control (Guide, Jayaraman, and Srivastava 1999; Guide, 2000).

An important difference between the production planning in traditional systems and in remanufacturing systems is that in the latter disappear the typical hierarchical relationship between the canonical stages of production planning.

The actual production volume, in fact, besides being dependent on market demand and production capacity constraints, depends also on the amount of materials (core, parts and components recovered, new components) that become available or necessary during the production process.

Hence, a proper production planning and control in remanufacturing systems (capacity planning, scheduling, monitoring the progress of orders, etc.) is strongly influenced by a sound planning and control of the recovered materials.

In remanufacturing systems a number of randomness related to the quantity, quality and timing of products or components recoverable, makes the definition of a good stock control policy more difficult (Fleischmann, Kuik, and Dekker 2002).

In Hybrid Manufacturing/Remanufacturing Systems literature comparisons between PUSH and PULL policies are usually performed but such policies typically control manufacturing orders release (Mahadevan, Pyke, and Fleischmann 2003).

In particular adopting the PUSH policy all returns are remanufactured as soon as possible (as soon as a batch of returns is available), while in the PULL policy returns are remanufactured as late as possible (remanufacturable returns are hold until the stock level of the finished product warehouse drops below a specific value).

All things considered, the choice is between remanufacturing or retain a core stock and then remanufacture them later. In these cases, the adoption of a PULL policy is due to the core holding cost increase as remanufacturing process flows by: it could be desirable retaining cores in the upstream warehouses to cut holding costs. In this way finished products stock is reduced but delays and stock out risk increases. Some interesting results are proposed by van Deer Laan, Salomon, Dekker and van Wassenhove (1999): as the difference between cores and finished products holding costs increases, the PULL strategy becomes more and more attractive. If these costs are equal (an unrealistic assumption) is better to use a PUSH strategy.

2.1. Stock level control strategies

In the following analysis the performances of three different stock level control strategies for secondary market supplying will be compared, namely: *PUSH* policy, *PUSH 2* policy and *PULL* policy.

In particular, the *PUSH* policy optimizes, in a specific time horizon, the percentage of high quality returns with which the secondary market will be supplied. This policy doesn't consider the warehouses' stock level in the system when an high quality return occurs, but, it pushes the product, with a certain probability (K_{PUSH}), or into the warehouse specifically intended for secondary market supplying or into the remanufacturing process.

The *PUSH 2* policy works like a disposal policy for returns in excess. A maximum value (K_2) for high quality returns (buffer 2) is defined and when it is exceed products are moved into the warehouse for secondary market supplying. So, without considering the other stock levels, the system pushes the high quality returns which exceed the above mentioned upper limit.

In the *PULL* policy, when an high quality return occurs, a control on the finished product stock level is carried out. Until this level is less than a specific threshold (K_{PULL}) the product is remanufactured, otherwise it is sold on the secondary market.

The product is "pulled" by the finished product warehouse intended for primary market supplying. K_{PUSH} and K_{PULL} values are considered for optimizing the objective function which will be conceived as a profit function taking into account both secondary market supplying opportunities and backorders or delayed orders on the primary market.

2.2. Logical model

We have considered the same multi stage inventory control model used in Gallo, Guerra and Guizzi (2009), but some features of the secondary market and the cost structure are quite different (fig.1).

Even though the secondary market on average is able to absorb the share of high-quality products not used by the primary market (working hypothesis in Gallo, Guerra, and Guizzi 2009)), it is necessary, in this case, to model the secondary market demand, taking into account the time between two subsequent requests in the secondary market.

We introduced, therefore, in the model a warehouse of high quality returns for the secondary market in order to meet demand from such market when it occurs. Such demand is modelled by a Poisson process with parameter $\gamma_{SM} = k_1 * r * \gamma$ (see Gallo, Guerra, and Guizzi 2009 for more details). Being such value the overall high quality returns fraction, on average the secondary market could be able to absorb all high quality returns.

The holding costs for the secondary market buffer are calculated using the "traditional method" (Teunter, van Deer Laan, and Inderfurth 2000).

Unlike the primary market, we don't consider backordered or lost sales on the secondary market.

Moreover to evaluate the effectiveness of the different policies considered, the holding costs are "amplified" in such a way their influence on overall costs is about 15% on average. This increase, in fact, makes more evident the impact of the control policies on the system operation.

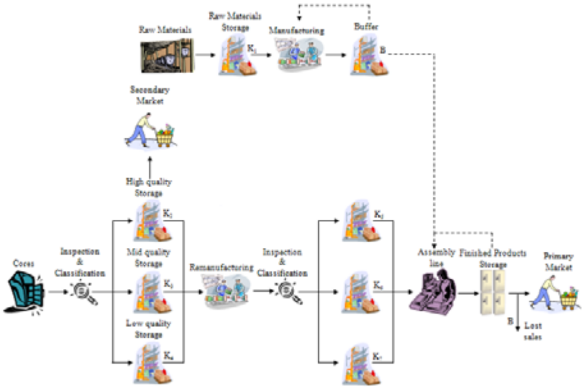


Figure 1 - Multi stage inventory control model with a secondary market

3. RESULTS AND DISCUSSION

3.1. Priority rules performance comparison for PUSH 2 strategy

The performance of the *PUSH 2* control policy is affected by the specific priority rule adopted in the remanufacturing process.

In particular, the remanufacturing process is fed by three different buffers containing high, medium and low quality cores and the choice about the first among these type of core to be remanufactured affects the way the secondary market is supplied with respect to:

- the number of units received in the warehouse intended for secondary market supplying;
- the time between two subsequent sales in the secondary market.

Therefore, to effectively compare the above mentioned policies performance, it is required to carefully choose among the various priority rules the one that allows the *PUSH 2* control policy to work as efficiently as possible.

As concern returns management policies comparison, for the *PUSH 2* policy, the priority rule which best performs according to return rate changes will be considered (Table 1).

Table 1: Adopted Priority Rule for PUSH 2 Policy

Return Rate	Priority Rule
0,7	2, 1, 3
0,8	2, 1, 3
0,9	2, 3, 1

3.2. Secondary market supplying policies performance comparison

Figure 2 summarizes the results obtained comparing the different secondary market supplying policies considered.

The *PULL* policy is the best one, while the *PUSH* policy performs the worst.

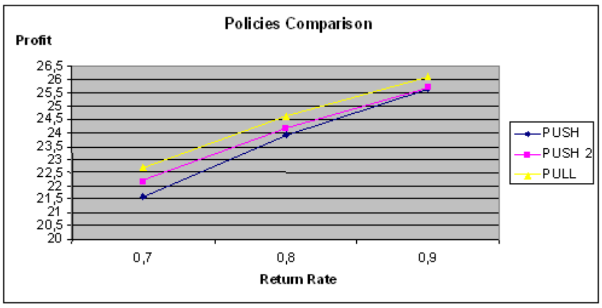
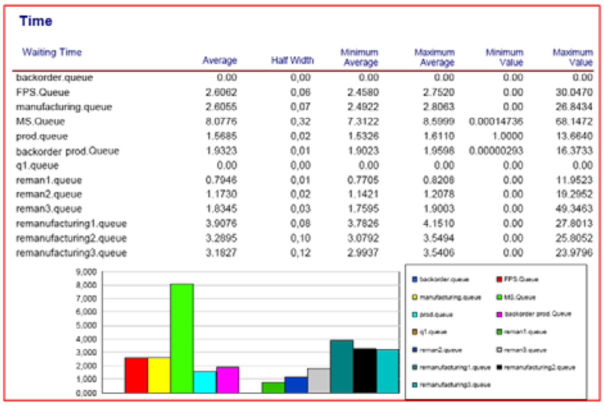


Figure 2: Policies Performance Comparison

Comparing PULL and PUSH policies it can be noticed that making the most of the information about finished products stock level, the PULL policy performs better in supplying the primary market by reducing the average waiting time of backorder sales (Figure 3, Figure 5) and the number of delayed sales (Figure 4, Figure 6). This reduction is really outstanding because of high backorder costs and loss of public image risks.

Moreover, the PULL policy supplies the secondary market with a larger amount of products (Figure 4, Figure 6) although the waiting time in the finished products store is significantly higher (Figure 3, Figure 5). However, the increased revenues from secondary market sales cover the increased holding costs. The PULL strategy, therefore, performs better in secondary markets managing.


Time

Waiting Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
backorder.queue	0.00	0,00	0.00	0.00	0.00	0.00
FPS.Queue	2.6062	0,06	2.4580	2.7520	0.00	30.0470
manufacturing.queue	2.6055	0,07	2.4922	2.8063	0.00	26.8434
MS.Queue	8.0776	0,32	7.3122	8.5999	0.00014736	68.1472
prod.queue	1.5685	0,02	1.5326	1.6110	1.0000	13.6640
backorder prod.Queue	1.9323	0,01	1.9023	1.9598	0.00000293	16.3733
q1.queue	0.00	0,00	0.00	0.00	0.00	0.00
reman1.queue	0.7946	0,01	0.7705	0.8208	0.00	11.9523
reman2.queue	1.1730	0,02	1.1421	1.2078	0.00	19.2952
reman3.queue	1.8345	0,03	1.7595	1.9003	0.00	49.3463
remanufacturing1.queue	3.9076	0,08	3.7826	4.1510	0.00	27.8013
remanufacturing2.queue	3.2895	0,10	3.0792	3.5494	0.00	25.8052
remanufacturing3.queue	3.1827	0,12	2.9937	3.5406	0.00	23.9796

Figure 3: PULL Policy – Waiting Time

Counter

Count	Average	Half Width	Minimum Average	Maximum Average
average	7215.20	51,19	7125.00	7378.00
average for assamble	4336.00	56,61	4213.00	4478.00
bad	4436.80	50,00	4369.00	4543.00
bad for assemble	1334.30	23,87	1279.00	1386.00
dispose average	2879.20	33,95	2799.00	2943.00
dispose bad	3102.50	54,95	3020.00	3264.00
dispose good	760.70	21,24	695.00	800.00
FP	26993.10	65,33	26843.00	27130.00
good	7568.40	91,57	7394.00	7785.00
good for assemble	6807.70	93,62	6629.00	7046.00
lost sales	2960.00	122,40	2649.00	3155.00
man for assemble	14513.90	29,09	14435.00	14561.00
ms1	4421.00	108,37	4181.00	4628.00
overflow average	12.9000	2,95	7.0000	19.0000
overflow bad	327.60	16,61	292.00	362.00
overflow raw	15487.00	29,13	15440.00	15566.00
sales	17467.50	196,23	17064.00	17908.00
sales ms	4420.80	108,98	4178.00	4629.00
backorder sales	9525.50	211,49	9110.00	10017.00
backorder sales ms	0.00	0,00	0.00	0.00
lost sales ms	7588.30	72,74	7444.00	7736.00

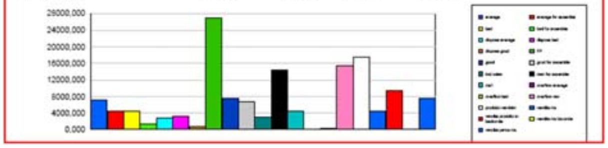


Figure 4: PULL Policy – Products Sold

Time

Waiting Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
backorder.queue	0.00	0,00	0.00	0.00	0.00	0.00
FP .Queue	2.6610	0,09	2.4234	2.8657	0.00	24.3171
manufacturing.queue	2.6801	0,05	2.5152	2.7955	0.00	28.3569
MS.Queue	3.7387	0,08	3.5143	3.9322	0.00018024	34.5719
prod.queue	1.6427	0,01	1.6211	1.6558	1.0001	16.4625
backorder prod.Queue	2.6358	0,02	2.5933	2.6768	0.00002461	16.7975
reman1.queue	0.7673	0,02	0.7221	0.8087	0.00	17.5437
reman2.queue	1.2114	0,04	1.1178	1.2996	0.00	23.6110
reman3.queue	2.7235	0,15	2.3210	3.0909	0.00	50.8204
remanufacturing1.queue	3.5126	0,08	3.2119	3.6604	0.00	27.0182
remanufacturing2.queue	3.5509	0,10	3.2544	3.8082	0.00	27.2694
remanufacturing3.queue	3.5291	0,10	3.2597	3.8366	0.00	24.5283

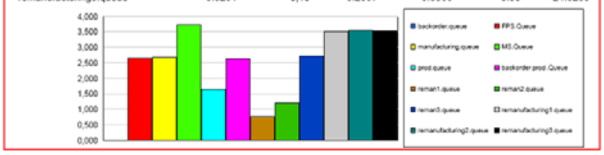


Figure 5: PUSH Policy – Waiting Time

Counter

Count	Average	Half Width	Minimum Average	Maximum Average
average	7194.50	45,79	7031.00	7289.00
average for assamble	4311.17	30,40	4222.00	4389.00
bad	4769.33	54,11	4559.00	4881.00
bad for assemble	1445.00	18,37	1418.00	1514.00
dispose average	2883.33	35,95	2809.00	3003.00
dispose bad	3324.33	51,28	3138.00	3409.00
dispose good	813.75	18,45	768.00	877.00
FP	26992.25	73,59	26803.00	27148.00
good	8082.33	52,14	7955.00	8225.00
good for assemble	7268.58	54,82	7134.00	7442.00
lost sales	2906.58	126,82	2650.00	3259.00
man for assemble	13969.50	44,00	13891.00	14115.00
ms1	3941.75	35,94	3847.00	4034.00
overflow average	3.1667	1,21	0.00	7.0000
overflow bad	4.4167	2,48	0.00	12.0000
overflow raw	16031.08	43,95	15886.00	16109.00
sales	15943.50	369,22	14794.00	16790.00
sales ms	3941.33	35,54	3846.00	4029.00
backorder sales	11048.83	348,83	10352.00	12133.00
backorder sales ms	0.00	0,00	0.00	0.00
lost sales ms	8057.92	68,20	7857.00	8222.00

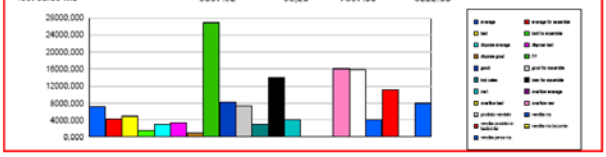


Figure 6: PUSH Policy – Products Sold

PUSH 2 and *PULL* policies have a quite similar performance with regard to the primary market, while *PUSH 2* policy has poor performance in secondary market management, as the number of sales (and hence the revenue) is significantly lower than in other control policies (Figure 7). This result shows the difficulties *of PUSH 2* policy in managing secondary market.

Counter

Count	Average	Half Width	Minimum Average	Maximum Average
average	7208.20	56,75	7046.00	7325.00
average for assamble	4341.30	58,62	4202.00	4434.00
bad	4823.60	52,74	4691.00	4923.00
bad for assemble	1458.30	29,54	1372.00	1534.00
dispose average	2866.90	30,13	2795.00	2943.00
dispose bad	3365.30	57,99	3196.00	3474.00
dispose good	913.50	16,67	885.00	959.00
FP	27064.10	34,11	26978.00	27119.00
good	9045.40	54,29	8927.00	9187.00
good for assemble	8131.90	56,33	8011.00	8274.00
lost sales	2964.20	115,63	2622.00	3218.00
man for assemble	13132.50	45,21	13034.00	13217.00
ms1	2978.20	43,88	2878.00	3068.00
overflow average	0.00	0,00	0.00	0.00
overflow bad	8.0000	3,00	2.0000	16.0000
overflow raw	16868.00	45,05	16784.00	16967.00
sales	18035.00	303,36	17331.00	18628.00
sales ms	2977.90	44,00	2878.00	3067.00
backorder sales	9030.70	290,50	8403.00	9708.00
backorder sales ms	0.00	0,00	0.00	0.00
lost sales ms	9029.50	72,04	8891.00	9166.00

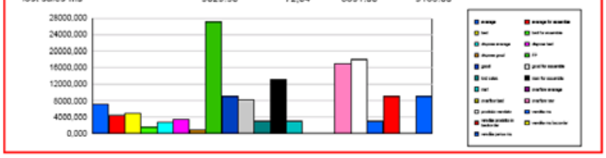


Figure 7: PUSH 2 Policy – Products Sold

The highest profit is achieved by the *PULL* policy while the *PUSH 2* policy gets the lowest cost (Figure 9). The *PUSH* strategy while getting higher revenues than the *PUSH* one, achieves the lowest profit because of the highest cost.

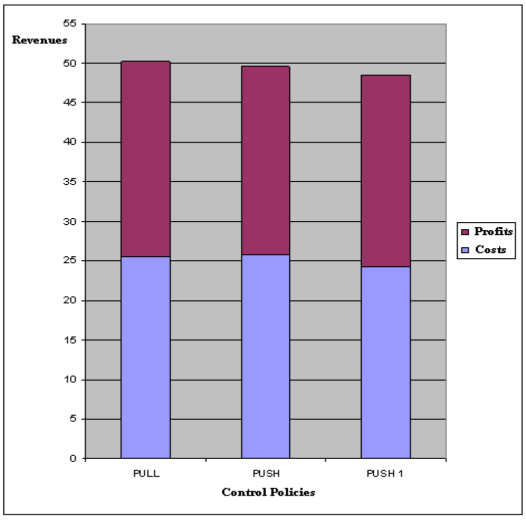


Figure 8: Control Policies – Profit/Cost Analysis

Costs incurred implementing respectively the *PULL*, the *PUSH* and the *PUSH* 2 policy are detailed in Figure 9, Figure 10 and Figure 11 and compared in Figure 12.

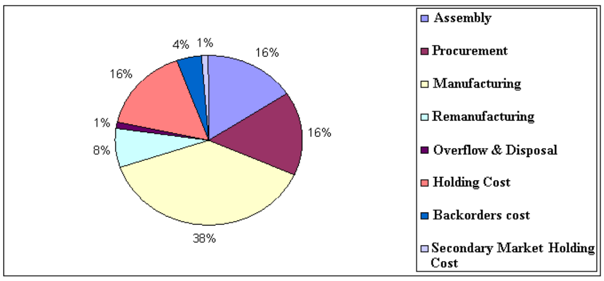


Figure 9: PULL Policy - Costs Detail

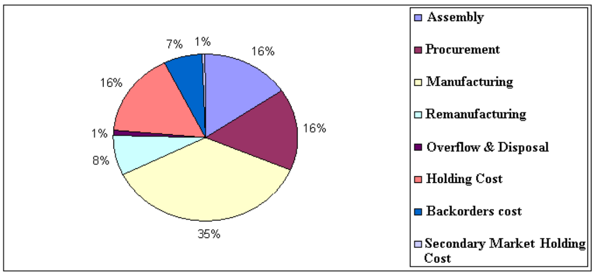


Figure 10: PUSH Policy - Costs Detail

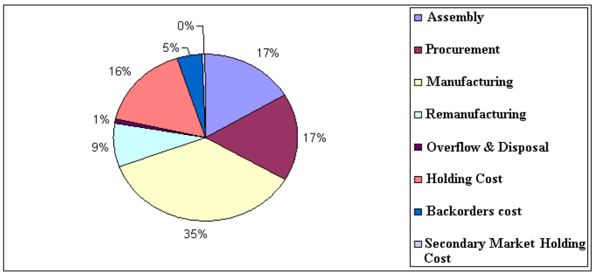


Figure 11: PUSH 2 Policy - Costs Detail

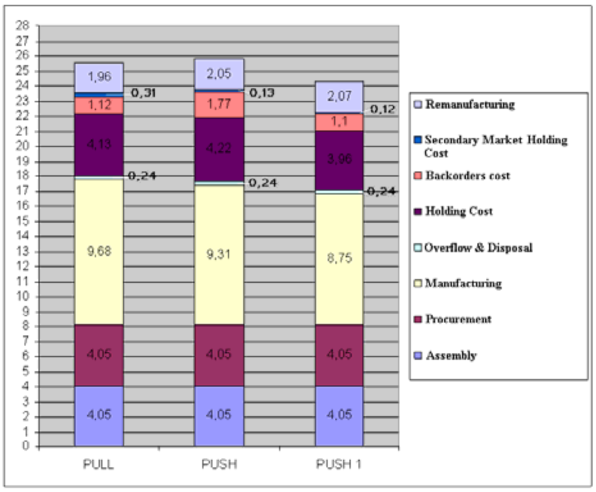


Figure 12: Incurred Costs Comparison

It can be noticed that adopting the *PUSH 2* policy the secondary market is fed with smaller quantities of high quality cores and so less products are manufactured (manufacturing costs are lower) and holding cost are lower (high returns holding cost is lower than manufactured products holding cost). Moreover, remanufacturing costs are slightly higher than those incurred with other control strategies because high quality returns have low remanufacturing costs. So the *PUSH 2* policy has the lowest operating costs. The *PUSH* and *PULL* strategies have a quite similar cost structure, the main difference concerns the high backorder costs incurred with the *PUSH* strategy.

4. CONCLUSIONS

In this paper, three returns management policies for secondary market supplying in Hybrid Manufacturing/Remanufacturing Systems are proposed, which are based upon different stock level control.

Particularly, *PUSH 2* control policy performance is affected by the specific priority rule adopted in the remanufacturing process and so, to effectively compare different policies performance, an analysis is made to carefully choose the rule which allows the *PUSH 2* control policy to work as efficiently as possible.

Summarizing, the *PULL* policy analyzes the inventory level of the warehouse intended for supplying the primary market and decides on a case by case basis how to use high quality returns. In this way the *PULL* policy allows for an improved system processes visibility: finished products stock level has an impact on workstations and upstream buffers state.

The *PUSH* policy is "myopic": it doesn't care about system buffer state but high quality cores are *a priori* used for the primary or secondary market.

The *PUSH 2* policy is a *middle way* solution, it has a "partial" insight of the system, high quality cores are used according to only the stock level of a certain buffer. So, as information quality increases, system performance increases too.

Note that even if the *PULL* policy economic advantages are noticeable, its implementation is more difficult from an organizational point of view because stocks cannot be independently controlled. So, according to the specific case, some decisions must be taken considering the trade-off between economic benefits and organizational difficulties.

REFERENCES

Daniel, V.D.R., Teunter, R.H., van Wassenhove, L., 2003. Matching demand and supply to maximize profits from remanufacturing. *Manufacturing & service Operations Management*, 5 (4), 303-316.

Fleischmann, M., Kuik, R., Dekker, R., 2002. Controlling inventories with stochastic item returns: A basic model. *European Journal of Operational Research*, 138 (1), 63-75.

Gallo, M., Guerra, L., Guizzi, G., 2009. Hybrid Remanufacturing/Manufacturing Systems: secondary markets issues and opportunities. *Wseas Transactions On Business and Economics*, 6 (1), 31–41.

Guide, V.D.R, Jayaraman, V., Srivastava, R., 1999. Production Planning and Control for Remanufacturing: A State-of-the-Art Survey. *Robotics and Computer Integrated Manufacturing*, 15 (3), 221-230.

Guide, V.D.R., 2000. Production planning and control for remanufacturing: Industry practice and research needs. *Journal of Operations Management*, 18 (4), 467–483.

Gungor, A., Gupta, S.M., 1999. Issues in Environmentally Conscious Manufacturing and Product Recovery: A Survey. *Computers & Industrial Engineering*, 36 (4), 811-853.

Inderfurth, K., Teunter, R., H., 2001. Production planning and control of closed-loop supply chains. *Econometric Institute Report*, EI 2001-39.

Mahadevan, B., Pyke, D. F., Fleischmann, M., 2003. Periodic review, push inventory policies for remanufacturing. *European Journal of Operations Research*, 151 (3), 536-551.
Srivastava, S.K., Srivastava, R.K., 2006. Managing product returns for reverse logistics, *International Journal of Physical Distribution & Logistics Management*, 36 (7), 524-546.
Teunter, R.H., van Deer Laan, E., Inderfurth, K., 2000. How to set holding cost rates in average cost inventory models with reverse logistics?. *Omega*, 28, 409-415.
Thierry, M., Salomon, M., Van Nunen, J., Van Wassenhove, L., 1995. Strategic Issues in Product Recovery Management. *California Management Review*, 37 (2), 114-135.
van Deer Laan, E., Salomon, M., Dekker, R., van Wassenhove, L., 1999. Inventory control in hybrid systems with remanufacturing. *Management Science*, 45 (5), 733-747.

AUTHORS BIOGRAPHY

M. Gallo is a PhD of “Production Systems and Technologies” (Department of Materials Engineering and Operations Management of the University of Naples “Federico II”). He also has a master degree in mechanical engineering from the University of Naples “Federico II”. At the moment he is a postdoctoral researcher at the aforesaid department. His research interests are several and related to the design and management of production systems for goods and services. In particular he has treated issues concerning quality, maintenance, production planning, environmental sustainability and soft computing techniques applied to industrial fields.

L. Guerra is currently PhD student at the Department of Materials Engineering and Operations Management of the University of Naples “Federico II”. He received his degree in mechanical engineering from the University of Naples “Federico II” (2002). His main research activities are in manufacturing systems organization, management, reconfiguration and dynamic control, artificial intelligence techniques applied to industrial fields, industrial plants safety and security.

G. Guizzi is researcher at the Department of Materials Engineering and Operations Management of the University of Naples “Federico II”. He is a PhD in Aerospace, Naval and Quality Engineering (2006) and he received his degree in management engineering from the University of Naples “Federico II” (2002). His main research activities are in operations management, project management, risk management, simulation techniques applied to industrial fields.

AN INFORMATION MODEL FOR INTEGRATED PRODUCTION SCHEDULING AND TRANSPORTATION PLANNING

K. Brauer[a], F. Zesch[b], M. Motta[c], A. Wagenitz[d]

[a-b]4flow AG
[c-d]Fraunhofer Institut für Materialfluss und Logistik

[a]k.brauer@4flow.de, [b]f.zesch@4flow.de, [c]marco.motta@iml.fraunhofer.de, [d]axel.wagenitz@iml.fraunhofer.de

ABSTRACT

Frequent changes in production programs of vehicle manufacturers cause fluctuating transport volumes. This leads to inefficient transportation processes because the two planning domains are not well coordinated. Consequently, transport capacities are poorly utilized and extra transports have to be scheduled. We have identified that well-structured, efficient and transparent planning processes are a key factor for improving the integration between production and transportation planning. Furthermore, information should be up-to-date and easily accessible. Two approaches for a better integration are proposed in this paper: First, we present concepts for the continuous adjustment of logistics to changing production programs. Second, we describe concepts for adjusting production sequences to logistics. Both approaches use a central information platform named Virtual Scheduling and Transportation Model (VSTM) as a basis. This model can be accessed by all relevant planning domains and serves as a common information basis, thus improving data quality.

Keywords: production scheduling, transportation planning, modelling, data quality

1. THE POTENTIALS AND OBSTACLES OF INTEGRATED PRODUCTION SCHEDULING AND TRANSPORTATION PLANNING

In an international setting, procurement and distribution networks expand globally. The automotive industry, which is the target industry of this study, uses widely spread supply and distribution networks.

Parts and materials are sourced from all around the world to make use of specialized suppliers and exploit cost efficiencies. The increasing concentration on core competences of vehicle manufacturers has lead to a high number of value creation partners that have to be coordinated and to high transport volumes that are shipped between plants. Customers are also located all around the world and finished vehicles are often shipped over long distances.

The ambitions of vehicle manufacturers to suit their customers' needs have lead to a high level of product customization and short product life cycles. Consequently, a growing number of product types and models have to be dealt with. This brings forth frequent changes in production programs, resulting in fluctuating transport volumes on the one hand and lower transportation lots on the other hand.

In order to keep transportation and related logistics costs at a low level, production and transportation should be coordinated. A lack of coordination between the two areas of planning causes

- disturbances in the short term horizon,
- low utilisation of transportation capacities,
- scheduling of extra transports,
- cancellation of transports at short notice, and
- increasing inventories.

If production scheduling and transportation planning are not integrated, the demand for transportation is difficult to estimate. Therefore shipments are often scheduled on short notice when the demand for transport becomes clear. With groupage transports for instance, suppliers decide for themselves when they place a shipment. Logistics service providers respond to an ABC or ABD scheme, day A being the day they are notified, day B the day they pick up the shipment, and day C the day they deliver the shipment. This procedure not only causes low transparency for the vehicle manufacturer, it also gives logistics service providers little time to react thus limiting their potential to exploit synergies. Pre-scheduled transport capacities are either not used efficiently if volumes are lower than expected or extra transports have to be arranged to cope with transport volumes that are larger than expected. A common reaction of vehicle manufacturers is to decouple production and transportation with buffer stocks. Inventory is built up until an efficient transport can be realized. While the balancing of inventory and transportation cost is vital to efficient logistics, in some cases both costs can be cut if for example vehicles for the same destination are produced within a short time period. Instead of waiting until eight vehicles for the same destination have been produced, the vehicles could be scheduled for production according to their ship-to location.

Other transportation concepts than direct transports and groupage transports are difficult to implement if little forecast information is available. Milk-runs and different regular services need to be planned ahead in order to ensure a smooth run and good utilization. Moreover, the use of road instead of rail transport is promoted, because of the requested flexibility and short response times. This conflicts with economic and ecological objectives as rail transports are often associated with lower transportation cost and less environmental impact.

Yet, despite of potential savings through integrated planning, logistics is often considered secondary to production planning whereby production is planned, and logistics has to cope with the inefficiencies caused by these decisions. Ertogral et al. (2007) show that when transportation is considered explicitly, production and inventory decisions are affected. Hence, transportation should be balanced with production decisions in order to achieve a better result. Even though it is recommended that companies can realize significant cost savings by integrating planning procedures (Chen 2004), most planners still tend to assume that production takes priority, because of the higher cost proportion. Thus they do not integrate decisions to optimize production and logistics overall.

The European research project InTerTrans meets the abovementioned challenges. It aims at the development of practical and marketable solutions in production scheduling and transportation planning and seeks to develop integrated planning processes and planning methods that allow the inclusion of transportation-related constraints in production planning and scheduling.

2. TOWARDS INTEGRATED PLANNING

2.1. Production scheduling in the automotive industry

Production scheduling takes place in three planning horizons (Boysen et al. 2006). In the long run, the total amount of the goods to be produced in a certain time period is determined through sales planning. Production program planning covers medium time ranges. Scheduling determines the production sequence in short planning horizons.

Scheduling of production orders is a very complex research field. It is characterised by the trade off between a multitude of restrictions and the objective of cost-efficient and on-time order fulfilment (Boysen et al. 2007). Often, high capacity utilisation is a superior goal to ensure the profitability of the capital investments made. In mixed-model assembly lines, the work progress in one cycle is limited. Therefore, it is necessary to combine tasks with higher and lower durations in order to achieve a production process with minimum idle times. After the balancing of the production line production and assembly requirements restrict the number of viable production sequences ensuring that no work station is overloaded. A production sequence respecting restrictions and pursuing objectives is created by applying sequencing algorithms.

Sequencing algorithms can be categorized into level-scheduling, mixed-model-sequencing and car-sequencing approaches. In level-scheduling, the objective is to create a smooth production flow and to minimize disruptions. Mixed-model approaches aim at reducing sequence-dependent work overload based on a detailed scheduling which explicitly takes operation times, worker movements, and station borders into account. Car-sequencing algorithms have the same objective but try to simplify these restrictions by formulating a set of sequencing rules of type H_o:N_o. These rules postulate that among N subsequent sequence positions at most H occurrences of a certain option o are allowed (Boysen et al. 2006).

H_o:N_o rules are a commonly used method for determining production sequences in the automotive industry. Many vehicle manufacturers formulate a large number of requirements taking production and assembly restrictions into account. Logistics requirements have not yet been considered in this process. Despite the high number of already existing rules, studies in the InTerTrans project have shown that room for integrating logistics requirements still exists (Hermes et al. 2009; Zesch et al. 2009).

Scholz-Reiter et al. (2009) also recognize the need for a better coordination of production scheduling and transportation planning. They favour a two stage approach. On the long run, transport relevant targets are made available to production planners (e.g. capacity), who in turn try to realize these targets and create a production sequence that also suits transport requirements. In the short run, information about disruptions in the logistics network (e.g. delays) are used to modify the production sequence and adjust it to new planning parameters.

2.2. ILIPT and the Virtual Order Bank

The aim of the European research project ILIPT was to define, validate and operationalise processes, structures and networks to fit the demand of a build-to-order production in the automotive industry (Parry and Graves 2008). According to Holweg and Pil (2001), one of the major requirements of a build-to-order environment are flexible processes. To support such flexible processes with IT solutions, the concept of a virtual order bank (VOB) was developed (Parry and Graves 2008). The VOB represents a system for collaborative capacity and order management. Capacity data of vehicle manufacturers and suppliers is stored in the VOB by the definition of available capacity buckets. The order management functionality then books customer orders on the defined capacity buckets. An exact determination of delivery dates and the selection of suitable production sites due to the available information increases transparency.

The described VSTM picks up the initial idea of the VOB, a holistic information model. As described in the following chapter, the VSTM currently focuses on the planning and execution domain of the vehicle manufacturer. In contrast to the VOB approach, the VSTM includes a functionality to determine an exact production sequence even taking logistic requirements into account. Therefore logistic capacity information is also part of the VSTM.

2.3. Transportation planning

The planning of transportation networks can be divided into two planning horizons. For longer periods transport network structures are determined through demand forecasts based on sales planning and settled in frame contracts with logistics service providers. The redesign of these structures takes place about every one to two years. Because of the dynamic business environment, these structures can become inefficient within a short time. This occurs for instance when shifting volumes of product flows diverge from the capacities planned earlier. A large number of models, modelling methods and software have been designed to support network design decisions (Beamon 1998).

In a short time horizon, logistics service providers are informed about changes of transport volumes and can react accordingly. They schedule transportation orders and plan tours for their vehicles. A number of algorithms have been proposed to solve the idealized vehicle routing problem (VRP) (Bankhofer et al. 2006). Buchholz and Clausen (2009) discuss solution approaches for practical problems and criticize, that idealized VRP fail to cover important industrial aspects whereas software applications tend to use out-dated algorithms. Moreover, the optimization potential is limited to adjustments within the given transport network structures and the short time span between notification and execution of the transport leaves little room for implementing measures. As coordination between order scheduling and short-term transportation planning does usually not take place, the use of road haulage with only low proportions of rail and sea transport is required in order to guarantee flexibility. This leads usually to higher environmental damage and elevated transport costs.

Consequently, a better visibility of production scheduling decisions is the basis for improved transportation planning. State-of-the art planning software that can cope with real-life transportation requirements is then needed to continuously optimize transportation networks (Brauer and Backholer 2009).

2.4. Approaches towards integrated production and transportation planning

The previous sections have shown that both, production scheduling as well as transportation planning, would benefit from a closer coordination. While this idea already has been addressed in a number of studies (Chen 2004), the focus has mainly been on models and problem formulations. A large number of problems include set-up-costs, which usually do not occur in automotive assemblies. Chen (2008) provides a review of „Integrated Production and Outbound Distribution Scheduling" problems (IPODS) and analyses their complexity. Scholz-Reiter et al. (2008) further illustrate, how capacities, lot sizes and delivery dates can be integrated into production scheduling. Jin et al. (2007) present an integrated approach to reduce production and distribution costs (stock, transport). They simplify the problem by introducing a new type of sequencing rule, which results in savings of about 7,5% compared to the original solution. While these publications have provided valuable insights into models and algorithms, little research has been carried out on how to transfer the results into industrial praxis.

3. THE IMPACT OF DATA QUALITY ON INTEGRATED PLANNING

One major issue in transferring research results into the industrial praxis and in developing solutions that are relevant for practitioners is the handling of planning data. During integrated production scheduling and transportation planning, different organizational entities have to provide, exchange and retrieve planning data.

Categories	Dimensions
Intrinsic	Believability
	Accuracy
	Objectivity
	Reputation
Contextual	Value-added
	Relevancy
	Completeness
	Appropriate Amount of Data
Representational	Interpretability
	Ease of understanding
	Representational Consistency
	Concise Representation
Accessibility	Accessibility
	Access security

Table 1: Data quality dimensions

It is widely acknowledged, that the quality of decisions made depends heavily on the quality of the information available at hand to make those decisions. Data quality has been previously described as a multi-dimensional construct. Wang and Strong (1996) have identified 15 relevant dimensions in an empirical study and classified them into the groups intrinsic, contextual, representational and accessibility data quality (Table 1).

Raghunathan (1999) shows that in addition to data quality the decision maker, especially his understanding of the problem and the relationships between its variables is a second important aspect. Eventually, it is important that users (Bovee 2001):

- are able to get information (**accessibility**)
- are able to understand it (**interpretability**)
- find it applicable to their specific domain and purpose of interest (**relevance**) and
- believe it to be credible (**credibility**).

4. THE VIRTUAL SCHEDULING AND TRANSPORTATION MODEL (VSTM)

4.1. Concept and Application

Due to the described increasing number of involved organizational units in the process of production planning and order scheduling, the demand of an integrated information model for these processes and the transport planning is obvious. The VSTM concept does not aim at the replacement or complete substitution of the existing system infrastructure. It should be seen as a virtual model that can be linked to the existing system infrastructure by interfaces (see figure 2). Apart from the big advantage of one central point of access to a valid data base, the VSTM can be extended by further functions, which are necessary for innovative scheduling considering logistic requirements.

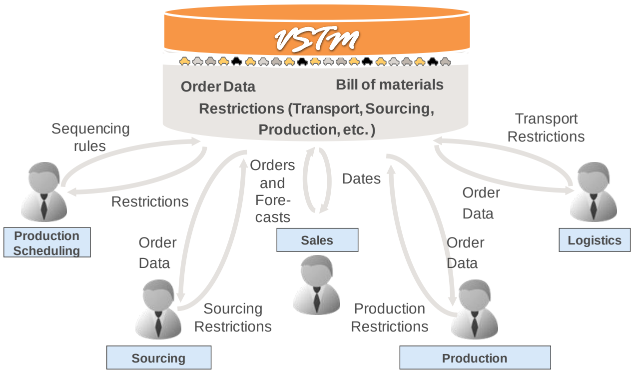


Figure 1: Concept of the Virtual Scheduling and Transportation Model

For gathering the information relevant for innovative scheduling with logistics constraints, interfaces to existing systems can be used. Additional data that are not available in today's processes emerge from new sub processes of individual organizational units during the scheduling. Fundamental information that must be provided by the VSTM is the description of the vehicles with their properties. This data basis covers the availability of individual products and properties in various markets. Moreover, it contains the bill of material and a list of properties that can be combined in a vehicle.

Based on the information provided about vehicles and properties and their availability in particular markets, sale forecasts can be made. These forecasts are the basis for the following production planning. An important aspect of the forecasting and planning process is to use all current information that is available in the network. Today, forecasting of volumes and property quantities is often carried out in different processes and by different organizational units. The initiative of the Californian government to advance small biodiesel vehicles is an example of available information that can be used for forecasting volumes and property quantities. It suggests an increase of volumes for small cars and also for the quantity of diesel engines.

Another opportunity is the generation of forecast orders. The forecast orders are derived from the forecast volumes and take property quantities into account. Additionally, the feasibility of property combinations has to be assured to generate valid and fully specified orders. In the VSTM concept, historical data are used to generate an automated forecast for properties, which are not forecast by the sales department. This procedure provides fully specified and valid forecast orders months before start of production. All further planning processes, such as the explosion of the bill of materials, benefit from this approach.

The use of forecast orders in production scheduling offers high information accuracy and allows the consideration logistic requirements. Within the VSTM the forecast as well as dealer or customer orders are booked against production and transportation capacities. Furthermore, a simulation mode for order booking is added to the VSTM. It enables the planner to asses the impact of a production plan and order sequence in advance.

The sequencing algorithm considers both, production and logistics restrictions and requirements (e.g. the fast formation of transportation lots in the distribution). It can be already applied to the forecast orders. Hence, excessive or inefficient capacity utilization of particular relations can be identified early, thus giving transportation managers more time to react.

Besides logistics capacities on different relations also production capacities are available in the VSTM. For the booking of production capacities it is necessary to describe the available maximum capacity as well as the current planned shift model. This includes information about shift duration and flexibility as well as the factory calendar. The booking functionality of the VSTM ensures that the maximum production capacities are not exceeded. However, in the course of a planning update the shift model is also considered flexible and can be adapted to the planned production volume.

After entering the first demand forecast, the forecast orders are continuously replaced by actual customer or dealer orders. Hereby, updates of the data base can take place at regular intervals or event-driven throughout the program planning and sequencing process. The objective in developing an adequate optimization algorithm for the order sequencing is to ensure a high stability of the original order sequence throughout the whole planning horizon.

In sum, the concept of the VSTM itself does not impose any restrictions on the workflow of production planning, order sequencing and transport planning. Therefore, it is applicable to various companies.

4.2. Impact on data quality

The concept of the VSTM provides access to required information for all relevant planning domains and serves as a common information basis. It has been designed to improve data quality and provide an interface for integrated production scheduling and transportation planning. The **accessibility** of information is increased by the VSTM concept. The VSTM provides a central virtual information model. Using interfaces, latest planning data will be transferred to the VSTM immediately. Each organizational unit and planning domain can then access the same data directly. Another benefit is, that the processes of production and transportation planning do not have to be synchronizes. Each entity can access and amend data according to its schedule. The **interpretability** is also enhanced because all planning data regarding the production program are derived from fully specified and valid forecast orders. This reduces inconsistencies and misunderstandings. Even though the VSTM provides a holistic data basis, the **relevance** of the data provided for a single organizational entity is high. By using customised interfaces and queries for different planning domains and organizational units, the available data can be focussed on the relevant data for a concrete planning process. An information overflow is avoided. The **credibility** of the planning data is maximised by the use of fully specified and valid orders. The shift to an order based planning increases the validity of all planning data as no inconsistence can occur. All entities work with the same data set and are informed immediately if changes occur or forecast volumes are reduced.

5. TWO LEVERAGES FOR INTEGRATED PLANNING

5.1. Adjustment of logistics to changing production programs

With the help of the VSTM current data can be accessed by logistics planners at any given time. This sections describes, how transportation planning can be coordinated with production planning by using the VSTM as an interface.

The order data retrieved from the VSTM are used to update the database for transportation planning. Hereby only relevant data at a suitable level of detail are processed. The estimated demand for production is then directly mapped on the current network. This innovative procedure yields transparency about the status quo of the transport network. Where are the major material flows? Will there be significant changes? Do the capacities suffice? These questions can be answered by automated analyses that depict the flows and capacities hence providing insights about current opportunities for transport consolidation, and possible cost-cutting measures. All possible measures can be compared and evaluated using state-of-the-art optimisation algorithms. They are then coordinated with logistics service providers, suppliers, and other parties involved. The benefit of continuously checking and adjusting the transportation network to production volumes is depicted in Figure 2. While in the left picture overcapacities and extra transports cause elevated transport costs, the right picture shows a coordinated scenario.

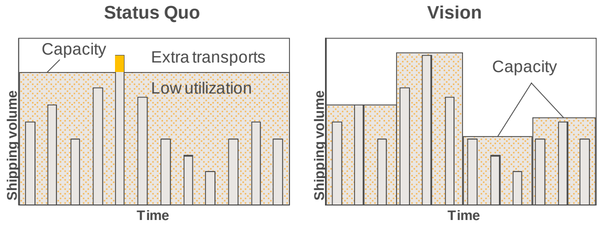


Figure 2: Adjustment of capacities to production volume improves utilization

The time needed ahead of transport execution for changing the transportation concept depends on the logistics flexibility of the firm and its service providers. The higher the logistics flexibility, the later the last dead line for changes in the transportation concept.

Figure 3: Logistics parameters in production scheduling

Measures that have been evaluated, coordinated, and approved are then implemented. This, of course, can result in different requirements for production scheduling (Figure 3). The information in the VSTM is updated to ensure that future production plans can incorporate the new requirements.

5.2. Adjustment of production sequences to logistics

The adjustment of production sequences to logistics requirements is designed as a two-staged process. The first step is the adjustment to requirements from distribution logistics. Feasible transports are determined before the order booking. Orders will then be booked on discrete transports. Several rules for the order allocation are imaginable. In a first application inside a case study, the main rule was to maximise the utilisation of train transports. Backward scheduling from the desired arrival at the dealer site is used to determine the latest possible production end that will ensure the on time distribution of the product. In case of a scheduled transport e. g. by train, all products that are booked on this train will have the same timestamp for the latest feasible production end. Orders for the same relation and transport are bundled. The bundled orders shall be produced immediately before the scheduled transport to avoid unnecessary stock keeping and capital costs during the build-up of a transport.

The second stage is the application of sequencing rules for building the production sequence to avoid capacity overload at single working stations and to ensure a smooth utilisation of capacities. The consideration of logistic requirements derived from the inbound logistics can also be realised by such sequencing rules.

The final production plan is built by applying optimisation algorithms on the available order pool. As it is not possible to respect every single sequencing rule at any time, the violation of sequencing rules has to be associated with penalties. Different penalties can reflect the relevance of a certain sequencing rule.

6. PROTOTYPAL IMPLEMENTATION AND NEED FOR FUTURE RESEARCH

The feasibility of the VSTM and the integrated production scheduling and transportation concept will be demonstrated by a prototype. Therefore the two standard software tools 4flow vista of the 4flow AG and OTD-NET of the Fraunhofer IML will be used together and equipped with additional features.

4flow vista is a widely applied integrated standard software for supply chain design. It automates and accelerates often-repeated and time-consuming activities of designing and planning logistics structures and processes. In the field of strategic transport planning the software helps to improve the transparency of the transportation network in order to identify and eliminate inefficiencies. Moreover it is possible to plan or re-schedule transports and thus create efficient and flexible transportation structures. In order to provide support for integrated production scheduling and transportation planning 4flow vista will be extended with further features for intermodal transportation planning. Moreover innovative methods for quickly adjusting transportation networks to changing production volumes will be added.

OTD-NET is a tool for event discrete simulation of value-added networks and order execution processes (Wagenitz 2007). The design of deployed models can vary in the level of detail depending on the aims of a study (Motta et al. 2008). OTD-NET will be used to demonstrate the functionalities of the VSTM. The required forecast orders will be generated by the simulation tool. Algorithms for order booking on transports and sequence optimization will be implemented. Overall, the simulation based approach allows determining the influences of logistic orientated production scheduling on the entire network.

Further steps in this research will include the specification and implementation of the prototypes and their application in case studies. The description of work-flows and planning hierarchies is another important aspect. Last but not least a road map for the adaption of these results into industrial praxis will be developed.

ACKNOWLEDGMENTS

This work has been partially funded by the German Federal Ministry of Economy and Technology (BMWi) and the Austrian Research Promotion Agency (FFG).

REFERENCES

Bankhofer, U., Wilhelm, M., Williner, G., 2006. Modelle und Methoden der Tourenplanung. *Ilmenauer Beiträge zur Wirtschaftsinformatik*, 5(2006).

Beamon, B.M., 1998. Supply chain design and analysis: Models and methods. *International Journal of Production Economics* 55(1998): 281–294.

Bovee M., Srivastava R.P., Mak, B., 2001. A conceptual Framework and Belief-Function Approach to Assessing Overall Information Quality. *Proceedings of the 6th Int.Conference on Information Quality*. MIT Boston-MA.

Boysen, N., Fliedner, M., Scholl, A., 2006. Produktionsplanung bei Variantenfließfertigung. *Jenaer Schriften zur Wirtschaftswissenschaft* 22.

Boysen, N., Fliedner, M., Scholl, A., 2007. Sequencing mixed-model assembly lines: Survey, classification and model critique. *European Journal of Operational Research* 2(192): 349–373.

Brauer, K., Backholer, C., 2009. A strategic view of transportation management in automotive supply networks. *International Journal of Automotive Technology and Management* 9(1): 69–86.

Buchholz, P., Clausen, U., 2009. *Große Netze der Logistik*. Springer.

Ertogral, K. , Darwish, M. and Ben-Daya, M., 2007. Production and shipment lot sizing in a vendor–buyer supply chain with transportation cost. *European Journal of Operational Research* 3(176): 1592–1606.

Hermes, A. et al., 2009. Integrierte Produktions- und Transportplanung in der Automobilindustrie zur Steigerung der ökologischen Effizienz. *Magdeburger Logistiktagung*. Magdeburg.

Holweg, M., Pil F.K., 2001: Successful build-to-order strategies: start with the customer. MIT Sloan Manag Rev 43(1): 74-83

Holweg, M., 2003. The three-day car challenge: Investigating the inhibitors of responsive order fulfilment in new vehicle supply systems. *International Journal of Logistics Research and Applications* 6(3): 165–183.

Jin, M., Luo, Y., Eksioglu, S. D., 2007. Integration of production sequencing and outbound logistics in the automotive industry. *International Journal of Production Economics* 113(2): 766–774.

Motta, M. et al., 2008. Design of Logistic Networks – A Case Study. *Advances in Simulation for Production and Logistics Applications Stuttgart, Fraunhofer IRB Verlag.*

Parry, G., Graves, A., 2008. Build to order. The road to the 5-day car. Springer, London

Raghunathan, S., 1999. Impact of information quality and decision-maker quality on decision quality: a theoretical model and simulation analysis. *Decision Support Systems* 26(4): 275–286.

Scholz-Reiter, B. et al., 2008. Transportorientierte Reihenfolgeplanung. PPS Management, 13(4): 15–17.

Wagenitz, A., 2007. Modellierungsmethode zur Auftragsabwicklung in der Automobilindustrie. *Dissertation, Technische Universität Dortmund.*

Wang, R.Y., Strong, D.M, 1996. Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems* 12(4): 5–33.

Zesch, F., Brauer, K., Schwede, C., 2009. Software-gestützte Produktionsterminierung und Transportplanung. In: Proceedings of the 11. Paderborner Frühjahrstagung.

MATHEMATICAL MODEL TO OPTIMIZE LAND EMPTY CONTAINER MOVEMENTS

S. Furió[(a)], C. Andrés[(b)], S. Lozano[(c)], B. Adenso-Díaz[(d)]

[(a)]Fundación Valenciaport
[(b)]Universidad Politécnica de Valencia
[(c)]Universidad de Sevilla
[(d)]Universidad de Oviedo

[(a)]sfurio@fundacion.valenciaport.com, [(b)]candres@omp.upv.es, [(c)]slozano@us.es mail, [(d)] adenso@epsig.uniovi.es

ABSTRACT
Empty container logistics is one of the most relevant costs for shipping companies. Thus, efficient empty container management is a key issue in the maritime business. This article analyzes the problem from the local maritime agent point of view, which controls and manages land container logistics. Specifically, it is presented a mathematical model to optimize land empty container movements among shippers, consignees, terminals and depots, along with minimizing storage costs. The mathematical model is defined, tested with real data and solved by using CPLEX. Obtained results confirm the benefits of implementing this kind of models. Finally, future research lines in empty container logistics are identified.

Keywords: empty container logistics, mathematical model, optimization

1. INTRODUCTION

Maritime container traffic has been growing very fast during the last 15 years with steady annual growth rates of over 10%, achieving a global port traffic throughoutput in 2007 of around 500 million TEUs.

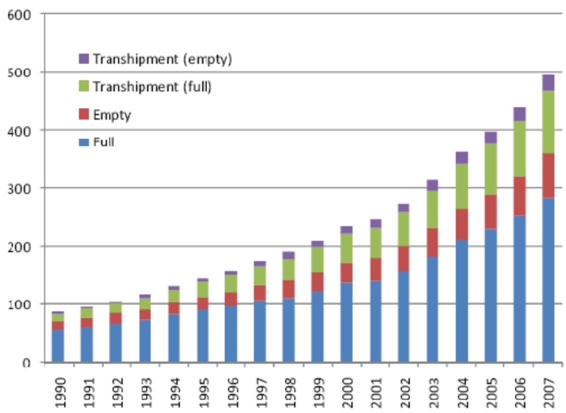


Figure 1: Evolution of world port container traffic

In order to meet this traffic there is a fleet of container ships with capacity for around 10 million TEUs and a maritime container fleet of around 25 million TEUs, which is owned by shipping and leasing companies. Concerning the type of maritime containers used in all these operations, there is a big variety of equipment (standard, highcube, reefers, platforms, etc.), but standard 20 and 40 foot dry cargo containers are almost 90% of the total fleet.

Container logistics has to do with efficient container fleet management minimizing transport costs and maximizing containers use. Empty container logistics deals with the movements, storage and distribution of empty containers, process which starts once the container is emptied at consignee facilities, and ends once the empty container is positioned to be loaded again at the next shipper facilities.

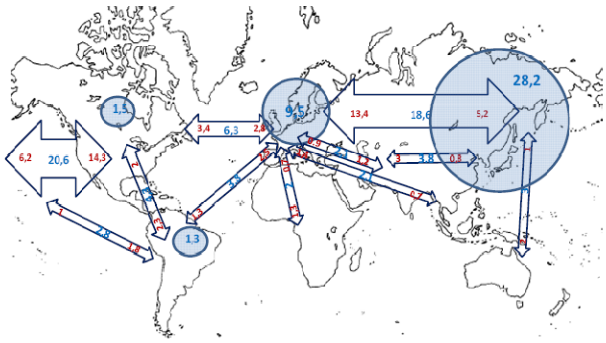


Figure 2: Main container traffic routes

The analysis of main world container traffics shows important eastbound/westbound differences at main trade routes. This way, shipping companies accumulate a large number of unnecessary empty containers at several import-dominant ports (mainly in Europe and USA) which need to be repositioned to different export-dominant ports (mainly in Asia) where they are required (Li et al, 2006).

Empty container incidence at main trade routes exceeds 21% of total traffic (Dekker, 2008), and containers spend more than a half of his life stocked or being transported empty to be repositioned (Crinks, 2000). All this makes empty container logistics one of the most relevant costs for shipping companies (Wang et al, 2007).

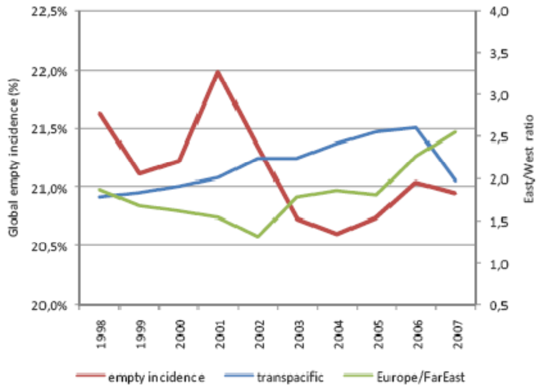


Figure 3: Empty container incidence

Two different levels can be defined when talking about empty container logistics: The international level and the local or regional level (Boile et al, 2008).

The former has to do with the movement of empty containers at global scale to reverse the imbalance problem at main world commercial axes. This is managed directly by shipping companies at global level.

The later deals with the land movement of empty containers between port terminals, empty container depots, consignees and shippers facilities. Local maritime agents, who are responsible of containers fleet management at their influence area, manage these kind movements.

Figure 4 shows the different container movement patterns that can be found.

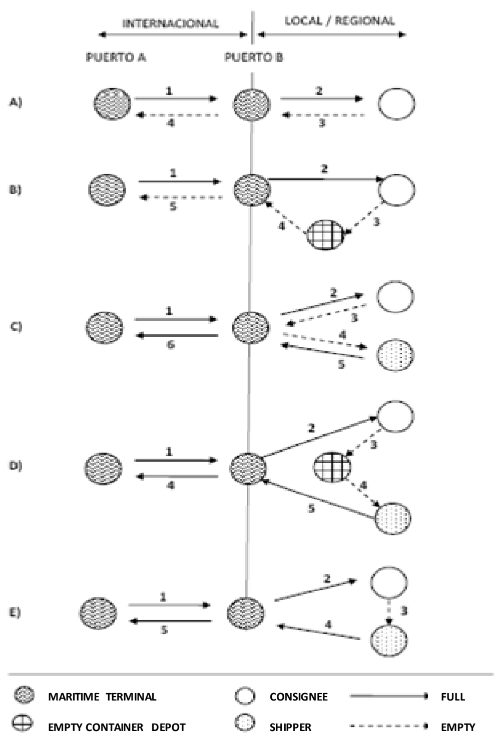


Figure 4: Container movement patterns

The first pattern presented (A) matches with the typical "**empty reposition"** operation. It could be an Asian import container which reaches a European port (movement 1), after this, the container is moved to the importer facilities to be unloaded (movement 2) and, once emptied, it is transported again to the port terminal (movement 3) to be repositioned or sent back to Asia empty (movement 4), in order to start a new cycle.

Pattern B is also an empty reposition operation. The difference is that the empty container is stocked in an empty container depot before being sent it back to the port terminal to be loaded in a ship for repositioning.

Next patterns C, D are known as "**match back**" operations, where an export load allows avoiding the empty maritime transport to reposition the container. The difference between C and D is the place where the container is stocked before being re-used in an export operation (port terminal for C, and empty container depot for D).

Finally, pattern E is a match back operation where the container once unloaded at importer facility (consignee) is moved directly to the shipper facility (exporter) to be loaded again. This operation is also known as "**street-turn**" but it has a lot of difficulties to be carried out:

- It needs an import and export operation coinciding at the same time, type of container, shipping company and place
- It needs also the coincidence of the road transport operator for complementary import and export operations
- Most of the times containers need intermediate operations after a cycle to be reused (brushing, cleaning, repairing, etc.)

The mathematical model presented next, deals with the local or regional level of empty container logistics which is managed by local maritime agents. Specifically it tries to serve as decision support system to make easier the fleet management minimizing transport and storage costs.

2. PROBLEM DEFINITION

Before defining the mathematical model to support empty containers management from the local maritime agent point of view, it is necessary to know the problem and to identify which are maritime agent decisions concerning empty containers management and movements. In order to achieve this knowledge lots of interviews and working sessions have been performed with different maritime agents and empty container depots operating at the Port of Valencia. Thus, it has been developed a deep analysis of the procedures and tools maritime agents use to manage their containers fleet.

The analysis result has arisen that well defined procedures are implemented. These procedures allow in most of the cases updated information and control of the equipment (containers), but there is a lack of operating tools to support decisions in order to minimize movements, storing and transport costs.

The mathematical model developed aims to support maritime agent decisions around container movements to achieve reduced costs for containers fleet management. It is important to remark that only daily operational decisions are being considered, being out of the scope other strategic decisions such as the selection of the container terminals and empty container depots to work with. These kinds of decisions are considered as a given input to the problem which is here addressed.

The daily operational decisions of the local maritime agent concerning empty container logistics are the following:

- Selection of empty container provider (from different terminals and empty depots) for every export operation. (This is needed to fill up the transport order)
- Selection of empty container final destination for every import operation. (This is needed to fill up the transport order)
- Selection and distribution of container movements among different storage places (terminals and empty depots)

And the main objectives to determine all these decisions are the following:

- Ensuring empty container at their disposal to attend export operations
- Maximizing ships capacity
- Minimizing empty container land movements or transport
- Minimizing empty container storage costs
- Minimizing empty container stock at port terminals in order to reduce their congestion and improve their productivity

Taking all these decisions and objectives into account, next comments will be considered when formulating the model:

- The objective of ensuring empty container availability could be guarantee by establishing a set of minimum stocks by type of container depending on demand forecasts
- Maximizing ships capacity could be approached by estimating the optimum number of empty containers to be loaded, having into account: Ship load situation at arrival, ship loading and unloading forecasts, specific instructions of the shipping company. To simplify the model the number of empty containers to be loaded in a ship could be considered as a given input data which is established at a higher level by the shipping company
- The minimization of empty container land movements would be achieved by selecting proper terminal and depots to provide empty containers at every export operation and to storage empty containers after every import operation. This should also reduce the need of empty containers movements between different storage facilities.
- The objective of minimizing empty containers stock at port terminals could be contradictory to the previous one, but a rational solution could be achieved by limiting the maximum number of days for an empty container at port terminals
- Minimization of transport costs is directly related to minimization of empty container land movements and total cost minimization (transport costs along with storage costs) could be the objective function of the mathematical model to be developed

Looking again through the problem definition the input data to the problem would be:

- Ship or line-services arrival forecasts: Date of arrival at port, ship load situation at arrival, loading/unloading forecasted movements, etc.
- Empty container stock situation by type of container at different storage facilities
- Export and import forecasted operations related to ships or line-services involved
- Storage costs of empty container at different terminals and depots. (These costs could be dependent on the number of days containers are stocked)
- Transport costs of moving containers between shippers, consignees and storage facilities
- Type of containers master file
- Storage facilities (terminals and depots) master file
- Consignees and shippers master file

3. MATHEMATICAL MODEL

The problem of allocating storage facility to every empty container supply or demand operation can be modeled as a typical transport problem (Network Flow Problem) to look for optimal solutions giving response to daily operational decisions of local maritime agents which have been discussed above.

Next, the model formulation is presented. (In order to simplify notation and understanding constant storage costs have been considered):

<u>Index</u>

$c= 1,2,\ldots, C$	Consignees
$s= 1, 2, \ldots, S$	Shippers
$d= 1,2, \ldots, D$	Empty container Depots
$j=1,2,\ldots, J$	Port Terminal
$t=1,2,\ldots, T$	Time periods
$r=1,2,\ldots,R$	Container types

<u>Demand / Supply Data</u>

ISBN 978-84-692-5416-5

L_{tsr} Number of type r empty containers to provide to shipper s at t time period

U_{tcr} Number of type r empty containers supplied by consignee c at t time period

R_{tjr} Number of type r empty containers to unload at terminal j at t time period

O_{tjr} Number of type r empty containers to load at terminal j at t time period

<u>Storage Capacity Data</u>

I_{jr}^{max} Upper limit of type r empty container stock at j terminal

$\hat{I}_{dr}^{max}$ Upper limit of type r empty container stock at d depot

$\hat{I}_{dr}^{min}$ Lower limit of type r empty container stock at d depot

<u>Storage Costs Data</u>

h_{jr} Unit storage cost of type r empty containers at j terminal

h_{jr} Unit storage cost of type r empty containers at d depot

<u>Transport Cost Data</u>

α_{cjr} Unit transport cost of type r container from consignee c to terminal j

β_{cdr} Unit transport cost of type r container from consignee c to depot d

γ_{jsr} Unit transport cost of type r container from terminal j to shipper s

δ_{dsr} Unit transport cost of type r container from depot d to shipper s

$\varepsilon_{jj'r}$ Unit transport cost of type r container from terminal j to terminal j'

$\varphi_{dd'r}$ Unit transport cost of type r container from depot d to depot d'

μ_{jdr} Unit transport cost of type r container from terminal j to depot d

<u>Initial Stock Data</u>

I_{jr}^{0} Initial type r container stock at terminal j

$\hat{I}_{dr}^{0}$ Initial type r container stock at depot d

<u>Variables</u>

x_{cjr}^{t} Number of type r containers supplied by consignee c at t time period and allocated to terminal j

z_{cdr}^{t} Number of type r containers supplied by consignee c at t time period and allocated to depot d

y_{jsr}^{t} Number of type r containers provided to shipper s from terminal j at time period t

w_{dsr}^{t} Number of type r containers provided to shipper s from depot d at time period t

$v_{jj'r}^{t}$ Number of type r containers moved from terminal j to terminal j' at time period t

$q_{dd'r}^{t}$ Number of type r containers moved from depot d to depot d' at time period t

p_{jdr}^{t} Number of type r containers moved from terminal j to depot d at time period t

f_{djr}^{t} Number of type r containers moved from depot d to terminal j at time period t

I_{jr}^{t} Number of stocked type r empty containers at terminal j and time period t

$\hat{I}_{dr}^{t}$ Number of stocked type r empty containers at depot d and time period t

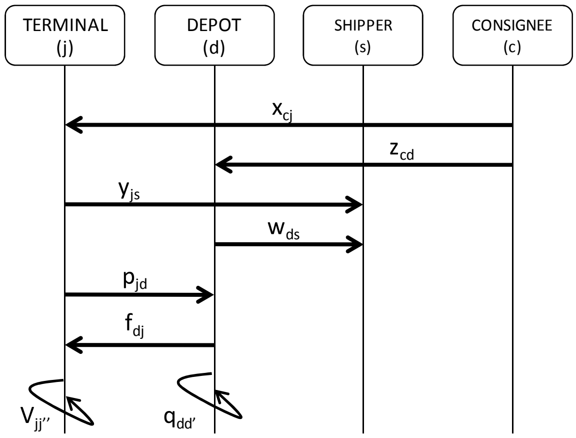


Figure 5: Decision variables associated to empty container movements

<u>Model</u>

$$Mimize \sum_t \sum_r \left[\sum_c \sum_j \alpha_{cjr} x_{cjr}^t + \sum_c \sum_d \beta_{cdr} z_{cdr}^t \right] +$$

$$+ \sum_t \sum_r \left[\sum_j \sum_s \gamma_{jsr} y_{jsr}^t + \sum_d \sum_s \delta_{dsr} w_{dsr}^t \right] +$$

$$+ \sum_t \sum_r \left[\sum_j \varepsilon_{jj'r} v_{jj'r}^t + \sum_d \varphi_{dd'r} q_{dd'r}^t \right] +$$

$$+ \sum_t \sum_r \left[\sum_j \sum_d \mu_{jdr} p_{jdr}^t + \sum_d \sum_j \mu_{djr} f_{djr}^t \right] +$$

$$+ \sum_t \sum_r \left[\sum_j h_{jr} I_{jr}^t + \sum_d \hat{h}_{dr} \hat{I}_{dr}^t \right]$$

Subject to

$$[1]\ \sum_j x_{cjr}^t + \sum_d z_{cdr}^t = U_{tcr} \quad \forall t\ \forall c\ \forall r$$

$$[2]\ \sum_j y_{jsr}^t + \sum_d w_{dsr}^t = L_{tsr} \quad \forall t\ \forall s\ \forall r$$

$$[3]\ I_{jr}^t \leq I_{jr}^{max} \quad \forall t\ \forall j\ \forall r$$

$$[4]\ \hat{I}_{dr}^{min} \leq \hat{I}_{dr}^t \leq \hat{I}_{dr}^{max} \quad \forall t\ \forall j\ \forall r$$

$$[5]\ I_{jr}^t = I_{jr}^{t-1} + R_{jr}^t - O_{jr}^t - \sum_s y_{jsr}^t - \sum_d p_{jdr}^t - \sum_{j'} v_{jj'r}^t + + \sum_c x_{cjr}^t + \sum_d f_{djr}^t + \sum_{j'} v_{j'jr}^t \quad \forall t\ \forall j\ \forall r$$

$$[6]\ \hat{I}_{dr}^t = \hat{I}_{dr}^{t-1} - \sum_s w_{dsr}^t - \sum_j f_{djr}^t - \sum_{d'} q_{dd'r}^t + + \sum_c z_{cdr}^t + \sum_j p_{jdr}^t + \sum_{d'} q_{d'dr}^t \quad \forall t\ \forall d\ \forall r$$

[1] The number of containers supplied from import operations match up with the number of containers sent from consignees to depots and terminals

[2] The number of containers provided for export operations match up with the number of containers sent from depots and terminals to shippers

[3] Empty container stock at terminals should be below the upper limit established

[4] Empty container stock at depots should be between the lower and the upper limits established

[5] Empty container stock update at container terminals

[6] Empty container stock update at empty container depots

4. RESULTS AND CONCLUSIONS

The model designed has been implemented in CPLEX for experimentation and result analysis. The model was firstly tested with different sets of random data and secondly with real data provided from a local maritime agent at the port of Valencia, which required the development of a database with the following information:

- Port terminals and empty container depots master files
- Shippers and consignees master files
- Type of containers master file
- Historic data of transport operations associated to different line services
- Ships arrival forecast
- Road transport costs information
- Storage costs information

Experiment results have shown us that an optimal solution to the problem is found rapidly. As it has been already stated, the model defined addresses a Transport Problem where all the decision variables are integer. The resulting Integer Programming problem (IP) can be successfully solved by using Linear Programming because all empty container supplies and demands *(L, U, R, O)* are also integer, and the optimal solution will have integer values for all the variables.

It is also interesting to remark that optimal solutions do not include movements between depots and terminals in most of the cases that have been tested, which looks logical due to the main objective of minimizing transport costs. This kind of movements only appear at the optimal solution if upper stock limits are going to be exceeded (in order to avoid it); if there are very high storage costs at any terminal or depot; or if there are punctual big demands of containers at a particular terminal.

The study presented provides a well structured approach to the empty container logistics problem, identifying different problems and basic lines to work in at different levels (optimizing sea movements, optimizing land movements, optimizing containers use, optimizing storage yards use, etc.), as well as different solutions or ideas to be developed and implemented (collaborative schemes, improved information systems, etc.) and the introduction to the use of mathematical models to face a particular empty container problem concerning surface transport and container equipment management by local maritime agents. Obtained results confirm the benefits of implementing this kind of models for operational decision making.

Finally, future research lines around empty container logistics can be pointed out, such as:

- Modeling empty container logistics introducing collaborative schemes among shipping lines
- Modeling container logistics and traffics introducing the “risk” concept
- Modeling international container traffic and different scenarios of container fleet composition
- Development of useful models to plan new infrastructures to support container logistics
- Development of new technologies to improve land use at container yards

REFERENCES

Boile, M.P. and Theofanis, S., 2008. Empty marine container logistics: facts, issues and management strategies. *GeoJournal*, 74, 51–65.

Crinks, P., 2000. *Assets Management in the Global Container Logistics*. International Assets Systems (IAS). Available from: http://www.interasset.com/docs/AssetManagementWP.pdf [accessed 01 May 2009]

Dekker, N., 2008. *Annual Container Market Review and Forecast 2008/09*. London: Drewry Shipping Consultants Ltd.

Li, J.A., Leung, S.C.H., Wu Y. and Liu K., 2006. Allocation of empty containers between multi-

ports. *European Journal of Operations Research*, 182, 400–412.
Wang, B. and Wang, Z., 2007. Research on the optimization of intermodal empty container reposition of land-carriage. *J Transpn Sys Eng & IT*, 7 (3), 29–33.

AUTHORS BIOGRAPHY

Salvador Furió is Director of Logistics at the Valenciaport Foundation. He is Industrial Engineering by the Polytechnic University of Valencia and holds a MoS in Ports Management and Intermodal Transport by Comillas University (ICADE). He has participated in many transport and logistics research projects at national, European and international level, always related with container logistics, maritime and intermodal transport. He collaborates regularly teaching at different master programs from the Universities of Valencia and Castellon and the Polytechnic Universities of Barcelona and Valencia, and participates at national and international congress concerning logistics and transport.

Dr. Carlos Andrés is an assistant professor of Management Science at Polytechnic University of Valencia in Spain. Professor Andres' research activities are focused on Scheduling, Operation Management, line balancing, cellular manufacturing and work study for disable workers. His work has been published in European Journal of Operation Research, Journal of Operation Research Society and Applied Mathematics and in several conference proceedings. Dr Andres has consulted with numerous automotive, furniture, tile and pharmaceutical firms.

Dr. Sebastian Lozano is Professor of Quantitative Methods in Management in the Department of Management at the University of Seville. He holds a Ph.D. from the University of Seville and a Master of Engineering in Industrial Management from the Katholieke Universiteit Leuven. He has published in different journal in the Management Science and Operations Research field such as the International Journal of Production Research, Robotics and Computer Integrated Manufacturing, Computers and Industrial Engineering, Computers and Operations Research, the Journal of the Operational Research Society, the European Journal of Operational Research, Journal of Productivity Analysis, etc. His main research interests are reverse logistics, closed-loop supply chains and performance/efficiency assessment including environmental impacts.

Dr. Belarmino Adenso-Díaz is a professor of Operations Management at the Industrial Engineering School of the University of Oviedo in Spain. He worked for the Spanish National Steel Company for seven years, and has been a visiting professor at the Universities of Colorado and Michigan. He has authored four books, and published articles in scientific journals such as the Operations Research, European Journal of Operational Research, International Journal of Production Research, Omega, Interfaces, Production and Operations Management, British Journal of Management, etc. He is associate editor of the Journal of Heuristics and Omega Int. J. of Management Science, as well as a member of INFORMS, EUROMA and POMS.

WHAT IS BETTER: 4 TIERS OR 5 TIERS IN THE CONTAINER STACKING PROBLEM?

Miguel A. Salido[(a)], Oscar Sapena[(b)], Federico Barber[(c)]

[(a)(b)(c)]Instituto de Automática e Informática Industrial, Universidad Politécnica de Valencia, Spain

[(a)]msalido@dsic.upv.es, [(b)]osapena@dsic.upv.es, [(c)]fbarber@dsic.upv.es

ABSTRACT
Nowadays, there exists a large competition between maritime ports, so that the improvement of customer service became a serious important problem within port container terminals which led to several sub-problems (Vis and Koster 2003). One of these sub-problems is the Container Stacking Problem. A container stack is a type of temporary store where containers await further transport by truck, train or vessel. The main efficiency problem for an individual stack is to ensure easy access to containers at the expected time of transfer. Since stacks are 'last-in, first-out', and the cranes used to relocate containers within the stack are heavily used, the stacks must be maintained in a state that minimizes on-demand relocations. In this paper, we study and compare the configuration of yard-bays with 4 tiers and yard-bays with 5 tiers.

Keywords: Container stacking problem, Planning, Artificial Intelligence

1. INTRODUCTION

Loading and offloading containers on the stack is performed by cranes. In order to access a container which is not at the top of its pile, those above it must be relocated. This reduces the productivity of the cranes.

Maximizing the efficiency of this process leads to several requirements. First, each incoming container should be allocated a place in the stack which should be free and supported at the time of arrival. Second, each outgoing container should be easily accessible, and preferably close to its unloading position, at the time of its departure. In addition, the stability of the stack puts certain limits on, for example, differences in heights in adjacent areas, the placement of empty and 'half' containers and so on.

Since the allocation of positions to containers is currently done more or less manually, this has convinced us that it should be possible to achieve significant improvements of lead times, storage utilization and throughput using improved techniques of the type indicated.

Figure 1 shows a container yard. A yard consists of several blocks, and each block consists of 20-30 yard-bays (Kim, Park and Ryu 2000). Each yard-bay contains several (usually 6) rows. Each row has a maximum allowed tier (usually tier 4 or tier 5 for full containers). Figure 2 shows a transfer crane that is able to move a container within a stacking area or moving to another location on the terminal.

Figure 1: A container yard (courtesy of Hi-Tech Solutions)

Figure 2: A Rubber-tired Gantry crane (courtesy of Kalmar Industries)

When an outside truck delivers an outbound container to a yard, a transfer crane picks it up and stacks it in a yard-bay. During the ship loading operation, a transfer crane picks up the container and transfers it to a truck that delivers it to a quay crane.

In container terminals, the loading operation for export containers is carefully pre-planned by load planners. For load planning, a containership agent usually transfers a load profile (an outline of a load plan) to terminal operating company several days before a ship's arrival. The load profile specifies only the container group, which is identified by container type (full or empty), port of destination, and size to be stowed in each particular ship cell. Since a ship cell can be filled with any container from its assigned group, the handling effort in the marshalling yard can be made easier by optimally sequencing export containers in the yard for the loading operation. The output of this decision-making is called the "load sequence list". In order to have an efficient load sequence, storage layout of export containers must have a good configuration.

The container stacking problem is known to be an intractable highly combinatorial optimization problem (Kefi et al. 2008). It is NP-complete so heuristic techniques are necessary to manage this kind of problems. Few studies are dealing with this problem. In (Kim and Bae 1998) the authors proposed dynamic programming to attain an ideal configuration while minimizing the number of containers to move and the follow-on travelled distance. The problem is divided into two sub-problems: Bay matching and move planning problem, and moving tasks sequencing problem. A mathematical model is used for the resolution of each subproblem. In (Kozan and Preston 1999) a genetic algorithm was used in their study to reduce container transfer and handling times and thereafter the berthing time of ship at port quays. In (Kim, Park and Ryu 2000) the authors dealt with the problem of determination of the best storage slot for containers with the aim of minimizing the number of expected relocation movements in loading operation. They deployed the weight of the container as criteria to define certain priority between the different containers to be stacked in the storage yard.

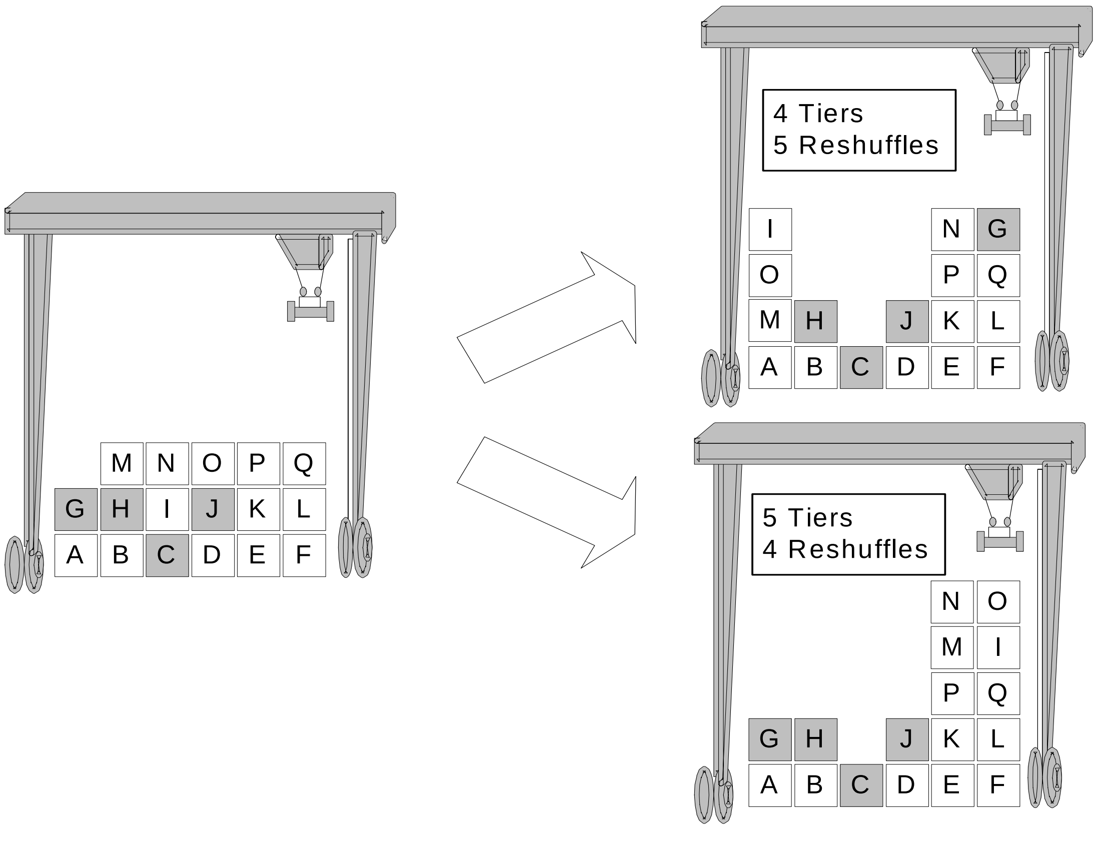


Figure 3: A container yard with 4 and 5 tiers

2. WHAT IS BETTER: FOUR OR FIVE TIERS? A COMPARATIVE STUDY

The main focus of this paper is to give response to harbor operator requirements regarding the best configuration of a yard-bay to minimize the number of reshuffles of export containers. We are interesting on comparing the number of reshuffles in yard-bays with 4 tiers against yard-bays with 5 tiers.

Thus, given a layout, the user selects the set of containers that will be moved to the vessel and the allowed tier. Our tool is able to organize the layout in order to allocate these containers at the top of the stacks in order to minimize the number of relocations. Thus a solution of our problem is a layout where all outgoing containers can be available without carrying out any reshuffle.

We have analyzed two configurations of yards: with 4 tiers and with 5 tiers. We have evaluated the minimum number of reshuffles needed to allocate all selected containers at the top of the stacks or under another selected containers in such a way that no reshuffles is needed to load outgoing containers.

Figure 3 shows an example of a bay with 6 stacks and 4 tiers. The grey containers (G,H,C,J) were selected as outgoing containers. The minimum number of reshuffles to achieve our goal with the restriction of 4 tiers is five. However, the minimum number of reshuffles to achieve our goal with 5 tiers is four. It can be observed that both solutions allow the yard crane to pick up all selected containers without any unnecessary reshuffle.

To evaluate the behavior of both configurations, the experiments were performed on random instances. A random instance is characterized by the tuple <*n*,*s*>, where *n* is the number of containers and *s* is the number of selected containers. Each instance is a random configuration of all containers distributed along the six stack with 4 or 5 tiers. We evaluated 100 test cases for each type of problem.

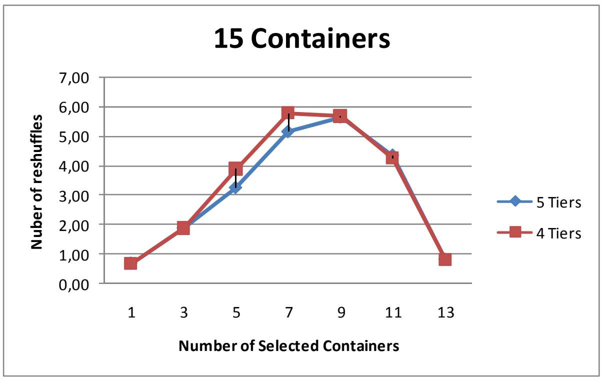


Figure 4: Number of reshuffles for problems <*15,s*>.

In Figure 4, we evaluated the number of reshuffles needed for problems <*15,s*> with 4 tiers and 5 tiers. Thus, we fixed the number of containers to 15 and we increased the number of selected containers *s* from 1 to 13. It can be observed that as the number of selected containers increased, the number of reshuffles increased until a threshold in which the number of reshuffles decreases due to the fact that the number of selected containers is closer to the number of containers (a Gaussian curve). The number of reshuffles was similar in both configuration for "easy" problems, that is, with few selected container or many selected containers. However, the number of reshuffles in problems <*15,5*> and <*15,7*> was lower in yard-bays with 5 tiers than in yard-bays with 4 tiers. In both cases, the upper bound was reached when the number of selected containers was 7.

In Figure 5, we evaluated the number of reshuffles needed for problems <*20,s*> with 4 tiers and 5 tiers. Thus, we fixed the number of containers to 20 and we increased the number of selected containers *s* from 3 to 19. As in the previous figure, it can be observed the Gaussian curve. Nevertheless, due to the tightest of the configuration, mainly for 4 tiers, there were a high number of unsolvable problems with 4 tiers, due to the fact that our tool returns "time out" after 300 seconds. We assigned an upper bound of 22 reshuffles for unsolvable problems. Thus, it can be observed that in problems <*20,13*>, most of the problems were unsolvable for 4 tiers. In all cases, the configuration of 5 tiers maintained a better behavior than the configurations of 4 tiers.

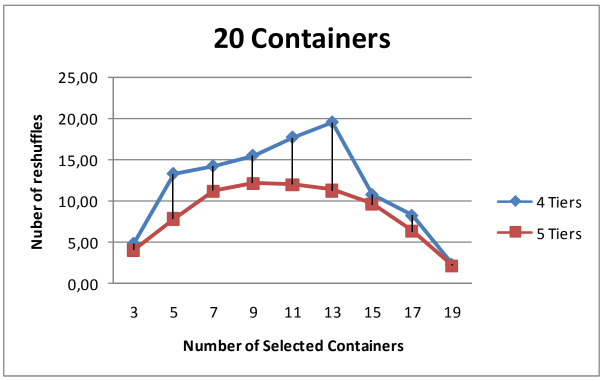


Figure 5: Number of reshuffles for problems <*20,s*>.

In Figure 6, we evaluated the number of reshuffles needed for problems <*n,4*>. Thus, we fixed the number of selected containers to 4 and we increased the number of containers n from 11 to 23. It can be observed that as the number of containers increased, the number of reshuffles increased. For small number of containers (low values of *n*) there is no difference between 4 tiers and 5 tiers. This is due to the fact that it is not needed the use of the higher stacks to achieve a solution because there exist many combinations to achieve a solution. However, the number of reshuffles with 5 tiers was lower than the number of reshuffles with 4 tiers for higher number of containers. Due to the fact that we consider yard-bays of 6 stacks, for problems with 4 tiers the maximum number containers is bounded to 24. In this case instances <*23,4*> for problems with 4 tiers generally has no solution. Thus, we can conclude that for low loaded yard-bays (<15 containers) there is not different between 4 tiers and 5 tiers, meanwhile for high loaded yard-bays 5 tiers is more appropriate for minimizing the number of reshuffles.

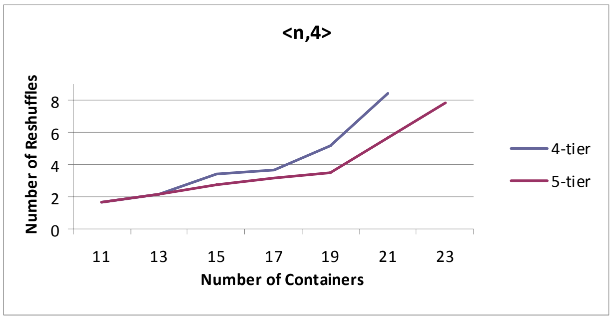


Figure 6: Number of reshuffles for problems <*n,4*>.

3. CONCLUSIONS

In this paper we study and compare the number of reshuffles in yard-bays with 4 tiers against yard-bays with 5 tiers. The obtained results recommend the use of stack with 5 tiers due to the fact that the number of reshuffles is reduced for outgoing containers. In further works, we will evaluate these configurations taken into account the distance between the source and the destination of each reshuffle with the aim of minimizing the number of reshuffles and the total distance.

ACKNOWLEDGMENTS

This work has been partially supported by the research projects TIN2007-67943-C02-01 (Min. de Educacion y Ciencia, Spain-FEDER) , P19/08 (Min. de Fomento, Spain-FEDER) and by the Technical University of Valencia.

REFERENCES

Vis, I.F.A. and Koster, R.D., 2003 Transshipment of containers at a container terminal: an overview. In *European Journal of Operational Research*, 147, 1-16

Kim, K.H., Park, Y.M. and Ryu, K.K., 2000. Deriving decision rules to locate export containers in container yards. *European Journal of Operational Research*, 124, 89–101

Kefi, M., Korbaa, O., Ghedira, K. and Yim, P., 2008 Heuristic-based Model for Container Stacking Problem. In *19th International Conference on Production Research.*

Kim, K.H., Bae, J.W., 1998. Re-marshaling export containers in port container terminals. *C&IE,* 35, 655-658

Kozan, E., Preston, P., 1999, Genetic algorithms to schedule container transfers at multimodal terminals. *In International Transactions in Operational Research*, 6, 311-329

AUTHORS BIOGRAPHY

Miguel A. Salido is an associate professor in Computer Science at the Technical University of Valencia, Spain. Most of his research is focused on techniques for constraint satisfaction techniques and railway scheduling problems. He is the recipient of some national and international awards. He is author of more than 70 papers published on international journals and conferences. He is editor of some books and guess editor of some international journals. He is member of several Scientific and Organizing Committees.

Oscar Sapena is an associate professor in Computer Science at the Technical University of Valencia, Spain. His research is focused on Planning form the Artificial Intelligence point of view. He is author of more than 20 papers published on international journals and conferences. He is member of several Scientific and Organizing Committees.

Federico Barber is Full Professor with the Department of Computer Science, where he leads a research team in artificial intelligence. He has worked on the development of temporal reasoning systems, Constraint Satisfaction Problems, planning and scheduling. He is the author of about 90 research articles which have been published in several journals and conferences. His research has produced several tools for solving real-world optimization combinatory problems. He has participated in and led several national and European research projects related to these areas. He is currently president of the Spanish Association for Artificial Intelligence, and member of several scientific associations.

MODELING AND ANALYSIS OF MARINE TRAFFIC FOR ANOMALY DETECTION (A POSITION PAPER)

Wen-Jing Hsu[(a)], Hui Fang[(b)], and Shell-Ying Huang[(c)]

[(a),(b),(c)]School of Computer Engineering
Nanyang Technological University
Nanyang Avenue, Singapore 639798

[(a)]hsu@ntu.edu.sg; [(b)]hfang@ntu.edu.sg; [(c)]assyhuang@ntu.edu.sg;

ABSTRACT
This paper presents a system for modeling and analyzing the patterns and anomalies in the vessel movement from collected sensor data. The major tasks involved are (a) Detection of macro-level traffic patterns, which includes identifying places of significance based on dwelling time or traffic density. It also includes construction of representative routes of real traffic as opposed to the nominal path on navigational maps. (b) Characterization of vessel movements: classification of journeys, transitional probability between significant places, ship move models, interaction model, etc. which describe the moves of the ships; (c) Tracking and prediction of ships, which enables prediction of the future physical position of a given ship, and the evolution of multiple interacting ships. (d) Anomaly detection: Vessels that behave anomalously in time, space, or mode are identified and monitored. We believe that the system developed will be a useful tool for marine safety and security.

Keywords: Anomaly detection, pattern analysis, machine learning, marine traffic

1. BACKGROUND AND MOTIVATIONS

Hundreds of thousands of vessels visit hub ports annually. At any point in time, there are more than 1,000 ships inside or around Singapore waters. Activities of these vessels are generally of high value to the nation's economy, and hence services must be provided to ensure their smooth operations; illicit or anomalous activities, on the other hand, must be prevented or minimized as much as possible.

To safeguard Singapore's interest and provide quality services, maritime and port authorities closely monitors the movements of these vessels. The authority typically deploys radar and vessel Automated Identification System (AIS) for this purpose and it collects large volume (multi-billion bytes annually) of data.

Upon judicial analysis, the vessel data can reveal many kinds of useful information. For instance, Fig. 1 shows a time-lapsed view of the vessel data, which reveals the spatial distribution of vessels and usage of the navigational channels and anchorages.

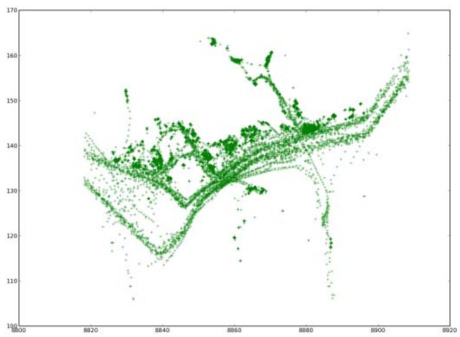


Figure 1. Time lapsed vessel data showing the spatial distribution of vessels in and around Singapore waters

Fig. 2 shows circular movements of a particular vessel, which reveals the actions of the tide and currents on the vessel in the anchorage.

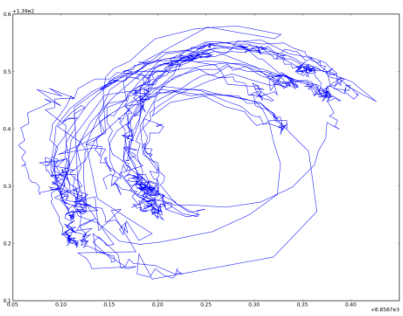


Figure 2. Vessel moving in circular fashion

It is also possible to trace particular vessels in rather fine-grained resolutions. Fig. 3 shows the trajectory of a particular vessel in a given visit, which enables closer examination, e.g. its nearest point to certain security-sensitive areas.

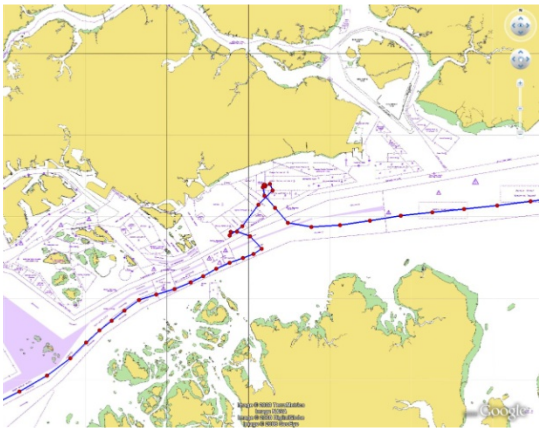


Figure 3. Trajectory of a particular vessel during a recent visit

Thus, it is possible to characterize the general traffic flows and movement patterns of various types of vessels, their temporal or spatial distributions, interactions and navigational behaviors over different segments of the navigational channels and their junctions. This information can be of use in the planning of navigational resources and marine operations. Given the normal activity patterns, it becomes possible to identify anomalous or suspicious behaviors which can be of use to marine security and safety.

2. RESEARCH OBJECTIVES

Presently the burden of analysis is placed mainly on specially-trained human operators and administrators. Machine-assisted data-analysis, however, are becoming inevitable because of the massive amount of data, short time to response, scarcity of qualified personnel, high cost of training. Machine tools are also valued for their consistence and dependability.

This project aims to develop computer-based tools to analyze the patterns and anomalies in the vessel movement from the archived sensor data. The following outlines the major tasks in this project.

- **Detection of macro-level traffic patterns**. This includes identifying places of significance (e.g. locations of long dwelling time or changes of speed or directions), reconstruction of actual trajectories as opposed to the designated path on navigational maps.

- **Characterization of vessel movements**: classification of journeys, transitional probability between SPs (Significant Places), ship move models, interaction model, etc. which describe the moves of the ships;

- **Tracking and prediction of ships**: In the simplest scenario, predict the future physical position of a given ship, and in the more complicated scenario, the evolution of multiple interacting ships.

- **Anomaly detection**: Vessels that behave anomalously in time, space, or mode should be identified and monitored. For instance, vessels near Jurong Island should be monitored automatically.

3. APPROACH AND METHODOLOGY

To meet the challenges, we propose the following methodology from machine learning:

- **Automated extraction of macro traffic patterns:** The data log could be divided into training set and verification set; where the training set is used to extract patterns by using clustering and classification algorithms. The verification set is used to test and verify the rules extracted by the algorithms. The boundaries and the representative pathways of the traffic flows can be inferred automatically to generate the actual or *de facto* navigational network (as opposed to the stipulated, on-paper layout), and the distributions of the ships inside the actual routes can also be derived. General unsupervised clustering algorithms will be adapted to automatically identify the feature values that exhibit high concentrations. This will allow discovery of, e.g., places and time points of significance.

- **Characterization of vessel journeys:** Given the uncertain sensor data, paths linking the significant places can be constructed; the resulting paths are simplified representation that minimizes the expected errors. A journey may be decomposed into higher level “motifs” such as piecewise linear polygonal lines and semi-circular curves that link SPs. Statistics about vessel speed, directions, transitional probability between SPs (Significant Places) over the paths can be gathered for normal operations. Instantaneous traffic volume, vessel spatial distribution, flow rate can be inferred to find places of possible bottlenecks and congestion.

- **Traction and Prediction of ship movements**: For simple non-interacting vessel behaviors, ship moves can be modeled by using known algorithms such as Kalman Filter; For interacting behaviors, a probabilistic Bayesian model such as Particle Filters will be more adequate. Vessel interaction model will require identification of symptoms and signatures of various scenarios; With sufficiently detailed historical data, blow-by-blow evolution of multiple interacting ships can be modeled and analyzed.

- **Anomaly detection**: Once the patterns of normal ship journeys and movements are found, they are used to identify exceptional cases that deviate from the norms. Vessels that behave anomalously in time, space, or mode can be identified automatically. Based on the high-level motif representation, journeys may be classified according to normal/suspicious categories for further analysis. Input from experienced ship captains and hydraulic experts will also be crucial.

- **Screening and cleansing of raw data:** Because of interferences to radar signals or disturbances in electro-magneto fields, the sensor data may be erroneous and less than precise or reliable. For instance, the recorded data may have spurious instances of records which show vessels at an unlikely location (e.g. on land) or vessels traveling at exceedingly high speed. Certain data filed may also be missing. This pruning phase will require careful filtering through statistically justified method and the data should be discarded only after verification with domain experts. The sheer volume of the logged data and the complexities involved in the movement patterns dictate that only efficient algorithms that exhibit lower computational time and space complexity will be useful. It may eventually require a solution that runs on multi-core computers.

4. SYSTEM ARCHITECTURE

The architecture and the system components for vessel movement analysis are further described below:

(a) Vessel Journey Data Mining Engine: This is the software for finding “norms” from massive amount of data, and summarizing into higher-level abstraction (e.g. envelop or skeleton”, list of significant places, paths and journeys); and automated construction of ship movement model.

(b). Vessel Tracking Engine: this software implements tracking of vessel movements by using Kalman filter or Particle filters. It handles inaccuracies in live sensor data and uncertainties in positions/locations by means of probabilistic distributions and recursive Bayesian inference methods.

(c). State Inference Engine: This software implements higher-level state inference schemes by using tools such as Partially Observable Markov Decision Process (POMDP).

(d). Anomaly Detection Engine: This software implements lower-level geometry matching (Digital Hough Transforms), sequence matching methods and recognition of activities by making reference to templates such as norms of same types of vessels or vessels with the same purpose of visit, or the vessel’s own historical activities.

(e). Query Processing Engine and Console: This software handles the human computer interactions and operator control. It allows user to activate subsystem functions such as querying for vessel information, specifying regions, time, vessels, events, behaviors of interest.

(f). Visualization Engine: Software that shows output generated by the other subsystems, e.g.,

- Show vessel activities over a period of time in the past (which is an output from Data-Mining Engine or Query Processing Engine), present (from live sensors), and projection into future(from Vessel Tracking Engine);

- Show vessel activities within regions of interest from HCI-human computer interactions;

- Select vessels of interests either automatically from State Inference Engine/Anomaly Detection Engine or manually from HCI, to display/highlight according to types (e.g. VLCC, large passenger vessels), size (e.g. LOA> *x* m, GT > *y* tons), sensitive areas (e.g. inside junctions or near an crude oil refinery), projected situations (e.g. vessels whose domains are projected to overlap), vessel interactions: display vessels that come into vicinity; special purposes (tracking particular vessels) etc.

5. CONCLUSION

This paper presented a system design and a Bayesian model for simulating and analyzing the patterns and anomalies in the vessel movement from collected sensor data. The system is able to identify places of significance based on dwelling time or traffic density. It can also construct representative routes of real traffic as opposed to the nominal path on navigational maps. The system can classify journeys, characterize vessel movements between significant places, and derive transitional probabilities. Bayesian models are used to describe the moves of the ships; ship move models, interaction models. Upon completion, they will be able to track and predict ship moves and the evolution of multiple interacting ships and detect anomalies. We believe that the system will be a useful tool for marine safety and security.

AUTHORS BIOGRAPHY

Dr. Hsu Wen Jing is a faculty member of NTU. He is also an Affiliate Faculty Fellow of Singapore-MIT Alliance Program. He received research awards such as Ministry of Transportation Innovation Award (Merit), Best Paper Awards (PDCAT/IEEE Automation/PDCS), Gold Award and Grand Champion of Crayquest. He has about 200 publications in reviewed journals and conferences. He is presently Deputy Director of the Maritime Research Center of NTU.

Fang Hui is a Ph.D. candidate of NTU. His thesis developed a prediction system by using Bayesian networks. He is also a Research Associate working on a project related to simulation of marine traffic. His thesis research is on mining positional data for patterns.

Dr. Huang Shell Ying is an Associate Professor of NTU. Her research interests include intelligent agents, multi-agent systems, simulation based optimization and intelligent decision support systems. She has been the Principal Investigator or Co-PI for many projects funded by Maritime and Port Authority of Singapore, NSTB (now ASTAR), GINTIC (now SimTech), PSA Int'l, Jurong Port Pte Ltd, Surbana Consultants, APL and NTU. She has published many papers in international journals and international conferences.

CHALLENGES ON ORE PORT SIMULATION

Marcelo Moretti Fioroni [(a)], Caio Eduardo Zanin [(b)], José Alexandre Sereno Quintáns [(c)] Lúcia Dini Pereira [(d)] , Isac Reis de Santana [(e)] , Paulo Savastano [(f)] , Luiz Augusto G. Franzese [(g)] , Vítor Luciano de Almeida Benevides [(h)]

[(a)(b)(c)(d)(e)(f)(g)]Paragon Tecnologia – R. Clodomiro Amazonas, 1435 5th floor – São Paulo, SP – Brazil - 04537-012
[(h)] Anglo Ferrous Brazil - Av Inf Henrique, 345 – Rio de Janeiro, RJ – Brazil - 20021-140

[(a)]marcelo@paragon.com.br, [(b)]caio@paragon.com.br, [(c)]alexandre@paragon.com.br, [(d)]lucia@paragon.com.br, [(e)]isac@paragon.com.br, [(f)]paulo@paragon.com.br, [(g)]augusto@paragon.com.br, [(h)]vitor.benevides@angloferrous.com.br

ABSTRACT
There are several challenges involving the representation or an ore loading port system on a simulation package. This kind of port handles bulky material, much more adequately represented by continuous flow then discrete flow, as in opposite to the case of a container-handling port. This paper addresses some challenges faced such as the navigation restrictions to enter/leave the berth, the stacker/reclaimer position in the ore pile, the impact of meeting product mix requirements in the delivery and the loading plan for the ships. When these aspects are not represented under the correct level of detail, the model can present poor results. Also, results reporting challenges are addressed in this paper. The modeling approach of each of these aspects is presented, and the experimentation in the case of Porto do Açu, located at Rio de Janeiro, is addressed with the purpose of evaluating the efficiency of the solutions proposed.

Keywords: simulation, port, iron ore, ore pile

1. INTRODUCTION

Some of the most favorable applications of simulation studies are port projects. Ports involve expensive equipment, large scale cargo movement and heavy fines for delays. Also, the efficient use of ports is critical to countries, since they are the most important structure demanded to drain the export production.

This kind of simulation study becomes even more important if the port is not yet constructed, and is in the design phase. The simulation of the chosen design can help find problems or improve performance. It can save millions only by avoiding some design or operation mistakes that would be much more difficult to correct with an operating port.

Countries that are great commodities producers and exporters have ports specially designed to load bulky products like grains or coal into ships that are also specially designed to carry these types of products.

An ore loading port system presents characteristics that differentiate it from other kinds of ports:

- The product to be loaded is bulk material. Thus, a very specific system is needed move the material from one place to another. Usually, this is done by conveyor belts.
- The bulk material must be stored in piles. These piles must be handled by equipment specially designed for this purpose: stackers, that store material and build the pile, and reclaimers, that remove material from it.
- Specially designed equipment is necessary to load ships: the shiploader. It is fed by a conveyor belt that transports ore from the pile into one of the ship's compartments.

Also, bulk material ports present other usual problems that are common to regular ports:

- Ship arrivals must follow traffic rules involving weather, channel occupation and tide variations.
- There must be balance between product arrivals and ship arrivals, or the storage space might not be enough.

To simulate this kind of system some techniques and modeling approaches were applied to face each challenge, and they are all presented in the following sections of this paper. These techniques were validated over a real case, which presents all the characteristics listed above.

2. MAJOR CHALLENGES

First of all, the simulation of an ore handling port is very different from other kinds of ports, at which discrete cargo like containers are received, stored and loaded. This class of problem has been extensively studied, and some reference work is that of Meisel & Bierwirth (2009) and Karlaftis, Kepaptsoglou and

Sambracos (2009). The solutions and algorithms proposed in these studies do not apply to a continuous cargo flow port. These being the case, new developments were required to achieve a correct process representation.

2.1. Rules for Ships Entering or Leaving the Port

When approaching or leaving the port, ships follow certain rules or restrictions. Some of these rules apply to any port, while others apply only to bulk material ports. The list below details a worst-case scenario for these rules and restrictions:

- There is only one waterway channel. Thus, it is not possible to have more than one ship entering or leaving the port at the same time. In this case the priority is for ships leaving the port.
- There are a limited number of berths. A berth must be free before the ship enters the channel.
- The weather conditions must be good enough to allow ship mooring.
- Large loaded ships must wait for the correct tide to leave the port.
- A ship may not enter the channel if there are not enough products on stock to allow loading.

These restrictions must be cautiously implemented, since they can cause either deadlocks or wrong behavior. An example situation is when one ship leaves the berth at the same time that a queued ship, which was waiting for berth availability, goes ahead and seizes the channel. Since the first ship did not leave the port yet, it needs the channel. When the second ship arrives to the berth, the model will be representing an impossible situation: two ships at the same berth.

Many of these rules and restrictions are dependent from each other, and the implementation of some of them requires the use of information from other parts of the system (yard, weather, etc.). All of this must be remembered when modeling.

2.1.1. Modeling Approach for the Ships

To implement the circulation of ships at the port, separated algorithms for arriving and departing ships were developed.

Arriving ships must check:

a. If the weather is good to sail.
b. If there is enough product in the yard to load this ship.
c. If one berth AND the channel are free.

Departing ships must check:

a. If the weather is good to sail.
b. If the tide variation allows circulation of a ship of this particular size.
c. If the channel is free.

2.2. Ore Pile Slot Division

The ore is stored in piles and handled by equipments named stackers and reclaimers (some are stacker and reclaimer at the same time). This kind of equipment works connected to the conveyor network, so it can remove/store the material directly to/from the conveyor belt.

These piles are normally long, and the stackers/reclaimers are normally located at one or both sides of the pile. The stackers and reclaimers move over an axis parallel to "b". Any position of the pile can be reached by these equipments, which move over a track system. To store different products on the same yard, different piles may be formed. Figure 1 shows the main parameters of a pile and its aspect.

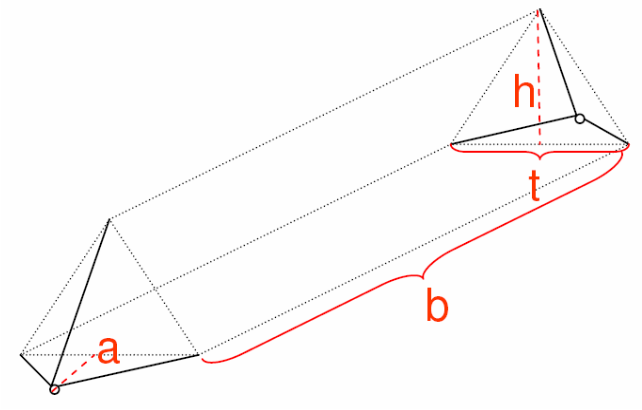


Figure 1: Parameters and aspect of an ore pile

Dimensions a, b, t and h define the pile's geometry, and depend on the kind of ore being stored. The "b" axis is divided into "slots" of 5 meters each, so that a record of each pile's position can be kept.

The challenge is to model the ore stored at each slot, and the stacker or reclaimer movements over the "b" axis. The movements and ore storage/consumption must be synchronized, and follow the company's policy for handling the ore.

2.2.1. Modeling Approach for the Pile

In this case, the most important aspect is to represent the storage of ore at the correct position in the pile. This can be done by using a variable with as many lines as the number of slots in the pile. In fact, the ore position will rule the stacker/reclaimer position. Therefore, once the pile is modeled, the stackers and reclaimers are impelled to reach the correct slot.

Every time the system needs to recover or store the ore, it follows the algorithm:

a. Check if the stacker/reclaimer is in the first slot to be handled in the pile.
    a. If the equipment is not there, move it to that slot.
b. Begin the storage/consumption of the ore on this slot.

c. If the material batch was accomplished, finish the operation.
d. When the slot is full (if storing) or empty (if recovering), change the first slot to the next one.
e. Go back to step a.

By doing this, the equipment movements will be considered in the simulation, conferring it with a high level of detail. Other ore yard studies, like Fioroni et. al. (2007) and Coelho et.al. (2005), do not go further on that detail level.

2.3. Balance between Ship Arrival and the Production

One of the problems when representing a sub-system of a much larger system, is the impact of the other sub-systems on the one represented.

In this case, the arriving ships must load ore that comes from a mine, passes through a beneficiation plant, a mining duct and a filtering process, a railroad or other sub-systems. The problem is that the type of ore sent from the mine must match the type of ore to be loaded into the arriving ship, days later. Lack of product type matching may cause an excess of an unrequired product type at the yard, while the system may be out of stock for the required product type.

In the worst situation, this can cause a deadlock in the simulation model. For example, there could be a situation where a ship has arrived to the port but cannot enter the berth because it needs product A, and there is only product B in stock. Product A is being received by the port but the yard is full, so product A cannot be stored. In the meantime, a second ship that needs product B has arrived, but remains in the queue because the first ship is on delay. In a real situation such as this one, the operational managers could decide on a number of alternatives to solve the problem. Hence, the simulation model must be prepared to respond to such cases.

2.3.1. Modeling Approach for Production and Dispatch Balance

In this case, the weakness of the model can become its strong point: the model lacks the representation of the entire production line of the real system, but it “knows” which ship will arrive, and when, and what product will be loaded into it. With this information, the product arrival at the yards can follow the ship arrivals, balancing the storage area.

It is important to note that this approach is based on the premise that the commercial team does not make mistakes, and the production of the entire system worked as expected, which is not always true, but valid to check the sub-system’s maximum capacity.

2.4. Loading Plan for Different Types of Ships

Ships that transport ore are loaded by certain equipment, the shiploader. This equipment drops ore continuously into the ship’s compartments, one by one, until the ship is completely loaded. However, the loading process cannot access the different compartments at any random order. If all compartments in one side of the ship are loaded in sequence, the weight of the material and the impact of its fall is likely to sink the ship to the side. There is also a risk of breaking the ship’s structure in two parts due to excess weight in the middle section.

In order to avoid these and other operational risks, the ships are loaded in a sequence of steps called loading plan. Each step of the loading plan is determined by a compartment to be accessed and a quantity of ore to be loaded.

The key challenge here is to account for the movements of the shiploader accessing each compartment, which are part of the loading process, but consist in a time spent with no actual ore dropping into the ship.

2.4.1. Modeling Approach for the Loading Plan

The loading plan can be controlled in the simulation model through the use of a matrix variable that represents the compartments and the sequence of loading steps.

The incoming ships vary in terms of their capacity. Since the difference of loading procedure can be significant between a very small or a very large ship, two or more loading plans can be inputted to the model, each one to be applicable to a different category of ship size. Thus, the idea developed is to minimize the number of different loading plans inputted to the model, as opposed to inputting as many loading plans as there are ships to it.

Furthermore, in the approach designed, the compartments are to be filled with a fixed quantity of ore at each step of the loading plan. As the ships vary in terms of capacity, the number of steps will also vary according to the ship capacity.

Both the minimized number of different loading plans and the fixed quantity per step approaches are simplifications that accurately represent the loading process in the ore port, but allow reducing the amount of data inputted to the simulation model.

3. CONCEPT APPLICATION

3.1. The Case of Porto do Açu, LLX

LLX is a company that belongs to the Brazilian holding EBX, which administrates a portfolio of businesses in mining, logistics, oil and gas, real estate, traditional and renewable energy and entertainment. LLX focuses on the development of logistic infra-structure, especially in the portuary area. Porto do Açu is an LLX port, currently in the design and construction phase, developed to receive very large ships and to handle several different types of cargo, including containers and also ore from MMX / Anglo Ferrous mines in the state of Minas Gerais.

The focus of the simulation developed for the case of LLX was:

- Validating the designed conditions for the equipments and operational rules of the port, verifying if the volume loaded is in accordance with what is being budgeted. The operational times of ships are also object of validation.
- Verifying the impact on the overall indicators of operating two different products and not only one.
- Verifying the possibility of, in a longer future, increasing the port's capacity only by installing larger shiploader equipment.
- Verifying the impact on the overall indicators of operating larger ships vs. smaller ones.

3.2. The Major Challenges Emerge From the Objectives of the Simulation

The major challenges that have emerged from the simulation of Porto do Açu are all results of the need to represent operational rules and details that have a significant impact on the overall indicators that are the main object of the simulation.

The navigation restrictions have a direct impact on the time that each ship spends in each phase of the process in the port. This has an impact not only over the total time spent by each ship in the system, but also over the occupation rates of resources such as berths and the total volume loaded annually at the port.

Furthermore, the correct representation of the movements of stackers, reclaimers and shiploaders across yards and ship compartments is essential for the correct assessment of the real hourly capacity of these equipments. Had the time spent on these movements been very few, they would be irrelevant. However, these movements have consumed up to 15% of certain equipments available time in some simulated scenarios, which has a clear impact on the annual loading capacity of the port, as well as in these equipments´ occupation rates.

Moreover, the need to test the ports indicators in situations where more than one product is being handled makes it inevitable for the simulation analyst to deal with the problem of meeting required product mix. While this aspect had not yet been effectively taken in hand, in the model validation phase, the simulation developed for LLX presented several scenarios of deadlock results due to the excess of one type of product and the lack of another. It is also important to note that although the stocking areas capacities were represented in the model, the goal was not to check if the stocking area had been correctly sized. If this was so, there might have been necessary to face the product mix challenge in a different way.

In short, it must be said that the approach chosen to represent the process at Porto do Açu allowed the correct assessment of the main indicators of the simulation.

Since the port is not on duty yet, and are being built while this paper is written, the regular validation process was not used since there is no real data to be compared with the simulation results. So all results was evaluated by the port designers, whose experience on the process was used to criticize the results and detect wrong behaviors.

4. RESULTS

Having adequately represented key issues in the process at the ore loading port, there are still some challenges to be faced on the reporting of the results obtained by a simulation model, built in the ARENA discrete-event simulation tool. The adequate reporting of the simulation results allows the interpretation and analysis of the simulation model and leads to the answers of the questions primarily asked in the project.

This simulation study followed the process proposed by Valentin et.al. (2005) to simulate maritime infrastructure, were a simulation expert builds the model and perform the experiments. Then, the results are analyzed by the port expert.

4.1. Result Reporting Challenges

The main challenges on the reporting of the results obtained by the ore loading port simulation are a result of two circumstances:

- When one rule is composed of several sub-rules or conditions, it may occur that the behavior of a system in a certain point in time cannot be explained by the occurrence of non-occurrence of one of these conditions but by an ensemble of them. In the case of the port, the navigation rules of the ships leaving the system state that three conditions must be met at the same time: The tide must be adequate, the weather must be good and the channel must be free. In a certain point in time, if both the channel is occupied and the weather is bad, a ship is retained in the berth, but the time retained can be attributed neither to the channel not to the weather.
- When a simulation model takes in consideration the preparatory movements of equipments (to begin a task), such as it is the case for the stackers, reclaimers and shiploaders, the total time spent in an operation cannot be attributed solely to equipments capacities to process a given material, but also to equipments capacities to prepare to process a given material. Thus, the equipments capacities must be described as a pair of

capacities: the processing and the preparation capacities.

The solutions adopted to solve these reporting challenges were as described below.

4.1.1. Reporting Results of Navigation Conditions

Navigation rules at ports are, as stated before, a combination of different conditions that must be met. To report results related to navigation conditions, the approach used started by the determination of an hierarchy of conditions, from the strongest condition to the faintest. The strongest conditions are the ones which are very unlikely to be altered by process managers. In opposition, a faint condition is one that can be suppressed by a change in infrastructure.

In the case of ports, a good example of a strong condition would be the climate restriction for ship navigation, since humans have, we can say, no control over the weather. On the other hand, the channel restriction is quite faint, since there is always the possibility of excavating a second channel for a port.

The logic adopted is to attribute all of the time when the strongest condition was present directly to it. Since it is the strongest condition, it should be very hard to get back this time. In the opposite situation, the time attributed to the faintest condition will be the time when only this condition was present, which could be gained if the condition could be suppressed.

To illustrate this principle once again the example of the port outbound navigation is applicable. In this case, the conditions are: The weather must be good (strongest), the tide must be adequate (middle), and the channel must be freed (faintest). Thus, the waiting time to leave the port is reported according to the diagrams that follow.

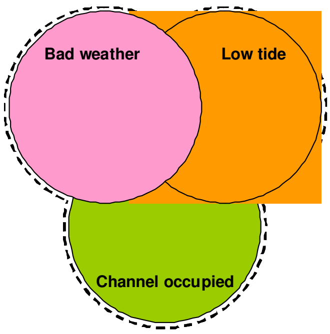


Figure 2: Outbound navigation waiting time reporting for ore handling ports

Table 1: Outbound navigation waiting times reporting for ore handling ports

Time reported	Colors
Total time	Total time
Bad weather time	Bad weather
Low tide AND good weather ->> Time that could be spared in case the surface of the ocean is dredged: useful to analyse the benefit of dredging the surface of the ocean.	Low tide AND good weather
Time of occupied channel only ->> Time that could be spared in case other channels are dredged: useful to analyse the benefit of dredging other channels.	ONLY occupied channel

4.1.2. Reporting Equipment Occupation

Equipment occupation, in a situation where the preparatory equipment movements are considered, must be reported separating these movements from the main equipment material processing. This separation allows the assessment of the impact of the two different classes of capacities of the equipments: the preparatory capacity and the processing capacity. Henceforth, it is possible to understand weather a better overall performance of equipment can be attained by increasing one or the other capacity.

In the ore handling ports, this difference becomes clear when we imagine a situation where 30% of the time of equipment is occupied with processing and 20% with preparatory movements, in an overall 50% occupation. In this case, if the processing capacity is doubled, a probable outcome would be a 15% occupation with processing and 20% with preparatory movements, summing up a global 35%. Although the processing capacity was doubled, the occupation did not drop by half. The analyst may explore, in these cases the possibility of increasing not only one of two capacities to obtain a better overall performance.

4.2. Results Obtained With the Simulation

The results obtained with the simulation showed that the overcome of the modeling challenges described in this paper were essential to the validation and the experimentation over the model of the ore handling port.

First of all, the complete treatment of the mixed product handling situation given to the model allowed a comparison of two extreme scenarios: a first scenario with only one product flow and a second scenario with a two product-demand for all ships. The demand of two products for a ship causes the shiploader equipment to perform additional preparatory movements, which increase dramatically the occupation of this resource. The chart that follows illustrates the occupation of resources in both scenarios.

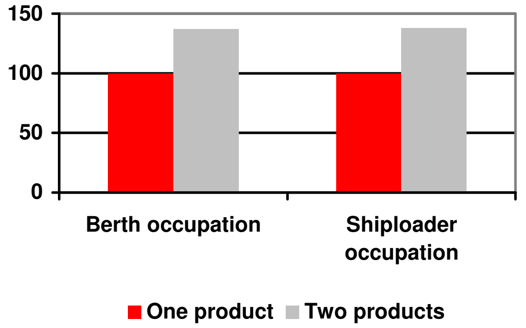


Figure 3: Occupation of resources in scenarios with one and two products per ship (indexed)

Also, the treatment given to preparatory movements allowed the comparison between two scenarios with different shiploader processing capacities. The comparison showed that a 30% increase in shiploader capacity provoked a 13% increase in maximum volume loaded annually.

Finally, the impact of operating larger ships vs smaller ones was the object of an assessment using the simulation model. The results showed that the use of larger ships in the port has very little impact on the volume of material that can be loaded annually. As ships get larger, the time spent by each ship in the port gets bigger, and the number of ships loaded per year gets smaller. In this situation, the occupation of berths does not grow significantly, and nor does the occupation of other resources such as shiploaders. Another way to read these results is to say that the variation of time spent by each ship at the port vs the size of the ships is fairly linear. In other words, the fixed time spent by ships in the port, which involves activities such as channel transit and docking, represents a very small percentage of the total operating time of each ship in the port.

5. CONCLUSIONS

In this paper the main challenges faced to simulate an ore handling port are described. These challenges consist of the development of an innovative approach for modeling certain situations, and may also apply to other simulation cases that present similar challenges.

The development of this approach has allowed an adequate representation of the operations of an ore port and the evaluation of several questions related to the port project.

The lessons learned on this study were:

- To simulate the process in high detail is not ever the best option. The ore piles representation considering every single slot of the yard and the pile geometry, gave a nice improvement on the precision. But at the cost of a hard work on coding the model. And, for a strategic study, the additional precision was not relevant. It would have better use on a tactical or operation study.
- When simulating a "non existing" system, the participation of specialists on the process is critical. In this case, the port designers were very experienced on other ports, ore yards, belt operations, ship queues and so on. It helped a lot to reach a reasonable conclusion for the study.
- The use of a good simulation tool is very important in these studies, when there is a phase of frequently updating the model code, while the port specialists pointed problems on the results. The flexibility and very nice documentation capabilities of the chosen tool, the Rockwell Software ARENA, was critical for the successful ending of the study.

ACKNOWLEDGMENTS

The Authors thanks the LLX by supporting this project and for authorizing the use of its information.

REFERENCES

Karlaftis, M. G., Kepaptsoglou, K., Sambracos, E. 2009. Containership Routing with time deadlines and simultaneous deliveries and pick-ups. In *Transportation Research Part E*, n.45, p210-221. ed. Elsevier.

Meisel, F., Bierwirth, C. 2009. Heuristics for the integrationof crane productivity in the berth allocation problem. In *Transportation Research Part E*, n.45, p196-209. ed. Elsevier.

Coelho, R.J.; Fioroni, M. M. 2005. Estudo de expansão do sistema de manuseio de matérias-primas da CST através de simulação. In: 60° Annual ABM Congress. Belo Horizonte, MG, 25 - 28 jul., 2005.

Fioroni, M. M., Franzese, L. A. G., Fúria, J., Perfetti, L. T., Silva, N. L. 2007. Simulation of Continuous Behavior Using Discrete Tools: Ore Conveyor Transport. In: Winter Simulation Conference, 2007, Washington DC.

Valentin, E. C., Steijaert, S, Bijlsma, R. A., Silva, P. 2005. Approach for Modeling of Large Maritime Infrastructure Systems. In: Winter Simulation Conference, 2005, Orlando FL.

AUTHORS BIOGRAPHY

MARCELO MORETTI FIORONI is a simulation consultant with an Electrical Engineering degree, MSc. in Manufacturing and PhD in Logistics at University of São Paulo (USP). Has participated in almost 150 successful projects with simulation. Co-founder of PARAGON Tecnologia in 1992, the pioneer and leading consulting company in simulation in South America. Teaches Simulation at Faculdades Metropolitanas Unidas (FMU) in São Paulo, Brazil. Has trained more than 1,200 professionals in simulation. He can be contacted by email at <marcelo@paragon.com.br>

CAIO EDUARDO ZANIN is a simulation consultant with an Electrical Engineering degree. Has participated in almost 10 successful projects with simulation and teaches simulation in Paragon. He can be contacted by email at <caio@paragon.com.br>

LUIZ AUGUSTO G. FRANZESE is a simulation consultant with a Production Engineering degree and MSc. in Logistics, who has completed almost 150 successful projects with simulation. Founded PARAGON Tecnologia in 1992, the pioneer and leading consulting company in simulation in South America. Has trained more than 1,200 professionals in simulation. He can be contacted by email at <augusto@paragon.com.br>

LUCIA DINI PEREIRA is a simulation analyst with a Production Engineering degree from the University of São Paulo (USP) and a diploma for specialized Logistics studies from École Centrale de Lyon, in France. Her background is in decision-making analysis and mathematical modeling in several organizations in Brazil. She can be contacted at <lucia@paragon.com.br>

JOSÉ ALEXANDRE SERENO QUINTÁNS is a simulation consultant. He has participated in almost 10 successful projects with simulation and teaches simulation in Paragon. He can be contacted by email at <alexandre@paragon.com.br>

ISAC REIS DE SANTANA is a simulation consultant. With Technical Degree, he has participated in almost 10 successful projects with simulation and teaches simulation in Paragon. He can be contacted by email at <isac@paragon.com.br>

PAULO SAVASTANO is a simulation consultant. He has participated in almost 10 successful projects with simulation and teaches simulation in Paragon. He can be contacted by email at <paulo@paragon.com.br>

VITOR LUCIANO DE ALMEIDA BENEVIDES Coordinator for the project of implantation of Porto do Açu. Mechanical Engineer degree from Universidade Federal do Espírito Santo – UFES. Doing masters degree in Management at the Fundação Getulio Vargas – RJ. Specialist in Internationl Commerce and Logistics by the same university. He can be contacted by email at < vitor.benevides@angloferrous.com.br>

OPERATIONAL PLANNING TOOLS FOR U.S. NAVY MARITIME COMMANDERS

Gerald G. Brown[(a)], W. Matthew Carlyle[(b)], David Kelton[(c)], Jeffrey Kline[(d)] and Javier Salmerón[(e)]

[(a)-(e)]Naval Postgraduate School, Monterey, CA, 93943, U.S.A.

[(a)]ggbrown@nps.edu [(b)]mcarlyle@nps.edu [(c)]wdkelton@nps.edu [(d)]jekline@nps.edu [(e)]jsalmero@nps.edu

ABSTRACT

This work presents recent research to develop a series of optimization- and simulation-based decision support tools to aid in operational maritime mission planning for integration into the Globally Networked Maritime Headquarters with Maritime Operations Center. We present three tools: Global Fleet Station Mission Planner recommends route and mission scheduling for a naval contingent embarked with training teams and sustainment requirements. Combat Logistic Force Planner recommends how to employ special transport ships that carry ship and aircraft fuel, ordnance, dry stores, and food, to sustain navy battle groups operating worldwide for 90-180 days. Container Port Simulation models individual container vessels approaching the west coast of the U.S., starting with their required notice of arrival 96 hours out in the Pacific and, in case of closure of the intended destination port, suggests an alternate destination port in accordance with the shipper's economic self-interest to minimize time and cost to the ultimate destination.

Keywords: maritime optimization models, port simulation models.

1. INTRODUCTION

The Operations Research Department at the Naval Postgraduate School has over fifty years of experience applying scientific, mathematical, and systems tools to planning military operations. In this paper, we introduce representative current planning tools in development by faculty and operationally-experienced military officer students. Specifically, we tap the most recent research in optimization and simulation to describe a series of decision support tools to aid in operational maritime mission planning for integration into a Globally Networked Maritime Headquarters with Maritime Operations Center (Department of the Navy 2008).

The specimen planning tools chosen for illustration are Global Fleet Station Mission Planner (GFSMP), "Navy Combat Logistic Force" (CLF) planner, and "Container Port Simulation Model" (CPSM). Spitz (2007), Brown and Carlyle (2008), and Pidgeon (2008), respectively, document earlier, prototypic versions of these planning tools, and references therein further document the needs leading to their development.

2. GLOBAL FLEET STATION MISSION PLANNER

2.1. Overview

GFSMP is a mission planning and scheduling optimization tool designed to provide fleet staffs the ability to examine the feasibility of future deployments and activities in support of the Global Fleet Station (GFS) concept, Maritime Domain Awareness, and the Global Maritime Partnership initiative.

GFSMP identifies how one naval ship with embarked Expeditionary Partnership Teams (EPTs) can best meet logistical and other requirements (such as berthing space, budget, re-supply needs, transit times, mission requirements, etc.) to provide navy strategic priorities in Maritime Security and Theater Security Cooperation (TSC).

A typical solution would include a deployment schedule and the combination of EPTs required to perform the missions. These results can guide planners in best utilizing current naval resources available in the region and provide insights for future planning, offering opportunities to understand where tradeoffs can be made.

GFSMP is applicable in all theaters using afloat basing to support engagement teams. Changing limitations on deployment time frame, ship or EPT availability, and budget, offer opportunities to understand where tradeoffs can be made. GFSMP can also be used by planners to understand how different ship types may be used to accomplish similar missions. Exploration into EPT constraints can help the navy to determine team requirements: GFSMP optimally allocates training teams to a ship, and schedules the ship route to achieve the maximum, aggregate TSC "value" for all executed missions. Here, mission value refers to an informed, quantitative assessment of contribution to TSC if the mission is performed.

2.2. The Problem

GFSMP posits a Navy vessel, equipped with EPTs, carrying out TSC activities in its area of responsibility during a continued period of time referred to as the planning horizon. GFSMP's primary objective is to devise an optimal route and mission schedule that maximizes the TSC value collected. As a secondary objective, GFSMP seeks to minimize the total port cost incurred while conducting the missions. (We use the term "country" and "port" interchangeably to refer to locations where the ship may conduct missions, and/or obtain fuel or provisions.)

During operations, the ship must maintain at least its minimum fuel level as well as maintain sufficient provisioning supplies for all personnel. A fixed amount of provisions is consumed every day, depending on the personnel deployed. However, the amount of fuel consumed, in barrels per day, depends on the ship's activity, and three burn rates are used: while underway (at a nominal transit speed), while at anchor (supporting "hotel load" power requirements), and moored pier-side receiving shore power. Travel times (in days) between any pair of ports are pre-calculated based on feasible sea routes and the nominal transit speed of the vessel.

Each ship incurs a port-dependent charge for each day it remains in-port. This charge also depends on whether the ship stays at anchor or docked. One or several fictitious "At Sea" ports represent locations at sea where the ship may, for example, conduct multi-country training missions, or have opportunities to schedule underway replenishments (of fuel or commodities) at select times when auxiliary fleet ships are present in the area (see CLF planner in Section 3).

Every candidate mission in each country has a pre-determined TSC value, a cost, a fixed duration, and can be completed by one (among potentially several) EPTs. Some missions have precedence requirements in relation to other missions such that one mission may require one or more other mission(s) to be completed before it can be carried out.

Every mission requires the ship to deliver a qualified EPT to the respective country to complete the mission and to pick up the team immediately upon completion of that mission. Additionally, some missions may require the ship to stay in-port for the duration of the mission. Also, some missions require the ship to be moored pier-side (e.g., in order to load or unload heavy equipment) during the first and last days of the mission, while other missions may allow the ship to remain at anchor. Each EPT can only participate in one mission at a time, but once finished it is assumed to become available to perform other missions immediately.

There is limited availability of each type of EPT. Each one has a varying number of personnel, and there is limited amount of berthing space to provide for all the personnel of all EPTs assigned to the ship. Thus, in addition to selecting the optimal routing and scheduling, GFSMP must determine the optimal EPT configuration for the deployment.

2.3. Modeling, Scenario and Results

GFSMP model's key decision variables are:

- EPT configuration, i.e., how many teams of each type should be embarked on the GFS
- Mission schedule by country
- GFS schedule: When the GFS stays in a given port (conducting missions or resupplying), or is underway between countries
- Fuel and supply levels

To evaluate GFSMP, we use notional data from the Fall 2007 Gulf of Guinea (GoG) African Partnership Demonstration developed by CNE-C6F GOG Regional Planning Team (Spitz 2007), which have also been utilized by Second Fleet during Exercise Trident Warrior 2009 (TW09). The original Demonstration assumed an amphibious ship LSD-43 and/or a High Speed Vessel HSV-2 as the platforms to accomplish 66 missions over six months. However, for TW09, LSD-43 is replaced by USS Kearsarge, LHD-3 (U.S. Navy 2009).

Figure 1 is an example of the recommended route and mission schedule for select dates. (The snapshot covers approximately two of the six month's deployment of the GFS.) Bars indicate the period of execution of each mission, with names inside indicating the team performing the mission (for example, "NCF" refers to a "naval construction force" team). We can observe, for example, that no more than the four available NCF teams are used, and that the GFS transit between ports while missions are still ongoing, in order to maximize mission accomplishment. Overall, GFSMP shows that the GFS can complete all missions in substantially less time (only five months) than originally planned by CNE-C6F (Spitz 2007), and at a much lower cost (by incurring in lower port costs).

In addition to its capability as a planning tool, during the TW09 Exercise GFSMP also proved useful by helping planners to deal with a variety of exigencies. These included rescheduling the GFS in the middle of the deployment due, for example, to the unexpected failure of a CLF asset originally scheduled to resupply the GFS, or to new requirements for the GFS to engage for an extended period (several weeks) in an anti-piracy emerging situation. These disruptions to the baseline schedule may require rescheduling or cancelling some of the original missions.

As U.S. Navy CAPT Lansing, Director, Experimentation Directorate, Navy Warfare Development Command states, "the use of GFSMP proved to be very timely as there is currently no other automated aid that can produce optimized courses of action for a Global Fleet Station deployment. GFSMP had a significant impact on the operational planning of Operational Level Command and Control and the [TW09] Experiment as a whole."

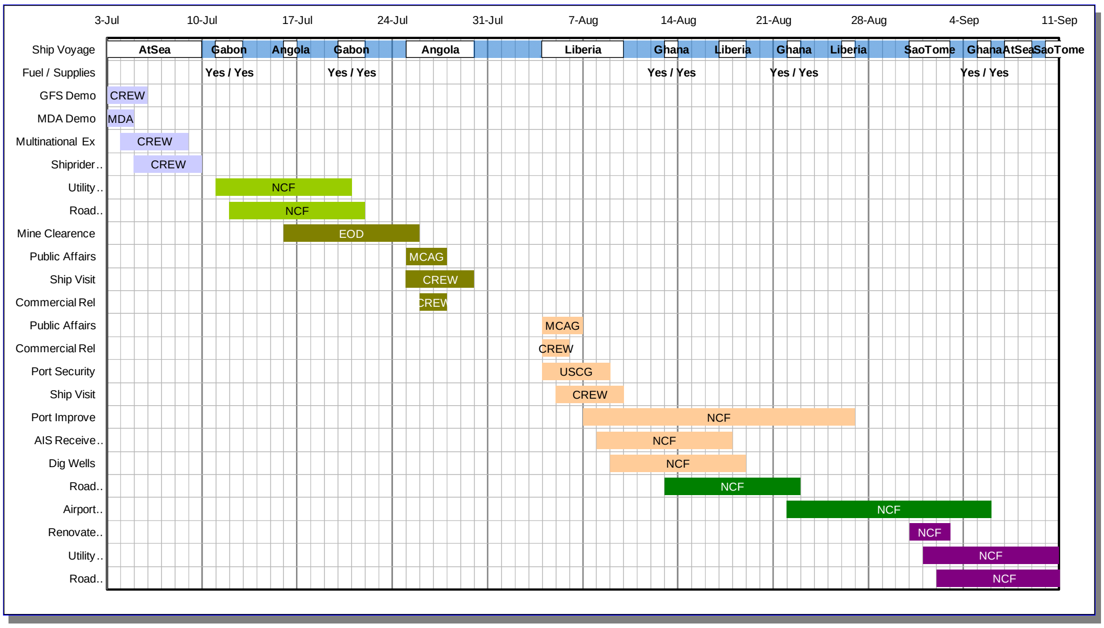


Figure 1: This is a snapshot of, approximately, two of the six months for which GFSMP plans the routing and scheduling of an LHD in the Gulf of Guinea. During the recommended deployment, the LHD transits through different ports ("Ship Voyage" line on top) while accomplishing missions listed on the left. Bars indicate the execution period of each mission and the performing team.

3. COMBAT LOGISTIC FORCE PLANNER

3.1. Overview

The Navy CLF consists of about 30 special transport ships that carry ship and aircraft fuel, ordnance, dry stores, and food, and deliver these to client combatant ships underway, making it possible for U.S. naval forces to operate at sea for extended periods.

The CLF is being transformed to a fleet with fewer different types of transports, and no more total transports, but it expects to have to serve more clients for a greater variety of deployments in the next decade. Conventional planning has relied on steady-state, average-rate-of-consumption models and rules-of-thumb to assess CLF ability to re-supply our fleet operations. Details matter, and we want to determine whether or not, and how, the new CLF can actually support its anticipated missions.

We have modeled CLF operations to evaluate a number of transforming initiatives that simplify its operation while supporting an even larger number of client ships for a greater variety of missions. Our input is an employment schedule for navy battle groups of ships operating worldwide, extending over a planning horizon of 90-180 days. We use optimization to advise how to sustain these ships.

Most U.S. Navy deployments are groups of ships assembled with a particular mission. Some frequent examples of these "Battle Groups" (BGs) are a Carrier Strike Group (a nuclear-powered aircraft carrier, CVN, a guided-missile cruiser, CG, two guided-missile destroyers, DDG, and a fast combat replenishment ship), an Expeditionary Strike Group (an amphibious assault ship, LHA or LHD, with a dock landing ship, LSD, amphibious transport dock, LPD, a CG and two DDGs); a Surface Strike Group (ships equipped with missiles and missile defense weapons, such as a CG and two DDGs); and a Littoral Combat Squadron (new class of small ships where larger ships cannot safely navigate, engaging in anti-surface warfare, mine counter measures, intelligence, surveillance and reconnaissance, homeland defense and maritime interdiction, and special operations forces support, each of which we treat as a frigate FFG). Combatant classes are available from Jane's (2008, pp. 873-943).

The Combat Logistics force is being consolidated to just three ship types, with 30 total ships: The TAO187 (Henry J. Kaiser) which can carry about 180,000 barrels of fuel oil, and 271 tons of cargo lube oil, dry stores, and refrigerated containers, at about 20 knots; The T-AOE6 (Supply) Class, which can carry 156,000 barrels of fuel oil, 1,800 tons of ordnance, 250 tons of dry stores, and 400 tons of refrigerated stores, at speeds exceeding 26 knots; And, the T-AKE1 (Lewis & Clark), which can carry 18,000 barrels of fuel oil, 5,900 tons of ordnance, dry stores, and refrigerated stores at up to 20 knots. Customarily, each storage hold is designed to carry either ordnance or dry stores, but T-AKE storage holds can be converted between the two. Other ship classes (e.g., T-AE, T-AFS, T-AOE(X), etc.) are either scheduled to leave active service for the reserve fleet, or are still on the drawing boards.

We evaluate new CLF ship designs, advise what number of a new ship class would be needed, test concepts for forward at-sea logistics bases in lieu of conventional ports, demonstrate the effects of changes to operating policy, and generally try to show whether and how the CLF can support planned naval operations.

3.2. Modeling, Scenario and Results

We use an integer linear program to plan optimal employment of CLF ships to minimize policy penalties accruing from any commodity shortage. Decision variables for the CLF model account for:

- Indicators of whether or not at least one shuttle visits a given BG on a particular day, and the port from where the shuttle is coming
- Taking stock of each shuttle's daily commodities, and the amounts delivered to a BG
- Taking stock of each BG's commodity levels, and categorizing these by levels of deficiency (if any)

The CLF model can schedule a single shuttle ship sortie from port to make many separate consolidation (CONSOL) visits, perhaps to different battle groups.

Our model takes into account changing daily consumption of each of four basic commodities classes for each client ship, navigational issues such as slow passages through canals, and the possibility of several client ships, or groups of ships, running low on the same commodity at the same time. Given a deployment scenario and a current configuration of the CLF fleet of transports, a solution to our model is a face-valid logistical plan for the CLF that minimizes shortfalls of any commodity for any customer, highlighting any unavoidable low-inventory events, and maximizing the utilization of transports by maximizing the total volume delivered over the scenario.

A typical scenario consists of multiple BGs, and for each BG specifies last-minute in-port preparations and/or pre-deployment workup training in preparation for deployment, a high-speed transit to an area where we will show our military presence, a surge into combat operations to achieve a given objective, a sustainment phase to hold that objective, and perhaps a post-combat period where we stand guard and provide humanitarian assistance while diplomacy and other non-military measures unfold.

An example planning scenario includes 13 battle groups served by nine TAO and seven T-AKE shuttle ships over a 90-day planning horizon. Some of the larger battle groups are accompanied by station ships (supply ships that serve as accompanying inventory vessels). Five of the battle groups are deployed on day one, and the rest deploy on day 10, 12, etc., until all are underway to their various areas of operation worldwide. Figure 2 depicts totally automatic world sea-route model for estimating times and positions of optimally-deployed CLF ships worldwide.

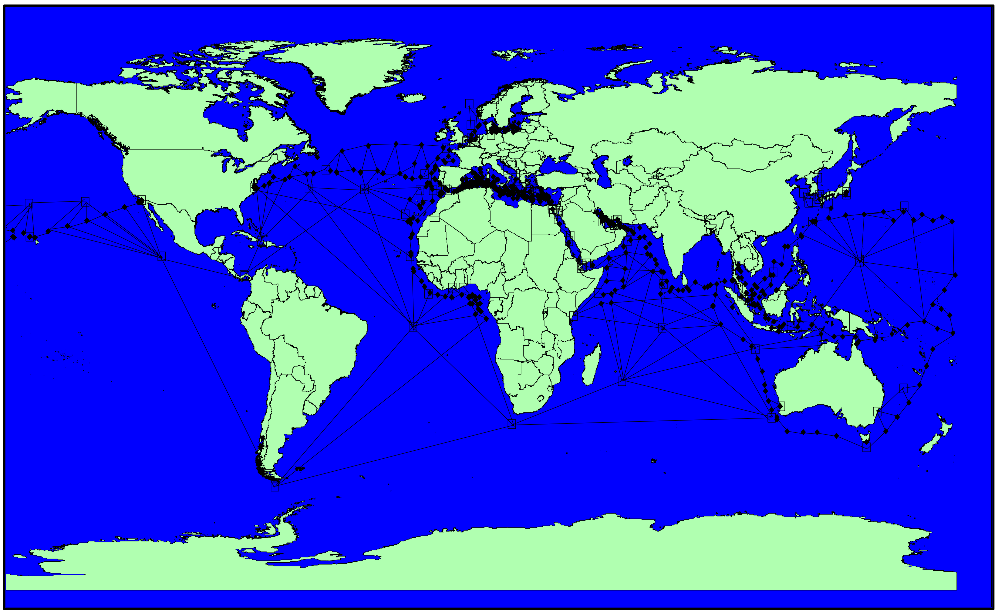


Figure 2: Worldwide sea route network. The static searoutes shown here (as a connected network of ports and at-sea location nodes and arcs indicating navigationally adjacent nodes) connect all ports and trans-oceanic routes a CLF ship might use at the discretion of the optimization (that the shuttle movements are *outputs* rather than inputs is a key innovation of CLF). Transit of most arcs can be made at any speed desired, but some arcs (e.g., canals) require a fixed transit time. CLF merges the tracks of combatant customers with these static searoutes to produce the full, worldwide maneuver network. When needed, local operating areas can be embellished with even more at-sea locations and arcs for finer fidelity. Each at-sea location can be coded with an administrative restriction that permits access only to compatibly-coded CLF ships. Each port may accommodate only a subset of CLF ships by type or by hull number, and each port may offer only a subset of commodities.

A monolithic (i.e., omniscient) solve of this situation generates an optimization model with about 367,000 constraints and about 23,000 variables. A cascade (i.e., myopic) solve using a 30-day planning windows advanced 15 days at a time yields a sequence of problems with between about 37,000 and 75,000 constraints and between 6,000 and 9,000 variables. The monolith takes several hours to solve on a 2-GHz laptop, while the cascade takes a few minutes. For this particular case, the two solutions compare closely. The CLF fleet conducts a total of 101 CONSOLs (i.e., deliveries). The availability in this scenario of forward (i.e., close to the areas of operation) supply base ports is key: our average cycle times from a CONSOL back to a port and on to the next CONSOL is only just over five days, two of which are spent pierside loading the shuttle ship.

A cumulant inventory constraint, combined with an elastic violation device for a shortage, and another for extremis shortage, combine to signal a deficiency, and carry this forward, paying a daily penalty, until this deficiency is remediated. The distinction between shortage and extremis penalties is important: CLF ships are equitably assigned to CONSOL all customers by maximizing deliveries, to preferentially serve needy customers by CONSOL to avoid shortage, and to energetically serve customers in extremis.

CLF (Brown and Carlyle 2008) is a strategic planning model to assess our ability to supply about a dozen large carrier battle groups. The Trident Warrior 09 Exercise includes many operations in the Gulf of Guinea, a long way from resupply ports, and CLF has been generalized to be an *operational* planning model serving individual ships worldwide. CLF shuttle employments can now include unlimited consolidations between port calls for reloading ─so may consolidations that a shuttle ship may end up resembling a station ship (i.e., a ship that accompanies a deployed battle group), though our shuttles may switch customers as advantageous). There are now many more ports, and each port may only offer some subset of commodities. Deployed combatants can now be scheduled to make port calls for resupply pierside. To enhance operational planning, CLF now runs faster, and features a graphical user interface offering geographic views of the world, hemispheres, or operating areas with high close-up resolution. The display now includes animation of moving shuttles and their moving customers to better help planners understand the prescribed ballet. CLF became a centerpiece for TW09, providing quick, trustworthy advice in response to the exigent events that inevitably disrupt operational plans already in execution.

4. CONTAINER PORT SIMULATION MODEL

CPSM models the effects of a transportation security incident on one or more of the U.S. West coast marine container ports. Operation of the container ports on the U.S. west coast is critical to ongoing national commerce. We built a simulation model of the seven major West-coast container ports to study their ability to process container throughput, and especially to measure what would happen system-wide were one or more ports taken out or degraded due to a natural or human-caused event.

We modeled individual container vessels starting with their Notice of Arrival 96 hours out in the Pacific. Vessels then travel to their intended port, or to an alternate port if the intended port is closed. In the case of closure of the intended port, an alternate port is advised in accordance with the shipper's own economic self-interest, with an eye toward minimizing time and cost to the ultimate destination. Once at a port, either intended or diverted, the vessel unload time is accounted for, and the shipment is broken into ten pieces bound for each one-digit ZIP code in the continental United States. These landside shipments then travel to their designated aggregate destinations.

We have collected data on vessel arrival patterns by intended port, unload time, port capacities (berths), landside travel times, and various costs, including demurrage costs for freight. These real data were then used to specify inputs to the model, both deterministic constants like distances and speeds, as well as probability distributions from which uncertain elements like arrival times, number of containers on a ship, and unload times were generated.

We have built several alternate versions of the model, treating different scenarios of port transportation security incidents. We built models both with and without the proposed port at Punta Colonet, Mexico, to see how the presence of that port might help maintain operations in the face of U.S. port closures. The model was run for a one-year planning horizon in each of several configurations, with thousands of replications to assure adequate statistical precision. The model has been streamlined to be general and scalable in the number of ports, and an animation has been used to help with model verification and establishing credibility CPSM was validated by modeling the ten-day closure in 2002 when members of the International Longshoremen's and Warehousemen's Union (ILWU) were locked out in retaliation for a labor slowdown on the part of the union.

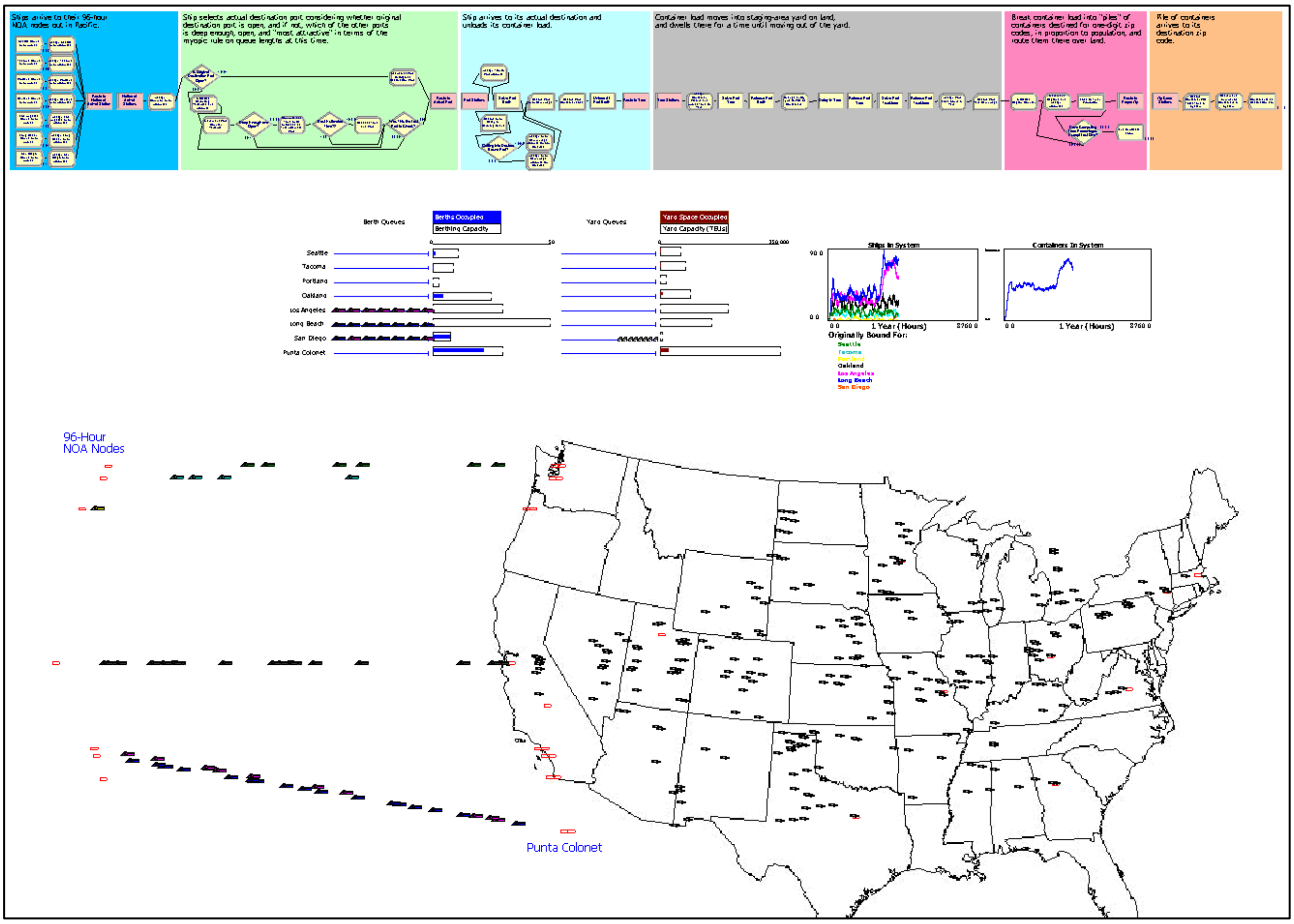


Figure 3: Screenshot of Arena CPSM (top flowchart) and animation (bottom) of West-coast container-port operations. At the time of this screen shot, both the ports of Los Angeles and Long Beach are closed, so vessel traffic is being diverted to other ports, mostly to the proposed port in Punta Colonet, Mexico.

Figure 3 is a screenshot of the model implemented in the Arena simulation software (Kelton, Sadowski, and Swets 2010) to illustrate both the logic (the flowchart at the top) and the animation at the bottom. The logic flowchart at the top proceeds from left to right: the first blocked-out area is arrival of ships to their Notice of Arrival positions, the next block is the port-selection logic including choosing the next-best port if the intended port is closed, next is unloading operations at the berth (including any necessary queueing), next is landside storage at the port, next is overland transit, and finally arrival of a shipment to its destination ZIP code. The map animation at the bottom shows movement of ships as well as the ZIP code loads once the ship is unloaded and broken into ten loads. The queue and plot animations in the middle depict queues of ships for berths, utilization levels of berths, and graphs of both the number of ships and number of containers in the system at any given time. This particular situation shows a point near the middle of the year a few weeks after the ports of Los Angeles and Long Beach were taken out, so ships bound for those ports are being diverted mostly to Punta Colonet to the south, and the graphs show significant spikes of both the number of ships and number of containers in the systems, representing increased congestion and delays due to the closure of these two major ports. Later in this scenario, the ports of LA and LB were gradually brought back up in groups of berths at a time, and congestion was eventually relieved as these full-capacity ports eventually worked off the backups. Other scenarios without the availability of Punta Colonet resulted in even worse congestion, delays, and costs. In this way, the model quantifies the effects of port closures, and can thus be used as a decision-making aid.

5. CONCLUSIONS

In this paper we have presented the GFSMP, CLF and CPSM optimization- and simulation-based decision support tools. We have illustrated application of GFSMP and CLF to exercises developed by U.S. Commander, Second Fleet, exhibiting significant improvements over manual planning. As these tools continue to be used by Navy staffs and modified with their suggestions, the longer-range intent is that they are integrated into a suite of decisions aids supporting maritime planning staffs.

ACKNOWLEDGEMENTS

The authors thank the Office of Naval Research for its support of this research.

REFERENCES

Brown, G., and Carlyle, W.M., 2008, Optimizing the US Navy's Combat Logistics Force. *Naval Research Logistics*, **55**, 800-810.

Department of the Navy, 2008, Maritime Headquarters with Maritime Operations Center. Online: http://www.doncio.navy.mil/EATool/Forcenet/mhqmoc.htm (accessed April 2009).

Jane's, 2008, *Fighting Ships 2007*-2008. 110th Edition. Cambridge University Press, U.K.

Kelton, W.D., Sadowski, R.P., and Swets, N.B., 2010, Simulation With Arena, 5th edition. McGraw-Hill, New York.

Pidgeon, E., 2008, *Modeling the Effects of a Transportation Security Incident upon the Marine Transportation System*. Thesis (M.S. in Operations Research). Naval Postgraduate School, Monterey, CA.

Spitz, G., 2007, *Mission Resource Allocation in the Gulf of Guinea*. Thesis (M.S. in Operations Research). Naval Postgraduate School, Monterey, CA.

U.S. Navy, 2009, USS Kearsarge. Online: http://www.kearsarge.navy.mil/default.aspx (accessed April 2009)

AUTHORS BIOGRAPHY

Gerald G. Brown is a Distinguished Professor of Operations Research at the Naval Postgraduate School, a member of the National Academy of Engineering, a recipient of the United States Navy Superior Civilian Service Award, an INFORMS Fellow, and a founding Director of Insight, Incorporated, the leading provider of strategic supply chain optimization-based decision support tools to the private sector.

W. Matthew Carlyle is an Associate Professor in the Operations Research Department at the Naval Postgraduate School. He joined the faculty in 2002, after five years as an Assistant Professor in the Department of Industrial Engineering at Arizona State University. He received his Ph.D. in Operations Research from Stanford University in 1997, and his B.S. in Information and Computer Science from Georgia Tech in 1992. His research and teaching interests include network optimization, integer programming, and network interdiction. Applications of this research have included attack and defense of critical infrastructure, delaying large industrial projects and weapons programs, theater ballistic missile defense, sensor mix and deployment, naval logistics, communications network diversion, underground mining, and semiconductor manufacturing

W. David Kelton is Professor of Quantitative Analysis and Operations Management at the University of Cincinnati, and Visiting Professor of Operations Research at the Naval Postgraduate School. He received a BA in mathematics from the University of Wisconsin-Madison, an MS in mathematics from Ohio University, and MS and PhD degrees in industrial engineering from Wisconsin. He was formerly on faculty at Penn State, the University of Minnesota, The University of Michigan, and Kent State. Other visiting posts have included Wisconsin, the Institute for Advanced Studies in Vienna, and the Warsaw School of Economics. His primary interests are in simulation modeling and analysis. He has co-authored two books on simulation, and was Editor-in-Chief of the INFORMS Journal on Computing for seven years. He is a Fellow of both INFORMS and IIE.

Jeff Kline, CAPT, USN (Ret.), is a Senior Lecturer in the Operations Research Department at the Naval Postgraduate School and Navy Warfare Development Command Chair of Warfare Innovation. He has over 26 years of extensive naval operational experience to include commanding two U.S. Navy ships and serving as Deputy Operations for Commander, Sixth Fleet. In addition to his sea service, Kline spent three years as a Naval Analyst in the Office of the Secretary of Defense. He is a 1992 graduate of the NPS's Operations Research Program, where he earned the Chief of Naval Operations Award for Excellence in Operations Research, and is a 1997 distinguished graduate of the National War College. Jeff received his B.S. in Industrial Engineering from the University of Missouri in 1979. His teaching and research interests are joint campaign analysis and applied analysis in operational planning.

Javier Salmerón is an Associate Professor of Operations Research at the Naval Postgraduate School, where he teaches several courses, and conducts applied research and student thesis advising in optimization. Before his time with NPS, Dr. Salmeron taught at the University Complutense of Madrid, Spain, and worked as a project engineer at the Department of Decision Support Systems for the electric utility Iberdrola in Madrid. Dr. Salmeron's most recent research has been sponsored by the Department of Homeland Security, the Department of Energy, the Office of Naval Research and the Joint Improvised Explosive Device Defeat Organization.

PORTSIM 6.0: A PORT SIMULATION MODELING MULTIPLE MODES OF OPERATION

Reejo Mathew[(a)], Thomas W. Mastaglio[(b)], James F. Leathrum, Jr.[(c)]

[(a)]Research Scientist, MYMIC LLC, Portsmouth, VA 23704, USA
[(b)]President & CEO, MYMIC LLC, Portsmouth, VA 23704, USA
[(c)]Associate Professor, Old Dominion University, Norfolk VA 23529, USA

[(a)]reejo.mathew@mymic.net, [(b)]tom.mastaglio@mymic.net, [(c)]jimleathrum@cox.net

ABSTRACT
The PORTSIM tool is used to analyze the movement of military cargo through seaports used by the Defense Transportation System (DTS). The current development version, PORTSIM v6.0, assists analysts within the Surface Deployment and Distribution Command – Transportation Engineering Agency (SDDC-TEA) to compare and select ports for military cargo movement based on their capabilities and develop deployment/redeployment strategies. It supports modeling both the Seaport of Debarkation (SPOD) and Seaport of Embarkation (SPOE) modes of operation within a single model. It can also model concurrent SPOD and SPOE flow through a single seaport. PORTSIM supports programmable processes, allowing analysts to input processes at simulation run-time without requiring a software modification or recompile. PORTSIM v6.0 is an important tool in the decision-making process, used to improve the efficiency and throughput of military deployment or redeployment operations. The capabilities of PORTSIM v6.0 are also flexible allowing the modeling of any cargo terminal (airport, rail yard, etc).

Keywords: Discrete-event simulation, Port simulation, Programmable processes, decision support system.

1. INTRODUCTION

PORTSIM (Port Simulation) is a discrete-event simulation used to analyze the movement of military cargo through worldwide seaports used by the Defense Transportation System (DTS). The PORTSIM tool supports analyses by the Surface Deployment and Distribution Command – Transportation Engineering Agency (SDDC-TEA), component of USTRANSCOM at Scott AFB, IL. PORTSIM assists SDDC-TEA analysts study cargo throughput and latency within a seaport, identifying bottlenecks in terms of resources and infrastructure allowing them to refine their strategies for movement of a set of materiel through a selected seaport.

The current development version, PORTSIM v6.0, developed by MYMIC LLC, integrates the various components of the PORTSIM tool, some developed by other organizations, into a single integrated model, while providing major enhancements in capabilities over previous versions. PORTSIM v6.0 models the resources and infrastructure within a seaport (including terminals and operational areas), and supports the configuration of an individual seaport according to its characteristics. PORTSIM models the competition for resources (cranes, drivers, etc.) and infrastructure (berth space, staging space, etc.) by cargo and transports within a seaport. PORTSIM supports programmable processes. Processes are provided as a run-time input to the simulation. Previous PORTSIM versions captured processes within the simulation software code, requiring a rewrite and recompile whenever modifications were required. PORTSIM supports modeling different modes of operation, such as Seaport of Debarkation (SPOD) and Seaport of Embarkation (SPOE) within a single model. It also supports concurrent SPOD and SPOE flow through the same seaport.

Section 2 identifies the problem under study, Section 3 provides a complete description of the PORTSIM system architecture, Section 4 identifies other application areas for PORTSIM and Section 5 briefly describes related work in this field.

2. PROBLEM STATEMENT

Transportation logistics planning is used extensively to prepare for the movement of military cargo, i.e., troops, equipment, and supplies, from a military installation to a theater of operations. The initial input of troops into a theater of operations is usually by air. After this initial input, the movement of military cargo, both to support military operations and to sustain the existing force, is by sea, usually through a commercial seaport. For a variety of reasons, both political and economic, only a portion of an entire seaport is available for military operations. Planning for these operations involve modeling the Seaport of Embarkation (SPOE) mode, which includes ships being loaded at a seaport with cargo arriving via convoys, trucks and trains as shown in Figure 1. These ships are then transported to a distant seaport where cargo is offloaded using the Seaport of Debarkation (SPOD) mode of operation, shown in Figure 2. Previously, separate tools were used by SDDC-TEA to study these 2 modes of operation and were inhibited by the lack of inter-operability between the tools, the disparate requirements for the input data, the level of fidelity within the tools and the output data produced after a simulation run. PORTSIM v6.0 provides SDDC-TEA with the ability to concurrently study the SPOD and SPOE modes of operation, using the same level of fidelity and identical input and output data formats.

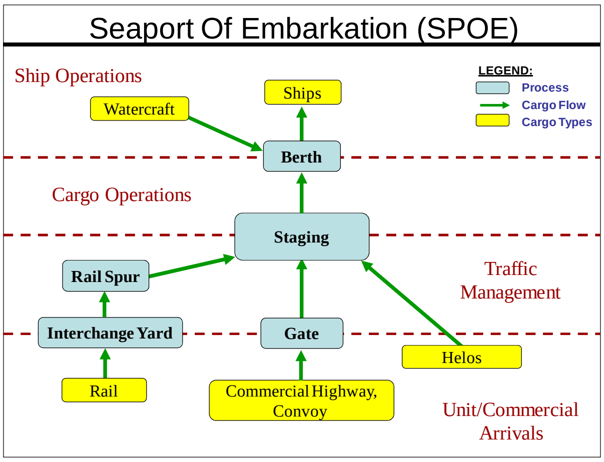


Figure 1. Seaport of Embarkation (SPOE) Operation

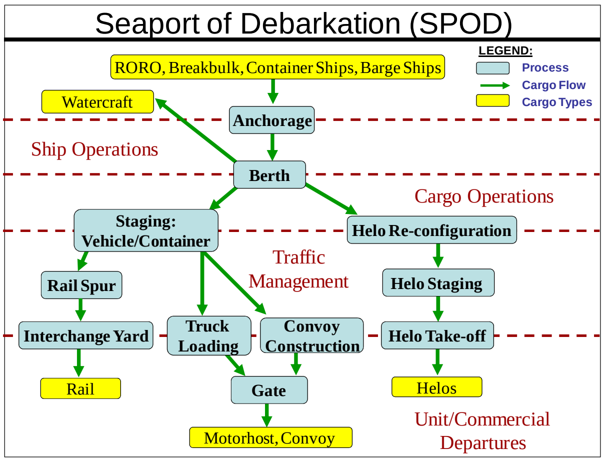


Figure 2. Seaport of Debarkation (SPOD) Operation

Moreover, for extended military operations, it is possible that the flow of cargo from the theater of operations back to an installation will begin before the flow into the theater of operations has ended, with some military units starting returning to the U.S., while other units are still being deployed. Additionally, following the initial surge of military equipment and personnel into a theater, cargo supplies needed to sustain the existing force often constitute a major portion of the continuing cargo flow. This results in a bi-directional flow of cargo and resources, with cargo moving in both directions competing for limited resources and infrastructure within a given seaport. Proper resource and infrastructure allocation is important in any transportation logistics operation. It is critical to ensure that the flow of cargo in one direction does not unnecessarily delay, or worse, block the flow of cargo in the opposite direction (Mathew 2002). Therefore, resource and infrastructure allocation between the concurrent flows of cargo, as well as the resolution of any contention that might arise, is vital. PORTSIM v6.0 facilitates this type of study condition through its ability to model concurrent SPOD and SPOE operations within a single port. This allows analysts to study the operational requirements for both modes and the impact of simultaneous execution of one mode on another.

3. MODEL DESCRIPTION

This section provides a complete description of the PORTSIM model. The model was developed to satisfy the following requirements identified by SDDC-TEA:

- ***Nodal Cargo Terminal Architecture:*** An individual cargo terminal should have the ability to be initialized independent of other cargo terminals. The cargo terminals should also be reconfigurable, that is, the infrastructure and the resources in the cargo terminal as well as the sequence of cargo flow through the areas in a single terminal should be changeable. The model should also support the initialization of multiple cargo terminals simultaneously as well as the flow of cargo between them, enabling the simulation of an end-to-end cargo flow from point of origin to destination.
- ***Programmable Processes:*** The processes within the model should be programmable by an analyst.
- ***Concurrent SPOD and SPOE operations:*** The model should support concurrent POD and POE operations through a single cargo terminal.
- ***Non-proprietary simulation executive:*** The model should utilize a non-proprietary simulation language as its underlying executive.
- ***Simulation of 1,000,000 pieces of cargo:*** The model should support the simulation of over 1,000,000 pieces of cargo within a reasonable execution time.

Moreover, SDDC-TEA identified the following performance requirements:

- Operate on a PC, minimally 2 GHz with 2 GB of RAM.
- Handle at least 30 multiple runs on the same scenario for statistically meaningful results.
- Capable of executing experiments involving on the order of 150 variations.
- No significant performance degradation compared to legacy versions.

The performance requirements mandated that a single PORTSIM run, with a standard cargo load of about 25000 pieces of cargo take no more than 19.2 seconds.

The PORTSIM v6.0 tool is shown in Figure 3. The tool accepts scenario data from different databases. The *Transports* database provides information about trucks, trains and ships, while the *GIS* database provides information about the seaport under study. Process flows, developed using the PPFN language (described in Section 3.2), are provided as input during simulation run-time.

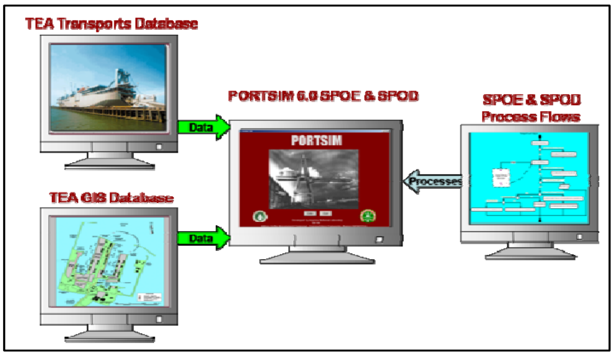


Figure 3. PORTSIM v6.0

The PORTSIM system architecture is shown in Figure 4. The characteristics of the seaport, including operational areas, infrastructure, and resources are provided as input to the simulation along with routing information to support simulating the movement of cargo through the seaport via XML files. Process flows are also passed to the simulation kernel via text files.

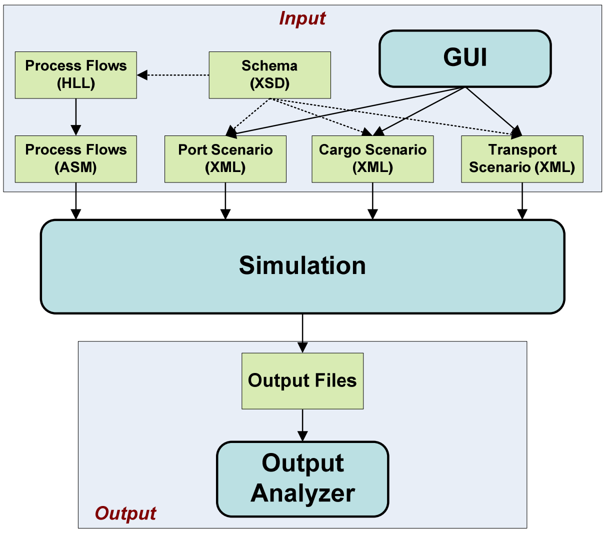


Figure 4. PORTSIM System Architecture

3.1. Nodal Cargo Terminal Architecture

The architecture is defined as a network of nodes, each node defined by processes. A node could also potentially represent a whole other network of nodes. This allows a hierarchical structure where individual nodes are defined either by a new network (network nodes) or by the internal processes performed on the cargo and transports. The capacities and capabilities of individual nodes are programmable, and so are the processes within them. The architecture is configurable to support different port types, different individual ports, different cargo types, and different transport types. A network node may define a seaport, a terminal within a seaport, or an operational area within either the seaport or a terminal. Sets of terminals in a seaport could handle ports that are made up of individual terminals, for example the Port of Hampton Roads which itself includes multiple terminals, such as Norfolk International Terminal, Portsmouth Marine Terminal, Lambert's Point Docks, etc. Each terminal itself is defined in terms of port areas: the berths, staging, and loading areas within a seaport. Multiple network nodes, each modeling a complete seaport in itself, can be connected together to model an end-to-end cargo flow from a point of origin to a destination.

The architecture also supports routes through the network. A route is defined by a starting node, a destination node, and a sequence of nodes traversed en route. An important approach to avoid infinite queuing in the interconnection segments is that a node must request the next node in the route before advancing. This provides a level of flow control to aid in the avoidance of deadlock.

3.2. Programmable Processes

Processes within PORTSIM are programmable, unlike many applications where processes are hard coded within the simulation. The Programmable Process Flow Network (PPFN) supports defining processes that are input into PORTSIM during simulation run-time (Mathew and Leathrum 2008). PPFN is a process flow programming language that describes the process activities at a cargo terminal. Previous versions of PORTSIM were implemented in the MODSIM simulation language and processes were an integral part of the simulation code. Once the processes were programmed and the code compiled, the processes were static in nature and any change in processes required modification to the simulation code; this could not be done by an analyst. PPFN provides an analyst the ability to modify processes independent of the simulation and provide it as an input to the simulation in the same way other data is input; this makes the tool more responsive to changes in requirements. Moreover, port Subject Matter Experts (SMEs) view operations within a cargo terminal in terms of processes that must be performed in a given situation. PPFN provides an intuitive approach to capturing these processes for PORTSIM. PPFN provides a process-oriented interface (in terms of flowcharts) for a discrete-event simulation, with processes being defined as sequences of activities and decisions.

PPFN is partitioned into a high level language and an assembly language. Processes are developed using the high level language, which contains programming control structures commonly found in standard programming languages. There is usually a one-to-one relation between the activities in the flowchart and the structures within the high-level process flows. The high-level language is translated into assembly language instructions by a *Translator*; these correspond to data structures internal to the simulation. During run-time, the assembly-level instructions are provided as input to the simulation; they are interpreted by a *Virtual Machine* within the simulation kernel. The *Virtual Machine* interacts with the cargo terminal architecture within the underlying simulation through a static declaration file defined in the XML Schema Definition (XSD) format; it contains definitions for types, modes, resources, and infrastructure. This structure is shown in Figure 5. The programmable processes provide analysts complete control of the simulation processes, enabling PORTSIM execution given an appropriate set of process instructions. This approach also allows PORTSIM to model military or commercial logistics, at any cargo terminal.

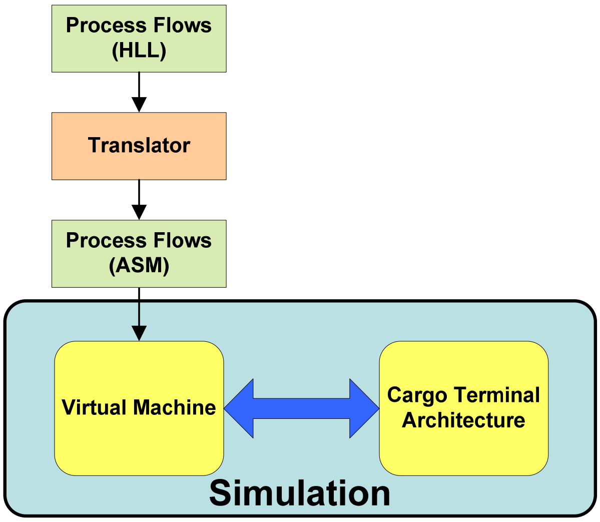


Figure 5. PPFN Structure

3.3. Programmatic Events

Programmatic events allow an analyst or an external disruption model to program unusual events to occur at specific simulation times. These events can change the available resources, capabilities of infrastructure and processing and transit times (specified as simulation input) within a seaport during simulation run-time. For example, an analyst can program a new berth to become available on day 20, instead of day 0. This allows PORTSIM to interface with other disruption models, to portray the effects of disruptive events on seaport operations, i.e., allowing the capability of the seaport to be degraded as a result of a disruptive event and then be gradually restored over time (Leathrum, Mathew, and Mastaglio 2009).

3.4. Output Analysis

PORTSIM models cargo and transports at the individual entity level and the data collected during a simulation run (or iteration) can be used for extensive after-action reporting and visualization. Data capture within PORTSIM is completely programmable and is part of the processes input at simulation run-time. Data can be captured at the entity-level as well as at an aggregate level. This data is captured during a simulation run using text, .csv and .xml files and passed to the PORTSIM Output Analyzer. The Output Analyzer generates various charts and graphs used by analysts to study the results of an iteration. The Output Analyzer currently captures information about the following:

- Throughput for the seaport, terminal and individual operational areas.
- Average processing times for cargo and transports in individual operational areas.

Since data capture is programmable, any relevant data can be captured during PORTSIM execution and the Output Analyzer can be modified to process and display the new data.

For any simulation, sufficient independent and identically distributed (IID) iterations of a particular scenario needs to be performed (30–60 runs are normal) to ensure the precision and reliability of the statistics being calculated. PORTSIM is a terminating simulation and inherently supports a multiple iteration capability so that a single scenario can be executed any number of times. PORTSIM captures pertinent information about each individual iteration and the total number of iterations as a whole. For every new iteration, PORTSIM calls a new instance of itself before terminating and passes on the last random number drawn from the predefined random number stream thereby continuing the random stream.

3.5. Simulation

PORTSIM is capable of simulating over a million pieces of cargo at an entity level of detail. The PORTSIM simulation kernel is a high-performance simulation model capable of simulating the flow of over 25,000 pieces of cargo through a seaport in under 6 seconds and a million pieces of cargo in under 50 minutes, comfortably meeting its performance requirements. PORTSIM execution times for different cargo loads are shown in Figure 6. The execution times are nearly linear, which implies that these times are a directly related to the number of pieces of cargo being processed. The PORTSIM simulation kernel was developed as an ANSI Standard C++ implementation. The underlying simulation engine used is a C++ implementation of the Distributed, Independent-platform, Event-driven, Simulation Engine Library (DIESEL) (Mathew 2007), a non-proprietary simulation executive model.

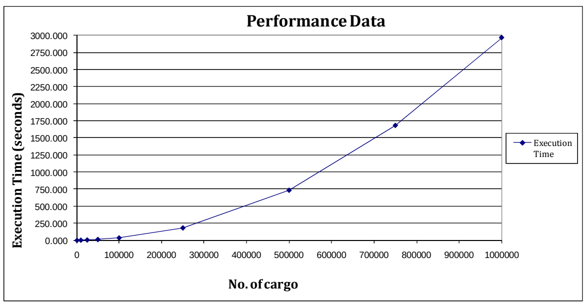


Figure 6. PORTSIM Execution Times

ISBN 978-84-692-5416-5

The PORTSIM system is designed for portability across alternative computing systems. The simulation kernel is separate from the I/O user interfaces. Simple text files and industry-standard XML and XSD files are used to handle data-passing between the kernel and the different interfaces. The user interfaces (the GUI and the Output Analyzer) are easier to develop with code specific to a computer platform. The simulation kernel can be ported to any platform using an appropriate C++ compiler.

4. OTHER APPLICATION AREAS

The capabilities of the PORTSIM system have applicability beyond modeling military cargo flow through a seaport. These capabilities can be utilized for a variety of studies as described below.

4.1. Modeling Commercial Operations

The programmatic process capability described in Section 3.4 can model commercial operations with a seaport. Commercial shipping activity differs from military operations in terms of types of cargo being moved and the methods for processing each cargo type. The programmable process capability allows different processes to be defined for different seaports. Additionally, concurrent SPOD and SPOE operations are a vital component of commercial operations, where ship loading activity has a direct impact on ship offloading capability.

4.2. Modeling Disruptive Events

PORTSIM can be used to study the effects of disruptive events -- natural ones, such as earthquakes, hurricanes, or man-made events such as terrorist attacks or chemical/biological accidents -- on normal operations within a seaport. The programmatic event capability described within Section 3.3 introduces events into the simulation that alter the capabilities of different resources and infrastructure within a seaport and/or change behavior of entities. Analysts can program these events to occur at specific times during a simulation run. The scenario defined can then be used to study different strategies, both proactive and reactive, to deal with such disruptive events in order minimize the effect of these types of events on seaport operations. An initial study using this capability is described in (Leathrum, Mathew, and Mastaglio 2009).

4.3. Evaluation of New Technologies

PORTSIM can be used to study the potential impact of new technologies that have not been developed as yet. The user-programmable and generic nature of the architecture makes it easy to define new technologies, such as new resources in terms of their capabilities. This capability will support the refinement of requirements for future technologies able to support a deployment or redeployment operation. A study using this capability to analyze the potential impact of a new Theater Support Vessel on a Joint Logistics Over the Shore (JLOTS) operations is described in (Leathrum, Mielke, Frith, and Mathew 2002).

4.4. Networks of cargo terminals

PORTSIM can model the end-to-end flow of cargo at an entity level, from multiple US-based installations to a set of destinations within a theater of operations located anywhere in the world in the form of a nodal network. PORTSIM models the resource and infrastructure allocations and competition within each node in the network, as well as between the nodes, to identify conflicts that may arise. Since the underlying PORTSIM architecture is hierarchical in nature and is completely configurable and programmable, the same architecture that is used to describe a seaport can be used to link together different cargo terminals, even those with different transportation modes, into a single integrated scenario. Used in this fashion, PORTSIM can model any type of cargo terminal (airport, rail yard, inland port, or distribution center), whether military or commercial. For an airport, the terminals could be partitioned into the cargo terminal and the passenger terminal, with multiple possible passenger terminals. The runways would then be defined as another port area at the port level. An initial prototype for modeling end-to-end logistics is described in (Mathew, Leathrum, Mazumdar, Frith, and Joines 2005).

5. RELATED WORK

Transportation logistics planning for military operations has been captured in a variety of simulation tools. These tools model the cargo terminals (points of origin and destination, intermediate transfer points for transportation mode changes and/or points of intermediate storage) within the DTS, as well as the transportation infrastructure connecting these terminals. Examples of cargo terminal models include *TRANSCAP* (Burke, Love, Macal, Howard, and Jackson 2000), which models offloading at installations, *TLoads* (Hamber 2001) which attempts to assess the capability of tactical and sea-based distribution systems and *POPS* (Snyder and Henry 1987) that estimates the capability of a seaport to handle military cargo at an aggregate level. Past work at modeling the actual transportation segments of the DTS, include *ELIST* (Braun, Lurie, Simunich, Van Groningen, Vander Zee, and Widing 2000) which models theater rail and highway infrastructure, and *MIDAS* and *JFAST* (Mckinzie and Barnes 2004) which model the strategic lift segments. However, none are stochastic models. These models study the restrictions of the transportation infrastructure rather than the flow of cargo through the cargo terminals. They also aggregate the cargo, often dealing with it as quantities in terms of weight, bulk size, or number of pieces, rather than individual items. The *AMP* model (Stevens, Tustin, and Key 2004) acts as a shell to interconnect these models into an end-to-end logistics model, but at a low level of fidelity, again focusing more on the links than the nodes within the network.

PORTSIM models the different cargo terminals represented by the models mentioned before at a very high level of fidelity. PORTSIM however, models the transportation infrastructure between terminals at a low level of fidelity, by representing the transit between terminals as a transit time. Outputs from the MIDAS, JFAST and ELIST models can be used to derive accurate values for these transit times.

6. CONCLUSION

The PORTSIM development has implemented a programmable process input, entity based model of cargo approach to satisfy the requirements of analysts planning a military deployment or redeployment. Its architecture, design, and implementation approach provide a scalable and reusable capability to model any type of port operations, irrespective of end user and type of port facility. PORTSIM is a flexible approach to modeling nodal logistics operations that can be used not just for individual ports but also to capture networks of ports while still providing detailed results down to the individual cargo items level.

ACKNOWLEDGEMENTS

This work is funded by the Military Surface Deployment and Distribution Command - Transportation Engineering Agency (SDDC-TEA), under USTRANSCOM at Scott AFB, IL.

REFERENCES

Braun M.D., Lurie G., Simunich K., Van Groningen C.N., Vander Zee H.J., Widing M.A., 2000. ELIST8: A Simulation System for Transportation Logistics Planning Support. *Proceedings of the Summer Computer Simulation Conference*, pp. 693–698, July 16–20, Vancouver (British Columbia, Canada).

Burke J.F., Love R.J., Macal C.M., Howard D.L., Jackson J., 2000. Modeling Force Development from Army Installations Using the Transportation System Capability (TRANSCAP) Model. *Proceedings of the Summer Computer Simulation Conference*, pp. 85–90, July 16–20, Vancouver (British Columbia, Canada).

Hamber B., 2001. TLoaDS abbreviated systems architecture. *Proceedings of the Winter Simulation Conference*, pp. 749–757, December 9–12, Arlington (Virginia, USA).

Leathrum, J.F., Mielke, R.R., Frith, T. and Mathew, R, 2002. Modeling New Technologies in a Joint Logistics Over The Shore (JLOTS) Operation. *Proceedings of the Summer Computer Simulation Conference*, pp. 165–169, July 14–18, San Diego (California, USA).

Leathrum, J.F., Mathew, R., Mastaglio, T., 2009. Modeling & Simulation Techniques for Maritime Security. *Proceedings of IEEE International Conference on Technologies for Homeland Security*, May 11–12, Waltham (Massachusetts, USA).

Mathew, R., 2002. *An Object-oriented Architecture for Concurrent Processes in a Port Simulation*. Thesis (MS). Old Dominion University.

Mathew, R., 2007. *The Distributed Independent-Platform Event-Driven Simulation Engine Library (DIESEL)*. Thesis (PhD). Old Dominion University.

Mathew, R., Leathrum, J.F., Mazumdar, S., Frith, T., and Joines, J, 2005. An Object-Oriented Architecture for the Simulation of Networks of Cargo Terminal Operations. *Journal of Defense Modeling and Simulation (JDMS)*, 2 (2), 101–116.

Mathew, R. and Leathrum, J.F., 2008. Programmable Process Language for Modeling Cargo Logistics. *Proceedings of MODSIM World 2008*, September 15–18, Virginia Beach (Virginia, USA).

Mckinzie, K., and Barnes, J., 2004. A Review of Strategic Mobility Models Supporting the Defense Transportation System. *Mathematical and Computer Modelling*, 39 (6–8), 839–868.

Snyder, A.L., Henry, D., 1987. Port Operational Performance Simulator (POPS) User Manual, Version 1.

Stevens, S., Tustin, J., Key, W.H., 2004. U.S. Transportation Command Analysis of Mobility Platform Federation. *Proceedings of the European Simulation Interoperability Workshop*, June 28–July 1, Edinburgh (Scotland).

AUTHORS BIOGRAPHY

Dr. Reejo Mathew is a Research Scientist at MYMIC LLC, Portsmouth, VA and the Chief Technical Lead on the PORTSIM project. He received his PhD in Electrical and Computer Engineering in 2007 and his MS degree in Computer Engineering in 2002, both from Old Dominion University. His research interests include Simulation Executives, Parallel and Distributed Simulation and Cargo Logistics Simulations.

Dr. Thomas W. Mastaglio is the PORTSIM Program Manager. He is the CEO and President of MYMIC LLC. Dr. Mastaglio holds MS and PhD degrees in Computer Science from the University of Colorado. He has worked in the computer simulation industry for over 20 years at the technical, program, and corporate management levels. Dr. Mastaglio has served as a Director for the Society for Computer Simulation, member of the U.S. Army Science Board, and an advisor to the Virginia Legislature Joint Committee on Technology and Science.

Dr. James F. Leathrum, Jr. is an Associate Professor in the Department of Electrical and Computer Engineering at Old Dominion University. He received his PhD and MS degrees in Electrical Engineering from Duke University in 1992 and 1989 and his BS degree in Computer Engineering from Lehigh University in 1987. His areas of interest include modeling and simulation, distributed simulation, formal methods, and computer architecture.

UVSIM: A HARBOR CRANES TRAINING SYSTEM

Ignacio García-Fernández[a], Marta Pla-Castells[b], Miguel A. Gamón[c] and Rafael J. Martínez-Durá[d]

[a] [b] [c] [d] Instituto de Robótica, Universitat de València
PO-Box 2085
46071 -- Valencia, Spain

[a]Ignacio.Garcia@uv.es, [b]Marta.Pla@uv.es, [c]Miguel.A.Gamon@uv.es [d]Rafael.Martinez@uv.es

ABSTRACT
This paper presents a simulation-based instructional system for harbor crane operators training. The configurable design of the simulator allows the training of several crane types: Quay-Side Gantry, Rubber Tired Gantry, Mobile, Reach-Stacker and Ro-Ro tractor. In this paper the simulator architecture, the Virtual Reality techniques, the projection system, the configurable cabin and the instructional design are described.
The virtual environment design includes the representation of harbor terminals, container ship load and unload operations, mobile crane driving and bulk material handling. The simulator includes an automatic evaluation system that measures the trainees' evolution in terms of their abilities. The instructor's panel gives full control on the exercise and on many parameters such as weather conditions, crane failures or load distribution during the simulation. The system also includes an advanced debriefing system for exercise revision and evaluation.

Keywords: Training Simulator, Virtual Reality, Harbor Training

1. INTRODUCTION
In the last decades, interactive simulation based on Virtual Reality techniques has become a powerful tool for training purposes, specially in activities which involve high risks and costs, as it is the manipulation of heavy harbor equipment (different types of cranes, trucks, etc.) (Serón et al. 1999; Kim 2005). The many benefits that can be obtained from the use of a simulator in training tasks, has motivated an increase of the use of such systems in the marine and port environment (Jiing-Yih et al. 1997; Wilson et al. 1998; Huang 2003; Daqaq and A.H.Nayfeh 2004; Rouvinen 2005; Korkealaakso et al. 2007).

The use of a simulator forces the trainee to practice the theoretical concepts that have been taught and shows the consequences of the actions in a very immediate and visual manner (Farmer 1999). Also, the simulator provides the instructor with a controlled environment where a large amount of data can be recorded and analyzed to evaluate the trainee's evolution. From the risk prevention point of view, it avoids the danger of an inexperienced user manipulating the real machine in the working environment.

Furthermore, it reduces the cost associated to training; it prevents inexperienced users from using the machinery, thus reducing machine breakdown due to misuse, and it releases machines for work that, without the use of a simulator, have to be used for training purposes.

In addition it provides with the possibility of working in any desired conditions. Arbitrary weather conditions can be reproduced and training for the climate conditions of a given harbor can be done thousands of kilometers away. Moreover, in the simulator any configurable situation can be easily reproduced, while it is not possible to modify the work flow of the terminal to train a given situation (emergency situations, unusual work load,. . .).

Many harbors around the world are currently profiting from the benefits of simulation technologies for training crane operators. Several companies have presented different systems for crane operator training during the last decade for different types of cranes (Drilling Systems 2009; Globalsim 2009; LAMCE 2009; L3-Ship Analytics 2009)

This paper presents a training simulator based on Virtual Reality technology that is currently being used for training tasks related with five different harbor equipment: quay-side cranes, rubber-tired gantry cranes, reach-stacker cranes, mobile cranes and ro-ro tractors. The main contribution of this development is the use of new modeling approach for complex subsystems, such as cable and pulleys hoists, bulk material or traffic, and a powerful analysis and evaluation system.

The paper is structured as follows. Section 2 describes the training system and its main elements. Section 3 gives an overview of the different hardware configurations that can be used with the current system. Section 4 describes the particular properties of the different equipments simulated. Finally, Section 5 gives concluding remarks and an overview of future work.

2. TRAINING SYSTEM DESCRIPTION
A harbor crane simulator is a tool aimed to the instruction of new operators and advanced training of

experienced ones. Within this premise, a simulator has to include a set of tools that help both the instructor and the trainee to obtain the maximum performance of the equipment and enables an optimum use of the instruction time. The system comprises four main elements.

- The Instructional Design.
- The User Management System.
- The Evaluation System.
- The Simulation System.

Next, these modules are described in detail.

2.1. Instructional Design

The Instructional Design includes a complete series of exercises, which guide both the instructor and the trainee through the process of learning and training.

These exercises are grouped into practice modules, which cover different aspects of the maneuverability of the harbor cranes and the set of skills that should be acquired in the simulator. The exercises gradually increase their complexity and the number of activities involved, from simple control of the payload to avoid excessive swing, to complex operations involving trucks and container-ships. Also the working pace is increased gradually through the course, including exercises aimed to advanced training of experienced operators.

The exercises include visual aids that are shown on screen during the simulation. All the indications to the user are given using the same sign convention that is used in the real port terminal. This helps the new trainees to get used to real operation, shortening the adaption time to the real environment.

An instructor's manual gives additional information to the instructor so that new training modules can be implemented by means of the selection of certain exercise sequences.

2.2. User Management System

The system includes a database system, the User Management System, that makes it possible to register the simulator users and to store every exercise that is done with the simulator. Thanks to the User Management System, every user of the training system can be tracked individually, storing all relevant data, in order to have an overview of the evolution of every trainee.

The system also includes the necessary database management tools to register or delete users, clean obsolete registers, make backups or archive old data to save disk space.

2.3. The Evaluation System

The Evaluation System analyses the data that is collected during simulations to provide the instructor with an objective and clear report of the progress of every trainee. This system is based on three tools:

- An automatic evaluation algorithm to monitor the evolution of every user along the training program.
- A report generator which summarizes in a printable document the results of the evaluation.
- A debriefing system, to visually reproduce and evaluate any finished exercise.

More detailed descriptions of these components are presented in the next sections.

2.3.1. Exercises Evaluation

One of the advantages of simulator-based training is the ability to perform an objective evaluation of the trainee, thanks to the possibility of capturing any system variable and the registration of the user actions during the simulation.

The Evaluation System is responsible for monitoring and store a set of parameters that will determine the performance and evolution of the student. For every exercise, the operator movements, the scenery details and any variation introduced in the exercise by the instructor are stored for later use in the report generation or exercise debriefing.

The Evaluation System stores the correct execution of the tasks and orders of the exercise, as well as the time taken to perform them. Every collision with trucks, containers, crane structure, etc. . . is recorded together with the load speed and the maximum accelerations suffered. The Evaluation System also detects and stores dangerous maneuvers, such as strong payload swing or heavy collisions, as well as the effectiveness of the user in the use of the crane controls, two-hand coordination, etc. . .

A set of mathematical algorithms have been implemented in order to transform the large set of raw data captured during the simulation into user readable parameters. These parameters have been chosen by experienced instructors during the simulator design process.

However, the captured raw data is stored after every exercise in the system's database. This enables the evaluation to be recovered at any time later on and also that the evaluation reports can be completely redefined in the case the instructor decides to include new parameters in the Evaluation System.

2.3.2. Report Generation

The evaluation results are shown by means of customized reports that can be seen on screen or printed. They can be used to give the trainees a summary of their progress and assessment of their skills during the operation of the machine. They can be also used as an examination system that objectively gives a report of whether a trainee has achieved the proposed goals.

The exercise report includes identification data, such as the exercise description, the academic profile of the student, the date and the time of the execution, the total time, etc. In addition, it shows the results of the exercise by means of the selected parameters and several graphics and data plots, becoming a powerful evaluation tool. The collision count rate, number of dangerous maneuvers, collision severity, spreader trajectory representation, etc. are also shown, providing the instructor with a complete summary of every exercise. One page of a report generated by the simulator is shown in Figure 1.

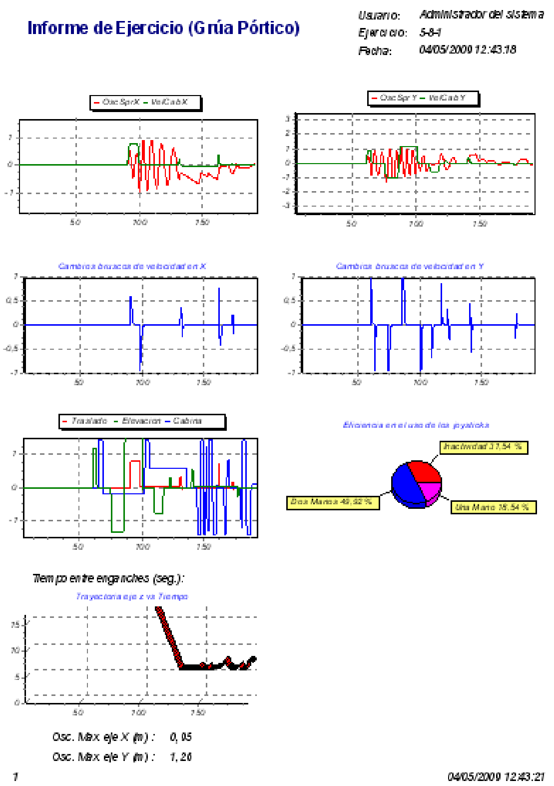


Figure 1: An example page of a report generated by the Automatic Evaluation System

The report of an exercise is available just after the exercise has been finished, so that the instructor can discuss with the trainee any issue that is considered of interest for the learning process. Additionally, it can be recovered from the database at any time on a later session.

2.3.3. Session Debriefing

The debriefing system is a powerful tool that enables the automatic, off-line reproduction of any exercise done in the past, with the possibility of pausing or moving to any instant its reproduction. Following the realization of an exercise, the debriefing mode can be activated and the exercise can be revised by the trainee and the instructor who can make any comments or remarks on it. Additionally, the debriefing of any past exercise can also be activated, providing the instructor with a demonstration system to support theoretical concepts and with a revision tool to evaluate the evolution of a trainee.

In order to do that, the system does not store a movie of the simulation, but instead, the minimum number of variables necessary to completely reproduce the simulation session are stored. This presents two main advantages. First, it reduces the amount of space needed to maintain the database, since the storage of full movies is unaffordable. And second, as the visual representation is being rendered again during the debriefing of an exercise, it is possible to change the point of view of the scene according at the instructor's discretion

2.4. Simulator Architecture

A crane simulator aimed to training tasks is a complex Virtual Reality system composed of several modules. Together, they make it possible to acquire the user's actions, to reproduce the behavior of the environment and to visually represent the simulated environment. The simulator considered in this work is composed of the following elements:

- Visualization system, which provides a graphical representation of the environment.
- Physically based model of the machine and the environment.
- Immersion system, including the projection system, Input/Output devices and 6-Degree of Freedom (DOF) mobile platform.
- Control panel.

During the simulation, the user provides the input data by means of the input devices, which reproduce the control panel of the simulated machinery. These input data are used by the dynamic models in order to determine the evolution of the machinery and of the environment. The dynamic models, which are discussed later in this work, are a combination of multibody models, implemented by means of the ODE programming library, and a set mathematical models of other complex systems.

The state of the simulated environment is reproduced by means of a 3D graphics application which stores a scene graph of the virtual world. This application is implemented by means of the OpenGL Performer programming library. The 6-DOF mobile platform is controlled using the dynamic model output, to simulate the accelerations experimented in the real machinery in order to provide a better feeling of *presence* (Reid and Nahon 1985; Reid and Nahon 1986).

The control panel is the Graphical User Interface of the application from which the instructor has full control of the system. Together with the virtual environment and the immersion system, the Instructional Design, the evaluation and debriefing system and the user database convert the simulator into a powerful training equipment.

2.4.1. Visual Subsystem

One of the most important elements of a training simulator is the visual subsystem. This module is responsible for rendering a visual representation of the virtual environment of the simulator. It reproduces a complex scenery which includes the harbor environment including containers, several vehicles, quays, etc. and the simulated machines.

This application comprises a scene graph, which stores a large database of geometrical data, and a logical model that computes the evolution the scene graph and other system outputs. The geometrical database contains all the data necessary to represent the different elements that appear in most international harbor terminals, while the scene graph is a complex data structure that permits an easy update of the positions of the different elements through the scene. The logical model is a state machine which determines whether the position, the topology or the visibility of an object, or a compound of objects, must change to represent an event, e.g. the attachment of a container to the spreader, the extension of the spreader, etc.

All the scenery models have been modeled using a 3D modeler environment, and stored in ASCII text files, which can be converted into OpenGL Performer's native format `.pfb`. All the models have been texturized with photographic textures obtained at several harbors. In addition, several techniques, such as *bump-mapping* or *environmental maps* have been implemented in order to increase the realism of the virtual environment. These techniques, together with the use of different Levels Of Detail (LOD), provide the simulator with a very realistic and efficient environment.

The virtual environment considers different types of cranes and vehicles (trucks, ro-ro tractors, etc.) that provides it with the realism needed. Furthermore, container yards, quay sides, container-ships, bulk material yards and other locations have been modeled using the information provided by port authorities of Valencia Harbor.

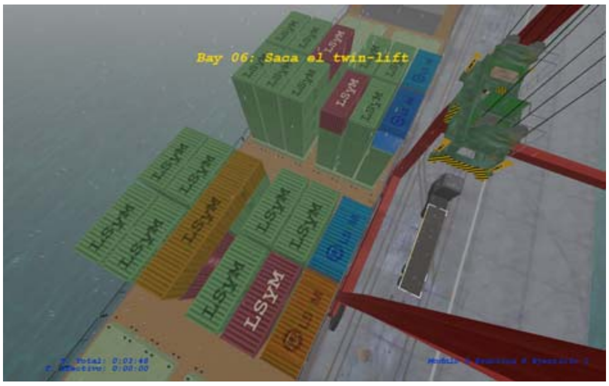


Figure 2: The virtual environment can reproduce different weather conditions that can be modified during the simulation

Also, several state-of-the art techniques have been used to reproduce different weather conditions that can be modified during the simulation; rain, fog and snow can be simulated, affecting the operator's visibility. Additionally, a *skybox* represents the aspect of the sky at different day hours and with different cloud conditions and, by night, OpenGL illumination techniques are used to simulate artificial light.

2.4.2. Dynamical Model

In order to obtain a realistic simulation environment, not only a high quality graphic application is necessary, but also the behavior of the machine, considering its subsystems, and its interaction with the user and with the environment has to be reproduced based on its physical properties. In order to reduce the computational cost of the final application the properties of the crane that are considered more relevant for training purposes are considered, discarding irrelevant behaviors . This leads to a very efficient, yet realistic simulator adequate for training purposes.

The dynamic model used for every crane is based on its mechanical properties, using the multibody system dynamics methodology. The crane is decomposed into different bodies, that are linked by means of kinematic constraints, and implemented using the Open Dynamics Engine Library (Smith, 2001). The characteristics that are considered in the crane models are summaryzed following.

- Collision detection and response against the elements of the scenery, using complex geometric models for accurate collision point determination.
- Flexible elements to simulate the deformation of the crane structure.
- Realistic modeling of the different engines, brakes, electric and hydraulic subsystems, etc. present on the simulated machinery, according to the real machine specification, including differentmachine failures.
- Ship movement due to load distribution.
- Effect of wind and rain over on the payload and on collisions friction.
- Liquid loads and other complex mass distributions.

All the models are defined using a set of parameters which determine their main properties (sizes, speeds, response curves,. . .). This enables to simulate a large set of different cranes through the adjustment of these parameters, adapting the system to the needs of most crane models.

In addition to the mechanical behavior of the cranes themselves, the physical properties of several complex systems have been also considered. More precisely, the developed simulators include new models for the simulation of bulk materials, hoists systems and traffic flow.

Granular Systems Simulation The simulation of a mobile crane requires the modeling of the bulk material that is being handled. Load and unload works, together with bulk material arrangement in the yard, need to be simulated in a realistic manner.

According to most common types of goods that are present in a harbor terminal (grain, granulate fertilizer, coal, etc.) bulk materials have been modeled as granular systems with different properties. However, modeling and simulation of granular material is a complex problem, which involves a very high computational cost

(Pöschel and Schwager 2004). For these reason, most common methodologies are not adequate for real time simulations. Then, new models have been developed in order to meet both the physical realism and the computational efficiency necessary in training simulators.

A model based on the Cellular Automata formalism has been developed for the evolution of the surface of the granular system (Pla-Castells et al. 2008). It represents the granular material by means of a grid, which covers the material surface, and a local rule, which determines the evolution of the system. The model also considers mechanical interaction between the granular system and the objects of the scene, permitting bulk material manipulation (Pla-Castells et al., 2006). In addition, the model uses an implementation strategy that exploits statistical properties of granular systems in order reduce the number of computations.

Cable and Hoist Simulation Another subsystem which shows a highly complex behavior is cable and pulleys systems that are used in most elevation equipments. Quay-side, gantry and mobile cranes use a set of pulleys and cables in order to hoist the load and also to move some of their structural elements.

Cable systems have been usually simulated as massless spring elements, discarding cable oscillation and complex pulley systems (Abdel-Rahman et al. 2003; Daqaq and A.H.Nayfeh 2004; Servin and Lacoursière 2007). However, this leads to an important loss of realism both in the visual aspect of the simulator and in the physical properties of the crane models.

In order to increase the quality of the simulation, providing the model with cable oscillation and with arbitrary pulley configurations, a new model has been developed (García-Fernández et al. 2008).

This model uses a combination of simple pulley models and cable segment models to efficiently simulate a hoist with cable oscillation. Cable behavior is reproduced by means of a series of layered models that represent different vibration features. The decision of which layers are used is taken during the simulation in an adaptive manner. The model is designed to be able to handle the high mechanical tension situations that appear in the cranes considered in the simulator (García-Fernández, 2009).

Traffic Simulation The simulator includes a Traffic Simulation Module. It controls the different mobile elements of the scene, such as trucks and other yard cranes, and its behavior can be modified during the simulation. It determines the container flow rate and the amount of goods that the operator must handle per unit time.

The Traffic Simulation module is redesigned for every harbor in order to reproduce their particular operation procedures, average and peak container-per-hour rate, types of goods, classification in the container yard, etc. . .

The combination of all these state-of-the-art models provide the Simulation System with a high degree of realism and interactivity, improving its training capabilities.

2.4.3. Immersion System

The Simulation System is usually formed by an immersive projection system and a cabin, where a seat and the Input/Output devices are installed. This cabin, and potentially the projection system, can be mounted on a 6DOF Stewart platform to produce inertial stimuli that simulate the accelerations experienced in the real crane.

However, the system presented here comprises a wide range of hardware configurations which make it more versatile. They are described in detail in Section 3.

2.4.4. Instructor's Control Panel

During a working session, the instructor has full control over the simulator through the Instructor's Control Panel (see Figure 3). It is an interactive panel from which the different aspects of the simulation can be controlled. It includes a 2D view of the state of the simulation, with the crane and the containers that are involved, and a set of frames and menus that give access to the following actions:

- Crane setup
- Weather conditions setup
- On-line failure injection and container modifications
- Traffic management

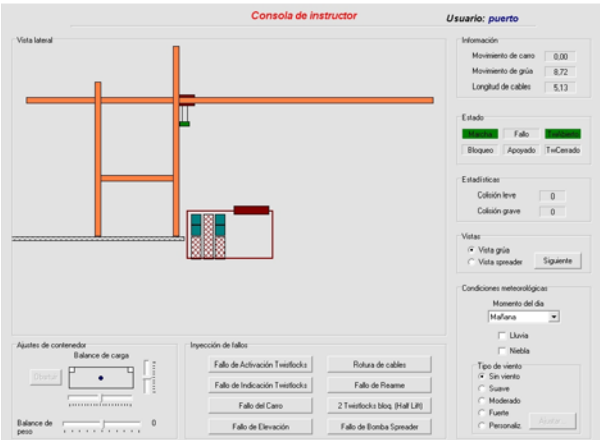


Figure 3: The Instructor's Control Panel during a Quay-Side Gantry Crane simulation exercise

Crane Setup Before the exercise starts, the instructor can access a graphical user interface which controls all the adjustable parameters of the crane. Through this panel, the crane model can be modified and many customizations can be done to introduce new situations that can improve training tasks.

Options such as anti-swing systems, automatic flippers, cell guides or other features can be activated through this panel. In addition, there are several parameters that provide the instructor with additional tools. E.g. joysticks sensitiveness can be increased to force the users to be more careful in their manipulation, a delay can be introduced to the controls in certain exercises to train quick responsiveness, etc.

Weather Conditions As it has been stated before, one of the benefits of using a training simulator is the control over weather during the simulation. Through theWeather Conditions panel, wind, precipitation and fog can be controlled, as well as the daylight properties or the use of artificial light during night. All these properties are reproduced in the virtual environment by means of the appropriate techniques, as described in Section 2.4.1. Also, dynamic models consider the parameters of weather conditions in order to modify the friction properties with wet floor or the payload behavior under the action of wind.

Failure Activation and Container Management During the simulation, the instructor can activate a series of failures that depend on the particular equipment that is being simulated. All of them correspond to failures that happen with certain frequency in the real cranes. When a failure is activated, the trainee is expected to detect it and to act accordingly. Failure activation is recorded in the stored data in order to determine whether the reaction was adequate.

In the case of container handling cranes, it is also possible to modify every container's properties individually through the Control Panel. The instructor can select a container through a graphic interface and modify its weight, type and load distribution. A container can also be marked to activate a failure when it is attached to the spreader.

Traffic Management Another frame of the Control Panel manages traffic properties in the corresponding terminal. The frequency at which trucks appear can be configured, and the appearance of a truck can be fired at any time. This forces the pace at which the trainee has to work, as traffic jams can appear if containers are not moved at the needed frequency, and permits to simulate work overload situations to train experienced operators.

3. SCALABLE HARDWARE CONFIGURATIONS

A training simulator is a complex system which needs a series of hardware components to achieve a high level of presence of the simulator. However, depending on the needs of the harbor, it can be implemented using different hardware configurations that range from a simple desktop PC to a fully immersive system installed on a 6-DOF platform.

3.1. Desktop System

Desktop based system configuration is intended to provide a virtual environment through an standard desktop computer. The main advantages of this configuration are its accessibility to any user, its low cost, and its small size. In this system, both a set of special controls that emulate the crane control desk or a set of standard game devices can be used. In order to improve the perception of distance in the virtual scenery, the use of guidance systems or trackers have been also included, as shown in Figure 4.

Figure 4: A desktop configuration, with several displays and conventional game devices as controls

Desktop systems are designed as a basic approach to the training of operators through virtual reality, but the effectiveness is much lower than that achieved in fullyimmersive systems.

3.2. Medium-Size System

This system incorporates a real seat with real instruments that is located in front of the projection system (see Figure 5). This one is usually compound by a big LCD or plasma or even by a single rear projection screen, although multiple projection screens or monitors can be also arranged to configure several views or a high resolution tiled display. Moreover, stereographic imaging can be achieved, using some type of shuttered glasses in synchronization with the graphics system.

Figure 5: Medium-Size System used for training in Guayaquil (Ecuador)

This system can be appropriate for the simulation of quay-side or gantry cranes, since they only need the single front view of the virtual environment and do not precise to look around.

3.3. Fully Immersive System
In this configuration, the user is fully immersed in the virtual environment, using a replica of a real crane cabin, where several projection screens and LCD projectors are attached to it. In turn, the cabin is mounted on a motion platform, which reproduces in the user the forces generated by the movements of the vehicle simulated. The motion simulation is a very important factor in immersive environments, since it simulates the accelerations caused by the real crane, improving realism, avoiding motion sickness caused by the absence of movement.

The sense of full immersion is achieved by providing the user with a environment very close to real. A replica of the crane controls, provided by the manufacturer, are used in the cabin desk. The types of controls present are: joysticks, pushbuttons, illuminated pushbuttons, LEDs (indicators), switches, combiners (lever position) switches, digital displays and buzzers. Several USB Data Acquisition Boards are used to capture the sensor's state.

In the case of Reach Stacker and Maffi crane, an special frame, integrating a steering wheel, levers and switches, direction indicators and a set of pedals has been built. This piece is attached in the cabin, between the desks and in front of the driver's seat, and is easily removable when it is not used.

Every crane simulator can use this fully immersive system, but Reach-Stacker and Ro-Ro tractors are particularly suitable for this configuration, since they need a panoramic visual environment in order to drive along the yard. In addition, since the movements of the vehicles are faster than the other cranes', a motion platform is recommended to avoid simulation sickness.

Figure 6: A CAD view of a Fully Immersive simulation system

3.3.1. 3D Sound
The Simulation System integrates spacial 3D sound that reproduces the environmental sound, the engine noise of the crane during normal operation and the different sounds of collision and alarms. The sounds have been implemented using the library *OpenAL* (Yuzwa, 2006) to render positional and multi-channel audio in three dimensions.

4. USE CASES
Based on the described architecture, five harbor crane simulators have been developed: quay-side gantry crane, rubber-tired gantry crane, mobile crane, reach-stacker crane and ro-ro tractor. After describing the main properties of the Simulation System, next an overview of the particular features of every crane simulator are given.

4.1. Quay-Side Gantry Crane
Quay-Side Gantry Crane simulator (see Figure 7) is based on Paceco cranes of types PANAMAX and OVERPANAMAX. The cranes have been modeled in detail to reproduce their physical behavior and their different subsystems.

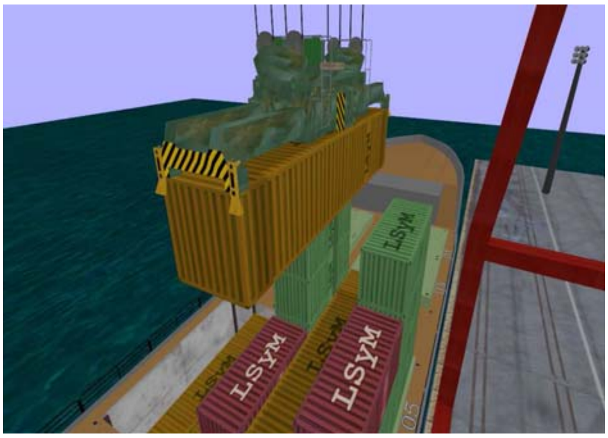


Figure 7 A view of the Quay-Side Gantry Crane simulator

One of the most important elements of the crane is the spreader, as it is the tool used to hold the containers, and is involved in most of the interactions of the user with the environment. For this reason, it has been modeled in great detail including its geometry, mobile flippers and twist-locks, for accurate collision detection.

Trim, list and skew movements are simulated and the hydraulic system is modeled in order to simulate its behavior and several malfunctions. Different spreader sizes and types are simulated, including twin-lift and over-height spreader, which can be used in different exercises included in the training course.

The simulator includes a library of container ships for training in many different situations. Load and unload operations can be simulated inside holds or on the deck, including the removal of hatch covers for operation inside holds. Cell guides can be included when the instructor considers that it is adequate.

The behavior of trucks that transport containers from the quay side to the container yard and vice-versa are also simulated. A discrete event model produces a realistic truck traffic that feeds the trainee with containers during loading operations and takes them away during ship unloading.

4.2. Rubber Tired Gantry Crane

The rubber tired gantry cranes stack and unstack containers in the container yard using an spreader (see Figure 8). The same spreaders that were considered in the Quay-Side Gantry Crane have been simulated in the Rubber Tired Gantry Crane simulator.

Figure 8 A view of the Rubber Tired Gantry Crane simulator

In this crane, an aspect that is of special relevance is the influence of tires in the gantry movement. During its displacement, differences in tire erosion and in the load distribution among the tires can lead to different friction in both sides of the gantry. As a consequence, the gantry can deviate from straight line during displacement. Rubber tired gantry cranes incorporate a control to change throttle of both tire groups. This system has been modeled, obtaining a very realistic simulation of gantry displacement.

The rubber-tired gantry also uses the truck and traffic simulation model. Trucks carry and take away containers that have to be manipulated by the crane operator.

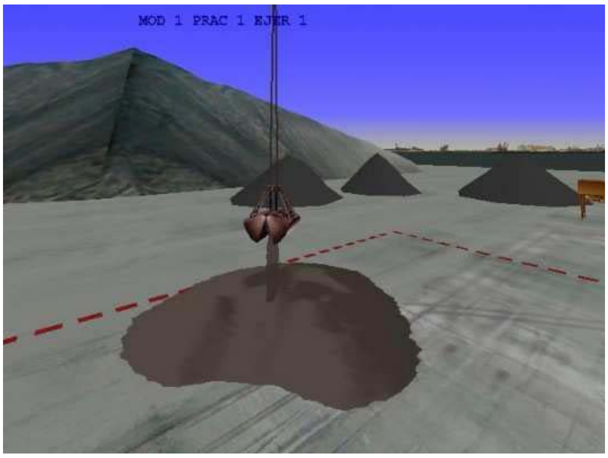


Figure 9 Bulk material handling in the Mobile Crane simulator

4.3. Mobile Crane

Mobile cranes, mainly used to load and unload bulk material, are also considered in the simulator (see Figure 9). The mobile crane consists of a mast and a moving lattice boom that is supported by a cable and pulleys system. The sceneries of this crane also consider the simulation of several ships, including hold operations and yard bulk movement. Different bulk materials are realistically simulated. Vessel load and unload operations can be done, as well as bulk material movement through the yard. In addition, the translation of the mobile crane, which is a complex and dangerous operation, is simulated.

4.4. Reach-Stacker Crane

A Reach-Stacker crane is a heavy vehicle with a boom and a spreader attached to its end (see Figure 10). The dynamic model of the crane involves, in this case, soil-tire interaction which is made by means of Pacejka models (Pacejka 2006). The spreader subsystem is simulated in a similar way as it is done in Gantry or Quay-Side cranes but, in this case, the hydraulic system is connected to the rest of the vehicle, thus influencing its behavior.

Figure 10 A view of the Reach-Stacker Crane simulator

Two hydraulic pumps are considered, separately feeding boom and vehicle operation. The simulation of the hydraulic subsystem includes the loss of power if too many operations are being performed at the same time, and also the possibility of simulating different failures. As well as in the previous cases, the Reach-Stacker simulator also includes the truck traffic simulation system.

4.5. Ro-Ro Tractor

A roll-on/roll-off (Ro-Ro) tractor (see Figure 11) is used to move road trailers onto and off roll-on/roll-off ships (ferries). It is a powerful machinery with high maneuverability, and the driver must have a good knowledge of its operation to properly park a trailer inside the vessels when it has to be left in a difficult position.

Several types of loads and trailers are considered in the simulator and the different exercises force the trainee to drive up ramps and through narrow corridors as those found in real ferries


Figure 11 A view of the Ro-Ro Tractor simulator. The trailers are taken with the tractor head and loaded into a vessel

5. CONCLUSIONS

An immersive crane simulator for the training of operators of five different cranes has been presented. The system is highly flexible, allowing several hardware configurations, and uses the latest Virtual Reality techniques available.

A highly realistic Virtual Reality environment is used to reproduce the harbor, using different techniques to reduce its computational complexity. A set of advanced dynamic models simulate the evolution of the different systems involved, including systems that had not been simulated before, such as arbitrary cable-pulley systems or interactive bulk materials simulation.

The system includes an Instructional Design with a large set of exercises, which guide the instructor through the training process to take the most advantage of the system. The simulator provides the instructor with a helpful automatic evaluation and report system and a powerful debriefing tool to review any exercise done in the past. Also, through a complete control panel the instructor has full control over the simulation parameters.

The Training System is currently being used in Valencia Harbor and in several South American international terminals, and it is in continuous revision by the instructors therein. The main improvements that are currently in progress are the develpment of an exercise designer and the extension of the physical and mathematical models to improve the realism of the virtual environment. Also, a series of tests and questionnaires are being designed to evaluate the users experiences in order to improve the system.

REFERENCES

Abdel-Rahman, E., Nayfeh, A., and Masoud, Z. (2003). Dynamics and control of cranes: A review. *Journal of Vibration and Control*, 9(7):863–908.

Daqaq, M. and A.H.Nayfeh (2004). A virtual environment for ship-mounted cranes. *International Journal of Modelling and Simulation*, 24(4):272–279.

Drilling Systems (2009). Drilling systems. Web Page. http://www.drillingsystems.com (Last visited June 2009).

Farmer, E., editor (1999). Handbook of Simulator-Based *Training*. Ashgate Publishing Limited.

García-Fernández, I. (2009). *Propuesta de nuevos modelos din´amicos de cables de elevaci´on para simulación en tiempo real*. PhD thesis, Escola Tècnica Superior d'Enginyeria. Universitat de València.

García-Fernández, I., Pla-Castells, M., and Martínez-Durá, R. (2008). Elevation cable modeling for interactive simulation of cranes. In *SCA '08: Proceedings of the 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 173–181.

Globalsim (2009). GlobalSim MasterLift simulation. Web page. http://www.globalsim.com (Last visited June 2009).

Huang, J.-Y. (2003). Modelling and designing a low-cost high-fidelity mobile crane simulator. *International Journal of Human-Computer Studies*, 58(2):151–176.

Jiing-Yih, L., Ji-Liang, D., H., J.-R.,Ming-Chang, J., and G., C.-Y. (1997). Development of a virtual simulation system for crane-operating training. In *Proceedings of ASME*.

Kim, G. (2005). Designing Virtual Reality Systems: The *Structured Approach*. Springer.

Korkealaakso, P. M., Rouvinen, A. J., Moisio, S. M., and Peusaari, J. K. (2007). Development of a real-time simulation environment. *Multibody System Dynamics*, 17:177–194.

L3-Ship Analytics (2009). L-3 Communications. MPRI Ship Analytics. Web page. http://www.shipanalytics.com/ (Last visited June 2009).

LAMCE (2009). Laborat´orio de M´etodos Computacionais em Engenharia. http://www.lamce.ufrj.br (Last visited June 2009).

Pacejka, H. (2006). *Tyre and Vehicle Dynamics*. Elsevier, 2 edition.

Pla-Castells, M., García-Fernández, I., and Martínez-Durá, R. J. (2006). Interactive terrain simulation and force distribution models in sand piles. *Lecture Notes on Computer Science*, 4173:392–401.

Pla-Castells, M., García-Fernández, I., and Martínez-Durá, R. J. (2008). Physically-based interactive sand simulation. In Mania, K. and Reinhard, E., editors, *Eurographics 2008 - Short Papers*, pages 21–24, Crete, Greece. Eurographics Association.

Pöschel, T. and Schwager, T. (2004). *Computational Granular Dynamics : Models and Algorithms*. Springer-Verlag Berlin Heidelberg, 1 edition.

Reid, L. D. and Nahon, M. A. (1985). Flight simulation motion-base drive algorithms: Part 1 – developing and testing the equations. Technical report, University of Toronto Institute for Aerospace Studies (UTIAS).

Reid, L. D. and Nahon, M. A. (1986). Flight simulation motion-base drive algorithms: Part 2 - selecting the system parameters. Technical report, University of Toronto Institute for Aerospace Studies (UTIAS).

Rouvinen, A. (2005). Container gantry crane simulator for operator training. In Publishing, P. E., editor, *Proceedings of the Institution of Mechanical Engineers, Part K: Journal ofMulti-body Dynamics*, volume 219, pages 325–336.

Serón, F., Lozano, M., Martínez, R., Pérez, M., Vegara,P., Casillas, J., Martín, G., Fernández, M., Pelechano, J., Brazález, A., and Busturia, J. (1999). Simulador de gruas portico portuarias. In *Congreso Español de Informática Gráfica (CEIG'99)*.

Servin, M. and Lacoursière, C. (2007). Massless cable for real-time simulation. *Computer Graphics Forum*, 26:172–184.

Smith, R. (2001). Open Dynamics Engine (ODE). http://www.ode.org (Last visited June 2009).

Wilson, B., Mourant, R., Li, M., and Xu, W. (1998). A virtual environment for training overhead crane operators: real-time implementation. *IIE Transactions*, 30:589–595.

Yuzwa, E. (2006). Game Programming in C++: Start to *Finish*. Charles River Media.

AUTHORS BIOGRAPHY

Dr. Ignacio Garc´ıa-Fern´andez. He received his BsC in mathematics by Univ. de Valencia in 2000, his MsC in Computer Science in 2003, and his PhD by Univ. de Valencia in 2009. His main research interest are physically based modeling and simulation for virtual environments.

Dr. Marta Pla-Castells. She received her BsC in mathematics by Univ. de Valencia in 2000, her MsC in Computer Science in 2004, and her PhD by Univ. de Valencia in 2009. His main research interest are physically-based simulation and computer graphics for virtual reality applications.

Miguel A. Gamón. He received his grade on Computer Engineering by Univ. de Valencia in 2002 and his MsC in Computer Science in 2005. His main research interest are computer graphics and Human-Machine interfaces

Dr. Rafael J. Martínez-Durá. He received his BsC in informatics by Univ. Polit´ecnica de Valencia in 1991, and his PhD by Univ. de Valencia in 1997. His main research interest are visualization systems for virtual reality applications and sensorization systems. He is the director of the Modeling and Simulation Laboratory.

A VIBRATION EFFECT AS FATIGUE SOURCE IN A PORT CRANE SIMULATOR FOR TRAINING AND RESEARCH: SPECTRA VALIDATION PROCESS

Agostino G. Bruzzone[(a)], Paolo Fadda[(b)], Gianfranco Fancello[(b)], Gianmarco D'Errico[(b)], Enrico Bocca[(c)], Marina Massei[(fc)],

[(b)]MISS DIPTEM University of Genoa
[(b)]C.C.S.Tra University of Cagliari
[(b)]Simulation Team

[(a)]agostino@itim.unige.it, [(b)]{fadda, fancello, derrico}@unica.it
[(c)]{enrico.bocca, marina.massei}@simulationteam.com

ABSTRACT
This paper is concerned with the loading and validation of vibration spectra, monitored in a ship-to-shore crane in operation at the Port of Cagliari, the distributed real-time interoperable simulator includes a motion platform. In fact Simulators are typically used in active safety applications and in this specific case, our simulator was designed for the purposes of both training and basic and applied research, as it is possible to monitor physiological parameters recorded using appropriate electromedical instruments. The onset of fatigue in quayside crane operators can be attributed not only to awkward posture, but also to the high vibration levels generated by the nature of crane operations, transmitted to the operator through the cab seat.

Key Words: vibrations, motion platform, fatigue, simulator, VV&A process

1. INTRODUCTION
The aim of this paper is to describe the methodologies used to evaluate and verify and validate, within the virtual reality environment, the intense vibratory action which the operator experience during training sessions.

Figure 1: Ship Size and Crane Interference

By identifying the key task phases and recording the synchronous vibration spectra, it was possible, with this highly versatile and flexible simulator, to repeat the procedure for all types of cranes operating worldwide, thus improving fidelity by tailoring to suit specific needs. The research database, compiled using information gathered during training sessions and also containing data on the effects of vibrations, a major source of latent stress in quayside crane operators, will also be more reliable. The authors developed a real-time simulator named ST_PT_1 (Bruzzone et al.2008); this is a HLA (high level architecture) distributed simulation implemented in a standard container shelter for guarantee mobility; the ST_PT_1 incorporate multiple vehicle simulation interacting in a federation, however it focuses on a quay crane full scope simulator including CAVE, 6 D.O.F. motion platform, sound 3D, panel controls and it represent and important step forward for research in port activities; in fact this system could provide significant contribution in researches devoted to improve the whole supply chain that is strongly affected by logistics node performances and port facilities capabilities; in this context safety and productivity issues are strongly related, while the necessity to handle in short time very large ship, characterized by a continuous trend in growth, represent a major challenge (Ircha 2001). This crisis situation is reducing the impact of the problem related to being able to unload/load large container ship, however it is expected that as soon as the traffic will be back to regular levels, this characteristics will correspond to a strategic competitive advantage (Merkuryev et al. 2009); so in fact contingency crisis is providing the opportunity to develop new solutions as well as to proceed in their test and evaluation for being ready for next economic phase.

2. SIMULATION AS PORT ENABLER
The figure 1 puts in evidence the critical aspect correlating ship size and port capabilities; obviously this picture is providing just a view of largest current ships versus gantry crane size, without paying attention to interference issues; however a basic computation considering:

S_c current largest container ships capacity (about 11'000 TEU)
C_p mean gantry crane productivity (about 25 containers/hour)
P_n nautical port operations (arrival and departure, about 0.75 hours each)
D cruise distance (i.e. Hong Kong-Los Angeles 6363 nautical miles)
S_s cruise speed (for latest generation ships about 25 knots)

P_{ltp} port time lapse (about 10% of full cycle time for effective use of the ship)

k_{tf} factor corresponding to the ratio between 20 feet and 40 feet container on board (for instance 1/3)

k_{xm} factor corresponding to the influence of extra movements (about 10% considering hatch operations)

k_{oe} factor corresponding to the percentage of containers unloaded for import (about 90% considering this large ships on oceanic cruises)

k_{oi} factor corresponding to the percentage of containers loaded for export (about 90% considering this large ships on oceanic cruises)

P_{ltp} port time lapse (about 10% of full cycle time for effective use of the ship)

$$C_n = \frac{(k_{tf}+1)\cdot(k_{xm}+1)\cdot(k_{oe}+k_{oi})\cdot S_c \cdot S_s \cdot (1-P_{ltp})}{C_p \cdot (k_{tf}+2)\cdot [D \cdot P_{ltp} - 2P_n \cdot S_s \cdot (1-P_{ltp})]} \quad (1)$$

In the hypotheses (pretty optimistic) presented the number of crane requested to complete in time the operation (C_n) resulted in about 20 concurrent gantry crane; this result is critical considering the mutual interference, the necessity to keep pretty good productivity; considering the ship size growth it is required to improve productivity of the cranes; current evolutions is moving forward improving the gantry cranes, i.e. new spreaders (i.e. multiple lift), very high speed and accelerations. However all these aspects introduce the necessity to simulate the whole process to create virtual prototypes to test the real effectiveness of proposed changes; in addition it becomes very critical to train people operating systems that run much faster with higher weights and in very high density operation. So it is evident that simulation represent an enabler for improving port capabilities (Bruzzone et al.2006)

In fact Simulation models are used efficiently by engineers as a decision support tool for dealing with strategic, tactical or operational decisions in logistic systems (production logistics such as warehousing and hospital logistics and transportation, Merkuryev Y., Merkuryeva, G., Piera, M.À., 2009).

Simulators are being increasingly used to research human factors optimization in the transportation sector as these tools enable to reproduce and evaluate, singly and in combination, all those factors contributing to fatigue (Fancello et al. 2008b). Task simulators are also models, but offer the advantage of having been specifically designed to simulate the distinctive characteristics of real world tasks in a virtual reality environment.

The quay crane simulator installed at Cagliari University (Bruzzone A. et al., 2008), which ultimately aims to enhance safety and productivity in container terminals, is an innovative system designed for the purpose of solving human factor related problems (ergonomics, anthropometrics, etc.) relative to tasks performed by quayside crane operators (D'Errico 2009) (Figures 2/3).

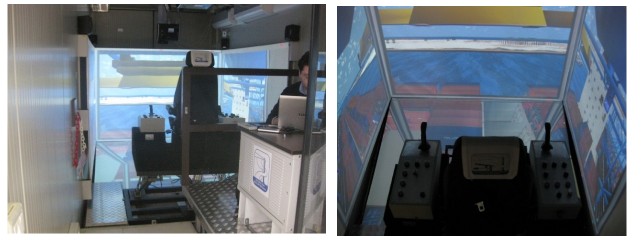


Figure. 2/3: Details of dynamic driver cockpit (6 degrees-of-freedom motion platform) and graphic interface of ST_PT_1 quay crane simulator.

The mobile simulator is housed in a 40 ft High Cube container, purposely chosen for its internal height of 2.70 m, higher than a standard ISO container, thus allowing more space for the main platform as will be described in detail later. This minimizes the time required to prepare training and research sessions, providing on site training services for container terminal operators.

The programme envisages three different types of activities:

- Research: studies of operator performance under various operating conditions using objective medical parameters (*EEG*, *ECG*, *EMG* plots, *goniometer, inclinometer, accelerometers, eye tracker*, etc.) and of the procedure already adopted in a real world environment for handling and analysing fatigue test data (performance curves, neural networks);
- Technological advance: applied research and validation of new design solutions for the crane control systems and commands, for enhancing operator performance and thus aimed at optimising the man-machine interface;
- Training and refresher training of crane operators in container terminals; this activity is important for two reasons: economic, in that research can be funded by selling training packages to container terminal operators, and functional, as the data gathered during training sessions will be used to create and build up the research database.

This entirely original virtual reality instrument, that is able to satisfy the real needs of the container transportation sector, is equipped with an extremely versatile and flexible hardware and software system that is unlike any other existing quayside crane training simulator (Bruzzone et al. 2008). This feature means that operator training can be adapted to all cockpit and quay crane types existing worldwide.[1]. In actual fact, the ultimate design of the simulator was the result of a major effort including detailed literature reviews, research, in depth studies and propedeutic analysis of crane operator tasks conducted in a real world environment by the CIREM-Human Factors research team at Cagliari University, as contribution to the realization of this innovative machine. (D'Errico G., 2009).

2.1 Study of human factors in quayside crane operations in a real world environment

CIREM (Interuniversity Centre for Economics Research and Mobility) at Cagliari University has provided some important contributions consisting of surveys, studies and subjective analyses conducted at CICT (Cagliari International Container Terminal) at Cagliari Industrial Port (Fancello et al. 2008a). The research conducted used an integrated database containing:

- the first arousal-performance curves to be plotted in this particular scientific context of quayside crane operators, were built using a familiar model (a study on human behaviour simulation using a dynamic stress model adopting the same model of experimental performance curves has been conducted by Seck et al., 2005), theYerkes-Dodson function (Fancello G. et al., 2008a/b), that relates work load with performance level;
- the results of a questionnaire completed by crane operators at Cagliari Industrial Port concerning a number of ergonomic aspects (physical complaints indicated) and task related fatigue;
- a series of anthropometric measurements taken using films of CICT quayside crane operators during task execution (Meloni M. et al., 2009);
- two vibration measurements recorded by means of accelerometers, placed in previously identified strategic points inside the crane cab during operation, the first for the duration of the work shift, the second for each key phase of the job task identified with the first..

Using these data, which are dealt with in this paper, it was possible to obtain in output the vibration spectra for the linear and angular accelerations over time. These were then loaded into the software used to handle the simulator's motion platform, making it possible to faithfully reproduce, except for a few corrections, the vibrations experienced by the operator during the simulation scenarios. In this way it will be possible, using the electromedical instruments with which the simulator is equipped, to evaluate fatigue, also taking into account operator exposure to the vibration effect generated by the motion platform (Burdorf et al. 1993).

3. BACKGROUND AND PRELIMINARIES

The containers are loaded/unloaded using ship to shore cranes by means of a "spreader", that is mechanically connected to the hoist motors via a beam suspended from cables and electrically connected to the crane. The container is hooked/unhooked by means of four corner flippers on the spreader (Ircha 2001).

The containers are transferred from ship to shore through a combination of two movements: the spreader-container system is hoisted to the maximum clearance height, and the crane then travels with its load along the bridge rails to the container stacking bay. This operation is usually repeated at least 20 times per hour, the cab travelling back and forth from the ship to the yard (Bruzzone et al. 2007).

Thus throughout the six hour shift the crane operator is exposed both to high vibration, due to cab movements, and to high noise levels generated by the very nature of the operation,. Added to this, is the discomfort caused by the bent forward posture and awkward head/neck positions that the operator is forced to assume to closely follow the movement of the container some 40 m below. These conditions give rise to psycho-physical stress that medically speaking can cause serious health problems over time, while from an operational standpoint they bring about a deterioration in performance that translates into reduced productivity.

3.1 State of the art on the combined effect of ergonomics-vibrations as a stress factor in dock workers

Ergonomics and in particular workstations re-design is a widely studied "human factors" branch that examines operator performance in relation to awkward and poor posture (Meloni et al. 2009). In a study on the effective ergonomic design of workstations, using a simulation model that recreates in a 3-D virtual environment industrial plant workstations, the specific design methodology used by researchers (Cimino et al. 2009). compares the present day workstations with alternative configurations, carrying out specific analyses supported by a well-planned experimental design (based on multiple design factors and multiple performance measures). There is a plethora of international scientific literature on the effects and permanent physical damage that awkward work postures can cause to dock workers, damage compounded by the high vibration levels generated by the machines they operate. These complaints are commonly referred to as *Work-related MusculoSkeletal Disorders"* (WMSDs). By way of example, a representative study on occupational risk factors for low back complaints in sedentary workers (Burdorf A, 1993), examined the relationship between low back disorders and sedentary work in yard crane operators (94 subjects), straddle carrier drivers (95 subjects) and a group of office workers (86 subjects), aged between 25 and 60. The information on the extent of low back complaints and of the subjective working conditions were extrapolated from a cognitive survey. The postural load on the back was evaluated by observing awkward trunk postures during normal work activities. Whole-body vibration exposure of crane and straddle carrier drivers was also measured. Fifty percent of crane operators suffered from low back complaints within the first 12 months of starting work, against 44% of straddle carrier drivers and 34% of office workers. These findings indicate that sedentary work involving poor and constrained trunk posture represents a serious risk factor for low back disorders. Concerning the correlation between poor posture and vibration exposure, a specific report (Lane R., 1999) compiled in British Columbia (Canada) provided evidence of a connection between body vibration exposure and back disorders, in a variety of motor vehicle driving tasks (Bianconi et al. 2006). The report consists of a collection of scientific literature and of a number of electronic databases (Medline, EMBASE, NIOSHTIC, Ergoweb and Arbline), collating pertinent data for each topic. The scientific investigations chosen adopted a standard epidemiological approach to show the association between back disorders and vibration exposure. The risk factor was found to be high for a wide variety of tasks performed by lorry drivers, agricultural workers, tractor drivers, bus drivers, helicopter pilots, bulldozer drivers, fork-lift truck and yard crane operators. Risk assessments indicated strong associations between back disorders and vibration exposure, the risk increasing with increasing of work shift and vibration exposure duration whereas no increase was observed for increasing vibration intensity. Twenty-five studies covering a variety of job tasks showed that the machines used exposed operators to vibration levels well above the threshold limit established by the British Columbia's new health and occupational safety directives. The findings show a strong correlation between back disorders and whole-body vibration, the most common complaints being low back pain, sciatica, generalised back pain, disc herniation and intervertebral disc degeneration.. The risk is higher after the first five years of exposure.

3.2 Motion platform in quayside crane simulators

The cockpit, or simply the work station in the less sophisticated simulators, is generally placed on a motion base, with a varying number of degrees of freedom (2, 3, 4 or 6 with different loads). Thus the simulator does not only generate visual and auditory stimuli but above all allows the operator to perceive the sensation of movement

The platform consists of a combination of electric/hyraulic or pneumatic actuators installed beneath the cockpit that impart movement in response to operator input and to the task phases executed. The motion system duplicates the vibrations and shocks generated in normal operating conditions, synchronised in the simulation software with the audio and video systems. Various types of motion base exist (Figures 4b/c), the most widely used being the *"Stewart Platform"* (Figure 4a), which like all simulators was originally devised for flight simulators in the aerospace and aviation sector. The six possible movements (maximum degrees of freedom) that the motion base imparts are divided into two categories of motion: linear translational (*heave*, translation along the z axis; *surge*, translation along the x axis; *sway*, translation along the y axis – Figures 4d-f) or rotational (*pitch*, rotation around the y axis; *roll*, rotation around the x axis;; *yaw*, rotation around the z axis; – Figures 4g-i).

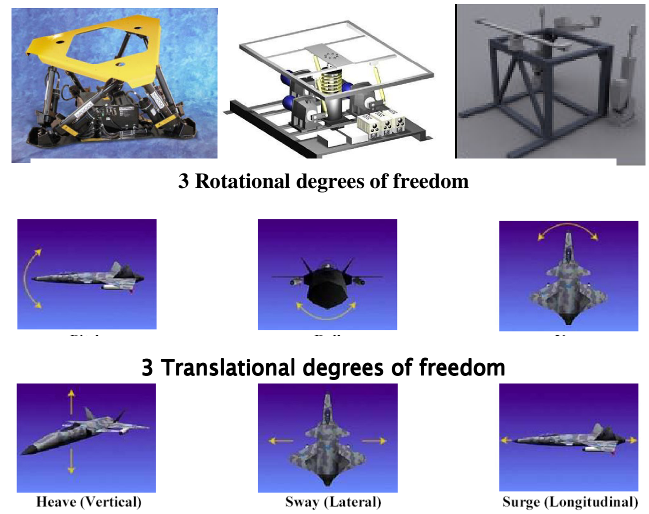


Figure 4a-i – Three different types of motion base (the first is a Stewart Platform) and the 6 degrees of freedom it achieves, depicted by the motion of a military aircraft in flight.

4. METHODOLOGY

Design the preparatory studies for the simulator design it was necessary to identify what type of real input data to load into the simulator software that also handles the motion platform so as to ensure correct response to the user commands, to the movements and to the task phases, as well as perfect synchronization with the audio and video systems.

The choice fell on vibration spectra. As mentioned in the introduction, a vibration measurement campaign was conducted preceded by a preliminary study to identify the strategic positions of the accelerometers inside the cockpit (based on previous experience with the CICT crane cabs). This campaign was carried out jointly with the Department of Mechanical Engineering at Cagliari University.) inside the cab of a CICT quayside crane in operation.

The accelerometers were positioned (Figure 5) such that the different linear and rotational acceleration components could be calculated using simple arithmetic; in particular:

- linear accelerations along X : Mean of signals $x1$ and $x2$;
- linear accelerations along Y : Signal $y1$;
- linear accelerations along Z : Mean of $z1$ and $z2$;
- angular accelerations around X : $(z2 - z1)/a$;
- angular accelerations around Y: $(z3-z2)/b$;
- angular accelerations around Z : $(x1 - x2)/c$

First of all (at Cagliari International Container Terminal on 21 September 2007) the work cycle of the quayside crane was examined, for a continuous 6-hour shift divided into the following key phases:

- Phase 0: Position spreader, secure container and hoist from quayside;
- Phase 1: Spreader travels with container along quayside to ship;
- Phase 2: Position container on ship and unlatch;

- Phase 3:Spreader returns to quayside starting position for next cycle.

The RMS values for each linear and angular acceleration component were calculated together with the maximum absolute values. Both values refer to consecutive samples of 1000 data, obtained at a scan rate of 2000 data/second, for a typical complete work cycle. In this way, it was possible to identify the critical acceleration values in the different phases.

Figure 6 shows the different phases identified, denoted by different colours. The same colours were used in Figs. 6a-c, which show, for the sake of simplicity, linear acceleration with respect to the three axes.

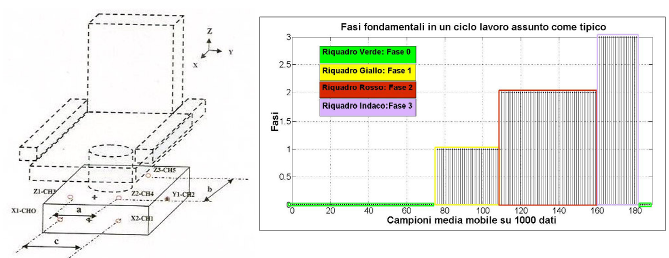


Figure. 5/6 – Position of motion sensors and work phases identified for a complete 6-hour shift

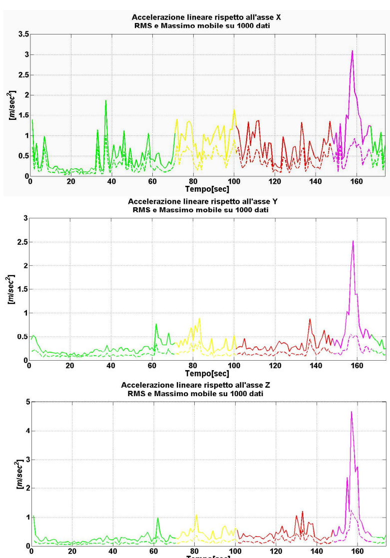


Figure 7a-c – Moving average over 1000 data of RMS values and maximum absolute values of measured linear accelerations

The vibration spectra proved to be essential for properly duplicating the psycho-physical stresses experienced by the operator through the motion base in question, during the training sessions.

However, the following difficulties and shortcomings were encountered during the first vibration measurement campaign:

- difficulties in loading the vibration spectra into the simulator software and especially in associating the recorded spectrum with the corresponding phase;
- difficulties in identifying the phases comprising the quayside crane operator's tasks to be duplicated singly with the simulator and in allowing the instructor to alternate simulation scenarios and to get the trainee to repeat the test.

In fact the original data sample was mixing different operation and it was very difficult to reattribute their characteristics to the simulation actions in addition it was necessary to consider the influence of boundary conditions; currently the approach is based on defining DRT (Data Reference Tables) related to each single operations and the relative boundary conditions, including current spreader weight (including container if present), crane speeds, wind speed, wind direction, etc.

The DRT include the reference to the pre-processed vibrations to be reproduced in the simulator corresponding to an original data sample; obviously pre-processing is required to adapt data sample to simulator device capabilities in order to guarantee fidelity.

ST_PT_1 analyze DRT during real-time operation, when a specific action is performed among that ones generating vibrations (i.e. movement of the crane); by mining the DRT the simulator finds the most appropriate configuration in the DataBase in relation to the current action and based on a weighted distance from the simulated boundary conditions and DRT.

For each action it is defined a default sample, in order to guarantee a basic vibration to be transferred to the simulator

So the proposed approach guarantees to being able to reproduce a specific crane just with few available data, and keep it open the possibility to extend this library over the time with new samples.

Based on these consideration it was necessary to conduct additional samples in order to correlate the vibrations with the boundary conditions.

These measurements did however bring to light one aspect that contributed to interfering with the crane operator's task. Vibration exposure limits are established by Decree Law No. 187 of 19 August 2005, implementing EC Directive 2002/44 on the minimum health and safety requirements regarding the exposure of workers to risks arising from physical agents. The vibration time histories generated during the work cycles, namely measures of acceleration along the three axes x, y and z, were recorded using motion sensors placed in appropriate points of the cab seat. Out of the 174 measurements along the x axis (corresponding to the back and forth motion of the cab for each operating cycle, roughly one every 1'30"-2') the RMS acceleration, which is in itself a mean value, exceeded the permissible daily action value of 0.5 m/s^2 for whole body vibrations 44 times, reaching as much as 1.00 m/s^2 and for groups even 5-7 continuous linear accelerations exceeding the limit.

RMS acceleration along the z axis only exceeded the daily action value five times, but on the other hand the peaks attained extremely high absolute values (one recording showed a RMS acceleration of 1.24 m/s^2

corresponding to a Max of 4.68 m/s^2). These operating conditions coincide with the cab position at the boom tip where the jerks caused by container locking/unlocking reverberate strongly at the free end of the gantry crane boom.

Accelerations along the y axis did not generate any significant whole body vibrations because of the nature of the operations in relation to the crane structure, which do not produce large shocks along the transverse direction.

In spite of the useful information gathered, it was decided to optimize operation of the vibration spectra loading process into the simulation software that also handles the motion base of the ST_PT_1 simulator, so as to duplicate the vibrational stresses associated with each phase of the task (Bruzzone et al. 2004).

4.1 Building a motion DataBase: Second non-continuous survey in a real world environment

The second series of measurements conducted at the CICT on 18 February 2009 was designed for the purpose of associating each recording taken with the accelerometers, once again placed on the cab seat (Figures 18-20), to the key task phases. This was done with the assistance of a crane operator who performed each task requested, with the technicians recording simultaneously.

The following activites were conducted by the team of collaborators from CIREM at Cagliari and DIPTEM at Genova, the latter responsible for constructing the quayside crane simulator:

- Identification of key operating phases to be monitored;
- Synchronised recording using accelerometers of each operating phase accomplished by the crane
- Creation of "talking" codes to load into the simulator software *(e.g. Cabt_S_Scenario 4.02_9_315_210_1000)* for perfectly reproducing the psycho-physical stresses on the operator's body (cyclic VV&V procedure).

For defining the "talking" codes, the name of the sample describing the scenario was indicated. For example

ID_K_Scenario 1_06_090_150_10

where:

- ID denotes the type of test;
- K denotes the load: S no-load, 4 4 tonnes, 30 30 tonnes, etc.;
- ex1 denotes the sample code (if measurements have been repeated for the same scenario then these will be numbered consecutively for example Scenario 1.01; 1.02 etc.);
- 06 indicates wind speed in knots;
- 090 indicates wind direction (degrees);
- 150 indicates travel speed expressed in m/min;
- 10 indicates instrument sampling rate in Hertz.

The talking codes were grouped into quayside crane travel, spreader movement and a number of additional manoeuvres that were identified on site during measurements.

Furthermore, the CICT crane operator was asked to repeat all the above movements for three different load configurations: Bromma spreader in standby (no load S:); spreader with container attached, medium load (load 4: 4 tonnes); spreader with container attached, full load (load 30: 30 tonnes). In this way it was possible to compile a database containing a wealth of information about the operating phases in the different configurations, in order to adequately duplicate the real world stresses experienced by the operator .

The measurements were taken in the six measuring points schematically indicated in Figures 8-10. The use of measurements taken along the three axes underneath the operator's seat, i.e. on the supporting structure, and on the seat frame itself, also allows to evaluate the vibration absorption effect exerted by the buffer placed between the support and the seat itself. During simulations it will be possible to evaluate other types of vibration dampening systems..

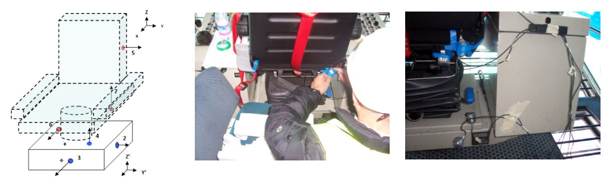


Figure 8-10 – Positioning the accelerometers on the CCIT crane operator seat for point measurements

Two types of measurements were obtained. Accelerometer recordings versus time sampled at a rate of 1 Hz, during the key phases of the work cycle and spectra recordings, consisting of consecutive series of 500 data representing the spectral components at a sampling step of 1 Hz. As the typical duration of operating phases is in the order of 10-15 seconds, each recording contains several thousand data.

Figures 14-16 show portions of some of the spectra recorded for the key operating phases:

- Talking code *Engc_4_Scenario 20.01_17_315_XX_1000* recording of phase Engage Container (Engc) load 4 tonnes (4) with 40 ft Bromma spreader, Scenario N°20, first recording (01), wind speed 17 knots, wind direction 315°, XX speed of the unknown operation for this phase, sampling rate 1000 (Figure 11);
- Talking code *Cabt_4_Scenario 4.02_9_315_210_1000* recording of phase Cab travel from shore to ship side max speed (around 15-20 m travel) (Cabt) with 4 tonne (4) container attached to 40 ft Bromma spreader, Scenario N°4, second recording (02), wind speed 9 knots, wind direction 315°, carriage travel speed 210 m/min

for QC type in operation at Cagliari container terminal sampling rate 1000 (Figure 12);

- Talking code *Jerkh_30_Scenario 23.01_6_315_85_1000* recording of phase Repeated movement (jerks) with cab at half boom operator manoeuvring cable take up drum joystick (Jerkh) with 30 tonne (30) container attached to 40 ft Bromma spreader, Scenario N°23, first recording (01), wind speed 6 knots, wind direction 315, cable winding/unwinding speed 85 m/min for QC type in operation at Cagliari container terminal , sampling rate 1000 (Figure 13);

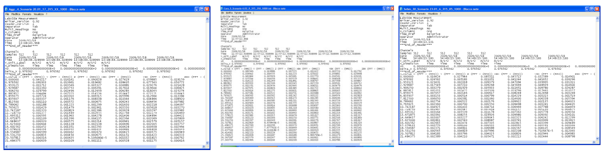


Figure. 11-13 – Portions of three recorded spectra extrapolated from the database created by the Department of Mechanical Engineering at Cagliari University.

Concerning the spectra, the sampling interval is a little less than 1 Hz because the Fourier transform is performed on 1024 and not on 1000 data in order to maximise algorithm efficiency.

It was also possible, as for the previous measurements, to easily evaluate the severity of vibrations experienced by the operator on the basis of existing legislation concerning occupational exposure to vibrations: extremely high acceleration levels were observed during emergency stops.

The most interesting aspect of the second series of measurements concerned the possibility of creating a universal database, that can be readapted to subsequent vibration measurements on other types of quayside cranes operating worldwide. The choice of vibration spectra as input data, and the current handling mode consisting of talking codes for associating vibrations to each operating phase, will facilitate loading, after measurements of vibrational stress data for all other types of quayside cranes. This makes the crane simulator at Cagliari University an extremely flexible and versatile tool for training activities but it can also be used to monitor and record fatigue, duplicating the real world psycho-physical conditions to which crane operators are exposed in the container terminals who apply for training courses.

4.2 Motion Generation

The original layout of the quayside crane simulator cockpit envisaged a Stewart type motion platform with six degrees of freedom, consisting of six prismatic actuators. This platform was able to move simultaneously in all six degrees of freedom (*Pitch, Roll, Yaw, Surge, Heave, Sway*), but proved too cumbersome to fit inside the shelter housing the simulator.

Thus, an alternative motion base had to be found that fitted comfortably onto the main platform considering the following:

- Maximum travel to be generated to reproduce the movements of the crane used;
- Maximum permissible dimensions;
- Easy access for maintenance.

It was also necessary to reduce the overall height of the main platform (including the workstation Figure 14) to enable the operator to perform tasks standing up (Figure 15).

The motion base chosen (LamceCoppeSIM40, Figure 16) consists of 5 pneumatic actuators (motion system that facilitates calibration operations) positioned as follows: 2 pairs of actuators on the platform sides each connected to the same base, the front ones fixed in a vertical position with the upper part free to rotate 360°, the others at an angle and oriented towards the back part of the platform, the fifth single actuator at the back moves almost 90°. The actuators are interconnected and connected to a console and require a starting power of up to 20 Kw. The platform is equipped with a safety valve for dampening the loud noise that consists either of a number of small sensors or of just one larger sensor that can also be installed outside the shelter.

The user friendly software for the check controls, such as joysticks, lights etc. is designed to check the control panels used by the operator, equipped with 12 Volt sensors, are functioning properly.

Figure. 14-16 –LamceCoppeSIM40 motion platform of the ST_PT_1 simulator

The technical specifications of the motion platform are as follows (Table1):

Table 1 – Technical characteristics of the LamceCoppeSIM40 motion platform

Degrees of Freedom	MIN	MAX
Pitch	- 11°	+ 11°
Roll	- 8°	+ 8°
Yaw	- 7°	+ 7°
Heave	- 0,25 mt	+ 0,25 mt
Surge	- 0,20 mt	+ 0,20 mt
Sway	- 0,20 mt	+ 0,20 mt
Dimensions		
Height (stand-by mode)	70 cm	
Length	135 cm	
Weight (including workstation) dell'operatore)	≈150 kg	
Maximum useful load	600 kg	

The maximum translational and rotational travel is somewhat less than the Stewart motion platform, The height in stand-by mode is roughly the same but with this configuration about half of the platform forms part of the operator workstation.

The psycho-physical stresses generated by the 6 DOF motion platform are: *shocks* (spreader colliding with container to be engaged or with ship's cells, rapid movements of the carriage along the rails and sudden stops at end stops, cab settling, abrupt release of container onto truck in buffer area, etc.) that the platform motion reproduces (pneumatic actuators extend or retract in the six degrees of freedom), vibration effect duplicated by a bus shaker, a kind of subwoofer placed inside the cube-shaped support on which the operator's seat is mounted that generates a sensation of slight shuddering (this system operates at low frequencies, 0-40 Hz) throughout the operator's body (caused for example by oscillations of the cables on which the spreader and container are suspended), that combined with the shocks generate a realistic sensation of the actual stresses experienced .

The principle of the VV&A, concerns the duplication of the innovative and technological data input into the simulator's conceptual (or descriptive) model that should ensure perfect duplicability of the real world environment (Bruzzone and Mosca 1998). This will be described in the next section. This cyclic procedure has also been applied to the vibrations generated by the motion platform, for reproducing in a virtual environment of the above psycho-physical stresses, i.e., those that contribute to the onset of operator fatigue.

The linear and angular accelerations recorded versus time in the second survey, loaded into the software that handles the simulator, with some corrections (raw data filtering and cleaning), allow to reproduce the combined (*shocks-shaker*) vibration effect, Thus the *loop* terminates with the *feedback* (actuator movements) from the platform to the simulator software.

The integration of a real time simulator with a motion system able to reproduce the vibration transmitted to the operator requires an heavy workload and involves different aspect (i.e. engineering of the platform, sampling and processing of data, enginnering of the vibration/motion system, software integration with the simulator). The author focused their research on the sampling and post processing of the signal.

4.3 Data collection

In order to have a significant number of samples, releted to the respective operation, the authors start creating DRT that permit to match (the rows) all the operation/movement that the crane operator do, during their job (i.e. lateral movement of the crane on the dock) with the operative and boundary condition (i.e without payload, with a 30 ton container on…). Each cells of the DRT represent a sample that must be collected, as additional information each sample was completed with wind speed and direction.

To collect data a real crane control cabin and seat was "customized" with 2 sets of accelerometers that permit to sample the acceleration/vibration on the 3 axis. The choice to use 6 accelerometers (2 set) is due to the fact that the authors want to know what is smoothing effect of the damper connect to the chair, then the first set detect the vibration transmitted to crane operator and the second one the vibration transmitted to cabin frame.

The data was sampled using a sample frequency of 1 kHz that permit, according to Shannon theorem, to detect 500Hz as max frequency (without aliasing).

4.4 Engineering of the motion system

Now the problem is reproduce the vibration sampled the system adopted is a mix of two different actuator:

- Pneumatic actuator
- Electromagnetic actuator

The idea is to reproduce low frequency vibration (less than 10 hertz) with the pneumatic actuator and the high frequency vibration with electromagnetic actuator. The electromagnetic actuators permit to reproduce vibration if connected to an audio amplifier driven directly by the sound card of the workstation used for run the simulator. Using a some of these actuators in a stereo configuration it becomes is possible to reproduce the vibration on the 3 axis. The problem is that the frequency response of the amplifier is linear on a bandwith 20-20000 Hz; this introduce a problem related to the window 10-20 Hz. The solution adopted was to boost the harmonic in that range according with the frequency response curve of the amplifier (i.e. 3dB/decade).

4.5 Post processing analisys

The data collected are represented as timeseries (the x axis is time in ms and the y axis is acceleration expressed in m2/s), applying the DFT (Discrete Fourier Transform) to each time series the authors obtained the frequency analysis of each sample. The next step was to correct the gain of each Harmonic in the window 10-20Hz according to the frequency response of the stereo amplifier used, as explained before. In our case for instance the DFT generates 2500 harmonics from 0 to 500 Hz, due to the fact that each sample have a duration of 5 seconds corresponding to 5000 data plots.

The authors analized the spectrum of each signal, for all samples it was decided to consider harmonics higher than 100Hz not significant and so based on the filter setting these component was void for rebuilding the vibration. The next step on the analysis was to identify the most significant harmonics for each sample, to make this the author apply a Pareto analysis focusing only most influent harmonics, contributing till the 90% of the whole vibration. This approach, obviously, generates a new pre-processed data set to be used in the simulator that result different respect original signals and introduces distortions; therefore in the proposed application this different resulted not significant for the main goal of this research that focus on the reproduction of vibration outside audible spectrum. As final pre-processing the reconstructed vibration based on modified harmonic modified spectrum was used to generate reference .wav files (22.1 kHz sampling frequency) using a multitrack recording software in order for being easily reproducible by sound cards and bass shakers

5 CONCLUSIONS

The analysis conducted highlighted needs and methodologies devoted to introduce specific vibrations in real-time simulators in order to reproduce specific real equipment. The case study proposed is related to ST_PT_1, a full scope simulator, adopted by C.C.S.Tra. for being used in training and R&D; in these context the requirement for effective reproduction of vibration spectrum becomes very hard.

The proposed approach is based on dividing phenomena attributing accelerations due to the cinematic and position of the vehicle from the vibrations. This approach emphasis the necessity to recreate realistic configurations of the environment in order to reproduce phenomena such as stress, fatigue, while analytical modelling of suspensions and vibrations requires to introduce many parameters that are very difficult to be estimated. The pre-processing of simulation data from real data sample, allows to tune the influence of all boundary conditions and to match the technical characteristics of the simulation infrastructure. The authors are working on additional developments for extending the database as well as for conducting researches on different kind of port cranes and terminal.

REFERENCES

Bianconi F.; Saetta S.A; Tiacci L.(2006) "A web-based simulation game as a learning tool for the design process of complex systems", Journal of Design Research, Vol.5,Iss.2;p.253-272

Bruzzone A.G., Mosca R. (1998) "Special Issue: Harbour and Maritime Simulation", Simulation, Vol.71, no.2, August

Bruzzone A.G., Frydman C., Giambiasi N., Mosca R. (2004) "International Mediterranean Modelling Multiconfernece", DIPTEM Press, ISBN 88-900732-4-1 Vol I e II (884 pp)

Bruzzone A.G., Guasch A., Piera M.A., Rozenblit J. (2006) "International Mediterranean Modelling Multiconference", Logsim, ISBN 84-690-0726-2

Bruzzone A.G., Bocca E., Longo F., Massei M. (2007) "Logistics Node Design and Control over the Whole Life Cycle based on Web Based Simulation", Int.Journal of Internet Manufacturing and Services, Vol. I, Issue 1,pp. 32-50

Bruzzone A., Fadda P., Fancello G., Bocca E., D'Errico G., Tremori A. (2008) "Ship-to-shore gantry crane simulator design: crane operator performance anlysis and assessment tool of Cagliari University", Proc. of the HMS, September 17-19, 2008, Campora San Giovanni, Amantea (CS), Italy

Burdorf A., Naaktgeboren B., de Groot HC. (1993) "Occupational risk factors for low back pain among sedentary workers", J Occup Med, 1213–1220, Volume 35, pp. 1213–1220, 1993

Cimino A. Longo F. & Mirabelli G. (2009) "A multimeasure-based methodology for the ergonomic effective design of manufacturing system workstations", International Journal of Industrial Ergonomics, Volume 39, Issue 2, pp. 447-455, March 2009

D'Errico G. (2009) ".L'analisi dei Fattori Umani: progetto e realizzazione di un simulatore fisico di gru portainer", Thesis (PhD). Consortium Caglairi-Palermo Universities;

Fancello G., D'Errico G. & Fadda P. (2008) "Processing and analysis of ship-to-shore gantry crane operator performance curves in container terminals", Journal of Maritime Research (JMR), Spanish Society of Maritime Research (SEECMAR) University of Cantabria, Santander, Spain, Volume 5, Number 2. 2008. ISSN: 1697-4840, pp. 39-54, 2008;

Fancello G., D'Errico G. & Fadda P. (2008) "Human factors involved in container terminal ship-to-shore crane operator tasks: operator fatigue and performance analysis at Cagliari Port", In D. de Waard, F.O. Flemisch, B. Lorenz, H. Oberheid, and K.A. Brookhuis (Eds.): Human Factors for assistance and automation, Maastricht, Shaker Publishing, pp. 479-492, September 2008;

Ircha M.C. (2001) "Serving Tomorrow's Mega-Size Container Ships The Canadian Solution", Int.Journal of Maritime Economics, Vol.3, 318-332 Lane R. (1999) "Whole Body Vibration and Back Disorders Among Motor Vehicle Drivers and Heavy Equipment Operators", Workers Compensation Board of British Columbia

Meloni M., Del Rio A., Setzu D., Cocco P., D'Errico G., Fancello G., Fadda P., (2009), "Analisi ergonomica, posturale e valutazione del rischio da vibrazioni nell'intero corpo in operatori di gru portainer del Porto di Cagliari". Proceedings of the XXV Giornate Internazionali di Medicina del Lavoro, May 20-22, 2009, Portofino Vetta, Genova, Italy;

Merkuryev, Y.; Merkuryeva, G.; Piera, M.À. (2009) "Simulation-Based Case Studies in Logistics: Education and Applied Research"; XXIV, 232 p. 125 illus., Hardcover, ISBN: 978-1-84882-186-6 Guasch, A. (Eds.)

Seck, M., Frydman, C., Giambiasi, N., Ören, T.I., & Yilmaz, L. (2005) "Use of a Dynamic Personality Filter in Discrete Event Simulation of Human Behavior under Stress and Fatigue", 1st International Conference on Augmented Cognition. Las Vegas: Nevada

AUTHORS BIOGRAPHY

Agostino G. Bruzzone since 1991, he has taught "Theories and Techniques of Automatic Control" and in 1992 he has become a member of the industrial simulation work group at the ITIM University of Genoa; currently he is Full Professor in DIPTEM.

He has utilized extensively simulation techniques in harbor terminals, maritime trading and sailboat racing sectors. He has been actively involved in the scientific community from several years and served as Director of the McLeod Institute of Simulation Science (MISS), Associate Vice-President and Member of the Board of the SCS (Society for Modelling & Simulation international), President of the Liophant Simulation, VicePresident of MIMOS (Movimento Italiano di Simulazione) and Italian Point of Contact for the ISAG (International Simulation Advisory Grouyp) and Sim-Serv. e has written more than 150 scientific papers in addition to technical and professional reports in partnerships with major companies (i.e. IBM, Fiat Group, Contship, Solvay) and agencies (i.e. Italian Navy, NASA, National Center for Simulation, US Army).

Paolo Fadda graduated in Civil Engineering (Transport) from the University of Cagliari in 1977.
He was a founder member of the International Center for Transportation Studies (ICTS) established in 1981 and member of the Scientific Committee. From 1981 to 1987 he was managing director of SST, a service company of the ICTS. He has authored numerous books, papers, treatises and articles on transport-related issues and has been granted patents in the area of urban public bus transport. From September 1994 to January 1998 he was Director of the Public Works Department of the Sardinian Regional Government. He also served as government vice-commissioner for Water Emergency during the period 1995-1997. From 1999 to 2003 he served as external expert on transport-related matters undertaking preparatory work for the Higher Council of Public Works projects of national importance. Appointed by the Ministry for Public Works, he has sat on the scientific committee supporting the Government Commission for Water Emergency in Sardinia since 2001. In February 2003 he was nominated Italian representative for the Commission of International Cooperation on Maritime Transport of PIANC (International Navigation Association) by the President of the Higher Council of Public Works. In September 2007 he was appointed Managing Director of Cagliari Port Authority.

Gianfranco Fancello was born in Nuoro (Italy) on 28 August 1965; and has been living in Cagliari since 1986. He is married and has two daughters. He is researcher at the Department of Land Use Engineering at Cagliari University (Italy). He obtained a 1st class honours degree in Civil Engineering (Transport) from the Engineering Faculty at University of Cagliari. In 1994 he was awarded a post-graduate diploma in town planning: land use and environment at Cagliari University; and in 1995 a master degree in assessment of information processes in the business sector, organised by the CIFRA consortium. In 1999 he completed his Phd research in Transport Technology and Economy at Cagliari University, discussing hos thesis “Retail trip generation for Italian case. Principles and methodologies for sector planning and for instructions predispositions”. His research interests lie in transport policy planning and assessment, safety and human factors in transport, logistics and freight transport, predictive models for impact of transport and mobility. He has co-authored 2 books and published some 40 papers in scientific journals and conference proceedings. He is a member of several local, national and international research groups.

Enrico Bocca achieved the Logistics and Production Engineering degree in 1999 with a thesis on "Automated Reorder System in Supermarket Chains", summa cum laude. During 2000-2001 he was involved in Customer Satisfaction analysis for Retail Business Sector. He completed during 2003 the thesis in Management Engineering working on "Innovative Technology Management in Retail Sector", summa cum laude. He participated to IEPAL experience coordinating an International Team of Engineering and MBA Students (from France, USA, Turkey) working in a Prefeasibility study related to Crowd Control in Shopping Centers; this initiative was sponsored by FIPSE US Dept.of Education, DGEAC EC, LSC. He had experiences as consultant on different initiatives: Business Plans related to R&D Projects; Logistics and Project Management for CFLI. He was visiting researcher in Kennedy Space Center working on a joint project DIPTEM/NASA for Modelling Test Lab Crane Operations, the related demonstrator was presented officially in I/ITSEC2005 Orlando. He worked as researcher in Simulation Team of DIPTEM with special attention to Logistics. Currently he is charge as Responsible of M&S R&D for MAST.

Gianmarco D’Errico was born in Cagliari (Italy) on 17 July 1977; he obtained his degree in Civil Engineering (Transport) from the University of Cagliari in 2004, discussing the thesis entitled “Hazardous arsenic containing materials transportation risk: the Baccu Locci case”. He is currently a PhD student in Transport Techniques and Economy organized by the Universities of Palermo and Cagliari. Since 2004 his research interests have concerned safety and human factors in container shipping, logistics and intermodal transport. He is also involved in a project on the design and construction of the ship-to-shore crane simulator to be used for researching human factors affecting crane operator performance.

Marina Massei is currently enrolled in DIPTEM University of Genoa as member of the Simulation Team of Prof.Agostino Bruzzone. She is involved in the organization of International Scientific Events (i.e. Summer Computer Simulation Conference 2003-Montreal, 2004-San Jose', 2005-Philadelphia) She served as Chair for Applications in Management, Planning & Forecasting Workshop in SummerSim 2004, 2005, 2006. She is Associate Director of the MISS Center at the University of Perugia She was Finance Control & Administration Director of MAST.

A GENETIC ALGORITHM FOR REAL-TIME OPTIMISATION OF DRAYAGE OPERATIONS

Alejandro Escudero Santana[(a)], Jesús Muñuzuri Sanz[(b)], Luis Onieva Giménez[(c)], Miguel Gutierrez Moya[(d)]

[(a)(b)(c)(d)] Grupo Ingeniería Organización. Dpto. Organización Industrial y Gestión de Empresas. Escuela Superior de Ingenieros. Universidad de Sevilla. Avd. de los Descubrimientos, s/n 41092. Sevilla, Spain.

[(a)]aescudero@esi.us.es, [(b)]munuzuri@esi.us.es, [(c)]onieva@esi.us.es, [(d)] mguti@esi.us.es

ABSTRACT

Proper planning of drayage operations is fundamental in the quest for the economic viability of intermodal freight transport. The work we present here is a dynamic optimization model which uses real-time knowledge of the fleet's position, permanently enabling the planner to reallocate tasks as the problem conditions change. Stochastic trip times are considered, both in the completion of each task and between tasks

Keywords: Intermodal transport, drayage, genetic algorithm, stochastic travel times

1. INTRODUCTION

Road transport has been and continues to be prevalent for the on land movement of freight. However, increasing road congestion and the necessity to find more sustainable means of transport has determined different governments to promote inter-modality as an alternative For inter-modality to become viable for trips shorter than 700 km a cost reduction is necessary. Final road trips (drayage) represent 40% of the intermodal transport costs. There is potential to overcome this disadvantage and make intermodal transport more competitive through proper planning of the drayage operation.

Originally, optimisation efforts focused on drayage operations concentrate on the cost and service quality improvements to be expected from the collaboration between drayage companies. Along this line, Morlok and Spasovic (1994) develop an integer programming model to plan truck and container movements in a centralised manner. They contemplate different payment options for drayage services and conclude that centralised management of drayage operations would result in savings between 43% and 63%, as well as an improvement in service quality.

Following the path opened by De Meulemeester et al (1997) and Bodin et al (2000), the number of references on centralised drayage management has increased significantly over the last years, but most of them consider the problem only from a static and deterministic perspective. The main objective is normally the assignment of transportation tasks to the different vehicles, often with the presence of time windows (Wang and Regan 2002). The first part of the work by Cheung and Hang (2003) develops a deterministic model with time windows, which is then solved by means of the discretisation of each task's start and end time, and by incorporating the concept of dummy tasks for the beginning and the end of the vehicle's day. Ileri et al (2006) cover a large number of task types, both simple and combined, and of costs involved in drayage operations, and solve the problem with a column generation method. Smilovik (2006) and Francis et al (2007) incorporate flexible tasks where either only the origin or the destination is precisely known.

Many works also allow for randomness in the generation of tasks (Bent and Van Hentenryck 2004; Bertsimas 1992; Gendreau et al 1995) or dynamism in their assignment (Bent and Van Hentenryck 2004; Psaraftis 1995; Wang et al 2007). However, it is hard to find randomness in trip times (Laporte et al 1992), which is appropriate when the intermodal terminal requiring drayage operations is close to a large urban centre. Cheung and Hang (2003) and Cheung et al (2005) do consider the dynamic and stochastic characteristics of the drayage problem and solve it with a rolling window heuristic, but this randomness affects only the duration of the task, and not the displacement time between different tasks.

The work we present here considers random trip times both in the completion of each task and between tasks. It also incorporates the real-time knowledge of the vehicle's position, which permanently enables the planner to reassign should the problem conditions change.

2. DESCRIPTION OF DYNAMIC DRAYAGE PROBLEM

The drayage operation can be modeled as a Multi-Resource Routing Problem with Flexible Tasks (MRRP-FT) (Smilowitz 2006). In a MRRP-FT multiple resources have to be used to complete a series of tasks. The MRRP-FT is defined as follows:

Given: A set of tasks (both well defined and flexible), that require some resources, with service times for each resource and time windows; a fleet of each resource type; operating hours at all locations; and a network with stochastic travel times.

Find: A set of routes for each resource type that satisfies all the tasks while meeting an objective function (minimize operation costs) and observing operating rules for both tasks and resources.

The region where the drayage operations are performed is represented by a graph ***G = (N, A)***. The nodes $i \in N$ represent the different facilities of interest for the problem: terminals, depots, loading/unloading points. To each of these nodes is associated a time to attach/detach the container to/from the vehicle, τ_i. Between each pair of nodes $i,j \in N$ there is an arc (i,j) characterized by the transit time τ_{ij}, unknown in advance. The transit time will have a discrete distribution, ***T*** if known.

Every day a series of drayage tasks $\mathcal{T}$ must be completed, and the failure to do so implies a given subcontracting cost. The drayage tasks can be classified as: well-defined tasks, $\mathcal{T}_w$, and flexible tasks, $\mathcal{T}_f$. To each $t \in \mathcal{T}$ is associated a time window [a_t^{ini}, b_t^{ini}] that limits the time period in which the task has to be completed.

Well-defined tasks represent movements between terminals and customers or vice versa, and both the origin $o_t \in \boldsymbol{N}$ and destination $d_t \in \boldsymbol{N}$ of the movement are known. Time windows for well-defined tasks can be relaxed, as shown in Figure 1: if the task represents the pickup of a container in the terminal, it can never start before the arrival of the train or vessel, while if the drayage driver is late the task can still be completed, but a given amount will have to be paid for the time the container spends waiting at the terminal. In a similar manner, if the task represents the delivery of a container to the terminal and it is completed before the allocated time, the container will also be subject to a waiting cost. The cost has been considered proportional to the waiting time.

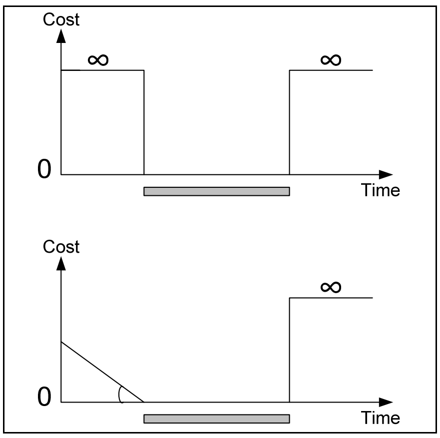


Figure 1: Types of time windows considered for well-defined tasks: hard (above) and relaxed (below).

Flexible tasks represent the movement of empty containers between customers and the depot. The movement of delivery or collection of an empty container can take place between a customer and the depot, but also from a customer who has requested the collection of an empty container directly to another who has requested the delivery of an empty container, given that their time windows overlap. Therefore, for flexible tasks only the origin or destination is known a priori, and therefore multiple scenarios, $\mathcal{R}_t$ are possible. The set of all movements, both well-defined tasks and different scenarios generated by possible flexible tasks, is $\mathcal{M}$.

In order to perform all the tasks a set of resources is available: containers, vehicles and drivers. The containers are linked to the movement of the tasks with no restrictions. Driver-vehicle pairs are considered, $\mathcal{V}$. Each pair is characterised by a location where the working day starts and ends. The different drivers have a time window for the start of their working day [a_v^{ini}, b_v^{ini}] and cannot work longer than MAX_v hours a day. The pairs driver-vehicle have different costs per unit of time depending on vehicle stopping, c^w or moving; and in the case of movement depending on the average speed the task is completed with.

A geographic positioning system by satellite (GPS, Galileo, Glonass) provides real time information about the position of each vehicle. This data is used to improve the solution dynamically.

3. METHODOLOGY

The static drayage problem is a NP hard problem extremely difficult to solve analytically. Exact solutions have been found for small problems, but computation time is high. The stochastic problem appears undoubtedly unsolvable analytically, even more so if flexible tasks are incorporated. Furthermore, the use of the real time information about the geographic position of the vehicles requires a high-speed procedure to find the solutions. An evolutionary algorithm has been used to solve the problem following the procedure shown in Figure 2.

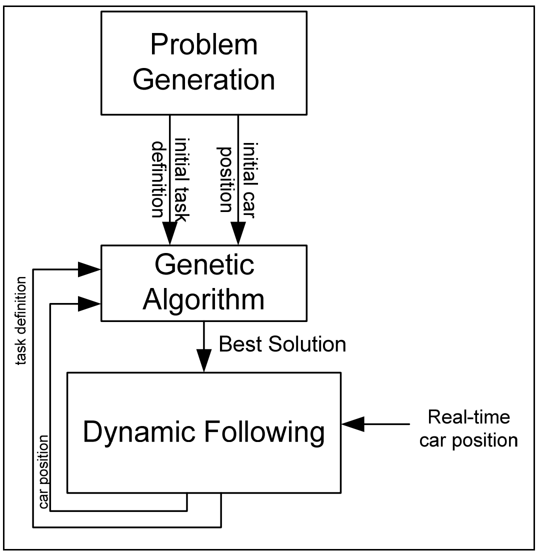


Figure 2: Schematic representation of dynamic drayage management with the proposed genetic algorithm.

The genetic algorithm used for solving the problem is as follows:

Genetic Algorithm

population = **InitPopulation**
for i=1:max_iter
    fitness = **Evaluation** (population);
    parents = **SelectionTOP;**
    child1 = **GeneticCross**(parents);
    child2 = **Mutate** (parents);
    population=population+child1+child2;
    dead=**SelectionBOTTOM**(population);
    population = population – dead;
    population=PopuGeneration
end

The chromosome which represents each solution is as shown in Wang et al (2007). In this representation, each chromosome is composed of some genes and each gene represents a task to complete. Each task is associated to a fixed gene. This gene is characterized by four features, first being the vehicle to which the task is associated, and is used to identify the order in which each vehicle completes the tasks. For example:

Table 1: Example of chromosome

1	2	3	4	5	6
1.123	1.673	2.234	1.942	2.440	2.294

The routes represented by the above chromosome would be:
Vehicle 1: task1→ task2→ task4.
Vehicle 2: task3→ task6→ task5

The parameters of the genetic algorithm were tested with a sample of problems, and no clear tendency was observed in its performance. The population size was finally set to 100 individuals, 99 of which were initially generated at random and the last one by an insertion heuristic, which also provided the base for comparison of the effectiveness of the algorithm. In each generation, 4 are selected out of the 10 best individuals, and they are then allowed to cross and mutate with probabilities of 0.9 and 0.1 respectively. 4 out of the 10 worst individuals are then eliminated from the resulting population. The repetition of individuals is allowed in the population, which speeds up the performance of the algorithm, and when the average fitness of the population is only 10% worse than the best individual the population is regenerated randomly except only for that best individual.

The crossover operator switches the genes of two parents between two tasks which are selected randomly, tasks 2 and 4 in the example of table 2.

Table 2: Crossover operator

Task	**1**	**2**	**3**	**4**	**5**	**6**
P1	1.123	**1.673**	**2.234**	**1.942**	2.440	2.294
P2	2.432	**1.721**	**2.325**	**1.987**	1.006	1.396
C1	1.123	**1.721**	**2.325**	**1.987**	2.440	2.294
C2	2.432	**1.673**	**2.234**	**1.942**	1.006	1.396

The mutation operator selects randomly a gene of the parent individual and changes its first digit to another possibility (See Table 3).

Table 3: Mutate operator

Task	**1**	**2**	**3**	**4**	**5**	**6**
P	1.123	1.673	**2.**234	1.942	2.440	2.294
C	1.123	1.673	**1.**234	1.942	2.440	2.294

The fitness of each individual represents the total costs of the resulting routes. The costs contemplated in each route are:

- Fixed cost per vehicle
- Distance cost
- Waiting cost of containers at the terminals due to early arrival or late collection
- Cost of task loss, assimilated to the subcontracting cost of that task to an external company

However, trip times are stochastic, so the fitness needs to be calculated as an estimation of the expected costs. An iterative algorithm was developed to complete that estimation, calculating the probability of reaching the next link of the route at a given time and the resulting costs involved. If the arrival time of the vehicle to the beginning of a given task is prior to the opening of its time window, the vehicle will wait, or else incur in a proportional cost. On the other hand, if the arrival is posterior to the closure of the time window, there is a higher penalty due to the waiting cost at the terminal or to the possible task loss (because of the departure of the train or vessel). If two tasks on the same route are both flexible and complementary, they will be combined and completed at the same time, thus avoiding the return to the depot.

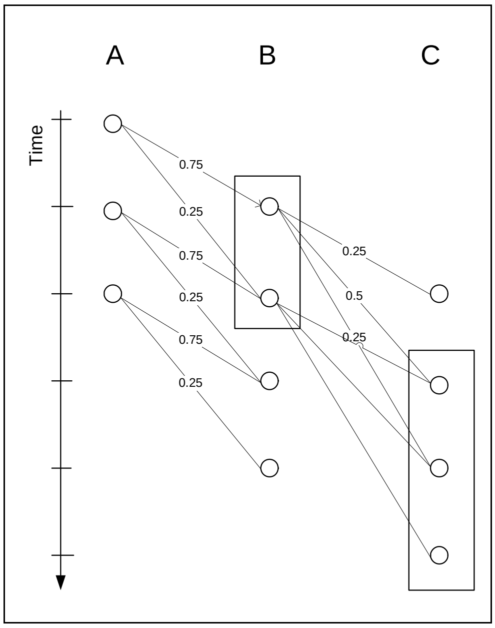


Figure 3: Random Transit Time

With the real time information about the position of the vehicles, the input data to the algorithm will be dynamically updated and used to find the best routes depending on the current circumstances. This update can be done:

- Every a fixed time, for example 15 min.
- When a task is finished
- When a car position is diverted of its expected position.

4. TEST PROBLEM AND RESULT

In order to test the performance of the algorithm for problems of different size and characteristics, we built a set of random drayage problems using the problem generator (see Table 4).

Table 4: Problem set

Problem code	Task number	No of well-defined tasks	No of flexible tasks	Fleet size
A1	20	0	20	5
A2		5	15	5
A3		10	10	5
A4		15	5	5
A5		20	0	5
B1	30	0	30	7
B2		5	25	7
B3		10	20	7
B4		15	15	7
B5		20	10	7
B6		25	5	7
B7		30	0	7
C1	40	0	40	9
C2		10	30	9
C3		20	20	9
C4		30	10	9
C5		40	0	9

The generator of problem randomly distributes the customers, the intermodal terminal and the depot in a 100x100 area (Example in Figure 4). The well-defined tasks consist, with equal probability, either of pickup or delivery of containers at the terminal, and flexible tasks will imply either collection or delivery of empty containers at the customers, also with equal probability.

Time windows for well-defined tasks range from 30 min. to 4 h. with a uniform stochastic distribution, and their start time is fixed randomly in the day. Time windows for flexible tasks will be open from the beginning of the day until the specified time for empty container deliveries and from the specified time until the end of the day for empty container collections. Those specified times are also generated randomly with a uniform distribution.

To simplify calculations, the time horizon is discretised in 5 minute intervals. Finally, to simulate in real time the position of each vehicle, a speed uniformly distributed between 45 and 55 km/h is calculated for each 5-minute period.

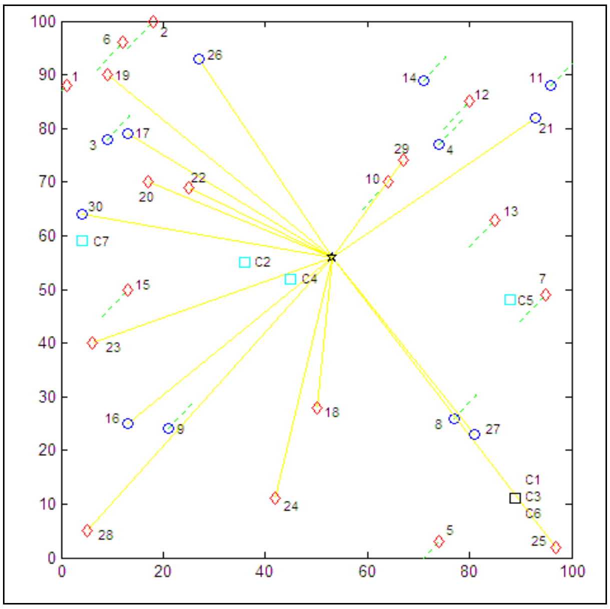


Figure 4: Random test problem example.

For each random problem, we determined the improvement of the genetic algorithm with respect to the insertion heuristic in the first iteration (see Table 5, column 2), the average improvement of the estimated cost for the best solution in iteration *i+1* with respect to the simulated cost on iteration *i* (column 4), and the estimated cost reduction between the first and last iteration of the genetic algorithm (column 5).

Table 5: Results obtained for the random problem set

Problem code	Genetic vs. Insertion Improvement (%)	No of iterations	Average improvement of the GA (%)	Dynamic Improvement (%)
A1	12.7	6	2.19	31.88
A2	0.8	7	1.47	16.48
A3	0	8	0.28	16.08
A4	2.88	11	1.41	31.16
A5	9.6	14	3.46	31.11
B1	3.61	7	1.21	22.75
B2	4.36	9	1.27	39.41
B3	2.26	11	1.04	30.69
B4	4.04	9	1.43	23.04
B5	0	12	0.21	32.91
B6	1.36	13	1.22	30.19
B7	7.74	16	1.45	38.32
C1	1.66	7	0.68	12.87
C2	0.17	9	0.82	16.34
C3	1.42	13	1.85	25.13
C4	4.84	17	2.39	37.34
C5	9.83	18	2.05	33.42

5. CONCLUSION

We have shown in this paper the importance of the exact knowledge of real-time vehicle locations in a drayage fleet, through the use of a satellite positioning system. This knowledge, together with an optimization algorithm based on metaheuristics, enables real-time management of the fleet in a changing environment, which reduces operation costs by as much as 30%. These results are especially valuable for intermodal operations in congested metropolitan areas, where travel times are stochastic due to congestion. Besides, given that we modeled the problem as a MRRP with flexible tasks, both intermodal drayage operations and the repositioning of empty containers can be optimized at the same time.

To solve the drayage problem, we developed a real-time optimization model based on a genetic algorithm that operates with stochastic cost estimations, and we tested it with a series of drayage problems generated randomly. The genetic algorithm improves the initial solution, provided by an insertion heuristic, with an average improvement of around 2% in each dynamic iteration for the type of problems considered.

ACKNOWLEDGMENTS

This research was supported from Era-Star Regions Program and regional government of Andalusia, grant SR-0197/2008 (Project Galileo-Drayage)

REFERENCES

Bent, R. W. & Van Hentenryck, P. (2004). Scenario-Based Planning for Partially Dynamic Vehicle Routing with Stochastic Customers. *Operations Research*, vol. 52, no. 6, p. 977.

Bertsimas, D. J. (1992). A vehicle routing problem with stochastic demand. *Operations Research*, vol. 40, no. 3, pp. 574-585.

Bodin, L., Mingozzi, A., Baldacci, R., & Ball, M. (2000). The Rollon–Rolloff Vehicle Routing Problem. *Transportation Science*, vol. 34, no. 3, pp. 271-288.

Cheung, R. K. & Hang, D. D. (2003). A time-window sliding procedure for driver-task assignment with random service times. *IIE Transactions*, vol. 35, no. 5, pp. 433-444.

Cheung, R. K., Hang, D. D., & Shi, N. (2005). A labeling method for dynamic driver-task assignment with uncertain task durations. *Operations Research Letters*, vol. 33, no. 4, pp. 411-420.

De Meulemeester, L., Laporte, G., Louveaux, F. V., & Semet, F. (1997). Optimal sequencing of skip collections and deliveries. *Journal of the Operational Research Society*, vol. 48, no. 1, pp. 57-64.

Francis, P., Zhang, G., & Smilowitz, K. (2007). Improved modeling and solution methods for the multi-resource routing problem. *European Journal of Operational Research*, vol. 180, no. 3, pp. 1045-1059.

Gendreau, M., Laporte, G., & Seguin, R. (1995). An exact algorithm for the vehicle routing problem with stochastic demands and customers. *Transportation Science*, vol. 29, no. 2, pp. 143-155.

Gendreau, M., Hertz, A., & Laporte, G. (1994). A Tabu Search Heuristic for the Vehicle Routing Problem. *Management Science*, vol. 40, p. 1276.

Ileri, Y., Bazaraa, M., Gifford, T., Nemhauser, G., Sokol, J., & Wikum, E. (2006). An optimization approach for planning daily drayage operations. *Central European Journal of Operations Research*, vol. 14, no. 2, pp. 141-156.

Laporte, G., Louveaux, F., & Mercure, H. (1992). The vehicle routing problem with stochastic travel times. *Transportation Science*, vol. 26, no. 3, pp. 161-170.

Morlok, E. & Spasovic, L. (1994). Redesigning rail-truck intermodal drayage operations for enhanced service and cost performance. *Journal of the Transportation Research Forum*, vol. 34, no. 1, pp. 16-31.

Psaraftis, H. N. (1995). Dynamic vehicle routing: Status and prospects. *Annals of Operations Research*, vol. 61, no. 1, pp. 143-164.

Smilowitz, K. (2006). Multi-resource routing with flexible tasks: an application in drayage operations. *IIE Transactions*, vol. 38, no. 7, pp. 577-590.

Wang, J. Q., Tong, X. N., & Li, Z. M. (2007). An Improved Evolutionary Algorithm for Dynamic Vehicle Routing Problem with Time Windows. In *Computational Science – ICCS 2007. Springer Berlin / Heidelberg*, ed., pp. 1147-1154.

Wang, X. & Regan, A. C. (2002). Local truckload pickup and delivery with hard time window constraints. *Transportation Research Part B*, vol. 36, no. 2, pp. 97-112.

AUTHORS BIOGRAPHY

Alejandro Escudero is Telecommunication Engineer at University of Seville from 2006. He finished a Industrial Management Master in 2008. Currently, He is part of Organization Engineering Group where he is doing his PhD. His research interests are the freight transport modelling and, the information and communication technologies applied to freight transport.

Dr. Jesús Muñuzuri is professor and researcher at University of Seville. He did his Industrial Engineer in 1997 and finished his PhD Thesis in 2003. His research interests are the city logistic and urban freight transport, the simulations, and the network models.

Dr. Luis Onieva is full professor and vice-rector at University of Seville. He is the principal research of Organization Engineering Group. The principal research lines of Organization Engineering group are production management, transport and traffic engineering, sequentiation and route design, and telecommunication network design.

Miguel Gutierrez is graduated in Economic and Business Sciences at the Universidad of Seville from 1997. He is teaching in School of Engineering from 1998.

ISBN 978-84-692-5416-5

AN EFFICIENT BLOCK-BASED HEURISTIC METHOD FOR STOWAGE PLANNING OF LARGE CONTAINERSHIPS WITH CRANE SPLIT CONSIDERATION

Xiantao Xiao[(a)], Malcolm Yoke Hean Low[(b)], Fan Liu[(c)], Shell Ying Huang[(d)], Wen Jing Hsu[(e)], Zhengping Li[(f)]

[(a,b,c,d,e)]School of Computer Engineering, Nanyang Technological University, Singapore 639798
[(f)]Singapore Institute of Manufacturing Technology, 71 Nanyang Drive, Singapore 638075

[(a)]xtxiao@ntu.edu.sg, [(b)]yhlow@ntu.edu.sg [(c)]liufan@ntu.edu.sg
[(d)]assyhuang@ntu.edu.sg [(e)]hsu@ntu.edu.sg [(f)]zpli@SIMTech.a-star.edu.sg

ABSTRACT
This paper presents an efficient block-based heuristic method for stowage planning of large containerships on a multi-port voyage, subject to a set of constraints. We describe in detail some issues in stowage planning problem, including ship profile, constraints, crane split and rehandle. A "block stowage" heuristic approach is developed to generate a set of stowage plans for a multi-port voyage automatically. Finally, we present the results with a practical test case, and analyze the tradeoff between crane split and rehandles in stowage planning.

Keywords: stowage planning, crane split, block stowage, rehandle

1. INTRODUCTION

Stowage planning is a crucial function for the container transportation business and it greatly affects a shipping line's operating cost. Existing stowage planning process in all major shipping lines worldwide is mainly carried out by human planners. The quality of the stowage plan generated depends very much on the experience of the stowage planners, who have gone through several years training onboard ships. With the capacity of containership rising from the relatively small 350 TEUs (Twenty Foot Equivalent Unit) to ten thousand TEUs, shipping lines are facing increasing challenge to generate workable and cost-saving stowage plans for their containerships as they move between ports.

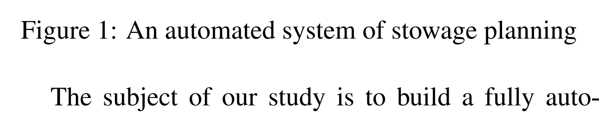


Figure 1: An automated system of stowage planning

The subject of our study is to build a fully automated system of stowage planning for large containerships on their multi-port voyages. As shown in Figure 1, the framework of the system is: given the input data, the stowage plan generator uses a list of heuristic strategies to generate a feasible stowage plan that fulfills a list of constraints. Then the stability module checks the stability of the feasible stowage plan and adjusts it to satisfy the stability requirements. Finally, the optimization engine takes the adjusted feasible stowage plan and optimizes it based on some specific objectives (such as minimizing the number of rehandles). In this paper, we present our current approach for the stowage plan generator. The work with reference to other parts of the system is still in progress and will be reported in the future.

The stowage plan generation process directly affects the number of rehandling and the berthing time of a containership at each port. These two components represent a significant part of the operating cost of the shipping operation. The ship berthing time at a port mainly depends on the crane split. In stowage planning, the requirement of perfect crane split may cause containers destined to a particular port to be stowed using many bays in a ship. This may result in a stowage plan that has unnecessary rehandles due to bays and blocks being used up partially to satisfy crane split requirement. The tradeoffs between rehandle and crane split will be analyzed in the later section of the paper.

The remaining part of the paper is organized as follows. Section 2 reviews the related literature. In section 3, we present some definitions and constraints of stowage planning. Our proposed block stowage algorithm is presented in Section 4. In Section 5, we give a simple test case and show some computational results. Section 6 concludes the paper.

2. LITERATURE REVIEW

Since the 1970s, the problem related to container stowage planning has been studied by shipping lines and researchers. The existing research works are mostly focused on the container loading problem, which can be formulated into a combinatorial optimization problem. The size of the container loading problem depends on the ship capacity and the shipping demand at each port. Even for a medium size containership, the problem is nontrivial due to the large number of variables. Moreover, the problem has been proven to be NP-hard, which is to say that it is very unlikely to guarantee an optimal solution in a reasonable processing time. Meanwhile, a few researchers try to develop heuristic methodology to provide workable solutions to the stowage planning. A brief review of recent research follows.

The early study about the container loading problem can be traced back to the work by Aslidis (1989) and Aslidis (1990). The author examined the stack overstowage problem of small size under certain assumption. Aslidis's work leads to a set of heuristic algorithms which were used to solve the container loading problem without considering stability. Another early work was carried out by Imai and Miki (1989) who considered the minimization of loading-related rehandles.

Avriel and Penn (1993) formulated the stowage planning problem into a 0-1 binary linear programming. They found that the general algorithm is too slow even after some preprocessing of the data. Also, Averiel et al. (2000) showed that the stowage planning problem is NP-complete and showed a relation between the stowage problem and the coloring of circle graphs problem.

Wilson and Roach (1999, 2000) developed a methodology for generating computerised stowage plan. The methodology embodies a two stage process to computerised planning. First they use branch-and-bound algorithms for solving the problem of assigning generalized containers to a bay's block in a vessel. In the second step they use a tabu search algorithm to assign specific locations for specific containers. Wilson et al. (2001) also presented a computer system for generating solutions to the stowage pre-planning problem using a genetic algorithm approach. Dubrovsky et al. (2002) used a genetic algorithm technique for minimizing the number of container movements of the stowage planning. The authors developed a compact and efficient encoding of solutions to reduce the search space significantly.

In the papers of Ambrosino et al. (1998, 2004, 2006), the stowage planning problem is called Master Bay Plan Problem (MBPP). Ambrosino and Sciomachen (1998) reported the first attempt to derive some rules for determining good container stowage plans, where a constraint satisfaction approach is used for defining the space of feasible solutions. Ambrosino et al. (2004) described a 0-1 linear programming model for MBPP. They presented an approach consisting of heuristic preprocessing and prestowing procedures that allow the relaxation of some constraints of the exact model. Ambrosino et al. (2006) presented a three phase algorithm for MBPP, which is based on a partitioning procedure that splits the ship into different portions and assigns containers on the basis of their destination. However they assumed that the ship starts its journey at a port and visits a given number of other ports where only unloading operations are allowed, which means that the loading problem is only considered at the first port.

Averiel et al. (1998) dealt with stowage planning in order to minimize the number of rehandles, without considering stability and several other constraints. They presented a 0-1 binary linear programming formulation and found that the optimal solution is quite limited because of the large number of binary variables and constraints needed for the formulation. Consequently, they developed a heuristic procedure called the suspensory heuristic procedure. However, they assumed that the ship only has a large cargo bay without considering the hatch covers and stability.

Since all these studies reviewed were carried out under some simplistic assumptions, they can hardly be applied by shipping companies in real life, especially for large containerships. In this paper, we describe a heuristic stowage planning algorithm that considers existing containership features and constraints to rapidly generate a set of feasible plans for a multi-port voyage.

3. PROBLEM DEFINITION

3.1. The Containership Structure

From the side view of the containership (see Figure 2), a containership contains a number of bays with number increased from bow to stern. In particular, each 40 foot (40') bay is numbered with an even number, i.e. bay 02, 06, 10, etc., while a 40' bay is associated with two 20' bays with two contiguous odd numbers, i.e. bay 06 = bay 05 + bay 07. Usually a bay is divided by hatches into two sections, below deck and above deck.

From the cross section view of a bay (see Figure 3), every bay contains a set of slots. Each slot is identified by three indices:

- *bay*, that gives the bay it is located in;
- *row*, that gives its position relative to the vertical section of the corresponding bay (counted from the center to outside);
- *tier*, that gives its position relative to the horizontal section of the corresponding bay (counted from the bottom to the top).

Usually, the containers are divided by size into two types: 20' container and 40' container. A 20' slot for the stowage of a 20' container (referred to as a Twenty-Foot Equivalent Unit or TEU) is indexed with the number of the corresponding 20' bay; while a 40' slot (usually is yield by two 20' slots) for the stowage of a 40' container is indexed with the number of the corresponding 40' bay. As for the second index, the location has an even number if it is located on the seaside, i.e. row 02, 04, 06, and an odd number if it is located on the yard side, i.e. row 01, 03, 05, etc. Finally, for the third index, the tiers are numbered from the bottom of the containership to the hatch with even number, i.e. tier 02, 04, 06, etc., while in the above deck from hatch to the top of the container ship, the numbers are 82, 84, 86, etc. Thus, for instance, slot 180406 refers to the slot in bay 18, row 04 and tier 06.

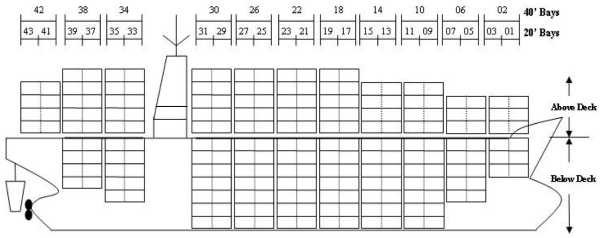


Figure 2: Side view of a containership

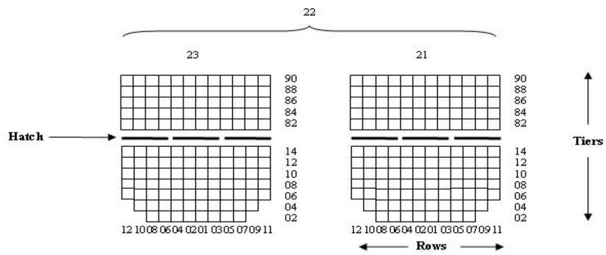


Figure 3: Cross section view of a bay

3.2. The Constraints of Stowage Planning

In stowage planning, a container contains many features, such as port of loading (POL), port of destination (POD), ISO, type, weight, etc. Together with the containership profile, the stowage plan is generated which is subject to many constraints related to the container characteristics and operational instructions. We list the main constraints of stowage planning in the following:

- *Standard container.* The dimension of a 40' container is equivalent to two 20' containers. When a 40' container is stowed into a 40' slot (for instance, slot 180406), the corresponding two 20' slots (slot 170406 and slot 190406) are also occupied and are not available for stowing 20' containers.
- *High cube container.* A 40' high cube container is almost identical to a standard 40' container, except that it is one foot taller. So if one or more 40' high cube containers are stowed into a row below deck, the topmost slot in this row must be left empty. To minimize killing of slots, for below deck, the high cube containers should be stowed in the deep rows (with more tiers).
- *Reefer.* A refrigerated container or reefer is a container used for the transportation of temperature sensitive cargo. Since a reefer relies on external power to maintain the required temperature, it must be stowed into a slot with electrical plug.
- *Hazardous container.* Hazardous container should be subjected to segregation constraints which is provided by the shipping company. Hazardous containers also should be stowed away from the accommodation area and heat source (e.g., engine and fuel tank).
- *Operational constraints.* No container can hang in the air, in other words, a slot below a container cannot be left empty. 20' container cannot be stowed on top of a 40' container. Above deck, no 40' container can be stowed on top of 20' containers.

3.3. Rehandle

Due to the structure of the containership, the containers are stowed in vertical stacks. When a container is unloaded, the containers above it in the same row must be unloaded first. Moreover, if the container is stowed below a hatch, to open the hatch, all containers above this hatch must also be unloaded. In stowage planning, a common situation is that, at port i, the container with POD j (after port i) must be unloaded and reloaded at port i in order to access the container below them with POD i. This is called "overstow" or "forced rehandle". Another situation is that, although a container with POD j does not block any container with POD i, to prevent costlier overstow in future ports or other reasons, the ship planner still decide to unload it and reload it at port i. This is called "voluntary rehandle". Usually, rehandling a container costs tens or hundreds of US dollars. A simple heuristic to reduce the number of forced rehandle is to load the containers in the order of their PODs, i.e. for stowing containers at port i, first load the containers with POD k (port k is the farthest port from port i), then load the containers with POD $k-1$ (the second farthest port), and so on. Finally load the containers with POD $i+1$ (port $i+1$ is the nearest port from port i).

3.4. Crane Split

At a port, the ship will be served by a given number of (usually 3-5) quay cranes to unload and load containers. The bays of the ship will be partitioned into several areas. Each area will be served by one quay crane. This is called crane split. For operating safety, there should be a separation between two adjacent working cranes. The distance of the separation is defined as follows: if a crane is working at bay i, the neighboring crane has to be working at bay $i+8$ or further (see Figure 4). Therefore, if the working areas of two adjacent cranes are too close, one crane has to wait until the other crane finishes its work and moves to the other bay further enough. The waiting time of a crane is called "idle time". A perfect crane split is that all cranes finish their work at the same time with no idle time.

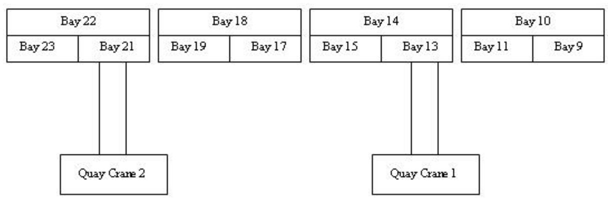


Figure 4: Minimum crane seperation distance

The quality of a crane split is measured by crane intensity (CI). CI is calculated by the following formula: CI equals to the total working time of all cranes divided by the longest crane working time. The duration a ship berthed in a port depends on the completion time of the longest crane and is mainly decided by the crane split.

3.5. The Objective of Stowage Planning

The containership profile containing the locations and the hatches, together with the lists containing all the characteristics of the containers to be loaded at every port in the voyage, are the input data required for generating the stowage plan. The evaluation of a stowage plan can be judged by many considerations, such as stability, handling cost, safety. In this paper, the objective is to generate a set of computerized feasible stowage plans that minimize the number of rehandles and maximize the crane intensity (CI).

4. BLOCK STOWAGE

This section presents a "Block Stowage" approach for the stowage planning problem. The main idea of "Block Stowage" is: first we partition all the locations of a containership into several blocks; then at every port, we divide all the loading containers into different groups by size, type and POD. Instead of stowing the containers one by one, we load the containers group by group into the blocks according to a set of rules. This approach is much more efficient especially for large containerships.

4.1. The Definition of Block

Based on the locations of the hatches in 40' bay, we define two types of block as follows (see Figure 5):

- *Blockbelow* contains the slots below a middle hatch or two symmetric side hatches in a 40' bay.
- *Blockabove* contains the slots above a middle hatch or two symmetric side hatches in a 40' bay.

In order to maintain the balance of the ship in stowage planning, we consider the slots affected by the two side hatches as one block.

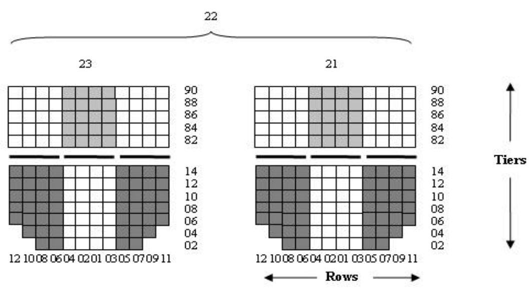


Figure 5: Blockbelow and Blockabove

4.2. The Block Assignment Algorithm

Before producing a stowage plan for a port, we first divide the loading containers into different groups according to their destination port, size and type. If a ship is on a loop voyage of n ports including the current port, there will be up to $n-1$ groups of containers to be loaded and transported to the $n-1$ ports. Group 1 includes the containers going to the farthest port in the loop. Group 2 includes the containers for the second farthest port and group $n-1$ is for the next port from the current port. For example, consider the containers at port C that are ready to be loaded to a ship with a loop voyage A-B-C-D-E-F-G-A. We divide the loading containers by destination port as:

- **Group 1** include the containers with POD B (the farthest port from current port in the voyage),
- **Group 2** include the containers with POD A,
- **Group 3** include the containers with POD G,
- **Group 4** include the containers with POD F,
- **Group 5** include the containers with POD E,
- **Group 6** include the containers with POD D (the farthest port from current port in the voyage).

For convenience of description, we define the "destination port of a block" to be the nearest destination port of the containers stowed in the block. Then we divide the blocks in the containership into different groups according to the destination port of the block in a similar way as we divide the loading containers. An additional group is added for empty blocks (blocks with no containers). When some new containers are stowed into a block, the destination port of the block will be updated. Using the same example of the loop voyage of A-B-C-D-E-F-G-A and port C as the current port, we will have,

Blockbelow [0]: include the Blockbelows which are empty,

Blockbelow [1]: include the Blockbelows whose destination ports are B,

Blockbelow [2]: include the Blockbelows whose destination ports are A,

⋮

Blockbelow [6]: include the Blockbelows whose destination ports are D,

Blockabove [0]: include the Blockaboves which are empty,

Blockabove [1]: include the Blockaboves whose destination ports are B,

Blockabove [2]: include the Blockaboves whose destination ports are A,

⋮

Blockabove [6]: include the Blockaboves whose destination ports are D.

We load the containers in order of group number, i.e. first load Group 1, then Group 2 and so on with Group 6 loaded last. When loading a group, we divide the group into two sets by size: 20' and 40'. According to the stowage constraints, 20' containers cannot be on top of 40' containers. So we load the 20' set prior to the 40' set. When stowing a set of containers, we subdivide the set into different subsets by type, such as normal, empty, reefer, hazardous and out of gauge (OOG), etc.

The sequence of loading different subsets in a 20' set or 40' set is: reefer subset, normal subset, hazardous subset, empty subset and OOG subset. The strategy of stowing a set of containers in the subsets (excluding the reefer subset and OOG subset, because the reefer containers have specific locations in the ship and OOG container cannot be stowed below any container) is expressed as follows (for example, loading a set of containers in Group k):

Step 0. Let $t = k$, go to Step 1.

Step 1. Stow containers into the blocks which are in Blockbelow[t] and above which the blocks are empty. If there are some containers left, go to Step 2.

Step 2. Stow containers into the blocks which are in Blockbelow[0] (empty) and above which the blocks are empty. If there are some containers left, go to Step 3.

Step 3. Stow containers into the blocks which are in Blockabove[t] and below which the blocks are in Blockbelow[t]. If there are some containers left, go to Step 4.

Step 4. Stow containers to the blocks which are in Blockabove[0] and below which the blocks are in Blockbelow[t]. If there are some containers left, go to Step 5.

Step 5. Let $t = t - 1$ (if $t - 1 \geq 1$), go back to Step 1. If there are some containers left and $t = 1$, go to step 6.

Step 6. Stow the remaining containers to those blocks which will cause the least number of rehandles. **End.**

4.3. Considerations for Crane Split

Based on the block stowage assignment algorithm, to generate an ideal crane split, we follow certain rules when stowing containers as follows:

1. *The total move count of two adjacent 40' bays should not exceed the average number of moves per crane.* This constraint is used to balance the loading workload of the crane in the current port taking into consideration that no two cranes can work concurrently in adjacent 40' bays. For example, if the total loading/unloading moves at a port is n with c cranes, the average moves per crane is n/c. Thus the total move count of two adjacent 40' bays should not exceed n/c.

2. *The containers with the same POD in two adjacent 40' bays should not exceed the average discharging moves per crane at POD.* This constraint is used to balance the unloading workload of containers among the cranes for future ports. For example, if the ship has k containers for unloading at port p with c cranes, then the average discharge moves per crane at port p is k/c. Thus the number of containers with $POD = p$ in any two adjacent 40' bays should not exceed k/c at all ports before p.

In stowage planning, the above rules may lead to a situation that many blocks are partially used up to stow the containers with same POD evenly. This will very likely cause unnecessary rehandles. To adjust the balance between cran split and rehandles, we introduce an imbalance tolerance factor k to relax the above rules. For example, in rule 1, "the total move count of two adjacent 40' bays should not exceed n/c" will be relaxed to "the total move count of two adjacent 40' bays should not exceed $(1 + k)n/c$". In Section 5, we will show the effect of the tolerance factor on stowage planning.

5. CASE STUDY

In our testing, we consider a containership with capacity of 5000 TEUs. The voyage of the containership is given as H-A-B-C-D-E-F-G-H. At every port, a number of containers are unloaded and loaded (see Table 1).

Table 1: Workload at each port

Port	A	B	C	D	E	F	G	H
Unload	904	236	57	770	1605	459	1266	254
Load	1146	1212	339	422	1194	651	6	399

To quantify the tradeoff between crane split and rehandle cost, we generated stowage plan using different tolerance level in the imbalance in crane workload. Tables $2 - 5$ show the statistics of the stowage plans generated for the set of ports for different tolerance level. The stowage plans generated are feasible with respect to the stowage planning constraints. The plans generated by our system are labelled as Plan A, whereas the plans produced by human planner are labelled as Plan B. In Tables $2 - 5$, t_{max} denotes the longest crane working time in each stowage plan. Each set of 8 stowage plans at different tolerance levels are generated within one minutes on an Intel Core2 PC with 2.66 Ghz CPU and 2 GB of RAM. As can be seen from Tables $2 - 6$, the crane split becomes worse and the total ship berthing time in the entire voyage (t_{total} in Table 6) becomes longer with the increase of tolerance factor, while the number of rehandles is reduced (actually no rehandle is needed when the tolerance factor k is greater or equal to 0.10 in this test case).

Given the cost of each rehandle and port stay per hour at each port, we can calculate the exact handling cost of a stowage plan. Consequently, we can generate a set of stowage plans using different tolerance level and select the minimum cost stowage plan. This will be indeed valuable for shipping company in practice.

6. CONCLUSION

In this paper, we present a stowage plan generator based on a "block stowage" heuristic approach to generate stowage plans for large containerships automatically.

Table 2: Stowage plans with $k = 0$

Port	A	B	C	D	E	F	G	H
Crane Numbers	5	4	3	4	4	4	4	4
t_{max}(Plan A)	875	804	354	671	1474	714	701	406
t_{max}(Plan B)	895	948	358	894	1480	703	751	732
CI(Plan A)	4.69	3.60	2.24	3.55	3.81	3.55	3.63	3.21
CI(Plan B)	4.58	3.06	2.25	2.70	3.78	3.24	3.39	1.78
Rehandles(Plan A)	0	0	0	0	29	43	0	0
Rehandles(Plan B)	0	1	3	8	0	14	0	0

Table 3: Stowage plans with $k = 0.05$

Port	A	B	C	D	E	F	G	H
Crane Numbers	5	4	3	4	4	4	4	4
t_{max}(Plan A)	875	804	374	671	1479	705	706	406
t_{max}(Plan B)	895	948	358	894	1480	703	751	732
CI(Plan A)	4.69	3.60	2.12	3.55	3.79	3.63	3.60	3.21
CI(Plan B)	4.58	3.06	2.25	2.70	3.78	3.24	3.39	1.78
Rehandles(Plan A)	0	0	0	0	15	74	0	0
Rehandles(Plan B)	0	1	3	8	0	14	0	0

Table 4: Stowage plans with $k = 0.10$

Port	A	B	C	D	E	F	G	H
Crane Numbers	5	4	3	4	4	4	4	4
t_{max}(Plan A)	875	794	389	701	1504	688	711	421
t_{max}(Plan B)	895	948	358	894	1480	703	751	732
CI(Plan A)	4.69	3.65	2.04	3.40	3.72	3.23	3.58	3.10
CI(Plan B)	4.58	3.06	2.25	2.70	3.78	3.24	3.39	1.78
Rehandles(Plan A)	0	0	0	0	0	0	0	0
Rehandles(Plan B)	0	1	3	8	0	14	0	0

Table 5: Stowage plans with $k = 0.15$

Port	A	B	C	D	E	F	G	H
Crane Numbers	5	4	3	4	4	4	4	4
t_{max}(Plan A)	870	794	429	676	1636	686	706	391
t_{max}(Plan B)	895	948	358	894	1480	703	751	732
CI(Plan A)	4.71	3.65	1.85	3.53	3.42	3.24	3.60	3.34
CI(Plan B)	4.58	3.06	2.25	2.70	3.78	3.24	3.39	1.78
Rehandles(Plan A)	0	0	0	0	0	0	0	0
Rehandles(Plan B)	0	1	3	8	0	14	0	0

Table 6: Total berthing time and rehandles

	Plan A				Plan B
k	0	0.05	0.10	0.15	
t_{total}	5999	6020	6083	6188	6761
r_{total}	72	89	0	0	26

Aiming to emulate the plans made by human planners, the stowage plan generator exhibits very good performance in terms of handling cost and computational efficiency. Moreover, we show the tradeoff between crane split and rehandle, which is useful to reduce the cost of the stowage plan.

The stowage plan generator is only the first part of the stowage planning system. Our next step is to develop the stability adjustment module and optimization engine to generate optimized stowage plans.

ACKNOWLEDGEMENTS

The study is supported by grants from the Maritime and Port Authority of Singapore and APL.

REFERENCES

Ambrosino, D. and Sciomachen, A., 1998. A constraints satisfaction approach for master bay plans, In: Sciutto, G., Brebbia, C.A. (Eds.), *Maritime Engineering and Ports*. WIT Press, Boston, pp. 155-164.

Ambrosino, D., Sciomachen, A. and Tanfani, E., 2004. Stowing a containership: the master bay plan problem, *Transportation Research*, 38, 81-99.

Ambrosino, D., Sciomachen, A. and Tanfani, E., 2006. A decomposition heuristics for the container ship stowage problem, *Journal of Heuristics*, 12, 211-233.

Aslidis, T., 1989. Combinatorial algorithms for stacking problems, Ph.D. Thesis, MIT.

Aslidis, T., 1990. Minimizing of overstowage in container ship operations, *Operational Research*, 90, 457-471.

Avriel, M., Penn, M., 1993. Exact and approximate solutions of the container ship stowage problem, *Computers and Industrial Engineering* 25, 271-274.

Avriel, M., Penn, M., Shpirer, N. and Witteboon, S., 1998. Stowage planning for container ships to reduce the number of shifts, *Annals of Operation Research*, 76, 55-71.

Avriel, M., Penn, M., Shpirer, N., 2000. Container ship stowage problem: complexity and connection capabilities, *Discrete Applied Mathematics*, 103, 271-279.

Dubrovsky, O., Levitin, G. and Penn, M., 2002. A genetic algorithm with compact solution encoding for the container ship stowage problem, *Journal of Heuristics*, 8, 585-599.

Imai, A., Miki, T., 1989. A heuristic algorithm with expected utility for an optimal sequence of loading containers into a containerized ship. *Journal of Japan Institute of Navigation*, 80, 117-124.

Wilson, I.D. and Roach, P.A., 1999. Principles of combinatorial optimization applied to container-ship stowage planning, *Journal of Heuristics*,5, 403-418.

Wilson, I.D. and Roach, P.A., 2000. Container stowage planning: a methodology for generating computerised solutions, *Journal of Operational Research Society*, 51, 1248-1255.

Wilson, I.D., Roach, P.A. and Ware, J.A., 2001. Container stowage pre-planning: using search to generate solutions, a case study, *Knowledge-Based Systems*, 14, 3-4, 137-145.

AUTHORS BIOGRAPHY

XIANTAO XIAO is a project officer at the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include stowage planning optimization and mathematical modelling.

MALCOLM LOW is an Assistant Professor at the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include parallel computing, modeling and simulation, planning and scheduling optimization.

FAN LIU is a project officer at the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include automated stowage planning and combinatorial optimization.
SHELL YING HUANG is an Associate Professor at the School of Computer Engineering, Nanyang Technological University, Singapore. Her research interests are in intelligent decision support systems, intelligent agents and agent-based simulations.
WEN JING HSU is an Associate Professor at the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include parallel and distributed processing and provable efficient algorithms.
ZHENGPING LI is an Associate Research Scientist at the Singapore Institute of Manufacturing Technology (SIMTech). His research interests include supplychain management and logistics optimization.

ANCHORAGE CAPACITY ANALYSIS USING SIMULATION

Shell Ying Huang[a], Wen Jing Hsu[a], Yuxiong He [a], Tiancheng Song[b], Charles De Souza [b], Rong Ye[c], Chuanyu Chen[c], Stuti Nautiyal [c]

[a] School of Computer Engineering, Nanyang Technological University, Singapore.
[b] Maritime and Port Authority, Singapore.
[c] Simplus Pte Ltd.

[a] {assyhuang,hsu}@ntu.edu.sg, [b] Tiancheng.SONG@mpa.gov.sg, [c] {yerong,chuanyu,stuti}@simplus.sg

ABSTRACT
With the substantial growth in marine traffic, anchorage space is now in high demand in certain hub ports. To provide decision support for port authorities, we have analyzed the usage of anchorages in recent years. The demands on the anchorages arise from the dynamically changing vessel mix and similarly complex service patterns. The utilization and the capacity of an anchorage space also depend heavily on the dispatching and allocation rules as well as the shape and areas of the anchorage. The complexity of the system studied is therefore beyond the current analytical tools, and hence simulation provides an effective means for the study. In this paper, we present a simulation-based capacity analysis on anchorages. The simulation model built is able to match the current scenarios well, and the analysis by simulation proves very useful in assessing anchorage utilization and capacity for future scenarios.

Keywords: Decision support system, Anchorage capacity, Circle packing, Model development

1. INTRODUCTION

The Straits of Singapore play host to an average of about 140,000 vessels annually. These include containerships, general cargo ships, oil tankers, chemical tankers, ferries, cruise ships and many more. Vessels of different lengths, tonnage and carrying different types of cargoes enter Singapore Straits and navigable sea space has to be allocated as anchorage space for vessels that require mooring space for various reasons. In recent years both vessel traffic volume and vessel size have shown an upward trend and are expected to increase significantly over the next few decades. It means that the demand on anchorage space will escalate as well. Similar to Singapore, in the Port of New York and New Jersey, the number of vessels has increased and their sizes have grown in the past few years (U.S. Environmental Protection Agency 2006). The anchorage space has frequently been filled to capacity and the Coast Guard proposes to revise the duration vessels are authorized to anchor in specific anchorage areas.

There are few published results on the capacity studies of anchorage space. In Bugaric and Petrovic (2007), the anchorage as part of a river terminal for bulk cargo unloading is simulated where the anchorage is an n-vessel holding area. The anchorage is abstracted as a First-In-First-Out queue with a capacity of n. When the queue reaches its capacity, incoming vessels are turned away. A vessel will leave the queue when a berth is free at the terminal. Fan and Cao (2000) presented a capacity model for an anchorage. The capacity of an anchorage is defined as the maximum number of vessels that can be accommodated by the anchorage over a period of time. The anchorage capacity is a function of the total area of the anchorage, the average percentages of each type of vessels coming to anchor, the average areas occupied by a vessel of each type and the average durations of vessel stays of each type of vessels in the anchorage.

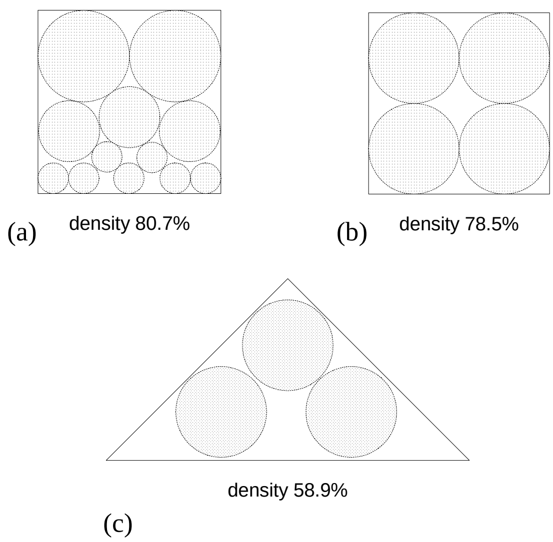


Figure 1: Different vessel mixes or different anchorage shapes can result in very different utilization results.

In actual operations, vessels of different types and sizes come and leave anchorages at all times and therefore the vessel mix in anchorages are changing dynamically. It follows that anchorage utilizations change all the time. Different vessel mixes with regard to sizes are able to effectively utilize the same anchorage space to different degrees. Figures 1(a) and 1(b) show two vessel mixes producing different utilization results in the same anchorage. A vessel mix may be accommodated into an anchorage with no problem but the same vessel mix may not be able to fit into another anchorage of the same area size but of a different shape. Figures 1(b) and 1(c) show two anchorages with the same area but different shapes result in different utilization figures. This means statistics of averages of various measurements will not present a clear picture about peaks and lulls of the demand and anchorage utilizations and therefore cannot be used as evidence of enough space meeting the demand.

A simulation-based tool will be the most effective way to assess quantitatively anchorage utilization levels and to evaluate anchorage capacity. Such a tool will be very useful for assessing whether the existing anchorage space is adequate for future scenarios. The tool will also be useful in evaluating whether changes in the configuration of the anchorages and/or changes in certain policies or practice are effective to satisfy the demand on anchorage space without allocating more physical space.

We define that an anchorage has reached its capacity if the probability of having n or more vessels not able to find anchorage space is greater than a limit p. Both n and p will be specified by the managing authority of the anchorage.

A computer simulation based planning tool was developed. Using this tool, we are able to provide quantitative information about instantaneous, average and maximum anchorage utilizations. We also analyze the probabilities of vessel overflows, that is, the probabilities of having vessels turned away because there is insufficient space in the anchorages. The management of the anchorage space may decide that the demand on anchorage space has exceeded its capacity when the probability of having at least a certain number of vessels turned away is more than a threshold value. The tool also allows us to conduct experiments to evaluate the effectiveness of different anchorage configurations when space allocated for certain types of vessels is found not enough. The tool allows the user to specify various vessel mixes, vessel arrival patterns, and current or future planned anchorage configurations. The simulation of the typical practice was validated by setting simulation parameters with suitable values and comparing simulation results with the corresponding statistics based on a set of historical data in Singapore.

The rest of the paper is organized as follows. Section 2 introduces a typical anchorage system and practice in port cities. It is followed by the description of the architecture of the simulation tool for anchorage capacity study in Section 3. Then Sections 4, 5 and 6 present the vessel arrival generator, the vessel dispatcher and the anchorage manager respectively. Section 7 describes the evaluation of the simulation tool. Section 8 presents a method to assess what the space utilization is like when an overflow occurs at an anchorage. Section 10 concludes our work.

2. A TYPICAL ANCHORAGE SYSTEM

Generally, vessels come to a port for various purposes, such as taking bunkers, going to shipyards for repair, loading and unloading of cargo at a terminal and/or a combination of these purposes. Some vessels go to an anchorage before visiting their terminal/shipyard and some do so after visiting the terminal/shipyard. Other vessels will visit an anchorage without going to any terminal/shipyard. Some vessels make multiple visits to anchorages for multiple purposes with or without visits to terminals or shipyards. Bunker vessels visit other anchorages to provide bunker services to other vessels and visit terminals to refill bunker supplies in between their stay in the home anchorages.

In the whole port there may be a few areas that are designated for anchoring vessels and these areas may be further divided into a number of anchorages of different shapes and sizes. Each anchorage also has a maximum depth that limits the vessels that can anchor in it. Some of the anchorages are reserved for vessels in certain gross tonnage (GT) groups. These anchorages are categorized to serve different types of vessels with different purposes of visit. Typically, shipping agents choose anchorages for their vessels before they arrive in the port. They take into consideration a few factors like the vessel's purpose of visit, the vessel's type, draft, gross tonnage, the location of the vessels' entry/exit points into/from the port and the locations of the terminals/shipyards they visit.

After a vessel arrives in the port, the pilot or vessel captain chooses an anchoring position in the anchorage when the vessel needs to go into the anchorage. If the anchorage of the vessel's choice is full, the pilot or the captain will choose another anchorage which is designated to provide space for the same type of vessels. There is no restriction on the duration a vessel is allowed to stay in an anchorage.

A careful examination of the historical vessel data of a port is carried out to extract patterns and statistical distributions of vessel calls, vessel visits to anchorages, arrival times and dwell times in anchorages.

3. SYSTEM ARCHITECTURE

Figure 2 shows the architecture of the simulation system. The system consists of a vessel arrival generator for generating vessel arrivals to anchorages, a vessel dispatcher for simulating incoming vessels' choice of anchorages, and for each anchorage there is an anchorage manager for emulating pilots' decisions to choose anchoring positions for incoming vessels in each individual anchorage. At the front end, there is a graphical user interface (GUI) for the user to specify the simulation parameters and the anchorage specification. There is also an output GUI for displaying the utilization status of individual anchorages during simulation runs and the output statistics after simulation runs.

Through the GUI, the tool allows the user to input the total number of vessel calls for the period of the simulation and the distribution of calls among the different vessel types. The user is also able to specify different anchorage configurations. The dimensions, depth, and usage for each anchorage and the distance between any two anchorages can also be defined. Existing anchorages can be removed and new anchorages can be added.

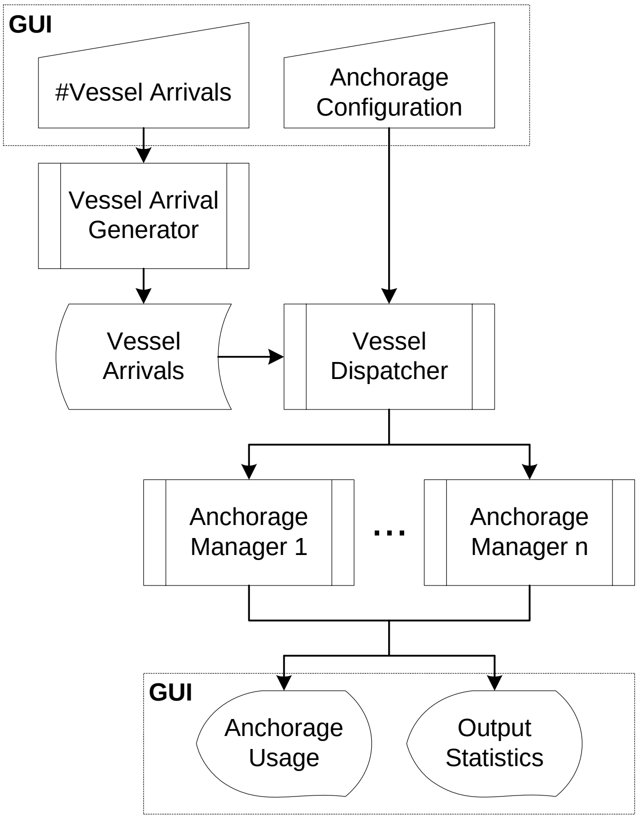


Figure 2: System Architecture

Output Statistics include

- Instantaneous utilization, indicating the occupancy of an anchorage at time instance *t*:

$$u_t = \frac{\sum_i Space_i}{Area} \quad (1)$$

where i is the index of vessels that anchor at the anchorage at that point of time; $Space_i$ denotes the space taken by vessel i and *Area* denotes the area of the anchorage.

- Average utilization, indicating the occupancy of an anchorage over the time period of interest:

$$u_a = \frac{\sum_i Space_i \times Dwell_i}{Area \times Duration} \quad (2)$$

where i is the index of vessels that anchor at the anchorage during the period; $Space_i$ denotes the space taken by vessel i and $Dwell_i$ denotes its corresponding dwell time at the anchorage; *Area* denotes the area of the anchorage and *Duration* denotes the simulation time period of interest, e.g. a day or a month.

- Number of overflows, indicating the occurrences when no suitable anchorage space can be allocated to a vessel. It is an indicator of the extent the anchorage space is packed and in order to accommodate such "extra" vessels, certain anchoring rules have to be violated in actual operations.

4. VESSEL ARRIVAL GENERATOR

The demand on anchorages of a port is driven by the arrivals of vessels calling at the port. However, not all vessels calling the port need anchorage space. Meanwhile, some vessels call at anchorages more than once during a visit. It is of practical importance to relate the number of vessel calls at a port to the number of vessel visits to its anchorages. The correlation between the two can be analyzed from historical data. Therefore the first component in modeling vessel arrivals to anchorages consists of a translation mechanism that maps the number of vessel calls to the demand on anchorages, that is, the number of arrivals at anchorages. The second component is a Non-stationary Poisson Process that assigns an arrival time to each generated vessel.

In the analysis of historical data of a port as an example, the relationship between the number of vessel calls and the number of visits to anchorages is established by regression analyses and we get, for each type of vessels:

$$y = a * x^b \quad (3)$$

where y is the number of vessel calls at anchorages and x is the number of vessel calls at the port of a particular type of vessels. With the knowledge of this correlation, it is possible to predict the demand on anchorage space based on the predicted number of vessel calls to the port.

Vessel attributes like the vessel length, gross tonnage and draft are expected to affect the anchorage utilization so they also need to be generated for each generated vessel according to historical distribution of each vessel type. The distributions of vessel lengths an be obtained from analysis of historical data. It is also found through data analysis that for each type of vessels, vessel gross tonnage and vessel draft are both related to the length of the vessel by Equation (1) with different parameter values for a and b. With the generated vessel length, the vessel draft and gross tonnage can be generated for each vessel, based on the correlation found between them in the historical data.

Table 1 shows the percentage differences between the generated numbers of arrivals to anchorages for each type of vessels and the numbers from the port's historical data in 2006. This indicates the success of the translation scheme that maps vessel calls to the demand on anchorages.

Table 1: Comparing generated demand on anchorage with historical data

Vessel type	% difference	Vessel type	% difference
TA	1.3	FR	2.0
VLCC	-1.4	CO	6.4
CH/LPG/LNG	-0.2	BA	-0.7
CTNR	0.7	BK	0.6
BC	-0.3		

Different distributions may also be specified by the user through the input GUI if it is necessary.

Arrival of vessels to a port is a complicated dynamic process that is dependent on shipping lines' planning and is very sensitive to economic activity fluctuations. Considering the whole port as a complete service facility, the arrival process would be best modeled as a Poisson process. This is based on the facts that: (1) even though the arrivals of individual vessels are scheduled, when the vessel traffic to the whole port is considered, the distribution of the inter-arrival times becomes random and fits well the exponential distribution; (2) the unpredictable weather conditions and possible delays in service by other ports of call further randomize the arrivals. Statistical tests on historical data confirmed that the inter-arrivals for each vessel type agreed reasonably well with the exponential distribution so the Poisson distribution for the arrival process is suitable in our study.

5. VESSEL DISPATCHER

As described in Section 2, shipping agents choose anchorages for their vessels. Therefore the role of the vessel dispatcher is to simulate the shipping agent's choice of anchorage for its incoming vessel. Modeling vessels' choice of anchorages is complicated, as a number of parameters have to be taken into consideration, for example, purpose of vessel call, vessel type, vessel gross tonnage and vessel draft. Moreover, each type of vessel may choose any of a few candidate anchorages. There are no fixed regulations and rules for the selection among these candidates so it may be based on agents' preferences. Agents' preferences are unknown to us, may be numerous and may not be general. It is unrealistic to summarize a comprehensive set of rules that emulate the preference of clients when they select among the candidate anchorages. There are also some unforeseen circumstances in actual operations that would affect the choice of anchorage.

Therefore, instead of applying rules to emulate their (agent, pilot, control centre) anchorage choices, we propose to mine the historical data of a port to figure out how its main anchorages were used as the end results of their choices. From the mining results, the probability (weightage) distribution for each type of vessels in choosing various anchorages can be obtained and the process of choosing an anchorage based on these weightages for a vessel is shown in Figure 3.

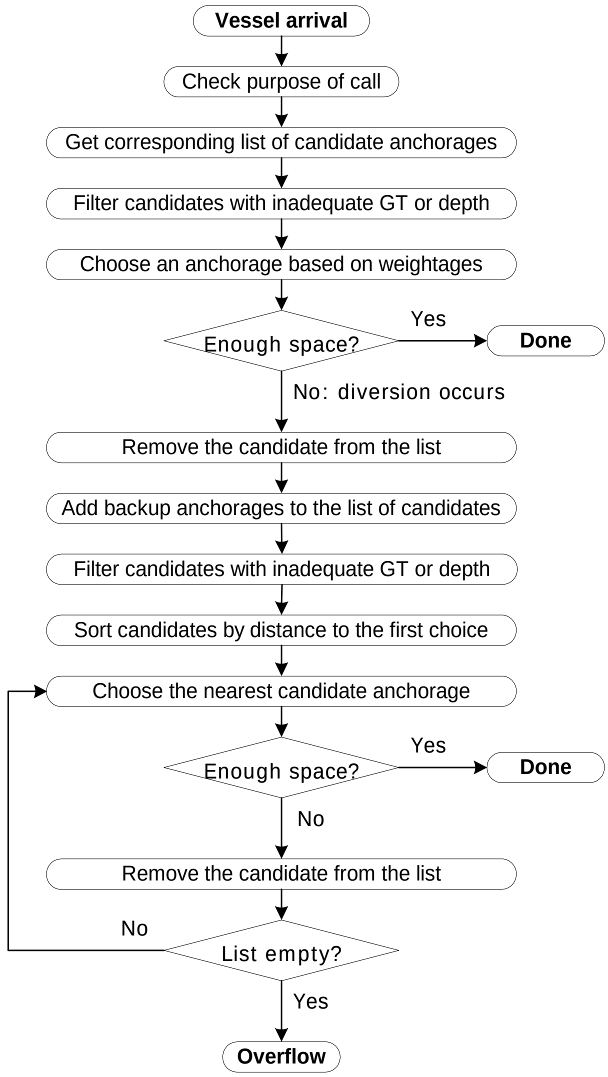


Figure 3: Process of choosing an anchorage for a vessel

6. VESSEL DISPATCHER

In most ports, pilots and captains are free to choose the anchoring positions for their vessels in an anchorage. Therefore the role of the anchorage manager in the simulation tool is not to assign anchoring positions for vessels but to simulate pilots' decisions in choosing an anchoring position. We present an algorithm here to simulate how pilots choose an anchoring position in an anchorage.

When a vessel anchors in an anchorage, the space it occupies is more than the width and length of the vessel. Due to wind and current, a vessel may be at different positions at different times within the space of a circle. A minimum safety clearance also needs to be maintained between any two anchoring points at all times. So each vessel will occupy a circular space with a radius of the length of the vessel plus half the minimum safety clearance. In this way, the distance between the two anchoring points is at least the sum of the two vessels' length plus the minimum safety clearance.

In a hub port where anchorage space is highly contested, pilots and captains usually anchor their vessels close to at least one of the anchored vessels, so that the anchoring space is better utilized than if a completely random position is chosen. Four typical anchoring scenarios are discussed for elaborating how anchoring positions are decided in the system based on the usual practice of the pilots and captains. Figure 4 shows these scenarios where a solid-boundary circle represents a vessel that has already anchored, and a dash-boundary circle represents the candidate choices for anchoring an incoming vessel. The four scenarios are

- Type 1: *vessel-side corner* formed by a border line of the anchorage and an existing vessel.
- Type 2: *single-vessel cut* is a position next to one existing vessel only.
- Type 3: *two-vessel corner* formed by two existing vessels.
- Type 4: *two-side corner* formed by two border lines of the anchorage.

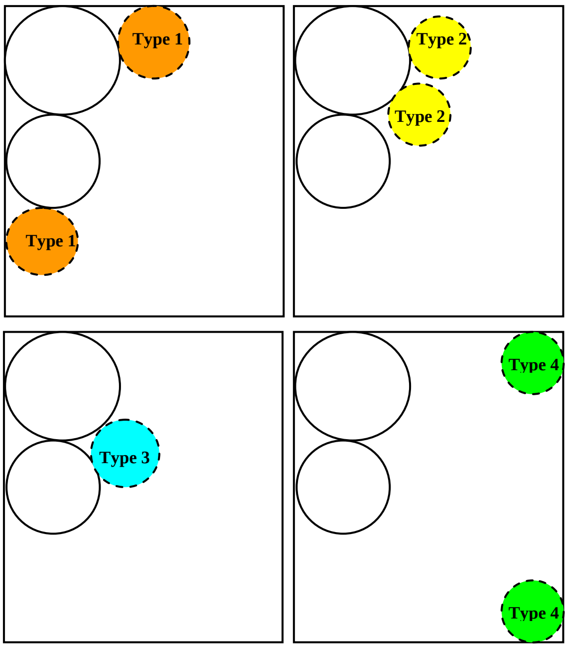


Figure 4: Four typical anchoring scenarios

In a survey conducted among 10 pilots (experienced and not so experienced), it was found that about 50% of them preferred Type 1, 20% preferred Type 2, another 20% preferred Type 3, and the remaining 10% preferred Type 4. Therefore the distribution was applied to the simulation tool when modeling the pilots' choice for anchoring positions for these typical scenarios. This is shown in Figure 5.

7. EVALUATION

To evaluate whether the system works correctly and accurately, we use the historical data of a port as an example. We set the simulation parameters to suitable values and configure the anchorages accordingly. The yearly average utilizations of individual anchorages were used as the indicators for comparing the outputs of the simulation tool with historical statistics. Figure 6 summarizes the results.

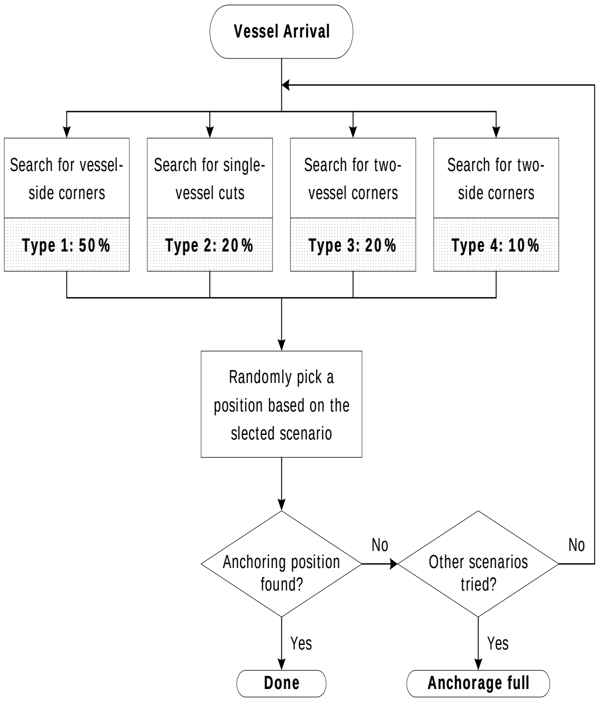


Figure 5: Anchoring algorithm

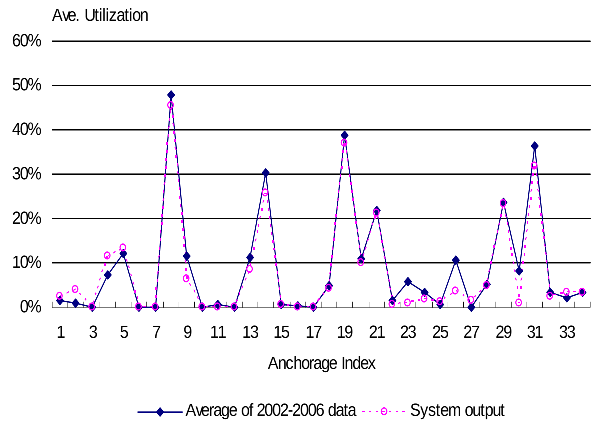


Figure 6: Comparing System Output and Historical Statistics

Most of the utilization figures from system output fit well with the historical averages. Note that in Section 4, we have confirmed that the generated numbers of arrivals to anchorages for each type of vessels match the numbers from the historical data. It was therefore accepted that if the parameters of the simulation tool is set correctly, it is able to emulate the operations of an anchorage system well.

8. ASSESSING SPACE UTILIZATION NEAR CAPACITY

The space taken by a vessel in an anchorage is represented as a circle with the vessel length plus a safety margin as the radius and to ensure safety, any two circles are not allowed to overlap. This means that the maximum utilization of an anchorage is lower than 100%. For example, in a scenario illustrated by Figure 7, the anchorage cannot accommodate any vessels longer than 120m, although its utilization at that point is 61.55%. However it is able to accommodate smaller vessels by filling them in the gaps in between existing vessels. Therefore, the achievable maximum utilization of an anchorage is decided by a few parameters, namely anchorage size, shape and vessel mix (of size), while vessel mix is the combined effect of anchorage dispatching rules and vessel arrival patterns. As these parameters vary in different anchorages, there is no single maximum utilization figure applicable to all anchorages. Assessing space utilizations near capacity for anchorages of different shapes and sizes, and receiving different vessel mix, is of great importance in evaluating anchoring space usage under future traffic scenarios.

Table 2: Space utilization when overflow occurs

Utilization	Percentage	Utilization	Percentage
5%	0.00%	55%	20.24%
10%	0.00%	60%	48.57%
15%	0.00%	65%	24.26%
20%	0.00%	70%	0.80%
25%	0.00%	75%	0.00%
30%	0.00%	80%	0.00%
35%	0.02%	85%	0.00%
40%	0.14%	90%	0.00%
45%	0.99%	95%	0.00%
50%	4.97%	100%	0.00%

Assessing space utilization near capacity mainly involves two tasks, namely, generating high but realistic vessel traffic to individual anchorages so as to create sufficient and representative overflow instances, and recording the instantaneous utilization at the points of overflow. One simple way to generate such traffic is by disabling all but one anchorage so that all associated vessels will be directed to that particular anchorage. As an example, Table 2 summarizes the overflow instances at one of the anchorages. Columns 1 and 3 in the table show the space utilization figures when an overflow occurs at the anchorage. Columns 2 and 4 show the percentages of overflows that occur with the corresponding utilization figures.

As shown in Table 2, for that particular anchorage and the vessel traffic, overflow could occur at the anchorage when utilization is as low as 35%, although most of the time it occurs only when utilization is at a higher level of around 60%. This confirms that even at the same anchorage, the maximum possible utilization varies with the size of incoming vessel and the vessel mix at that time. On average, overflow can happen when utilization reaches 57.10%. It should be noted that for a different anchorage and different vessel mix, overflows may occur at a different utilization level. From our experiments for 17 anchorages, we found that this utilization figure varies from 35.1% to 61.1%

9. CONCLUDING REMARKS

We have developed a reconfigurable tool for assessing the capacity of anchorages. It allows the user to input various vessel mixes and volumes in vessel arrivals for current or future scenarios. It also allows the user to specify current or planned anchorage configurations, e.g. to add, remove or change the sizes and shapes of the anchorages. The user can also change the types of vessels that use an anchorage.

Our system proves to be a useful decision support tool for assessing the impacts of different vessel demands, anchorage configurations, anchoring practices and policies. We used this tool to assess the anchorage utilization figures when vessel overflow occurs and we can also use the tool to compute the probability that a certain number of vessels cannot be accommodated in an anchorage. For future work, we plan to design and evaluate algorithms for improving anchorage utilizations.

ACKNOWLEDGMENTS

We thank the Maritime and Port Authority of Singapore for supporting the project.

REFERENCES

Bugaric, U. and Petrovic, D., 2007. Increasing the capacity of terminal for bulk cargo unloading. *Simulation Modelling Practice and Theory*, 15 (10) , 1366-1381.

Fan, H. S. L. and Cao, J., 2000. Sea space capacity and operation strategy analysis system. *Transportation Planning and Technology*, 24 (1), 49 – 63.

U.S. Environmental Protection Agency, 2006. Anchorage Regulations: Port of New York. DOI= http://www.epa.gov/fedrgstr/EPA-IMPACT/2006/November/Day-16/i19314.htm

AUTHORS BIOGRAPHY

Shell Ying Huang and **Wen Jing Hsu** are currently Associate Professors in SCE, NTU. **Yuxiong He** was a PhD candidate in SCE, NTU. They have done many projects in maritime R & D funded by Port Authority, terminal operators and a shipping company. **Tiancheng Song** is Assistant Director in Technology Division of MPA. **Charles De Souza** is controller in Pilotage Exam and Marine Projects of MPA, Singapore. **Rong Ye, Chuanyu Chen** and **Stuti Nautiyal** are postgraduates and graduate from NTU and did projects in maritime R & D in NTU. They are now the co-founders and directors of SimPlus Pte Ltd.

COMBINING LOGISTIC CONTAINER TERMINAL SIMULATION AND DEVICE EMULATION USING AN OPEN-SOURCE JAVA FRAMEWORK

Dipl.-Inf. P. Joschko[(a)], Dipl.-Wirt.-Inf. C. Brandt[(b)], Prof. Dr.-Ing. B. Page[(c)]

[(a), (c)] University of Hamburg, Department of Informatics
[(b)]Hamburger Hafen und Logistik AG

[(a)]joschko@informatik.uni-hamburg.de, [(b)]brandt-c@hhla.de, [(c)]page@informatik.uni-hamburg.de

ABSTRACT
In this paper we intend to show that the open-source simulation framework DESMO-J can be used not only for logistic investigations - as it was being used by many during the last years - but also for evaluating and testing control systems. Therefore we propose to implement model components of container terminals in such a way that they can be used in logistic simulation as well as in device emulation test beds. This paper outlines our experience in extending DESMO-J for a broad range of applications in the context of container terminals and our experience in developing reusable model components for simulation and emulation. While Java-based DESMO-J is a product of the University of Hamburg, the functional container terminal extension, called COCoS, we present in this paper has been developed by the Hamburger Hafen and Logistik AG.

Keywords: discrete simulation framework, logistic simulation, device emulation, container terminal

1. INTRODUCTION

1.1. Hamburg Container Terminals

With a container handling rate of 9.7 million TEU in 2008, Hamburg accommodates one of the ten largest container ports worldwide and accordingly the second largest within Europe, right behind Rotterdam. This capacity is to be expanded to 18 million TEU in the next couple of years.

Hamburger Hafen und Logistik AG (HHLA) is one of the leading port logistics groups in the European North Range. With its Container, Intermodal and Logistics segments, HHLA is positioned vertically along the transport chain. Efficient container terminals, high-capacity transport systems and a full range of logistics services form a complete network between the overseas port and its European hinterland.

HHLA operates three container terminals which were responsible for more than two thirds of the overall container handling in Hamburg in 2008. One of them is the HHLA Container Terminal Altenwerder (CTA) which is considered to be a state of the art terminal worldwide. The northern section of the CTA is shown in Figure 1. Since the container dispatching processes at the CTA are almost fully automated, a particularly significant relevance is attached to the integration of container handling technology and IT-based control systems.

Increasing handling rates require highly efficient dispatching strategies for container vessels and hinterland traffic. Not only that every deceleration would bring financial damage to terminal carriers and forwarding companies but additionally, an interruption in container handling process would immediately cause congestions in metropolitan area traffic because of nearby autobahns and city center.

These requirements stress the importance of using efficient methods for improving and validating the evolution of terminal layout, logistic handling strategies or control systems. The University of Hamburg and HHLA have been collaborating on developing simulation techniques to solve such terminal-specific problems for many years.

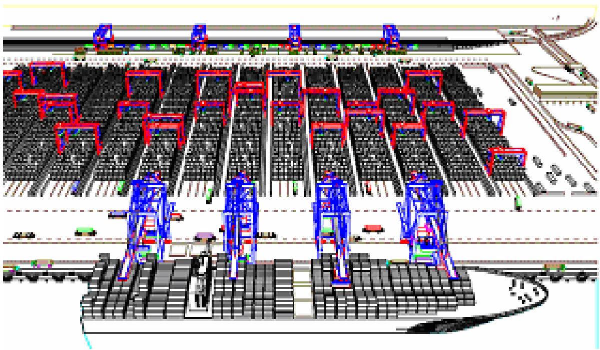


Figure 1: Northern section of Container Terminal Altenwerder (Brandt 2007)

1.2. Container Terminal Simulation

According to (Steenken et al. 2004) a multi modal container terminal can be considered as an open system. The waterside (ships) and the landside (trains, trucks) have always been the two main interfaces to its environment. But distinct container terminals may significantly vary in types and quantity of handling equipment, terminal layout or grade of automation of handling processes. Besides these physical aspects there are also many individual types of supporting and surrounding IT-Systems as well as planning and control strategies integrated therein. The fact that these and further aspects need to comply with various specific requirements of terminal operations finally leads to a considerable number of planning and optimization problems that emerge and evolve throughout a container terminal's life cycle. According to (Saanen 2001) this life cycle may be divided into three main phases:

1. Conceptual, functional and technical design
2. Implementation and Realization
3. Commissioning and Operation

In practice, the phases may overlap significantly and may be interconnected by iterative feedback loops, for instance if an existing terminal receives a redesign due to new requirements.

Simulation is to be used if experimentation with the real system is too expensive, complex or dangerous, which applies to container terminals in particular. Therefore simulation has become an important method in analysis and optimization of container terminal operations in the recent past (Steenken et al. 2004). Just like the planning problems mentioned above, the application of simulation may be attributed to the phases of a container terminal's life cycle and may be differentiated accordingly as illustrated in Figure 2.

Strategic simulation mainly points at problems of terminal design and is applied in phase (1) correspondingly. Different logistic concepts, decision rules or optimization algorithms that address problems of operational terminal planning within phase (2) and (3) deserve close comparison by means of operative simulation studies before they can be put into production on an existing terminal. Tactical simulation stands for an integration of simulation models within the terminal's IT-systems in order to generate solutions for operational and on line planning problems in phase (2) and (3). This integration is also referred to as simulation optimization or optimization via simulation. While strategic, operational and tactical simulation are directly focused on, mainly logistical, planning problems within a terminal's life cycle, phase (2) particularly requires simulation techniques for the purpose of development and validation of terminal operation systems by emulating terminal equipment, IT-systems or human behavior.

1.3. Outline

Section 2 details the comparison between simulation and emulation in order to extract differences and similarities. Section 3 concentrates on the simulation framework DESMO-J and explains its main functionality, its software architecture, how it can be extended and how it has been used in harbor simulation so far. Section 4 concentrates on COCoS, the extension HHLA uses in order to develop simulation models based on DESMO-J. Section 5 describes how DESMO-J and COCoS fit the requirements presumed in section 2 by referring to concrete model implementations. Section 6 finally discusses the experiences made in combining logistic simulation and device emulation by presenting benefits and limitations.

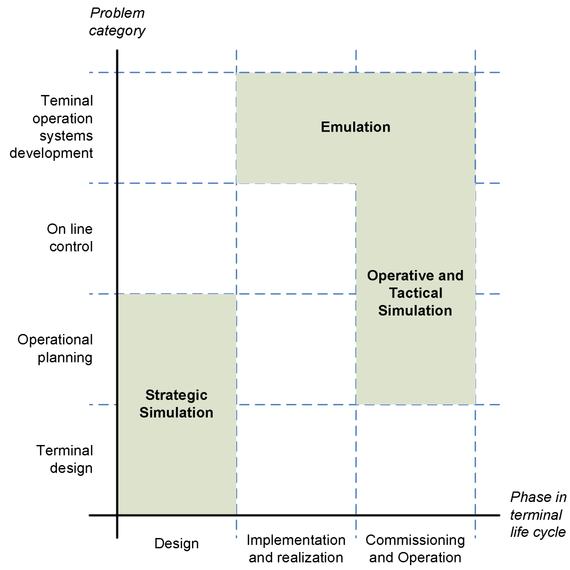


Figure 2: Correlation of simulation and emulation with container terminal planning problems and life cycle

2. SIMULATION VS. EMULATION

In principle there are two different application scenarios for using simulation in context of container terminals. The traditional one is executing logistic simulation experiments to compare different handling strategies or to investigate the terminal layout. A completely different approach is to use simulation technique for testing terminal operating systems (TOS). We speak of emulation in this context. We depict differences of simulation and emulation in this section.

2.1. Definition

The main difference between logistic simulation and device emulation is that not the modelled system – the terminal – is subject of the investigation, but a control system, in this case a real TOS which is linked to the model. Emulation is used to imitate a real system; it offers an interface so that the control system cannot distinguish between a model and reality. This intends to already examine operability of a TOS before the employment at a real terminal. We denote emulation to be a subset of simulation.

(Auinger et al. 1999) proposed using simulated systems in combination with real systems in different ways. Figure 3 illustrates the different possibilities to combine simulation models with reality. Arrow 1 describes testing all components in live respectively prototype mode, not using simulation at all. This is the most realistic way of testing, but also the most expensive, furthermore dysfunctions can cause maximum damage. Arrow 2 is near to classical simulation view, without interaction with reality. This can be used for logistical simulation described above. Arrow 3 shows a kind of simulation called Reality-in-the-loop or Real-Time-Control. Here the material terminal hardware is tested with a model of a control system. (Schütt and Hartmann 2000) described how HHLA tested AGVs (Automated Guided Vehicles) with a virtual control system during CTA's planning phase. Arrow 4 describes the way of coupling simulation with reality which we call emulation and this paper is about: A simulated terminal system is piloted by a real TOS, so that the TOS can be tested with help of a terminal model.

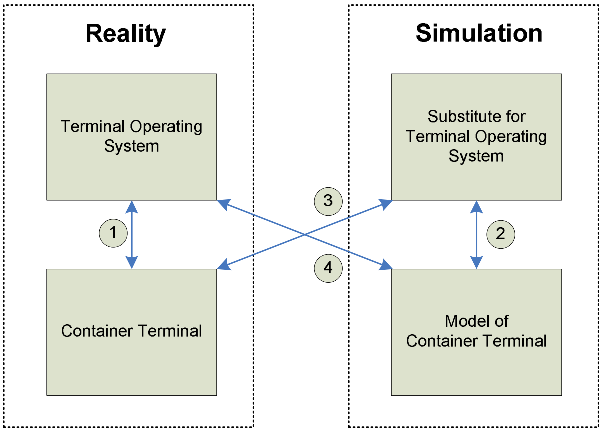


Figure 3: Possible combinations of reality and simulation (based upon Auinger et al. 1999)

2.2. General Requirements

Logistic simulation is often embossed by bird's eye view; we look at the behavior of a terminal in general, so we can abstract from many details and implement the model with a top-down approach. In contrast, device emulation always depends on worm's eye view. We have to implement models with a bottom-up approach that is to say that all activities on the terminal have to be mapped from the view of emulated components, for instance quay cranes or AGVs.

Beyond this, the needed level of detail for emulation requires more time and effort than conventional simulation models. For example, in logistic simulation models, only the time span is relevant, which a quay crane needs for picking up a container on landside and setting it down onto the ship, but when testing a TOS with help of emulation models, a lot of intermediate steps have to be modeled auxiliary. For example, a dual-trolley quay crane owns two trolleys, which get individual briefings by TOS and have to interact for dispatching the container (see section 5.2 for details). The TOS needs information about the exact position of these trolleys, for this reason modeling the kinematic behavior of mobile devices becomes necessary. Therefore, the technical characteristics of the devices have to be investigated and mapped exactly into the model, which is unnecessary in simulation models at all.

Fundamentally in emulation, interfaces for communicating with the Control System (in case of harbor simulation: the TOS) are needed. In general, these are the same interfaces, which are implemented by local device drivers of real terminal components. In most cases, we deal with TCP/IP-based socket communication, and there is a set of telegram types, which have to be implemented. These telegrams may brief the device for doing a selected task or may request the status of the device. In other direction, telegrams which have to be sent from the device to the TOS may contain the status of the device, signal finished jobs or announce error states.

This leads us to another main difference between simulation and emulation. In simulation experiments, time spans that do not include any events will be skipped; duration of a simulation run depends on CPU-speed. By contrast in emulation, the advancement of simulation time has to be adjusted to real-time. Since the TOS is a real-time system, which assumes that the processing of jobs consumes time, and since we do not want to change this behavior for testing the TOS, we need to slow down simulation. The challenge of doing this is to find an approach, which delays the advancement of execution time, and nevertheless allows the insertion of new processes which were commissioned by the TOS over the communication interface; i.e., if the next scheduled process is estimated to start in a selected time span, the emulation model may not advise the thread to sleep until just before the end of this time span, because it is possible that a new upcoming task will have to be started prior to the process the scheduler is currently waiting for.

2.3. Manual- or automatic-driven experiments

The way of executing emulation experiments, differs fundamentally from logistic simulation experiments. Because of real-time conditions, the number of executed experiments is substantially smaller, and the analysis of results is extended by methods that are inappropriate in logistic simulation. In the following, we distinguish two types of experiments: manually driven and automatically driven tests.

During the course of the manual experiments, a human tester takes the role of the dock workers and operators of package items. He achieves certain tasks, and examines whether the TOS is reacting correctly to his interactions. Therefore, the model has to provide another interface, which is used for user interaction. All the actions, which dock workers and operators may accomplish on the terminal and thus are not induced by the TOS, like lashing a container or entering a danger area, have to be mapped to this interface. Additionally, all information about component states, which could be in interest for the tester, has to be accessible via this interface. We call it the manual interaction interface.

The tester should get access to this interface in scope of a Graphical User Interface (GUI). This can be realized via buttons, dropdown boxes and input fields, which should give feedback to user action. In this context, the visualization of temporary state of model becomes substantial. This can either be table-based summing-ups of model components. Or, which is quite more ergonomic, via 2d- or 3d-animations enriched by additional information boxes. Since the status of emulated devices has to appear just in time, a concurrent animation is urgently needed, in contrast to logistical simulation experiments, where a post-process animation is adequate.

Beside these manual tests, also automatic tests are to be possible. In case of testing a new version of a TOS, each time a lot of standardized workflows have to be validated. This needs not to be done by human testers. Therefore, another component has to be added, which automatically acts on the manual interaction interface. The behavior of this component has to be described by an easy changeable set of rules. We do not need any animation here, but we have to insert logging mechanisms, which outruns the trace logging of conventional simulation loggings in the level of detail, so that dysfunctions of the TOS can be found via an additional analysis component.

2.4. Combination

There are many intersections of the requirements of logistic simulation and device emulation, so a reuse of model components can help saving time and money. (Vorderwinkler et al. 1999) already proposed reusing simulation components from planning phases for device emulation. But some fundamental differences should not be disregarded.

According to section 2.3 device emulation and logistic simulation differ in their requirements. Table 1 summarizes the most important differences. However simulation and emulation models may share many aspects of model structure and implementation and accordingly requirements on underlying simulation frameworks. Note, that the behavior and the tasks of terminal components like quay cranes are basically the same in simulation and emulation. The aim of our investigations was to investigate whether a reuse of model components in logistic simulation as well as in device emulation is advantageous particularly in container terminal domain.

Table 1: differences between emulation and simulation

Logistic Simulation	Device Emulation
bird's eye or worm's eye perspective	worm's eye perspective
keep model as simple as possible	high level of detail
closed system	needs interfaces for interacting with other systems
simulation time advancement depends on CPU speed	simulation time advancement depends on real time
huge number of experiments	smaller number of experiments
automatic-driven experiments	manual- or automatic-driven experiments
analysis depends on reports, post-process animation	concurrent animation and/or exhaustive logging

3. DESMO-J

3.1. Main Functionality

DESMO-J (Discrete Event Simulaton Modeling in Java) is a simulation framework targeted at developers of discrete-event simulation models (Page and Kreutzer 2005). It offers an exhaustive environment to ease implementation of event-oriented, process-oriented or transaction-oriented models using the object-oriented high level language Java. As specialty, a combination of these different simulation world-views is possible, too. DESMO-J was developed by Department of Informatics at University of Hamburg and published under APL 2.0, which is a common open-source license. Binaries, source code, API and a tutorial can be found on http://www.desmo-j.de.

DESMO-J clearly separates between model and experiment. Experiments are realized via several black box elements, which are ready-for-use for executing simulation runs, like the Experiment class which encapsulates experimentation functionality, the Simulation Scheduler or the Simulation Clock. Therefore the modeler only has to implement the model itself and does not have to care about implementing the technical framework. DESMO-J also allows creating hierarchical model composition, since models can be embedded into other models as so-called sub-models.

Models are implemented via deriving several white-box components. The main task the modeler has to solve at this, is mapping model activity to events (in event-oriented worldview) or to processes (in process-oriented worldview). Events and processes are to describe activity of model components, and can be scheduled on simulation scheduler. Entities describe dynamic model components. In event-oriented worldview, an entity's state can be changed by events. In process-oriented worldview, processes are derived from entities, which means a process is an active entity itself, which can interact with other entities, and change its state on its own.

Additionally, there are a lot of helpful components, which can be used in the model building process without derivation, e.g. there are different kinds of queues for waiting entities, or a set of stochastic distributions, which can be used to generate pseudo-random numbers. Since the user can specify a seed for random-number generation, a replay even of stochastic-driven simulation runs is possible. There are several objects, which can be used for reporting and analysis of experimentation runs, e.g. there are statistic classes like counter, time series or histogram, and report classes which generate reports in a selected format (e.g. HTML or XML) for model components, which are of interest for the user (e.g. entities or queues). The user also may implement statistic or reporting-classes on its own.

3.2. Implementation Language

DESMO-J bases on the object-oriented language Java, which addresses a large community of programmers. Firstly, all of our students have a Java-programming expertise, so teaching simulation with help of DESMO J is in line with their curriculum. Furthermore, many companies focus on the usage of Java, too, because advantages of Java are e.g. the portability across diverse platforms, the huge amount on reusable libraries (e.g. the Java API, GUI-libraries like SWING or even third party libraries like DESMO-J) or the robustness since there are no dangling pointers or memory leaks like e.g. in C++. Furthermore, because of its handy syntax and the large amount of developing tools, implementing simulation models using Java-language reach a satisfactory tradeoff between ergonomics and flexibility. For example there is the Javadocs-tool for easy creation of documentations or there are open-source, extendable Integrated Development Environments like Eclipse, which offer more features for manipulating code than most code-editors of commercial simulation tools do.

Although, the Java Virtual Machine cannot deliver the performance as models compiled to machine code for the underlying platform, note that commercial simulation tools often use scripting languages, which are definitely slower than the pre-compiled java-byte code of DESMO-J.

The current version 2.1.4b of DESMO-J bases on the newest Java version 1.6, so it includes state-of-the-art programming methods like Generics (e.g. Queues can act as specialized queues for selected entities, which primarily induces type safety) or other advanced concepts.

3.3. Expandability

A very important aspect in using DESMO-J is its expandability. It is possible to integrate DESMO-J into other software frameworks, as well as to extend it by additional class libraries. In order to ease modeling and simulation in special application domains, it makes sense to develop domain-specific extensions, which provide objects for that particular domain. These can be quite general, more technical extensions like FAMOS, which permits Multi-Agent-Based simulation with DESMO-J (Knaak 2006), or very specialized approaches, for example to use a DESMO derivate as simulation component to simulate business production processes and calculate ecological input/output-balances (Wohlgemuth 2005). There are several DESMO-J extensions, which are beyond the scope of this paper, however. Because of the flexibility of ASL 2.0, DESMO-J is licensed under, it is possible to implement extensions or changes without licensing it under an open-source license. Therefore, companies like HHLA have the possibility to develop their own extensions, which have not to be published obligatorily, just like the extension presented in this paper.

3.4. Harbor simulation with DESMO-J

At HMS 2003, we presented our first DESMO-J harbor extensions for our research projects in harbour logistics (Page and Neufeld 2003). This class library offers three types of objects: There are dynamic, mobile, temporary objects like ships, trucks and trains; dynamic, mobile, permanent objects like cranes and van carriers; and stationary, permanent objects like holding areas, gates, jetties and yards. While the stationary objects are mostly implemented as derivation of DESMO-J's queues, the mobile objects are derived from DESMO J's processes. The framework distinguishes between the material flow, which deals with transportation and turnover over material goods, and the information flow, which deals with handling of job briefings. We examined the practicability of these early concepts in an empirical investigation within Containerterminal Burchardkai in Hamburg in co-operation with HHLA. Because these extensions are good for logistic investigations but too abstract for the purposes of emulating container terminals devices, we present a new harbor extension for DESMO-J called COCoS in this paper. However since HMS 2003, DESMO-J was used to implement harbour models by several authors. We give a summary on selected articles of foreign parties in the following.

(Asperen et al. 2004) investigated the impact of arrival processes, using the example of a chemical plant's jetty in Rotterdam (Netherlands). They assumed that a company can affect the ship arrival times with help of tactical sales plans and compared four different arrival processes. Initially they used a commercial simulation tool for their investigations, later on in the project they changed to DESMO J. They assigned the following reasons: Limitations in scripting language of the commercial tool made it very hard to implement complex models, while Java-based DESMO J offered much more flexibility. In addition, communication with other components did not work satisfyingly, while DESMO J did not have any problems in that, which was very important for our intention of connecting a TOS to an emulation model, too. Last but not least they stressed the better runtime performance of DESMO-J (Aspeeren et al. 2004).

A team at the Christian-Albrecht-University in Kiel (Germany) compared different dispatching strategies for AGVs at CTA in co-operation with HHLA. With help of DESMO J, they built up a "simple model with short run times which can easily be configured and produces realistic results" (Briskorn and Hartmann 2005).The subject of their investigations were two variants of job assignment. In the first case, only AGVs were considered which were currently idle; in the second case, they also considered AGVs which would finish their task within a certain time. These strategies were tested within five different scenarios, with 100 simulation runs per scenario and strategy. (Briskorn and Hartmann 2005; Briskorn et al 2007)

A working-group at the Bleking Institute of Technology (Sweden) was very experienced with simulating logistic aspects of horizontal transport of container terminals. For example they identified the number of cassette-based AGVs that were required to achieve an optimal workload for quay crane for a given terminal (Henesey et al 2006a); they evaluated dispatching strategies for AGVs (Kosowski and Persson 2006); and compared cassette-based to traditional AGV-Systems (Henesey et al 2006b). They declared that they preferred DESMO-J, since it allowed the suitable approach to implement terminal components as process entities, which had their own life-cycles, properties and behavior, and which coordinated their tasks using Contract Net protocol. Furthermore, simulation experiments were able to be executed via command prompt and reports were created automatically (Kosowski and Persson 2006, p. 31; Henesey et al 2006a).

All these authors used DESMO-J for logistic simulation; we did not find any indications that DESMO-J had been used in emulation context so far. Therefore, the intention in our paper is to show, that using DESMO-J as emulation engine is definitely possible.

4. COCOS

As well as the examples for a usage of DESMO-J in harbour-simulation mentioned above, the COCoS-framework developed in (Brandt 2007) has originally been designed for logistic container terminal-simulation. This relationship and correlating application-domains are illustrated in Figure 4.

As an extension of DESMO-J for the development of integrated container terminal simulation models, COCoS can be used to compare different terminal-layouts or to evaluate the impact of different handling equipment or control strategies on key figures of terminal productivity. COCoS provides black and white box concepts for modelling and implementing such models.

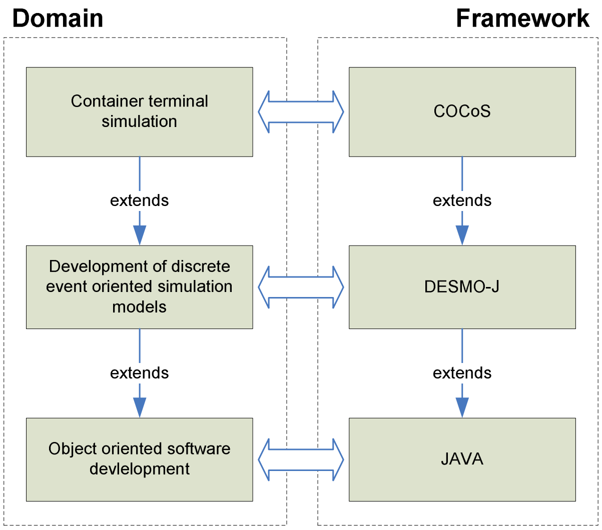


Figure 4: COCoS and Application-Domains (Brandt 2007)

4.1. Framework-architecture

A container terminal can be conceptually reduced to a collection of enclosed subsystems that are connected by an information and a material flow. Accordingly container terminal simulation models should be composed of enclosed model-components representing theses subsystems (Simulation Building Blocks, see Verbraeck (2004)). To achieve this aim COCoS provides an integration architecture for the combination, connection and communication between reusable model components as well as a basic architecture for the components themselves.

Basic relationships within the framework-architecture are shown in Figure 5. Components representing enclosed terminal subsystems are characterized as functional components (see section 4.2). Unlike these, the framework's technical components do not represent parts of the simulated system but provide services for pure technical aspects of a simulation model such as visualization or statistics. These reusable functional and technical components have been built upon the framework's core which contains all basic framework elements and upon several system libraries which form the framework's technological foundation.

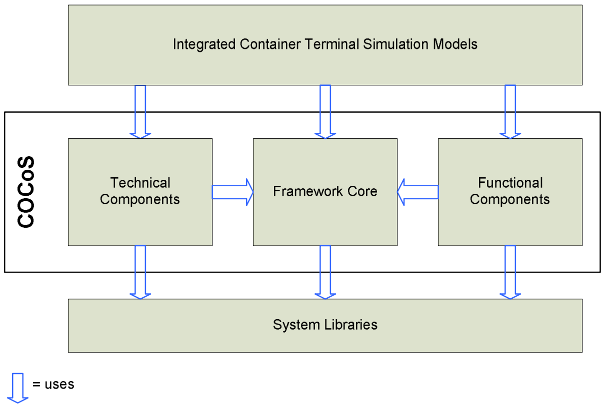


Figure 5: Framework architecture of COCoS (Brandt 2007)

4.2. Functional components

Functional components represent the container terminal subsystems such as quay-cranes, gantry cranes or horizontal transport systems. COCoS defines the architecture of functional components to assure their integrability and reuse in the context of different container terminal simulation models.

Multi-tier architectures can lead to advantages in loose coupling and reuse of software components. COCoS's functional components follow this important principle by implementing a three-layered architecture in order to transfer the mentioned advantages to the context of container terminal simulation. Hence each layer encapsulates a prominent scope typically existing in a container terminal simulation model, especially to refine a model wide separation of logical and physical aspects as described in (Pidd and Castro 1998).

The logical layer is responsible for planning decisions, resource management and task-handling mechanisms such as task-splitting, task-optimization or task-allocation. Through this the logical layer connects a functional component to a model's information flow.

In contrast, the material flow layer implements the physical aspects of a functional component. In terms of container terminal simulation, the material flow layer is mainly responsible for the handling and transportation of containers, thereby consuming simulation time for these physical operations. The communication layer serves as an extension point for adding additional interfaces to a functional component that are not supported by the standardized model architecture. This feature becomes especially important when embedding a functional component within a device emulator as described in section 5.2. But a functional component doesn't require all of the layers described above, given that pure information flow components that only provide the coordination of functionality, that may also be required in constructing a container terminal simulation model.

According to (Pidd and Castro 1998) the fast and secure development of complex simulation models through combination of model components especially requires a coupling scheme which specifies the communication between model components as well as the way they are connected. Therefore, COCoS contains mechanisms for the loose coupling of of functional components within an integrated container terminal simulation model or more precisely within the model's information and material flow. These aspects are to be described within the next two sections.

4.3. Information flow modeling

More than simply providing the basic elements needed for the development of container terminal simulation models, COCoS also supports the conceptual modelling process itself by providing a generic task-handling concept that standardizes the model's information flow as a flow of generic task-objects carrying sets of task-attributes.

Functional model components serve as task-handlers and primarily communicate by assigning tasks to each other accordingly. The underlying hierarchical model structure is defined by a model-wide XML-based task-model which can be considered as an executable coupling scheme within the model's information flow. More precisely a task-model specifies how super-tasks are split into sub-tasks and how task-attributes have to be passed on or generated thereby. The task-model also specifies targets for the generated sub-tasks as well as intermediate sequential relations. During their processing, the start and the completion of tasks are reported to their original sender. In order to execute a task model, COCoS provides black-box components which process the XML-based task-model-definition during experimentation.

By that task-models can be used to conceptually model the main information flow within an integrated container terminal simulation in an early phase of model development, which has also been proposed in (Robinson 2006).

4.4. Material flow modeling

The model's material flow is finally responsible for physical execution of tasks generated by the information flow. COCoS complements the described principles of information flow modeling by providing black- and white-box components for modeling and implementing these physical processes within a functional component's material flow layer. COCoS also contains basic resource classes like horizontal transporters routed on a graph based path net or crane components that can be combined to complex crane systems including a precise kinematic simulation. All these components follow DESMO-J's process oriented modeling style as it is set as default for implementation of functional components in COCoS. Material flow resources and their operational states are managed by the component's material flow layer making them available to the logical layer which then assigns tasks to these resources. Besides this connection to the information flow, material flow layers are interconnected by linking physical component layouts within an integrated terminal layout in order to facilitate container transportation in between components. This is realized by synchronous and asynchronous handover mechanisms that serve as the only coupling mechanism between functional components or their material flow layers respectively.

4.5. Technical components and model integration

Technical components are not part of the model logic and therefore do not follow the described architecture of functional components. Nevertheless COCoS defines architectural principles for important technological model aspects as well as for the support of the development of reusable and exchangeable technical components. The most important aspects are model visualization, model statistics, graphical user interfaces (GUI) or logging. In addition COCoS makes use of aspect oriented programming techniques (AOP) to separate technical and logical concerns as far as possible during the development process.

Functional and technical components are integrated to a container terminal simulation model by extending a centric COCoS model class. This model class is a direct descendant of the DESMO-J model class and thus provides the connection to DESMO-J's experimentation environment and simulation engine. It also serves as a model context for included functional components by offering common technical and functional services such as visualization, object and layout management or the input of scenario data.

5. COCOS APPLIED

At first this section briefly summarizes how COCoS has been used for the development of logistic simulation models and then proceeds with a more detailed description of the framework’s application in developing device emulators.

5.1. Logistic Simulation with COCoS

As mentioned above COCoS has originally been developed for modeling and implementing integrated container terminal simulation models for the investigation of logistic problems. The process of developing such models can be summed up as follows:

1. Modeling the information flow in the form of a model-wide task-model
2. Selection or development of appropriate functional components for the task model's execution
3. Selection or development of technical components for statistic analysis, visualization, user-interface et cetera.
4. Integration of functional and technical components within a COCoS-Model that is controlled by the task-model.

By iterating through these steps and varying model components and model configuration accordingly, different options of terminal design or control strategies can be explored.

Figure 6 shows a screenshot of a simple proof-of-concept-model during eperimentaion. The model features a simple pathnet based horizontal transport component, a single trolley quaycrane component and a storage-crane component as well as technical components that provide services for animation, logging and collecting task-based model statistics in a relational database. While the COCoS-Control-Center serves as graphical user interface to the model, DESMO-J's simulation engine is driving model execution inside.

Although most of these components are kept very simple, the model is able to show that COCoS's integration concepts work excellent for modeling, assembling and execution of logistic simulation models.

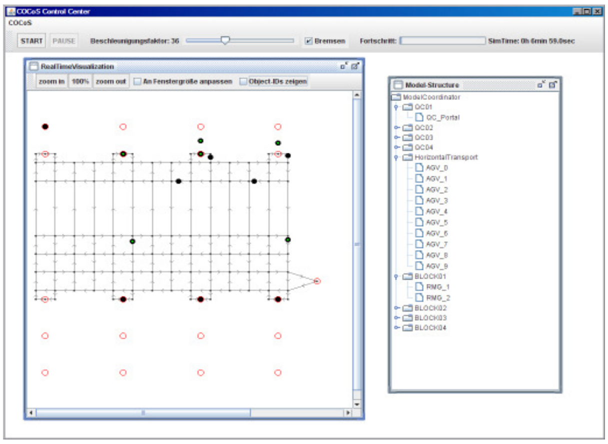


Figure 6: Logistic simulation model implemented with COCoS

5.2. Emulation with COCoS

Beyond this usage on logistic problems, functional terminal-components such as quay cranes or gantry cranes can be extended to device emulators by increasing detail level and adding specific communication- or user-interfaces. Related concepts have been proved in practice and will be described in this section.

5.2.1. General concepts

Different general, mainly technical features needed for device emulation had to be integrated directly into COCoS. Slowing down simulation time is a prominent example, since implementation of real-time functionality into DESMO-J has been in progress, but is not yet part of the official published DESMO-J release. To bridge this, COCoS was extended by a simple but adequate real-time control process that synchronizes real-time and simulation time by means of an adjustable time-synchronization cycle. Derived from DESMO-J's white-box simulation process, the simulation time control process basically linearises the discrete event oriented simulation time variation by constantly (re-)scheduling real-time consuming sleep-events. These sleep events are completely independent from the remaining functional model logic, rather their only purpose is to fill up the models discretionary time gaps. Within each cycle the advances of simulation time and real time are compared measuring the time divergence that is generated by computing time for model execution which is not reflected by the simulation time at all. This divergence is factored into the calculation of the next sleep delay in order to prevent simulation time and real time from continuously diverging. Cycle time and accuracy of this process as well as an arbitrary acceleration factor can be configured through model parameters. These parameters indicate the maximum response-time for the emulated devices for messages on their communication layer. For logistic simulation experiments real-time functionality can be turned off, but it can also be used to observe model behaviour by enabling real time visualization.

The second fundamental feature of an emulation framework is a concept for communication with an control systems and a human tester who want to manipulate a device emulator during run time. Since, we neither wanted to predefine the technical interface nor message types supported, we aimed for an abstract solution here. In addition to information flow layer and material flow layer, we implemented optional communication layers, which could be freely assigned when instantiating the device emulator. For example, there are communication layers for communicating via TCP-connection or via JMS-messages. These layers interpret incoming messages; depending on that, they choose a method to call on a device-specific task adapter, which creates the tasks according to the message. Logging the traffic on communication interface can be very helpful for debugging both the model and the TOS. The task adapter is also used by manual interaction interface, so that the human tester can directly insert tasks. If an event occurs which possibly postulates an outgoing message, the information flow layer will have to detect if a communication layer is allocated, and in this case call the according method on its interface.

5.2.2. Emulating a dual-trolley quay crane

As an instance, we outline the model-building process of a dual-trolley quay crane in this section. In Hamburg, these crane types are located at CTA. The two trolleys for handling containers are characteristic.

There is a manually driven main-trolley, which can reach every position under the quay crane, i.e. the ship or the quay lanes. Thus, the main-trolley is able to work autonomously without interacting with the second trolley. The driver of the main-trolley is also able to move the whole quay crane along the quay, which allows him to pick up all the containers from a ship in sequence. The second trolley is an automatic-driven portal-trolley, which gets its instructions directly from the TOS. It can only reach quayside transfer lanes and a special platform, called lashing platform, where it can set down the containers, which he picked up from the transfer lanes. From there, the main-trolley can retrieve them to set them down on ship. On the lashing platform, there are working people, who are preparing containers for transport; none of the trolleys is allowed to enter this danger area, if people are left on the platform.

This model component depicts the need of the manual interaction interface (see section 2.3). The behaviour of the lashing platform crew, the driving of the main-trolley and the possibility of deactivating the portal-trolley, everything has to be mapped in scope of this interface. Figure 7 shows the GUI of the COCoS' experimentation environment. Currently, there is a simple concurrent animation window showing a dual-trolley quay crane, one window containing some status information, and several windows providing access to the manual interaction interface. Thus, there are two possibilities to create tasks: Creation by the TOS via communication layer or by the human tester via GUI.

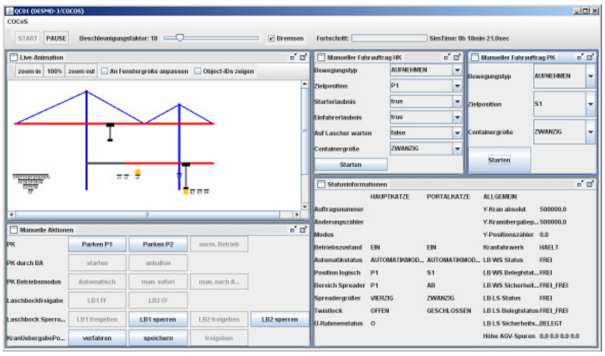


Figure 7: COCoS User Interface for device emulation)

A device like a dual trolley quay crane consists of several device parts, such as trolleys, lifts or spreaders. These device parts can be found on other terminal components, too. For example a van carrier contains a lift as well as a spreader, but their kinematic characteristics differ from the quay crane's lift and spreader. These detail information often is irrelevant in simulation studies, but in scope of emulation they are of much importance. That is why, we designed our model components of reusable and exchangeable device parts, which can be configured via XML-files. So there is the possibility to use very detailed device parts mapping the exact kinematic behaviour for emulation, or more simple device parts, whose time consumption depends on stochastic distributions for simulation experiments.

6. CONCLUSION

Combining logistic simulation and device emulation without overcharging the COCoS framework and the developed quay crane component was a challenging task. At first, it was important to clearly separate the component's layers for an efficient embedding of different communication interfaces. We supported this approach by developing a generic task adapter that could be used to translate arbitrary messages to COCoS task objects. This task adapter was also used to connect the manual interaction interface that was required for manually driven experiments. To satisfy the real-time requirements of device emulation we implemented an adequate time control process as a part of COCoS.

With real-time functionality a concurrent animation of simulation state became possible. This was urgently needed for manual device testing, but also pointed out to be very helpful for building, debugging and observing logistic simulation models.

Besides these more technical aspects, the main problem in combining logistic simulation model building and device emulation model building was the level of detail to decide on. To combine logistic simulation and device emulation, a precondition was to take worm`s eye perspective and build up model components in process-oriented world view, which is adequate in such cybernetic models at all. Nevertheless, there remain some substantial differences in requirements on detail level. For example, in logistic simulation the movements of components could be approximated via calculating time consumption by means of stochastic distributions, while exact kinematics are required in emulation context. Since devices managed by the material flow layer of a functional model component are exchangeable, it is advisable to use less detailed devices in simulation context, but more detailed devices in emulation context.

By bridging the described differences and combining logistic simulation and emulation it becomes possible to share the advantages of component based software development between the building processes of logistic container terminal simulation and device emulation models. Synergy effects emerging from exchanging and reusing functional and technical model components reduce effort and costs of model development and lead to improved model quality and a consolidated model architecture. It could be shown that DESMO-J and COCoS provide an extendible basis for this synergistic combination.

REFERENCES

Asperen, E. van, R. Dekker, M. Polman, H. de Swaan Arons. 2004. *Arrival processes in port modeling: insights from a case study.* Available from:

http://publishing.eur.nl/ir/repub/asset/1275/ei2004 16.pdf [accessed 31. March 2009]

Auinger, F., Vorderwinkler, M., Buchtela, G., 1999. Interface driven domain-independent modelling architecture for „soft-commissioning“ and „Reality in the Loop“. *Proceedings of the 1999 Winter Simulation Conference*, pp. 798-805, December 5-8, Phoenix (Arizona, USA).

Brandt, C., 2008. *Entwurf und Implementierung eines Frameworks zur Entwicklung von Containerterminal-Gesamtmodellen mit DESMO-J*. Thesis (Diploma), University of Hamburg.

Briskorn, D., Hartmann, S., 2005. Simulating dispatching strategies for automated container terminals. *Proceedings of the Annual International Conference of the German Operations Research Society (GOR).* September 7-9, Bremen (Germany).

Briskorn, D., Drexl, A., Hartmann, S., 2007. Inventory-based dispatching of automated guided vehicles on container terminals. In: Kim, K.H., Günther, H. eds. *Container Terminals and Cargo Systems*. Berlin, Heidelberg, New York: Springer.

Henesey, L., Aslam, K., Khurum, M., 2006a. Task Coordination of Automated Guided Vehicles in a Container Terminal. *Proceedings of 5th International Conference on Computer Applications and Information Technology in the Maritime Industries*. 2006, Oud Poelgeest (Netherlands).

Henesey, L., Davidsson, P., Persson, J.A., 2006b. Comparison and Evaluation of Two Automated Guided Vehicle Systems in the Transhipment of Containers at a Container Terminal, In: Henesey, L., 2006. *Multi-Agent-Systems for Containter Terminal Management*. Thesis (PhD), Blekinge Institute for Technology, 204-226.

Knaak, N., Kruse, S., Page, B., 2006. An Agent-Based Simulation tool for modelling sustainable logistic systems. *Proceedings of the iEMSs Third Biennial Meeting 2006*. July 9-13, Burlington (Vermont, USA).

Kosowski, P., Persson, O., 2006. *Development and evaluation of dispatching strategies for the IPSI™ AGV system.* Thesis (Master), Blekinge Institute of Technology.

Page, B., Kreutzer, W., 2005. *Simulating Discrete Event Systems*. Aachen: Shaker Verlag.

Page, B., Neufeld, E., 2003. Extending an object-oriented discrete event simulation framework in Java for harbour logistics. *Proceedings of International Workshop on Harbour, Maritime & Multimodal Logistics Modelling and Simulation*, pp. 79-85. September 18-20, Riga (Latvia).

Pidd, M., Castro, B., 1998. Hierarchical modular modelling in discrete simulation. *Proceedings of the 1998 Winter Simulation Conference*, pp. 383–389. December 13-16, Washington D.C. (USA).

Robinson, S., 2006. Conceptual Modeling For Simulation: Issues And Research Requirements. *Proceedings of the 2006 Winter Simulation Conference*, pp. 792–800. December 3-6, Montery (California, USA).

Schütt, H., Hartmann, S., 2000. Simulation in Planung, Realisierung und Betrieb am Beispiel des Container-Terminals Altenwerder. In: Möller, D. ed. *Simulationstechnik, 14. Symposium.* Hamburg: SCS, 425–430.

Steenken, D., Voss, S., Stahlbock, R, 2004. Container terminal operations and operations research – a classification and literature review. In: *OR Spectrum 26*, pp. 3–49. Berlin, Heidelberg, New York: Springer.

Vorderwinkler, M., Eder, T., Steringer, R., and Schleicher, M., 1999. An Architecture for Soft-Commissioning – Verifying Control Software by Linking Discrete Event Simulators to Real World Control Systems. *Proceedings of 13th European Simulation Multiconference*, pp. 191-198. June 1-4, Warsaw (Poland).

Verbraeck, A., 2004. Component-based distributed simulations: the way forward? *Proceedings of the 18th workshop on Parallel and distributed simulation,* pp. 141–148. May 16-19, Kufstein (Austria).

Wohlgemuth, V. 2005. *Komponentenbasierte Unterstützung von Methoden der Modellbildung und Simulation im Einsatzkontext des betrieblichen Umweltschutzes*. Thesis (PhD). University of Hamburg. Aachen: Shaker.

AUTHORS BIOGRAPHY

Bernd Page holds degrees in Applied Computer Science from the Technical University of Berlin, Germany, and from Stanford University, USA. As professor for Applied Computer Science at the University of Hamburg he researches and teaches in the field of Computer Simulation as well as in Environmental Informatics.

Philip Joschko studied Computer Science and Psychology at the University of Hamburg. Since he holds a diploma (MSc) degree in Computer Science he works as a research assistant and PhD candidate within the Center of Architecture and Design of IT-Systems at the workgroup of Prof. Dr. Page. Research interests are discrete event simulation, cooperative model building processes and the EU emission trading system.

Christopher Brandt studied Information Systems at the University of Hamburg with an emphasis on logistic simulation and distributed information systems. He holds a diploma (MSc) in information systems received for his thesis in which he developed the COCoS-framework for HHLA. Right now he works as a software architect for HHLA in the area of simulation, operations research and business intelligence.

INTEGRATION OF LOADING AND ROUTING IN LOGISTICS OUTSOURCING

S. Lisi[(a)], J. Egeblad[(b)], A. C. Garavelli[(c)], D. Pisinger[(d)]

[(a) (c)] DIMeG, Politecnico di Bari, Viale Japigia 182, 70126 Bari, Italy
[(b)] DIKU, University of Copenhagen, Universitetsparken 1, DK-2100 Copenhagen, Denmark
[(d)] DTU Management Engineering, Technical University of Denmark, Produktionstorvet 424, DK-2800 Kgs.Lyngby, Denmark

[(a)]s.lisi@poliba.it, [(b)]jegeblad@gmail.com, [(c)]c.garavelli@poliba.it, [(d)] pisinger@man.dtu.dk

ABSTRACT
Transportation management is a crucial issue in today's business environment. Firms pay attention to their core business and focus their competences and resources to improve and strengthen their competitive advantages. Consequently, firms find themselves integrated in a network of agents, since they purchase products and services to complete their processes. Logistics activities and services are committed to specialized companies. This changes the standard formulations of delivery and routing problems. This paper focuses the attention in a new formulation of the transportation problem accounting the supplier relationships of freight and groupage and the transportation problem. Transportation Management includes load planning and delivery route planning, respectively referred as Container Loading Problem and Vehicle Routing Problem. The two problems are very investigated but deal with optimization separately. This paper focuses the attention on an integrated approach considering also the outsourcing of logistic.

Keywords: container loading, vehicle routing, outsourcing, decision support system

1. INTRODUCTION
Outsourcing is a strategic and tactical tool of management. In the last decade of the last century, managers have increased its use in the business. The trend on the future is similar, and it is not related to a particular kind of firm, size or business.

Logistics activities and services are committed to specialized companies. In the literature, these actors of the supply chain are known as "third-party logistics (3PL) providers" or "logistics service providers (LSPs)" (Hertz and Alfredsson 2003; Rabinovich and Knemeyer 2006). They provide different types of logistics services (Larson and Gammelgaard 2001; Hong, Chin, and Liu 2004; Lai and Cheng 2004), and the carriage of goods seems to be the crucial service for the supply chain efficiency and effectiveness. More than 50% of the firms use logistics providers for transportation goods (Dapiran, Lieb, Millen, and Sohal 1996; Lieb and Bentz 2004).

The main aim of logistics outsourcing is the minimization of the transportation costs, without forcing the final client to receive a lower level of service. On the contrary, managers have to find ways to increase the customer service level. Generally, two types of managerial problems are related to logistics decisions. The strategic decisions deal with location, production, transportation and inventory problems, according to the specific configuration of the logistics networks. The operational decisions deal with scheduling, lead time, routing, truck loading problems, according to the coordination in the logistic networks.

Different costs could affect the optimization of the shipments. Inventory costs affect the frequency of the shipments, and they are also related to the production process. Transportation and delivery costs affect the design of the flow of goods and the number of travels. The importance of transportation is underlined by the effort made on the networks design and management - point to point, corridor, hub and spoke systems (Tavasszy, Ruijgrok, and Thissen 2003; Lapierre, Ruiz, and Soriano 2004; Hesse and Rodrigue 2004).

The use of logistics outsourcing is sensitive to the mechanisms of coordination among the actors in the supply chain. Logistics is a typical activity where coordination is very important, and the consolidation of orders (Galbraith 1977) is a suitable technique to perform it. Consolidation is the process of combining different items, produced and used at different locations (spatial consolidation) and/or at different times (temporal consolidation) into single vehicle load (Mintzberg 1979).

2. LITERATURE REVIEW

2.1. Three-dimensional Container Loading
The container loading optimization problem is a central problem in the industry. Common problem formulations are Bin-Packing, Knapsack Packing, Container Loading and Multi-Container Loading (Pisinger 2002). Many variants of the Container Loading Problem have been studied.

We consider the three-dimensional knapsack problem with irregular shapes. While two-dimensional packing of irregular shapes is a topic which has been thoroughly investigated, three-dimensional packing has received far less attention.

Generally, polygons are often used to represent the contour of shapes in the two dimensional problems, whereas triangles mesh structures are often used in three dimensional problems.

Moreover there are two overall placement strategies: relaxed placement and legal placement. In legal placement no overlapping is allowed.

Ikonen et al. were among the first to consider optimization problems with irregular three-dimensional shapes. They used genetic algorithms with a relaxed placement method based on triangle intersection (Ikonen, Biles, Kumar, Wissel, and Ragade 1997; Ikonen, Biles, Lewis, Kumar, and Ragade 1998).

Dickinson and Knopf used a legal placement method where every item is sequentially placed through an individual optimization heuristics (Dickinson and Knopf 1998; Dickinson and Knopf 2002).

Cagan et al. used a relaxed placement method (Cagan, Degentesh, and Yin 1998). Simulated annealing and spatial octrees (de Berg, Van Kreveld, Overmarks, and Schvarzkopf 2000) were used to quickly determine overlap.

Egeblad et al. generalize their 2D relaxed placement method to three-dimensions, improving the results of Ikonen et al. and Dickinson and Knopf both in speed and quality (Egeblad, Nielsen, and Odgaard 2006).

Egeblad et al. use a several consecutively applied heuristics to optimize container loading of furniture (Egeblad, Garavelli, Lisi, and Pisinger 2009).

2.2. Integration of Routing and Loading

Managing the distribution of goods is an important issue in Transportation Management. The vehicle routing problem is a topic which has been thoroughly investigated, but only in the last years integrated approaches have been presented to jointly solve the two problems.

Iori et al. used an exact approach to solve a two-dimensional loading capacitated vehicle routing problem (2L-CVRP) based on a branch and cut algorithm, for the minimization of the routing cost and on a branch and bound algorithm for checking the feasibility of the loadings. Actually, it was a three dimensional problem, but the items were often transported on top of rectangular bases (e.g., large pallets of suitable size), and, due to their fragility or shape, they could not be stacked one over the other. In this case, the general three-dimensional loading problem reduces to a suitably defined two-dimensional loading problem of the rectangular items on the floor of the vehicle (Iori, Gonzalez, and Vigo 2006).

Gendreau et al. developed, implemented, and tested a Tabu Search algorithm to solve a 3L-CVRP (Gendreau, Iori, Laporte, and Martello 2006). The authors considered a combination of capacitated vehicle routing and three-dimensional loading, with additional constraints frequently encountered in freight transportation. They proposed a Tabu search algorithm that iteratively invokes an inner Tabu search procedure for the solution of the loading sub-problems. The vehicle had a rectangular loading space. The items were parallelepipeds and had a fixed orientation with respect to the height. Moreover, there were fragility constraints. A 2l-CVRP is solved in (Gendreau, Iori, Laporte, and Martello 2008).

Moura and Oliveira considered an integrated approach to solve a three-dimensional loading and a vehicle routing problem with time windows (Moura and Oliveira 2004). They proposed two different approaches. The first one was a sequential approach; the second used a hierarchical approach to integrate the two problems.

3. PRELIMINARY CONSIDERATIONS

Consolidation requires that the placement of the items in a vehicle is verified before the real loading. Otherwise, some items could be left out with a negative impact on the backorder and service level for the final client. Consequently, a three dimensional container loading problem needs to be solved.

To carry out an internal consolidation in the shipment, two ways can be followed. First, goods of different orders and clients could be loaded in the same vehicle to carry out a route with many destinations. Second, goods of different orders and clients could be loaded in the same vehicle to carry out a route to a hub point where groupage services are asked to 3PL. In both cases a three dimensional container loading problem must be solved to verify that the selected route can be conducted.

Transportation and groupage services are given in outsourcing, and an optimization of the network utilization needs to be carried out.

In this paper, we present a model of distribution where there are three actors: a manufactory company that produce items for the clients, a provider of freights, and a provider of groupage. In the model, the company is both shipper and consolidator, while the providers are the carriers.

In the model we solve a vehicle routing problem integrated with a three dimensional container loading of items with irregular shapes. The cost formulation of the Vehicle routing problem takes into account an innovative objective function, consistent with the specific outsourcing of the logistic services. An Adaptive Large Neighbourhood Search (Pisinger and Ropke 2007) is then used to solve the distribution model, and some real tests are carried out with the distribution logistics of an Italian furniture manufacturer.

3.1. Furnishing Company Loading

The container loading of furniture problem has special features that were considered in Egeblad et al. 2009. Conventional three-dimensional container loading problems consider placement of boxes with fixed dimensions inside a box-shaped container.

Since sofas and armchairs have irregular shapes, they are usually coupled in order to create cuboids to mimic rectangular placement. Items can be squeezed slightly during transport, which makes a prediction of the final cuboid dimensions difficult. Moreover the shapes can be very different and the items are not rigid.

These conditions affect the strategy of integration of routing and loading, because the evaluation of volume of set of items is affected by combination of the products in the loading configurations. Let's say v_a the volume of item a and v_b the volume of item b, $v_{(a,b)}$ the volume of a configuration with item a and b. Then,

$$v_{a,b} \leq v_a + v_b. \tag{1}$$

Generally the volume of configuration is lower than sum of volumes of single items.

Let's say v_{c1} the volume of bounding box of configuration c1 and v_{c2} the volume of bounding box of configuration c2. $v_{(c1,c2)}$ the volume of the two configuration placed one nearest the other.

$$v_{overlap(c1,c2)} = v_{c1} + v_{c2} - v_{c1,c2} \geq 0. \tag{2}$$

Generally, sofas and armchairs are loaded with their arms resting on the container floor. To solve the container loading problem we use the heuristics presented in (Egeblad, Garavelli, Lisi, and Pisinger 2009) in which more details of the difficulties involved with container loading of furniture are also presented.

3.2. Delivery Features

In the model of shipment that we have analyzed there are four kinds of actors:

- Shopkeeper (consignee)
- Manufacturer (shipper and internal consolidator)
- Transportation suppliers (carrier)
- Groupage suppliers (carrier and external consolidator)

Shopkeeper specifies the orders of products to manufacturer. After the production manufacturer plan the shipments, define the service mode of shipments and buy the services they need from the logistic providers. The freights are transported by road. Transportation suppliers execute the routes organized by the manufacturer. They need to supply the trucks and all of the resources required to execute the routes, and to plan appointments with shopkeepers or groupage suppliers for delivery. Groupage suppliers organize the delivery of the products not sent by the manufacturer directly to the shopkeepers via the transportation suppliers.

Manufacturers can choose different service modes to deliver the products to the shopkeepers:

- Direct
- Multi Drop
- Groupage
- Mix

Direct shipment consists in a route supplied by carrier. It takes the products from the warehouse of manufacturer and delivers them to warehouse of the shopkeeper. The cost of service is fixed and it is related to a subarea of the shopkeeper's country. It's not related to effective distance and time: so if there are two warehouses in the same subarea and one is more distant than the other, the cost of service is the same for both warehouses. However there is there is a variability in the profits of carrier. This transport is used in Full Truck Load and Full Container Load.

Multi Drop shipment consists in a route supplied by carrier. It takes the product from the warehouse of manufactures and delivers them to warehouse of shopkeepers. The shopkeepers can be localized in the same subarea or in a different subarea. So the cost of service is related to the cost of single direct shipments for every shopkeeper. In fact, it is the biggest of all. Fixed costs are provided for each un-load location. Additional costs are related to specific travel and are related to distance and time of travel.

Groupage shipment consists of two steps. In the first step a carrier supplies a route from warehouse of manufacturer to warehouse of a groupage provider. The cost of service is fixed and it is related to a subarea of the groupage provider's country. This transport is used in Full Truck Load and Full Container Load. After the groupage supplier provides delivery of the products to shopkeepers, in according to Less than Truck Load deliveries. The cost of service is related to volume and weight of items and on locations of destinations.

Mix shipment is a combination of groupage shipment and multi drop; exactly one unload point is a groupage point.

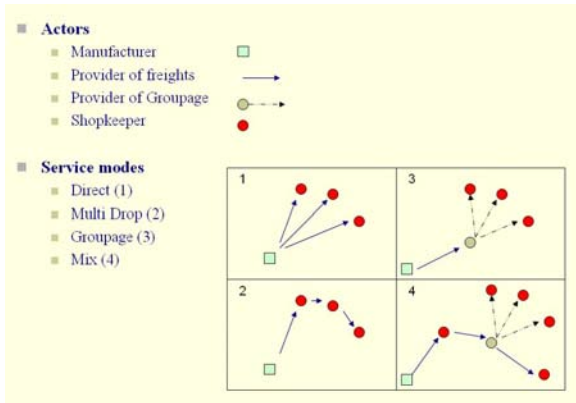


Figure 1 – Service modes in the model

Often, the routing optimization is "static": for each customer is a-priori defined the service mode (point-topoint, groupage, direct-to-unload, etc…). Therefore, each customer is considering lying on a pre-defined path or, at least, each one belongs to a specific area and no optimization of the route is made during the optimization referring to the loading. Likewise no optimization of loading is made during the optimization of routing. Generally the problem is solved in more steps: first of all the service mode is defined. After the route is organized and finally a container loading is performed. This could lead to an inefficient optimization because:

- it is easy to define the optimal service mode for a single travel but it's not easy to choose the optimal service mode for all orders and customers;
- not considering the loading during optimization of routing can lead to infeasible truck loadings or truck loadings with not good level of filling;
- if the loading is organized by a LIFO rule, not considering the routing during the optimization of loading can lead to not feasible truck loading.

4. MODEL OF ANALYSIS

4.1. Working hypothesis

Consolidation can be performed in internal way or in external way. Internal consolidation is performed with products of same company shifted in the time or in the space. External consolidation is performed with products of different firms. In the model we describe how the consolidation is performed by a shipper and by a groupage provider. The shipper performs an internal consolidation group in same trucks products of different shopkeepers and deliveries them buying freight services. An external consolidation when shipper acquires groupage service from logistic provider.

More logistic providers can be selected by the shipper. This means that we can consider that there are enough resources in the markets to supply service to shippers. In the analysis of this paper, no attention is involved in vendor rating and supplier selection. A hypothesis is that all suppliers are equal; this means that we can consider just one carrier and one provider of groupage service.

The previous hypothesis permits to classify the routing problem as a routing problem with an unlimited fleet of vehicles. One kind of vehicles at the time is considered in the problem.

The planning of delivery is built by shippers, only execution is performed by carrier. From this point of view, the vehicle routing problem can also be classified as an open vehicle routing problem.

Only capacity of vehicles is take in analysis, so the problem is a capacitated routing problem. There are no explicit time windows constraints.

To verify that vehicles are not over-filled a three dimensional container loading problem is solved. Loading constraints are considered in the model and products with irregular shapes are taken in account. The routing is a three dimensional loading capacitated vehicle routing problem.

To choose the products to deliver, some rules and constraints are considered. There isn't the capacity to ship all products of warehouse, because the loading operations are manual. Oldest produced orders have a higher priority in the delivery as orders to seasonal sales. In the model, only two kinds of priority are considered.

Generally the warehouse costs affect the priority of shipments. Since the delivery is a daily activity, the warehouse costs are not taken in account in the objective function. Besides an unfilled vehicle can be stopped in the warehouse for few days to allow filling with work in progress products of same shopkeepers. The problem is an open problem and to manage it we assume that two average filled trucks are worse than one filled truck and one no filled truck. In this way an objective will be to find solutions with more filled trucks.

4.2. Objective function and constraints

Since the transportation is in outsourcing the routing problem has to reflect a different formulation of the objective function. There is a transaction between carrier and shipper based on master contracts to delivery items to shopkeepers. The planning of deliveries performed by shippers has to contain more objectives.

The principle objective of the manufacturer is to minimize the cost of the delivery plan. The delivery plan is defined by the all routes need to perform the shipments. R is the whole of route and r_i is the i-th route.

$$C(R) = \sum_{i=1}^{n} c(r_i) \tag{3}$$

The cost of single route is made up of three parts: link cost, unload cost and groupage cost. Link costs and unload costs are defined in the contracts between manufacturer and freight providers, groupage costs are defined in the contract between manufacturer and groupage providers.

A route r_i is a sequence of points (source, d_{1r}, d_{2r}, …, d_{jr}, ….,d_{nr}), d_{jr} is the j-th point of delivery in the route r. $d_{j.}$ is a location and for each location a subarea l_i is defined.

$$d_j \in l_i \subset \{l_1, l_2, ..., l_i, ..., l_n\} = L \tag{4}$$

$$d_{j,r} \in l_{ir} \subset \{l_{1r}, l_{2r}, ..., l_{ir}, ..., l_{mr}\} = L(r) \tag{5}$$

$$m \le n \tag{6}$$

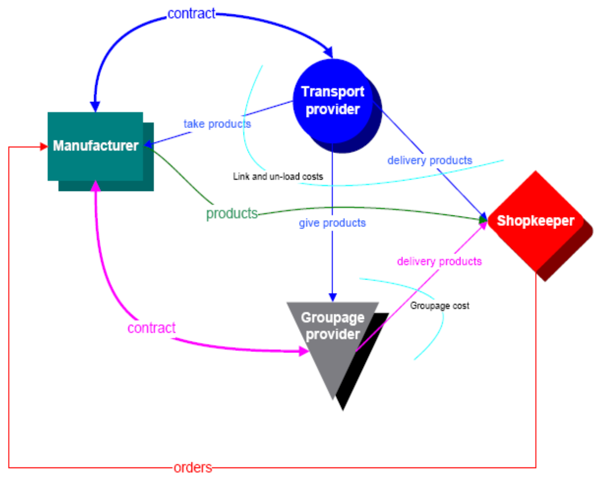


Figure 2 – Supply Chain in the outbound of logistic of the model: shopkeeper orders product to manufacturer. After the production, manufacturer delivery the products buying transport and groupage services from logistic providers

$C(l_i)$ is the cost of transport service between warehouse of manufacturer and subarea l_i. Carrier starts from his depot, goes to warehouse of manufacturer. After loading of products in the trailer, the carrier goes to destination d_j in subarea l_i.

$$c(l_i) = f(source, l_i) \quad (7)$$

In real contracts the cost of link is also due to specific carrier and kind of vehicle. Since we consider just one kind of vehicle and no selection of carrier, we can assume valid the expression (7).

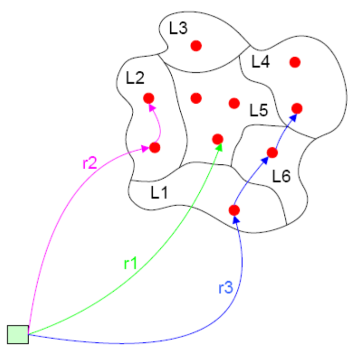


Figure 3 – Link Cost

In direct service mode and in groupage service mode the cost is fixed and it is related to a subarea of the shopkeeper's country or the country of the groupage warehouse. The route r1 in figure 2 is an example.

In multidrop or mix service modes the link cost is related to different subarea. If all delivery points are in the same subarea, the cost is defined as the cost of direct service mode. The route r2 in figure 2 is an example.

Otherwise the link-cost is the maximum link-cost of link-costs of subareas. The route r3 in figure 2 is an example.

$$C_l(r_i) = \max_{i \in L(r)} c(l_i) \quad (8)$$

The second part of route's cost is related to number of delivery points in the route. Since that the distance and the time to perform a route with more delivery point are bigger than in a route with one delivery, appropriately costs are imputed. The cost is made up of two parts. One is related to number of stops: for each stop except one a fixed cost c_{stop} is imputed. This cost is imputed because the carrier has to plan more appointments and queues and has to cover more distance. The second one is a variable cost and is related to the extra distance involved in a multi point travel (r_i) compared to a single point travel ($r^*(r_i)$). Since that some kilometres are covered in fixed part, a threshold k is defined.

$$C_u(r_i) = C_{u1}(r_i) + C_{u2}(r_i) \quad (9)$$

$$C_{u1}(r_i) = (n_{stop}(r_i) - 1) \times c_{stop} \quad (10)$$

$$C_{u2}(r_i) = (d(r_i) - d(r^*(r_i)) + \\ - (n_{stop}(r_i) - 1) \times k) \times c_{km} \quad (11)$$

$r^*(r_i)$ is the travel between source and the last delivery point in the route.

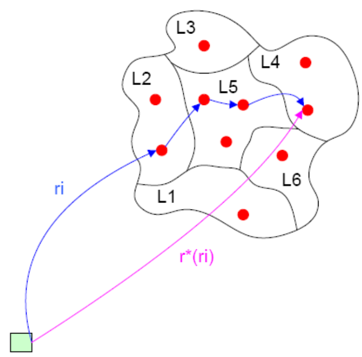


Figure 4 – A multi point delivery route and correlate route $r^*(r_i)$ in unload cost definition

The last part of route's cost is the groupage cost. It is applied to all products if the service mode is groupage or to products in the groupage point of mix. Exactly this part of cost is related to second phase of groupage service. The groupage provider organizes the shipment to shopkeepers. It benefits from filling a vehicle with products of a specific couple shopkeeper- manufacturer with products of other couples supplier-customer, generally in the same subarea of shopkeeper or nearest. In this sense the manufacturer benefits from external consolidation. The cost of service is related to subarea of the shopkeeper's location, volume and weight of products. Here the structure of costs is related only to volume of products and a not standardized unit of measure is used; it is typical in Italian furniture manufacturer and it measures the number of seats in the products. Let's say s to be the number of seats in the route to ship at groupage for location l_i. The cost per seats is:

$$\begin{array}{lll} c_0^g(l_i) & for & 0 \le s < d_1 \\ c_1^g(l_i) & & d_1 \le s < d_2 \\ .... & .... & .... \\ c_n^g(l_i) & for & d_n < s \end{array} \qquad (12)$$

Then the total cost of groupage in the route is:

$$C_g(r_i) = \sum_{l_x} s_i(r_i) \times c_j^g(l_z, s_i) \qquad (13)$$

Since that the groupage provider and the transport provider can be different, not necessarily the location defined in the groupage formulation are equal in the link cost formulation.

The total cost of route is:

$$C_t(r_i) = C_l(r_i) + C_u(r_i) + C_g(r_i) \qquad (14)$$

Minimizing the total cost of route is one of objectives of manufacturer. The shipment is travelled on long distance, so only truck with high utilization are used to delivery the products. High levels of consolidation are required. Often high level of consolidation means more use of groupage service. Since the delivery lead time using groupage service is larger, a trade-off between utilization of volumes and delivery lead time occurred.

The model has not a time windows formulation, but the orders are split in two groups:

- High priority of delivery
- Normal priority of delivery

Orders with high priority have to be placed in full vehicles and no groupage service has to be used. The first condition is modelled as a soft constraint: if the products of a high priority order are placed in an unfilled vehicle, a penalty cost is added to objective function. If the volume of truck is v_t, the available volume v_{ft} is:

$$v_{ft} = v_t \times f_r \qquad (15)$$

Since that not all space can be used to fill the truck, f_r is estimation on real useful space. It is related to variety of product's dimensions. The utilization of space is:

$$u(r_i) = \frac{v(r_i)}{v_{ft}} \qquad (16)$$

t_f is a threshold value to identify a full vehicle. If the $u(r_i)$ is bigger than t_f, the route is a full truck. Then priority cost is:

$$pc(o, r_i) = \begin{cases} 0 & t_f < u(r_i) \\ c^p & u(r_i) \le t_f \end{cases} \qquad (17)$$

If order o is a priority order y_o is 0, otherwise is 1. If $O(r_i)$ is the set of orders o in the route r_i, the penalty cost of priority in the route r_i is:

$$C_p(r_i) = \sum_{o \in O(r_i)} y_o \times pc(o, r_i) \qquad (18)$$

Since the delivery lead time using groupage is bigger than direct or multi drop service mode, the second condition is a hard constraint. Planning products of high priority orders with groupage service mode makes the solution not feasible.

As explained in Section 4, a sub-objective is to planning more filled trucks. If two solutions are equal as regard the total cost of route (expression 14th), the solution with more filled trucks is preferred. Besides, unit cost transport is generally larger in an unfilled vehicle than in filled vehicle. There is a soft constraint in the objective function, adding a penalty for whenever an unfilled truck is used.

Let' say t''_{nf} and t'_{nf} to be two threshold values to identify an average filled truck. Then if $u(r_i)$ is between the two threshold values, the cost of not full vehicle is:

$$\begin{aligned} C_{nf}(r_i) = Min[&(t''_{nf} - u(r_i))^2 \times cp_{nf}, \\ &((u(r_i) - t'_{nf})^2 \times cp_{nf}] \end{aligned} \qquad (19)$$

Otherwise it is zero. The penalty cost has a quadratic function and it penalize the solutions with many average full trucks in comparison to solution with filled and empty trucks.

The described objective function takes in account the aims of manufacturer. Since the formulation is different in comparison to standard vehicle routing problem formulation, the planning of the route could be far from objective of transport provider and not feasible. In the contract there are some constraints. They refer to:

- max number of stops in a route;
- max distance between first and last stop in the route;
- max distance in the route;
- max distance between stops.

The constraints allow at transport provider to plan a feasible execution of travels performing the appointments with shopkeepers or groupage providers to delivery the products. Since that the cost are related to locations and not on distances, nothing ensures that the route is cheapest for the carrier. He would like to change the order of visits to shopkeeper but he has to unload and load all products at every stop. Since a LIFO rules is used to load and unload the products, the manufacturer have to take in account it in the routes' planning. A standard formulation of vehicle routing problem needs to meet provider's objectives.

The provider's cost to perform a route is made up of three parts. One is related to fixed costs as assurances, hires, amortizations, etc. The second one is related to distance and the other one is related to time, as expressed in the formula below.

$$C_p(r_i) = C_{fp} + \sum_{i=0}^{n-1} c_{km} \times d(i,i+1) + \\ + c_h \times t(i,i+1) + e(i,i+1) \quad (20)$$

In the previous formula the return travel is not considered because the provider generally performs a not empty return travel, supplying services to other companies.

4.3. Heuristics description

To solve the model some heuristics are used to solve:

- the vehicle routing problem with outsourcing of logistic formulation (we call it Capacitated Vehicle Routing Problem with Contract Optimization, CCVRPCO);
- the Three Dimensional Container Loading of furniture with Irregular Shapes;
- the integration between Container Loading and Vehicle Routing.

Due to the intrinsic hardness of three-dimensional (3D) packing, it is intractable to verify feasibility of a packing in every step of the algorithm in reasonable time. Besides, it's not possible a pre-computation of all combinations of packages of products. Then a sequential integration approach can not be used, a hierarchical method is used. In the hierarchical approach the CLP is considered as a sub-problem of CVRPCO. The three dimensional CLP is solved with heuristics showed in (Egeblad et al. 2009).

To solve the CVRPCO an Adaptive Large Neighbourhood Search was implemented.

A heuristic was implemented to solve the integration problem. The principle problem on integration is that the volumes of packages of products is related to real placement as explained in §3. In the figure 5 the interaction of the three heuristics is showed.

ALNS combines various functions of removal and insertion of package of orders allowing an exploration of large neighborhoods of solution. After a definition of starting solution, the insertion heuristics and the removal heuristics are selected using a roulette mechanism of selection.

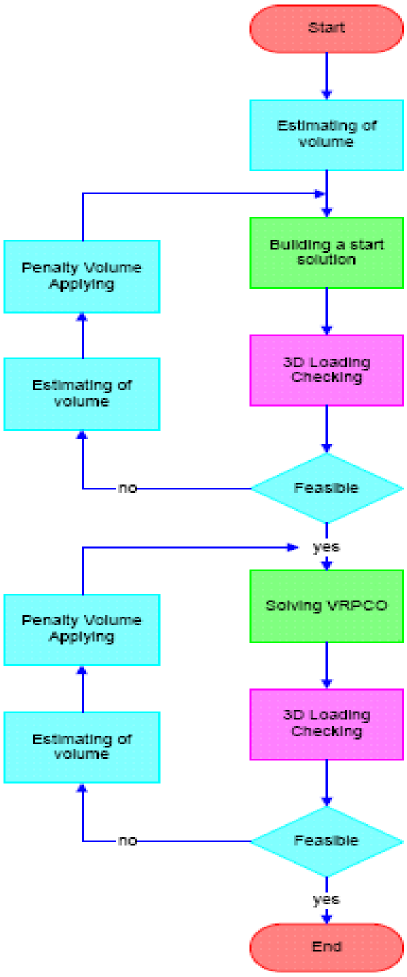


Figure 5 - A Heuristics for a 3DL-CVRPCO. The green boxes represent the 1L-CVRPCO activities, the pink boxes represent the loading modules and the cyan boxes represent the integration of routing and loading.

Randomly the size of action is defined and the removal and insertion heuristics are applied to define a new feasible solution of 1L-CVRPCO. The new solution is accepted with a Simulating Annealing Criteria. Exactly the probability is:

$$\begin{cases} 1 & f(s') \leq f(s) \\ e^{-\frac{f(s')-f(s)}{T}} & f(s') > f(s) \end{cases} \quad (21)$$

To manage the removal of orders three kinds of heuristics are used: random removal, shaw removal, worst removal.

In random removal q orders are randomly selected and removed from routes, allowing an escape from a local optimum and introducing a diversifying effect.

In shaw removal q orders are selected if they are closely related. Different criteria define the closeness. The idea is that no closely related orders could generate worse solutions. The q orders that give the minimum sum of relatedness value are selected:

$$\min \sum_{m \in O} \sum_{n \in O} r_{mn} x_m x_n$$
$$\sum_{o \in O} x_o = q \tag{22}$$
$$x_o \in \{0,1\} \quad \forall o \in O$$

r_{mn} is the measure of relatedness. In the model there are four different measurements:

- Distance: it measures the distance between orders. The idea is that the management of nearest orders could generate better solutions;
- Route: it measures if the orders are placed in the same truck. The idea is to move more orders from the same route to generate new good solutions.
- Volume: it measures the capability of orders to fill a truck. The idea is to use the orders from the least filled trucks to fill the most filled trucks.
- Groupage: it measures the service mode of orders. Orders that are placed in same truck and are delivered to groupage will be selected.

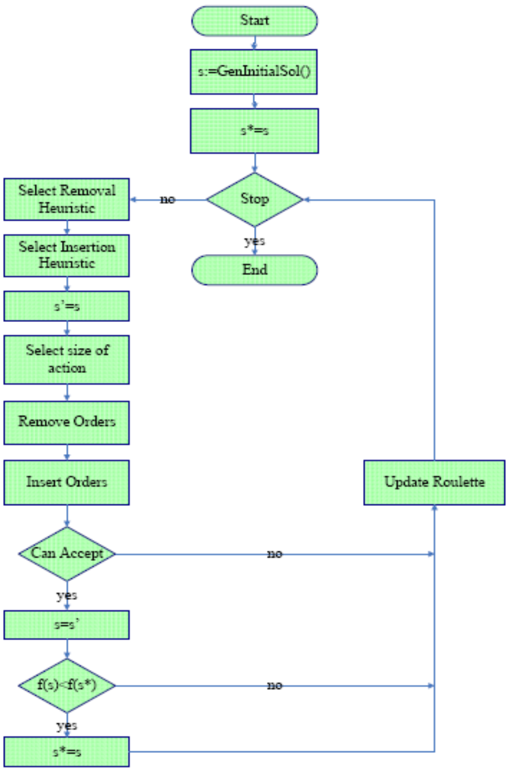


Figure 6 - algorithm of first level of ALNS heuristics

The worst removal is used to select the q orders that reduce more the objective function of solution. To select an order, for each order the objective function of solution s' is calculated. Solution s'(o) is the solution without the order o. The saving is:

$$\Delta(o) = f(s'(o)) - f(s) \tag{23}$$

The saving is negative if removal of order decreases the objective function. So the order with lowest saving is selected. After removal, the procedure is iterated starting from solution s'.

Similarly soma insertion heuristics are defined: random, best first insert, regret insert groupage and regret insert.

Random insert is a random insertion of the removed q orders and the advantages are the same of random removal.

Best first insert is an insertion heuristics that works as worst removal. The difference is that also the position in rouse has to be defined. Solution s'(o,r,i) his a new solution in which order o is inserted in the route r at position i. The saving is:

$$\Delta(o,r,i) = f(s'(o,r,i)) - f(s) \tag{24}$$

If an insertion of an order decreases the objective value of solution s', the saving is negative. So the order o in the route r at position i is selected. After insertion the procedure is iterated until all orders q are re-inserted.

The regret insert is similar to best first insert, but it is based on regret value. The regret value is the difference between the saving as defined in (24). Exactly the difference is between the saving of 2nd best position of order o and his best position:

$$\delta^o = \Delta_+^{0(ri)_2} - \Delta_+^{0(ri)_1} \tag{25}$$

High value of regret value means to insert order in his best position is priority. It's possible to extend the regret insert, considering more insertions. The aim is to preview the difficulties.

$$\delta_k^o = \sum_{j=1}^{k} \Delta_+^{0(ri)_j} - \Delta_+^{0(ri)_1} \tag{26}$$

In all iterations of ALNS a removal heuristics and an insertion heuristics are selected to modify the solutions. A roulette mechanism is designed to select the heuristics. Each heuristics has a score. The scores of couple of heuristics used in iteration are updated with:

- $\sigma 1$: a new best solution is found;
- $\sigma 2$: a new solution better than current solution is found;
- $\sigma 3$: a worse solution is found and it is accepted du to simulated annealing criteria.

After x iterations the scores are updated with the formula below:

$$w_i(1-r)+r\frac{\varphi_i}{\phi} \quad (27)$$

The probability to choose heuristics is related to the scores since that p_i is:

$$p_i=\frac{\omega_i}{\sum_{j=1}^{h} w_j} \quad (28)$$

Every time a new solution of 1LCVRPCO has been generated, the loading heuristics is used to verify the feasibility of loading of trucks. If all trucks are feasible the procedure stops and the loadings define the 3LCVRPCO. Otherwise a new solution 1LCVRPCO will have to be generated after a penalty volume has been added to product volumes. As explained in §3.1, there are two reasons on difference between sum of volume of single products and volume of bounding box of placed products: product in cuboids arrangements and cuboids arrangements in the package of products.

A utility function estimates the volume used by a product in cuboids arrangements. Since that the loading heuristics uses different cuboids arrangements, each time a three-dimensional problem is solved, statistical data are collected. Let C(i) be the set of encountered configurations with product i from all previously solved three-dimensional problems. Then the utility of product i is:

$$\frac{1}{|C(i)|}\sum_{c\in C}\sum_{j\in c}\frac{w(j)\times h(j)\times d(j)}{w(c)\times h(c)\times d(c)} \quad (29)$$

In the next stage of 1LCVRPCO the estimate of the amount of volume that will be occupied by i will be:

$$v_e(i)=u(i)\times w(i)\times h(i)\times d(i) \quad (30)$$

Waste space and occupied space in the truck are affected by lost space above and between products due to the related positions between cuboids arrangements. To consider it, a penalty factor is applied to volume of package of products. The penalty factor is a measurement of how hard it is to place an individual order of products inside the truck.

At end of every iteration and after the loading checking, if the truck is overfilled the package's penalty factor is increased.

If the truck is not overfilled a ratio r is calculated:

$$r=\frac{OccupiedVolume+WasteVolume}{ItemVolume+PenalVolume} \quad (31)$$

$$0.9\times IterationCount\le r\le 1.02 \quad (32)$$

If expression (32) is verified the penalty factor is not updated, otherwise it will increase or decrease due to value of r.

5. EXPERIMENTS

To test the model some instances were randomly built. Precisely, a random instance generator was used to select a subset of clients from a large set of clients with same probability of extraction. In the same way a group of orders is selected until to the specified total number of seats of products. Randomly the orders are assigned to clients. The only condition is that each client has at least one order.

Table 1: Experiment's Results

total seats #		19968,1		
service mode - prov. constraints	data	capacity criteria		
		Full3D	Seats	Vol
Only direct service mode - indifferent	Total # route	1411	1413	1405
	average filling index %	8	10	7
	# unloads	1411	1413	1405
	Manufacturer's cost in objective function [€]	2637385.00	2641280.00	2626555.00
Only Groupage Service Mode - indifferent	Total # route	144	144	105
	average filling index %	80;84;85 (all; >60; >100)	96	95
	# unloads	144	144	105
	Total %group.	100	100	100
	Manufacturer's cost in O.F. [€]	513108.89	513108.95	449928.90
Direct and Multi Drops Service Modes - no	Total # route	171	145	105
	average filling index %	69;73;76 (all; >60; >100)	95	95
	# unloads	1411	1413	1405
	Manufacturer's cost in O.F. [€]	401240.00	349870.00	272220.00
Direct and Multi Drops Service Modes - yes	Total # route	303	304	291
	average filling index %	39;62;78 (all; >60; >100)	45	34
	# unloads	1411	1413	1405
	Manufacturer's cost in O.F. [€]	645358.84	645787.24	619484.40
Direct and Multi Drops and Groupage Service Modes - Yes	Total # route	153	145	110
	average filling index %	76; 79; >80 (all; >60; >100)	95	90
	# unloads	552	566	461
	Total %group.	39	38	44
	Manufacturer's cost in O.F. [€]	407772.20	394483.39	340012.81

The parameters of the random instance generator are the number of clients and the number of seats. The number of clients is 25 or 50 or 75 or 100, the number of seats is close to 500 or 1000 or 1500 or 2000. Sixteen instances are so defined.

The instances are solved considering three different capacity criteria:

- three dimensional loading, using the integrated model
- one dimensional loading considering the volume of single products;
- one dimensional loading considering the seats of products.

The aim of the tests is to analyze the performance of models changing the service mode and the consolidation opportunity:

- no consolidation with only direct service mode;
- maximum level of consolidation with only groupage service mode;
- internal consolidation with more destinations in the routes and without the provider's constraints and objectives (multi drop service mode);
- internal consolidation with more destinations in the routes and managing the provider's constraints and objectives (multi drop service mode);
- internal and external consolidation giving freedom to model to select the product to delivery using groupage and accounting the provider's constraints and objectives.

In total 240 instances are solved using two Quad Core Intel Xeon E5430 2.66 Ghz processors. All routes defined by 1L-CCVRPCO, with both seats and volume capacity criteria and with all allowed service modes are been separately solved by Loading Heuristics (see table 2). Some tests are conduced in an Italian Furniture Company just using the seats capacity. Nine randomly generated instances are been solved by both experts and 1L-VRPCO using seats capacity criteria (see table 3).

Table 2: Loading Checks on the 16 instances with all allowed service modes and provider's constraints and objectives solved with 1L-CVRPCO

Capacity criteria	# cases	3D Average filling index %	# cases loading check: ok [%]	3D loading Average filling index %	# cases loading check: not ok [%]	3D loading Average filling index %
Vol	110	84	11	49	89	88
Seats	145	79	74	77	26	84

Table 3: a comparison between heuristics and expertise solutions for 9 randomly generated instances

Instances	# Clients	Seats	Improvements
t01	67	2761.4	7%
t02	63	2211.1	2%
t03	14	554.3	7%
t04	36	986.1	4%
t05	99	1141.6	6%
t06	35	274.7	4%
t07	242	2965.6	5%
t08	48	1754.0	7%
t09	82	3970.0	3%

6. CONCLUSIONS

The experiments show how the developed meta-heuristics correctly solves the model.

Allowing only direct service mode in the model, it performs a consolidation between orders of same clients. An upper-bound of the number of routes has been defined but the filling indexes are lower. The filling indexes are calculated considering the relative applied capacity criteria. If a capacity criterion is Full3D, a bounding box volume of placed products is defined and it is divided per available volume of truck. If capacity criterion is seats, the total number of seats of products is divided by the seats capacity of truck. If the capacity criterion is Vol, the sum of volume of each product is divided per volume of truck. In all cases the average filling indexes are lower, a different use of shipment network need to be lower the cost of shipment. Since that the route has only one delivery, the objective of transport provider is achieved. Allowing only groupage service mode, the max level of consolidation could be achieved. In fact a lower number of routes and high level of utilization are achieved. Since that the route has only one delivery, the objective of transport provider is achieved. The cost of shipment is very high and also the delivery lead time is high considering other service modes as direct and multi drop service modes. Allowing direct and multi drop service modes, without to accounts the aims of transport providers, a reasonable delivery cost and high filling indexes are achieved in 1L-CVRPCO model. In 3L-CVRPCO the filling index is lower but this is reasonable considering the effective loading of products. However the index increase seeing only at route with more seats (>60 or >100 in table1). Anyway the routes are not feasible because the aims of transport providers are not respected. This is confirmed by applying the aims of providers to the model: the number of routes increases as the costs of the routes, the filling indexes decreases. Finally good and reasonable solutions are achieved allowing internal and external consolidation using all service modes. The heuristics finds the optimal assignment of orders, routes and service mode.

It's interesting to notice the impact of loading in the solutions. In the tests with all allowed service modes the cost of 3L-CVRPCO solutions increases about 20% compared to only using volume as a criterion and 3% compared to using seats as a criterion. The seat is a no standard measurements unit that experience has suggested in the furniture companies. It looks to be a better measurement than volume of single items: the difference between the 3L-CVRPCO and 1L-CVRPCO(seat) is lower. Besides, seeing at number of routes in the solutions with high consolidation (only groupage service mode) it is equal in both Full3D and seats capacity criteria. This suggests using 1L-CVRPCO to solve the model, without the integration of loading. However, this is not true, since some routes could be infeasible. In the table 2 we report data from the verification process of the routes. Using seats about 26% of routes don't pass the loading verification. Clearly using vol, the percentage is bigger.

Finally the solutions of model are been compared with solutions of companies' experts. Since that the expert cannot check the loading, the comparison is between 1L-CVRPCO(seat) and experts. In all instances there are significant improvements and the previous tests suggest that 3L-CVRPCO warrant good improvements with certainty in loading.

Further developments of this study seem to be promising. An improvement could be made on the working hypothesis about warehouses' costs and delivery times required by customers. Other improvements could be made considering splitting of orders in overfilled routes.

REFERENCES

Cagan J., Degentesh D., Yin S., 1998. A simulated annealing-based algorithm using hierarchical models for general three-dimensional component layout. *Computer Aided Design*, 30(10):781–790.

Dapiran P, Lieb R, Millen R, Sohal A., 1996. Third party logistics services usage by large Australian firms. *International Journal of Physical Distribution & Logistics Management*, 26: 36-45.

de Berg M., Van Kreveld M., Overmarks M., Schvarzkopf O., 2000. *Computational Geometry: Algorithms and applications*. Berlin, Germany: Springer Verlag.

Dickinson J. K.,Knopf G. K., 1998. Serial packing of arbitrary 3d objects for optimizing layered manufacturing. *Intelligent Robots and Computer Vision XVII*, 3522: 130–138.

Dickinson J. K., Knopf G. K., 2002. Packing subsets of 3d parts for layered manufacturing. *International Journal of Smart Engineering System Design*, 4(3):147–161.

Egeblad J., Nielsen B. K., Odgaard A., 2006. Fast neighborhood search for two- and three-dimensional nesting problems. *European Journal of Operational Research,* 183(3): 1249-1266.

Egeblad J., Garavelli A.C., Lisi S., Pisiger D., 2009. Heuristics for Container Loading of Furniture. *European Journal of Operational Research,* forthcoming.

Galbraith JR., 1977. *Organization Design*. Reading, MA: AddisonWesley.

Gendreau M., Iori M., Laporte G., Martello S., 2006. A Tabu Search Algorithm for a Routing and Container Loading Problem. *Transportation science*, 40(3): 342-350.

Gendreau M., Iori M., Laporte G., Martello S., 2008. A Tabu Search Heuristic for the Vehicle Routing Problem with Two-Dimensional Loading Constraints, *Networks*, 51(1): 4-18.

Hertz S, Alfredsson M., 2003. Strategic development of third party logistics provider. *Industrial Marketing Management*, 30: 139-149.

Hesse M, Rodrigue JP., 2004. The Transport Geography of Logistics and Freight Distribution. *Journal of Transport Geography*, 12: 171-184.

Hong J, Chin ATH, Liu B., 2004. Logistics outsourcing by manufacturers in China: a survey of the industry, *Transportation Journal*, 43:17-25.

Ikonen I., Biles W.E., Kumar A., Wissel J. C., Ragade R. K., 1997. A genetic algorithm for packing three-dimensional non-convex objects having cavities and holes. *Proceedings of the 7th International Conference on Genetic Algortithms*, 591–598, July 19-23, 1997, East Lansing, Michigan. Morgan Kaufmann Publishers.

Ikonen I.T., Biles W. E., Lewis James E., Kumar A., Ragade R. K., 1998. Garp: genetic algorithm for part packing in a rapid prototyping machine. *Proceedings of SPIE, Intelligent Systems in Design and Manufacturing*, 3517: 54–62, 1998. Bhaskaran Gopalakrishnan and SanMurugesan, editors.

Iori M., Gonzalez J.J.S., Vigo D., 2006. An exact approach for the vehicle routing problem with two-dimensional loading constraints, *Transportation Science*, 41(2): 253-264.

Lai KH, Cheng TCE. 2004. A Study of the Freight Forwarding Industry. *Hong Kong International Journal of Logistics: Research and Applications*, 7: 72-84.

Lapierre SD, Ruiz AB, Soriano P., 2004. Designing distribution networks: formulation and solution heuristic. *Transportation Science*, 38: 174-187.

Larson PD, Gammelgaard B., 2001. Logistics in Denmark: A Survey of the industry. *International Journal of Logistics: Research and Applications*, 4: 191-206.

Lieb R.C., Bentz B.A., 2004. The Use of Third-Party Logistics Services by Large American Manufacturers: The 2003 Survey. *Transportation Journal*, 43: 24-33.

Moura A., Oliveira J.F., 2004. An integrated approach to the Vehicle Routing and Container Loading Problems. *Proceedings of EURO XX-20th European conference on Operational Research*, July 4-7, 2004, Rhodes, Greece.

Mintzberg H., 1979. *The Structuring of Organizational Structures*. Englewood Cliffs, NJ: Prentice-Hall.

Pisinger D., 2002. Heuristics for the container loading problem. *European Journal of Operations Research*, 3(141):382–392.

Pisinger D, Ropke S., 2007. A genral heuristic for vehicle routing problems. *Computers & operations research* 34: 2403-2435.

Rabinovich E, Knemeyer AM., 2006. Logistics service providers in internet supply chains. *California Management Review*, 48: 84-108.

Tavasszy LA, Ruijgrok CJ, Thissen MJPM, 2003. Emerging global logistics networks: implications for transport systems and policies. *Growth and Change*, 34: 456-472.

THE CONTAINER STACKING PROBLEM: AN ARTIFICIAL INTELLIGENCE PLANNING-BASED APPROACH

Miguel A. Salido[(a)], Oscar Sapena[(b)], Federico Barber[(c)]

[(a)(b)(c)]Instituto de Automática e Informática Industrial, Universidad Politécnica de Valencia, Spain

[(a)]msalido@dsic.upv.es, [(b)]osapena@dsic.upv.es, [(c)]fbarber@dsic.upv.es

ABSTRACT
Large container ports around the world are major hubs in the global cargo transport system. A container stack is a type of temporary store where containers await further transport by truck, train or vessel. The main efficiency problem for an individual stack is to ensure easy access to containers at the expected time of transfer. In this paper, we propose a planning tool for finding the best configuration of containers in a bay. Thus, given a set of outgoing containers, our planning tool minimizes the number of relocations of containers in order to allocate all selected containers in an appropriate order to avoid further reshuffles. Furthermore, we compare the number of reshuffles in yard-bays with 4 tiers against yard-bays with 5 tiers. The obtained results recommend the use of stacks with 5 tiers in high loaded yard-bays, due to the fact that the number of reshuffles is reduced.

Keywords: container-stacking, artificial intelligence, planning

1. INTRODUCTION
Loading and offloading containers on the stack is performed by cranes. In order to access a container which is not at the top of its pile, those above it must be relocated. This reduces the productivity of the cranes.

Maximizing the efficiency of this process leads to several requirements. First, each incoming container should be allocated a place in the stack which should be free and supported at the time of arrival. Second, each outgoing container should be easily accessible, and preferably close to its unloading position, at the time of its departure. In addition, the stability of the stack puts certain limits on, for example, differences in heights in adjacent areas, the placement of empty and 'half' containers and so on.

Since the allocation of positions to containers is currently done more or less manually, this has convinced us that it should be possible to achieve significant improvements of lead times, storage utilization and throughput using improved techniques of the type indicated.

Figure 1 shows a container yard. A yard consists of several blocks, and each block consists of 20-30 yard-bays (Kim, Park and Ryu 2000). Each yard-bay contains several (usually 6) rows. When an outside truck delivers an outbound container to a yard, a transfer crane picks it up and stacks it in a yard-bay. During the ship loading operation, a transfer crane picks up the container and transfers it to a truck that delivers it to a quay crane.

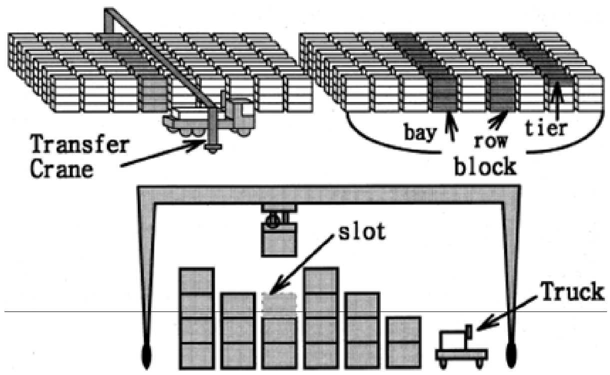


Figure 1: A container yard (Kim, Park and Ryu 2000)

In container terminals, the loading operation for export containers is carefully pre-planned by load planners. For load planning, a containership agent usually transfers a load profile (an outline of a load plan) to terminal operating company several days before a ship's arrival. The load profile specifies only the container group, which is identified by container type (full or empty), port of destination, and size to be stowed in each particular ship cell. Since a ship cell can be filled with any container from its assigned group, the handling effort in the marshalling yard can be made easier by optimally sequencing export containers in the yard for the loading operation. In sequencing the containers, load planners usually pursue two objectives:

- Minimizing the handling effort of quay cranes and yard equipment.
- Ensuring the vessel's stability.

The output of this decision-making is called the "load sequence list". In order to have an efficient load sequence, storage layout of export containers must have a good configuration. The main focus of this paper is optimally reallocating outgoing containers for the final storage layout from which a load planner can construct an efficient load sequence list. In this way, the objective is therefore to plan the movement of the cranes so as to minimize the number of reshuffles of containers.

Given a layout, the user selects the set of containers that will be moved to the vessel. Our tool is able to organize the layout in order to allocate these containers at the top of the stacks in order to minimize the number of relocations. Thus a solution of our problem is a layout where all outgoing containers can be available without carrying out any reshuffle. Furthermore, due to harbor operator requirements, we are interesting on comparing the number of reshuffles in yard-bays with 4 tiers against yard-bays with 5 tiers.

2. THE PROBLEM MODELLED AS AN ARTIFICIAL INTELLIGENCE PLANNING PROBLEM

A classical AI planning problem can defined by a tuple <*A*, *I*, *G*>, where *A* is a set of actions with preconditions and effects, *I* is the set of propositions in the initial state, and *G* is a set of propositions that hold true in any goal state. A solution plan to a problem in this form is a sequence of actions chosen from *A* that when applied transform the initial state *I* into a state of which *G* is a subset.

The container stacking problem is a slight modification of the *Blocks World* planning domain (Slaney and Thiébaux 2001), which is a well-known domain in the planning community. This domain consists of a finite number of blocks stacked into towers on a table large enough to hold them all. The positioning of the towers on the table is irrelevant. The *Blocks World* planning problem is to turn an initial state of the blocks into a goal state, by moving one block at a time from the top of a tower onto another tower (or on a table). The optimal *Blocks World* planning problem is to do so in a minimal number of moves.

This problem is closed to the container stacking problem, but there are some important differences:

- The number of towers is limited in the container stacking problem: a yard-bay contains usually 6 rows, so it is necessary to include an additional constraint to limit the number of towers on the table to 6.
- The height of a tower is also limited. In this paper we analyze the effect of limiting the number of levels in a tower on the number of required relocations to reach the goal configuration.
- The main difference is in the problem goal specification. In the *Blocks World* domain the goal is to get the blocks arranged in a certain layout, specifying the final position of each block. In the container stacking problem the goal state is not defined as accurately, so many different layouts can be a solution for a problem. The goal is that the most immediate containers to load are in the top of the towers, without indicating which containers must be in each tower.

We can model our problem by using the standard encoding language for classical planning tasks called *PDDL, Planning Domain Definition Language* (Ghallab et al. 1998). Following this standard, a planning task is defined by means of two text files: the domain file, which contains the common features for all problems of this type, and the problem file, which describes the particular characteristics of each problem. The contents of both files are described in the following subsections with more detail.

2.1. The container stacking domain

The main elements in a domain specification are (1) the types of objects we need to handle, (2) the types of propositions we use to describe the world and (3) the actions we can perform to modify the state of the world. In the container stacking domain we have the following elements: *object types*, *propositions* and *actions*.

Object types. In this domain we only need to define two object types: *containers* and *rows*, where the rows represent the areas in a yard-bay in which a tower of containers can be built.

Types of propositions. We need to define the following types of propositions:

- *on ?x - container ?y - (either row container)*
  This predicate indicates that container *?x* is on *?y*, which can be another container or, directly, the floor of a row (stack).
- *at ?x - container ?r – row*
  This indicates that the container *?x* is in the tower built on the row *?r*.
- *clear ?x - (either row container)*
  This predicate states that *?x*, which can be a row or a container, is clear, that is, there are no containers stacked on it.
- *crane-empty*
  This indicates that the crane used to move the containers is not holding any container.
- *holding ?x – container*
  This states that the crane is holding the container *?x*.
- *goal-container ?x – container*
  ready ?x - container
  These predicates are used to describe the problem goal. The first one specifies the most immediate containers to load, which must be located on the top of the towers to facilitate the ship loading operation. The second one becomes true when this goal is achieved for the given container.
- *height ?s – row*
  num-moves
  These are numerical predicates. The first one stores the number of containers stacked on a given row and the second one counts the

number of container movements carried out in the plan.

Actions. In this domain there are four different actions to move the containers from a row to another:

- *pick (?x - container ?r - row)*
With this action the crane picks the container *?x* which is in the floor of row *?r*.
- *put (?x - container ?r - row)*
The crane puts the container *?x*, which is holding, in the floor of row *?r*.
- *unstack (?x - container ?y - container ?r - row)*
With this action the crane unstacks the container *?x*, which is in row *?r*, from the container *?y*.
- *stack (?x - container ?y - container ?r - row)*
The crane stacks the container *?x*, which is currently holding, on container *?y* in the row *?r*.

As an example, we show in Figure 2 the specification of the stack operator in *PDDL* format. Preconditions describe the conditions that must hold to apply the action: crane must be holding container *?x*, container *?y* must be clear and at row *?r*, and the number of containers in that row must be less than 4. With this constraint we limit the height of the piles. The effects describe the changes in the world after the execution of the action: container *?x* becomes clear and stacked on *?y* at row *?r*, and the crane is not holding any container. Container *?y* becomes not clear and the number of movements and the containers in *?r* is increased in one unit.

```
(:action stack
 :parameters (?x - container
              ?y - container
              ?r - row)
 :precondition (and (holding ?x)
    (clear ?y) (at ?y ?r)
    (< (height ?r) 4))
 :effect (and
    (clear ?x) (on ?x ?y)
    (at ?x ?r) (crane-empty)
    (not (holding ?x))
    (not (ready ?y))
    (not (clear ?y))
    (increase (num-moves) 1)
    (increase (height ?r) 1)))
```

Figure 2: Formalization of the *stack* operator in *PDDL*.

Finally, we have defined two fictitious actions that allow checking whether a given (goal) container is ready, that is, it is in a valid position:

- The container is clear, or
- The container is under another (goal) container which is in a valid position.

2.2. A container stacking problem

A container stacking problem file contains the elements which are specific to the particular problem. These elements are:

- **Objects**: the rows available in the yard-bay and the containers stored in them.
- **Initial state**: the initial layout of the containers in the yard.
- **The goal specification**: the selected containers
- to be allocated at the top of the stacks or under other selected containers.
- **The metric function**: the function to optimize. In our case, we want to minimize the number of relocation movements (reshuffles).

3. DOMAIN-INDEPENDENT PLANNING FOR SOLVING THE CONTAINER STACKING PROBLEM

Since the container stacking problem can be formalized in *PDDL* format, as we have shown in the previous section, we can use a general planner to solve our problem instances. Currently we can found several general planners which work well in many different domains, such as *LPG-TD* (Gerevini, Saetti and Serina 2003), *MIPS-XXL* (Edelkamp 2003) and *SGPlan* (Chen, Hsu and Wah 2004). However, and due to the high complexity of the domain we are handling, these planners are not able to find good plan solutions efficiently. *LPG-TD*, for example, spends too much time in the preprocessing stages, so it takes a long time to provide a solution. On the contrary, *MIPS-XXL* and *SGPlan* can compute a solution rapidly, but the quality of the obtained solution is not good enough, including some additional relocation movements to achieve the goal configuration.

In order to solve this problem efficiently, we have developed a new general planning algorithm with several interesting properties for the container stacking problem:

- It is an anytime planning algorithm (Zilberstein and Russell 1996). This means that the planner can found a first, probably suboptimal, solution quite rapidly and that this solution is being improved while time is available.
- The planner is complete, so it will always find a solution if exists.
- The planner is optimal. It guarantees finding the optimal plan if there is time enough for computation.
- The main bottleneck while solving the container stacking problem is the large number of local minima (or plateaux) found during the search. We have developed a new search method which combines the use of two heuristic functions. This feature allows the planner to escape from a local minimum very efficiently.

The planning approach is a combination of an *Enforced Hill-Climbing* (Hoffman and Nebel 2001), which allows to find fast solutions, with a standard A search, which guarantees finding the optimal plan, that is, the plan that minimizes the number of reshuffles.

The plan is returned by the planner as a totally ordered sequence of actions. Figure 3 shows the obtained plan for a given problem. This plan shows the movements the transfer crane must carry out to achieve our objective.

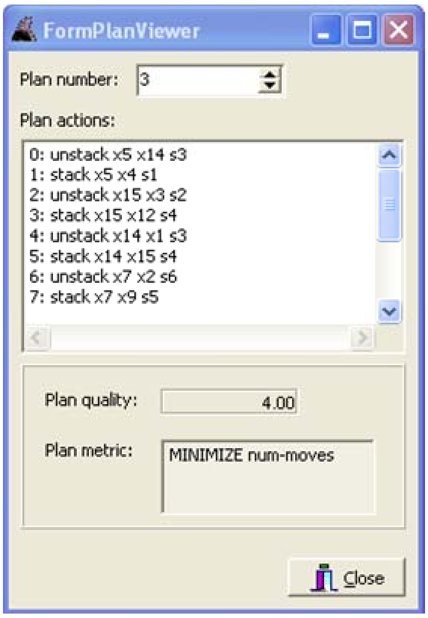


Figure 3: Plan to be carried out by the transfer crane.

4. EVALUATION

To evaluate our tool, we have analyzed two configurations of yards: with 4 tiers and with 5 tiers. We have evaluated the minimum number of reshuffles needed to allocate all selected containers at the top of the stacks or under another selected containers in such a way that no reshuffles is needed to load outgoing containers. Thus, the experiments were performed on random instances. A random instance is characterized by the tuple <*n*, *s*>, where *n* is the number of containers and *s* is the number of selected containers. Each instance is a random configuration of all containers distributed along the six stacks with 4 or 5 tiers. We evaluated 100 test cases for each type of problem.

In Figure 4 we evaluated the number of reshuffles needed for problems <*n*, 4>. Thus, we fixed the number of selected containers to 4 and we increased the number of containers *n* from 11 to 23. It can be observed that as the number of containers increased, the number of reshuffles increased. For small number of containers (low values of *n*) there is no difference between 4 tiers and 5 tiers. This is due to the fact that it is not needed the use of the higher stacks to achieve a solution because there exist many combinations to achieve a solution. However, the number of reshuffles with 5 tiers was lower than the number of reshuffles with 4 tiers for higher number of containers. Due to the fact that we consider yard-bays of 6 stacks, for problems with 4 tiers the maximum number containers is bounded to 24. In this case instances <23, 4> for problems with 4 tiers generally has no solution. Thus, we can conclude that for low loaded yard-bays (< 15 containers) there is not different between 4 tiers and 5 tiers, meanwhile for high loaded yard-bays 5 tiers is more appropriate for minimizing the number of reshuffles.

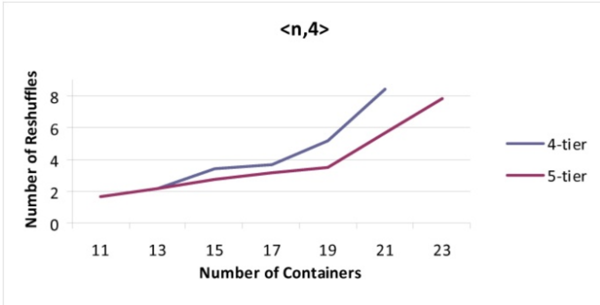


Figure 4: Number of reshuffles for problems <*n*, 4>.

In Figure 5 we evaluated the number of reshuffles needed for problems <19, *s*>. To this end, we fixed the number of containers to 19 and we increased the number of selected containers *s* from 2 to 6. The figure shows that as the number of selected containers increases, the number of reshuffles also increased. It can be also observed that the number of reshuffles for tiers 5 was lower than the number of reshuffles for tiers 4 in all cases. Furthermore, as the number of selected containers increases, the difference of reshuffles with 4 and 5 tiers became higher, so we can conclude that configurations of yards with 5 tiers is more appropriate to minimize the number of reshuffles with this configuration.

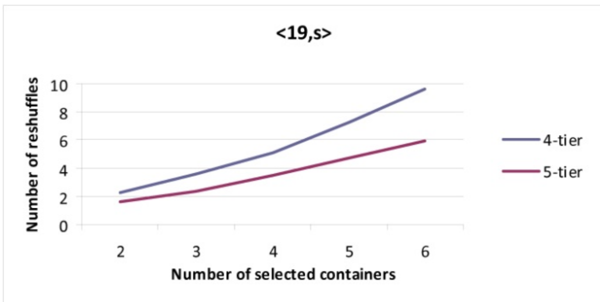


Figure 5: Number of reshuffles for problems <19, *s*>.

5. CONCLUSIONS AND FURTHER WORKS

This paper presents the modeling of the container stacking problem form the Artificial Intelligence point of view. We have developed a domain-independent planning tool for finding an appropriate configuration of containers in a bay. Thus, given a set of outgoing containers, our planner minimizes the number of necessary reshuffles of containers in order to allocate all selected containers at the top of the stacks or under another selected containers in such a way that no reshuffles is needed to load these outgoing containers.

Furthermore, we compare the number of reshuffles in yard-bays with 4 tiers against yard-bays with 5 tiers. The obtained results recommend the use of stack with 5 tiers for high loaded yard-bays, due to the fact that the number of reshuffles is reduced for outgoing containers.

In further works, we will focus our attention in the development of domain-dependent planning heuristic to include new hard and soft constraints for solving this problem.

ACKNOWLEDGMENTS

This work has been partially supported by the research projects TIN2007-67943-C02-01 (Min. de Educacion y Ciencia, Spain-FEDER) , P19/08 (Min. de Fomento, Spain-FEDER) and by the Technical University of Valencia.

REFERENCES

Chen, Y., Hsu, C.W., Wah, B.W., 2004. SGPlan: Subgoal Partitioning and Resolution in Planning. *IPC-4 Booklet (ICAPS)*.

Edelkamp, S., 2003. Taming Numbers and Durations in the Model Checking Integrated Planning System. *Journal of Artificial Intelligence Research (JAIR)*, 20, 195–238.

Gerevini, A., Saetti, A., Serina, I., 2003. Planning Through Stochastic Local Search and Temporal Action Graphs in LPG. *Journal of Artificial Intelligence Research (JAIR)*, 20, 239–290.

Ghallab, M., Howe, A., Knoblock, C., McDermott, D., Ram, A., Veloso, M., Weld, D., Wilkins, D., 1998. PDDL - The Planning Domain Definition Language. *AIPS-98 Planning Committee*.

Hoffman, J., Nebel, B., 2001. The FF Planning System: Fast Planning Generation Through Heuristic Search. *Journal of Artificial Intelligence Research*, 14, 253–302.

Kim, K.H., Park, Y.M., Ryu, K.R., 2000. Deriving decision rules to locate export containers in container yards. *European Journal of Operational Research*, 124, 89–101.

Slaney, J., Thiébaux, S., 2001. Blocks World revisited. *Artificial Intelligence*, 125, 119–153.

Zilberstein, S., Russell, S.J., 1996. Optimal Composition of Real-time Systems. *Artificial Intelligence*, 82(1-2), 181–213.

AUTHORS BIOGRAPHY

Miguel A. Salido is an associate professor in Computer Science at the Technical University of Valencia, Spain. Most of his research is focused on techniques for constraint satisfaction techniques and railway scheduling problems. He is the recipient of some national and international awards. He is author of more than 70 papers published on international journals and conferences. He is editor of some books and guess editor of some international journals. He is member of several Scientific and Organizing Committees.

Oscar Sapena is an associate professor in Computer Science at the Technical University of Valencia, Spain. His research is focused on Planning form the Artificial Intelligence point of view. He is author of more than 20 papers published on international journals and conferences. He is member of several Scientific and Organizing Committees.

Federico Barber is Full Professor with the Department of Computer Science, where he leads a research team in artificial intelligence. He has worked on the development of temporal reasoning systems, Constraint Satisfaction Problems, planning and scheduling. He is the author of about 90 research articles which have been published in several journals and conferences. His research has produced several tools for solving real-world optimization combinatory problems. He has participated in and led several national and European research projects related to these areas. He is currently president of the Spanish Association for Artificial Intelligence, and member of several scientific associations.

INTEGRATING EMPTY CONTAINER ALLOCATION WITH VEHICLE ROUTING IN INTERMODAL TRANSPORT

Kris Braekers[(a)], Gerrit K. Janssens[(b)], An Caris[(c)]

[(a)(b)(c)] Transportation Research Institute (IMOB) – Hasselt University
Universiteit Hasselt, campus Diepenbeek,
Wetenschapspark – gebouw 5, 3590 Diepenbeek, Belgium

[(a)]kris.braekers@uhasselt.be, [(b)]gerrit.janssens@uhasselt.be, [(c)]an.caris@uhasselt.be

ABSTRACT

In an intermodal transportation network, empty containers need to be repositioned in order to be able to fulfil empty container demands. At a regional level this repositioning takes place between importers, exporters, intermodal terminals, inland depots and ports. Repositioning movements with lowest costs may be determined by an empty container allocation model. Vehicle routing models may be used to find vehicle routes for performing loaded containers transports and container allocations determined by the container allocation model. Recently, approaches to integrate empty container allocations and vehicle routing have been discussed in literature. This paper shows the advantage of an integrated approach for the full-truckload problem with a vehicle capacity of a single container. Mathematical formulations for the container allocation and vehicle routing model are proposed. A model integrating the separate models is presented and numerical experiments are performed. Results show that the integrated model results in the lowest costs.

Keywords: empty container allocation, vehicle routing, pre- and post-haulage, intermodal freight transport

1. INTRODUCTION

Ever since the introduction of containers, containerization of freight transport, especially in international maritime shipping, is rising. The use of containers for freight transportation leads to a number of planning problems, such as fleet sizing and management, decisions about container ownership or leasing and repositioning needs (Dejax and Crainic 1987). This paper considers the last aspect, repositioning needs, in an intermodal transportation network consisting of maritime main haulage and pre- and post-haulage over land by truck, rail or barge transport.

Due to the natural imbalance of trade, certain areas in the network develop a surplus of containers while others have a deficit. As a consequence, there is a need for carriers to reposition their empty containers in order to be able to fulfil future demand for empty containers. Although it is a costly and non-revenue generating activity, empty container repositioning is an integral part of an overall efficient transportation system. In an intermodal transportation network, empty container repositioning takes place at a global level as well as at a regional level. At a global level, empty containers are repositioned between ports. At a regional level empty containers are repositioned between importers, exporters, intermodal terminals, inland depots and one or more ports within a relatively small geographical area, namely the hinterland of ports. (Boile, Theofanis, Baveja, and Mittal 2008)

This paper focuses on empty container repositioning at a regional level. Currently, empty containers are often immediately transported back to a port. Several improvements to current practice are proposed in literature. Empty containers may be transported to inland depots for fulfilling future empty container requests in the hinterland. Street turns, transporting empty containers directly from an importer to an exporter, can reduce empty movements dramatically. Another option is to allow container substitutions (fulfilling the request for a certain type of container by supplying another type of container). Finally, container leasing may be used to reduce repositioning needs.

When addressing empty container repositioning at a regional level, several decisions have to be taken. These decisions belong to several planning levels: strategic, tactic and operational. In this research, operational decisions are considered. Optimization at this level means making sure that demand for empty containers is satisfied everywhere and that the most effective routes and transport modes are chosen. The underlying intermodal transportation network is assumed to be fixed. To account for interactions between the different decisions to be made, Crainic, Gendreau and Dejax (1993) note that ideally a single mathematical model should be developed. Considering the available Operations Research techniques at that time, the authors state that developing such a model is not feasible due to the complexity of the problem. Therefore, the operational planning problem is divided into two separate optimization problems, namely a container allocation and a vehicle routing problem.

Recently, some authors have proposed approaches to integrate both models. These approaches will be discussed in section 2.

The objective of this paper is to investigate the integration of container allocation and vehicle routing models. Mathematical models are formulated to show the benefit of an integrated approach. In section 2 of this paper a literature review concerning container allocation and vehicle routing models is given. Integration approaches proposed in literature are discussed. Section 3 contains the model formulations. Numerical results are presented in section 4. Finally, in section 5 conclusions are drawn and future research prospects are identified.

2. LITERATURE REVIEW

This section gives an overview of container allocation and vehicle routing models proposed in literature. Integration approaches are discussed.

2.1. Container Allocation Models

The objective of container allocation models is to determine the best distribution of empty containers, while satisfying both known and forecasted demand. (Crainic, Gendreau, and Dejax 1993) Empty container demand and supply should be taken into account. Besides, repositioning empty containers in order to be able to satisfy empty container requests in future periods should be considered. The most realistic model would be a stochastic, dynamic, multi-commodity model including container substitution, street turns and interdepot movements. Formulating and solving such an elaborate model is a challenging task. To our knowledge no such model is described in literature.

As mentioned in the introduction, this paper focuses on the repositioning of empty containers at a regional level. Therefore, only container allocation models related to this problem are discussed. For container allocation models concerning global or maritime repositioning of empty containers, often using simulation, the reader is referred to the corresponding literature. (Lai, Lam, and Chan 1995; Cheung and Chen 1998; Li, Leung, Wu, and Liu 2007; Lam, Lee, and Tang 2007; Di Francesco, Crainic, and Zuddas 2009; Dong and Song 2009)

A general framework for the regional allocation of empty containers is offered by Crainic, Gendreau and Dejax (1993). The authors describe a dynamic deterministic model for both the single and multi-commodity case. A stochastic model for the single commodity case is also formulated. In a subsequent work, Abrache, Crainic and Gendreau (1999) discuss a decomposition algorithm for the deterministic multi-commodity model formulated by Crainic, Gendreau and Dejax (1993).

Other mathematical models for empty container allocation are proposed by Chu (1995). Firstly, a single and a multicommodity dynamic deterministic model are described. Secondly, a dynamic two-stage and multi-stage stochastic model for the single commodity case are formulated.

Olivo, Zuddas, Di Francesco and Manca (2005) develop an operational model for empty container management on a continental or interregional level by formulating it as a minimum cost flow problem. The model is dynamic and deterministic. Di Francesco, Manca and Zuddas (2006) take a similar modelling approach for the empty container allocation problem at a regional level. In a subsequent work, Di Francesco (2007) introduces the opportunity of short-term leasing into the model of Di Francesco, Manca and Zuddas (2006).

Jula, Chassiakos and Ioannou (2003) propose a static and a dynamic deterministic model with and without street turns and inland depots. Results show that when street turns are allowed and inland depots are used, costs drop significantly. Chang, Jula, Chassiakos and Ioannou (2006) introduce container substitution into the models of Jula, Chassiakos and Ioannou (2003). Next, the authors propose a model for the stochastic static single commodity problem, without container substitution.

A real-life application is discussed by Jansen, Swinkels, Teeuwen et al. (2004). The authors describe an operational planning system for the German company Danzas Euronet. Repositioning of empty containers is modelled as a minimum cost flow problem.

2.2. Vehicle Routing Models

Vehicle routing models aim to minimize overall transportation costs of both loaded and empty containers. The result of such a model is a set of vehicle routes which completely describe the loaded and empty movements to be executed during the next period. (Crainic, Gendreau, and Dejax 1993)

Literature on vehicle routing is extensive. Several sorts of problems exist and nomenclature is not always used in the same way. In this paper the classification of Parragh, Doerner and Hartl (2008) is followed. The authors distinguish two problem classes. The first class, called Vehicle Routing Problems with Backhauls (VRPB), is concerned with the transportation of goods from depots to linehaul customers and from backhaul customers to depots. The second class comprises models for the transportation of goods among customers. This class is denoted as Vehicle Routing Problems with Pickups and Deliveries (VRPPD). This research focuses on the second class of models since street turns cannot be considered by models of the first class. More precisely, this paper focuses on the classical Pickup and Delivery Problem (PDP), a subclass of VRPPD. With this type of problem, goods have to be transported between paired pickup and delivery locations. (Parragh, Doerner, and Hartl 2008)

A distinctive characteristic of this research is that full-truckload transportation, instead of the more extensively studied less-than-truckload problem, is considered. All vehicles are homogenous and have a capacity of a single container. This means that the pickup and delivery activity of the same request should be performed immediately after each other. As a consequence, when all transportation requests are known, the problem can be modelled as an asymmetric Multi-Travelling Salesman Problem (m-TSP). (Mitrovic-Minic 1998; Ioannou, Chassiakos, Jula, and Unglaub 2002). However, when considering the integration of container allocation and vehicle routing, not all requests are known in advance. This increases integration complexity seriously. Therefore, in this paper the routing problem is formulated as a full-truckload Pickup and Delivery Problem. Similar problems are considered in literature. (Savelsbergh and Sol 1995; Mitrovic-Minic 1998; Cordeau, Laporte, Potvin, and Savelsbergh 2007; Huth and Mattfeld 2009)

2.3. Integration Approaches

Dejax and Crainic (1987) already suggested that the independent consideration of container allocation and vehicle routing neglects possible synergies arising from an integrated view on these problems. However, Crainic, Gendreau and Dejax (1993) stated that a single mathematical model comprising container allocations and vehicle routing would be computationally intractable.

Erera, Morales and Savelsbergh (2005) are of the opinion that, with current Operations Research techniques, a single model optimizing the operational planning is feasible. To verify this statement, the authors propose a deterministic multi-commodity model integrating routing and repositioning decisions for tank container operators. It is shown that the model may be solved by commercially available software for real-life cases. (Erera, Morales, and Savelsbergh 2005)

Huth and Mattfeld (2009) study the integration of container allocation and vehicle routing for the swap container problem (SCP). The swap container problem considers routing loaded swap containers and allocating and routing empty swap containers between hubs in a hub-to-hub transportation network. The problem is very similar to the problem considered in this paper. The main difference is that Huth and Mattfeld (2009) assume a truck capacity of two containers, while in this paper truck capacity is assumed to be a single container. Furthermore, the authors consider allocation and routing between hubs while here it is considered between depots, terminals and individual importers and exporters.

Three modelling approaches for the swap container problem are distinguished by Huth and Mattfeld (2009). First, with sequential planning (SP), no integration takes place. Empty container allocations are modelled by an allocation model and then inserted into a routing model. Loaded container transport requests are routed separately. The other two approaches represent different levels of integration and are based on the integration approaches of Geoffrion (1989, 1999). Functional integration (FI) combines given models through a coordination mechanism. Empty containers allocations are modelled by an allocation model. Next, empty container allocations and loaded container transport requests are routed together. With the deep integration approach (DI), existing models are combined into a new model. Empty container allocations are modelled simultaneously with loaded and empty container routing. (Huth and Mattfeld 2009)

The authors propose formulations for all three approaches. Requests are represented by single containers and truck capacity is assumed to be two containers. It is shown that deep integration provides the best results, then functional integration and finally sequential planning. Huth and Mattfeld (2009) attribute the greater part of the advantage of integration to the opportunity of detouring and entrainment in the routing model (when transporting a single container, making a detour to include the transportation of another (empty) container may save costs).

Because in this paper, a truck capacity of a single container is assumed, detouring and entrainment are not feasible. Our objective is therefore to investigate whether in this case functional and deep integration still provide better results than sequential planning.

3. MODEL FORMULATION

In this section a basic container allocation and vehicle routing problem are formulated. Next, an integrated model is proposed. The objective is to illustrate the benefit of an integrated approach. Future research will focus on the extension of the models.

3.1. Container Allocation Model

The empty container allocation model proposed in this paper is a single commodity deterministic static model. Street turns are allowed. Only a single period is considered and thus no repositioning movements in order to be able to fulfil future requests are considered. The formulation is based on the static street turn depot-direct model of Jula, Chassiakos and Ioannou (2003). Parameter indices are slightly adapted to facilitate integration with the routing model and the constraints prohibiting interdepot movements are left out.

The network consists of consignees, shippers and depots. Consignees supply and shippers demand empty containers. Depots represent inland intermodal terminals. They may supply empty containers to shippers and may receive empty containers from consignees. For simplicity, it is assumed that depots have a sufficient stock of empty containers to fulfil all demands and do not request any empty containers themselves. Transportation costs are assumed to be proportional to transportation distances. Finally, the triangle inequality holds for the whole network. This means that a direct transport route between two nodes is at least as short/cheap as a route via an intermediate node. The following notation is used:

$C =$ set of n consignees (index $1,\ldots,n$)
$S =$ set of m shippers (index $n+1,\ldots,n+m$)
$D =$ set of p depots (index $n+m+1,\ldots,n+m+p$)
$I = C \cup S \cup D =$ set of nodes
$i, j \in I$
$(i, j = 1,\ldots,n, n+1,\ldots,n+m, n+m+1,\ldots,n+m+p)$
$c_{ij} =$ cost per unit of transport from node i to node j
$d_j =$ empty container demand at j $(j = n+1,\ldots,n+m)$
$s_i =$ empty container supply at i $(i = 1,\ldots,n)$
$x_{ij} =$ number of empty containers transported from i to j

The model may be formulated as follows:

$$Min \sum_{i=1}^{n+m+p} \sum_{j=1}^{n+m+p} c_{ij} x_{ij} \qquad (1)$$

Subject to

$$\sum_{i=1}^{n} x_{ij} + \sum_{i=n+m+1}^{n+m+p} x_{ij} = d_j \qquad \forall j = \{n+1,\ldots,n+m\} \qquad (2)$$

$$\sum_{j=n+1}^{n+m+p} x_{ij} = s_i \qquad \forall i = \{1,\ldots,n\} \qquad (3)$$

$$x_{ij} \geq 0 \text{ and integer} \qquad \forall i, \forall j \qquad (4)$$

The objective function (1) minimizes variable costs related to distance travelled. Fixed vehicle costs are not considered. Constraint (2) makes sure demand of every shipper is satisfied by empty container coming from consignees and depots. Constraint (3) ensures that all empty containers supplied by consignees are allocated to be transported either to a shipper or to a depot. Finally, decisions variables are restricted to non-negative integer values by constraint (4).

3.2. Vehicle Routing Model

In this section the vehicle routing model is presented. All transportation requests, for both loaded and empty containers, are assumed to be known in advance. Multiple homogenous vehicles, initially located at a single depot, are considered. Vehicle capacity is a single container. Therefore, an arc-based as well as a node-based formulation may be used. Because integration with the allocation model is intuitively simpler for an arc-based model, this type of formulation is chosen. A disadvantage of this type of formulation is that each vehicle can visit each node at most once. Therefore, Huth and Mattfeld (2009) propose to introduce dummy nodes at the same location when multiple requests at a node exist.

The network underlying the vehicle routing model is very similar to the one proposed for the container allocation model. Only a single node, the vehicle depot, is added and the decision variables are now restricted to binary values. The notation is as follows:

$C =$ set of n consignees (index $1,\ldots,n$)
$S =$ set of m shippers (index $n+1,\ldots,n+m$)
$D =$ set of p depots (index $n+m+1,\ldots,n+m+p$)
$V =$ vehicle depot v (index 0)
$K =$ set of q vehicles
$k \in K$ $(k = 1,\ldots,q)$
$I = C \cup S \cup D \cup V =$ set of nodes
$i, j, l \in I$
$(i, j, l = 0,1,\ldots,n, n+1,\ldots,n+m, n+m+1,\ldots,n+m+p)$
$R =$ set of z transport requests (loaded and empty)
$r_{ij} =$ number of requests from node i to node j

$$z = \sum_{i=0}^{n+m+p} \sum_{j=0}^{n+m+p} r_{ij}$$

$c_{ij} =$ cost per unit of transport from node i to node j
$t_{ij} =$ transport time from node i to node j
$FC =$ fixed cost per truck used

$$y_{ij}^k = \begin{cases} 1 & \text{if vehicle } k \text{ travels from node } i \text{ to node } j \\ 0 & \text{otherwise} \end{cases}$$

$T_i^k =$ time epoch that vehicle k leaves node i
$T_{\max} =$ maximum tour duration

The vehicle routing model may be formulated as follows:

$$Min \sum_{k=1}^{q} \sum_{i=0}^{n+m+p} \sum_{j=0}^{n+m+p} c_{ij} y_{ij}^k + \sum_{k=1}^{q} \sum_{j=1}^{n+m+p} y_{0j}^k FC \qquad (5)$$

Subject to

$$\sum_{k=1}^{q} y_{ij}^k \geq r_{ij} \qquad \forall i, \forall j \qquad (6)$$

$$\sum_{i=0}^{n+m+p} y_{il}^k = \sum_{j=0}^{n+m+p} y_{lj}^k \qquad \forall k, \forall l \ (i \neq l, j \neq l) \qquad (7)$$

$$\sum_{j=0}^{n+m+p} y_{0j}^k = 1 \qquad \forall k \qquad (8)$$

$$\sum_{i=0}^{n+m+p} y_{i0}^k = 1 \qquad \forall k \qquad (9)$$

$$\sum_{j=0}^{n+m+p} y_{ij}^k \leq 1 \qquad \forall i, \forall k \qquad (10)$$

$$\sum_{i=0}^{n+m+p} y_{ij}^k \leq 1 \qquad \forall j, \forall k \qquad (11)$$

$$y_{ij}^k (T_i^k + t_{ij} - T_j^k) \leq 0 \qquad \forall i, \forall j \neq 0, \forall k \qquad (12)$$

$$T_i^k + y_{i0}^k t_{i0} \leq T_{\max} \qquad \forall k, \forall i \qquad (13)$$

$$T_0^k = 0 \qquad \forall k \qquad (14)$$

$$y_{ij}^k \text{ binary} \qquad \forall i, \forall j, \forall k \qquad (15)$$

This formulation assumes a truck capacity of a single container. The objective of the model is to minimize both fleet size and variable transportation costs. A large fixed cost per truck used is introduced to first minimize fleet size. Next, variable transportation costs are minimized (5). Constraint (6) ensures that at least as many vehicles travel between two nodes as there are requests between these nodes. No equality sign is used since vehicle are allowed to travel between two nodes for other purposes than fulfilling a request. Constraint (7) verifies that each vehicle entering a node also leaves that node. A vehicle should leave and enter the vehicle depot exactly once (constraints (8) and (9)). Each other node can be visited at most once by the same vehicle (constraints (10) and (11)). When a vehicle is not used, it stays at the vehicle depot , at a cost of zero ($y_{00}^k = 1$). Constraint (12) makes sure that when a vehicle leaves a node at a certain time, it cannot leave the following node after this time augmented with the travel time between the nodes. The objective of this constraint is to prevent loops in the tours and to keep track of tour duration (Parragh, Doerner, and Hartl 2008). A maximum tour duration is imposed by constraint (13). This maximum tour duration may represent a maximum working shift duration for the drivers. Constraint (14) sets the starting time of each vehicle at the vehicle depot to zero. Finally, constraint (15) makes sure that the decision variables only take on binary values.

Before the model can be solved efficiently, constraint (12) has to be linearized. This may be done as represented by constraint (16), with M a big number (for example the maximum tour duration).

$$(M + t_{ij}) y_{ij}^k + T_i^k - T_j^k \le M \qquad \forall i, \forall j \ne 0, \forall \mathrm{k} \qquad (16)$$

3.3. Integrated Model

For the integrated model the same notation as for the vehicle routing model is used, except the following modifications and additions:

$R =$ set of z *loaded* container transport requests

$r_{ij} =$ number of *loaded* container transport requests from i to j

$d_j =$ demand at node $j \quad (j = n+1, \ldots, n+m)$

$s_i =$ supply at node $i \quad (i = 1, \ldots, n)$

$x_{ij} =$ number of empty containers transported from i to j

Only loaded container transport requests are now known in advance. Empty container allocations are modelled by the integrated model, together with loaded and empty container routing. The empty container allocations made by the integrated model are shown by decisions variables x_{ij} .

The formulation of the integrated model is very similar to the one of the vehicle routing model. Constraint (6) is replaced by constraint (17) which ensures that at least as many vehicles travel between two nodes as the sum of loaded container transport requests and empty container allocations. Constraints (2-4) from the container allocation model are added. This results in the following formulation:

$$Min \sum_{k=1}^{q} \sum_{i=0}^{n+m+p} \sum_{j=0}^{n+m+p} c_{ij} y_{ij}^k + \sum_{k=1}^{q} \sum_{j=1}^{n+m+p} y_{0j}^k FC \qquad (5)$$

Subject to

$$\sum_{k=1}^{q} y_{ij}^k \ge r_{ij} + x_{ij} \qquad \forall i, \forall j \qquad (17)$$

$$\sum_{i=0}^{n+m+p} y_{il}^k = \sum_{j=0}^{n+m+p} y_{lj}^k \qquad \forall k, \forall l \ (i \ne l, j \ne l) \quad (7)$$

$$\sum_{j=0}^{n+m+p} y_{0j}^k = 1 \qquad \forall k \qquad (8)$$

$$\sum_{i=0}^{n+m+p} y_{i0}^k = 1 \qquad \forall k \qquad (9)$$

$$\sum_{j=0}^{n+m+p} y_{ij}^k \le 1 \qquad \forall i, \forall k \qquad (10)$$

$$\sum_{i=0}^{n+m+p} y_{ij}^k \le 1 \qquad \forall j, \forall k \qquad (11)$$

$$(M + t_{ij}) y_{ij}^k + T_i^k - T_j^k \le M \qquad \forall i, \forall j \ne 0, \forall \mathrm{k} \qquad (16)$$

$$T_i^k + y_{i0}^k t_{i0} \le T_{\max} \qquad \forall k, \forall i \qquad (13)$$

$$T_0^k = 0 \qquad \forall k \qquad (14)$$

$$\sum_{i=1}^{n} x_{ij} + \sum_{i=n+m+1}^{n+m+p} x_{ij} = d_j \qquad \forall j = \{n+1, \ldots, n+m\} \quad (2)$$

$$\sum_{j=n+1}^{n+m+p} x_{ij} = s_i \qquad \forall i = \{1, \ldots, n\} \qquad (3)$$

$$y_{ij}^k \text{ binary} \qquad \forall i, \forall j, \forall k \qquad (15)$$

$$x_{ij} \ge 0 \text{ and integer} \qquad \forall i, \forall j \qquad (4)$$

4. COMPUTATIONAL RESULTS

The models formulated in the previous section are used to perform computational experiments on a small network. This network represents the pre- and post-haulage part of a larger maritime intermodal network and consists of nine nodes: three consignees (nodes 1, 2 and 3), three shippers (nodes 4, 5 and 6), two depots (nodes 7 and 8) and one vehicle depot (node 0). The two depots may represent intermodal terminals or container depots. The vehicle depot is assumed to be located near one of the depots, although this is no requirement. A graphic representation of the network is shown in figure 1. Maximum tour duration is set at 200. Dummy nodes are created when necessary. All models are solved with Lingo 10.0.

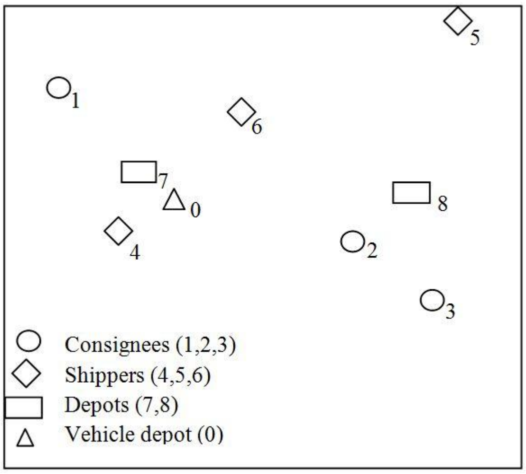


Figure 1: Network Representation

In total, ten problem instances are generated. The number of loaded container transport requests and empty containers supplied and demanded is kept relatively small to ensure limited computation times. The number of empty containers supplied and demanded by each consignee and shipper respectively, are generated randomly with a probability of 0.25 to be zero, 0.5 to be one and 0.25 to be two. Besides, six loaded container transport requests per problem instance are randomly generated. Because this paper focuses on the land transportation part of intermodal maritime container transport, all of these requests start or end at a depot. This means that no loaded containers are transported directly between shippers and consignees in the network. An overview of the problem instances is presented in table 1 and 2. Table 1 contains the empty container demand and supply. Columns two to four show the number of empty containers supplied by the three consignees. Columns five to seven show the number of empty containers demanded by each shipper. Table 2 presents the six loaded container transport requests for every problem instance.

Every problem instance is modelled according to the three integration approaches proposed by Huth and Mattfeld (2009) and discussed in section 2.3. Detailed results for the first problem instance are presented in the next section. Results of the other problem instances are discussed in section 4.2.

Table 1: Supply and Demand

No.	Supply			Demand		
	1	2	3	4	5	6
1	2	1	0	1	2	1
2	0	0	0	1	2	1
3	0	2	1	1	0	0
4	0	1	2	1	1	1
5	1	1	1	0	0	2
6	0	1	1	1	1	2
7	1	2	1	2	0	1
8	0	1	0	1	1	2
9	1	1	1	1	1	0
10	1	0	0	1	1	1

Table 2: Loaded Container Transport Requests

No.	Loaded requests					
1	4-7	4-8	5-8	6-7	8-1	8-2
2	4-7	4-8	6-7	6-8	7-2	8-3
3	4-7	6-8	7-1	7-2	7-3	8-1
4	4-8	5-7	6-8	7-2	8-3	8-3
5	4-7	4-8	6-7	6-8	8-1	8-3
6	4-7	6-7	7-1	7-3	8-1	8-3
7	4-7	5-8	7-1	7-3	8-3	8-3
8	4-8	7-1	7-1	7-2	7-3	8-2
9	6-8	6-8	7-1	7-2	7-3	8-2
10	5-8	6-7	6-7	6-8	7-2	8-2

4.1. Results of First Experiment

In this section results of the first problem instance are discussed in detail. The first step for the sequential planning (SP) and the functional integration (FI) approach is to model empty container allocations by the allocation model. Table 3 shows the results of this step. The first three allocations represent street turns or direct allocations between consignees and shippers. The last allocation is an empty container transport from a depot to a shipper.

Table 3: Empty Container Allocations (SP+FI)

Origin	Destination	# Containers
1	4	1
1	6	1
2	5	1
8	5	1

With the sequential planning approach, empty container allocations and loaded container transport requests are routed separately. For the routing of empty container allocations, dummy nodes are introduced for nodes one and five. Routing the loaded container requests requires dummy nodes for nodes four, seven and eight. Results are shown in table 4 and 5. Together, four vehicles are required. Total variable transportation costs are €666.56.

Table 4: Vehicle Routes for Empty Containers (SP)

Vehicle	Route	Cost
1	0-2-5-8-5-0	193.24
2	0-1-4-1-6-0	149.52

Table 5: Vehicle Routes for Loaded Containers (SP)

Vehicle	Route	Cost
1	0-4-8-1-6-7-0	167.96
2	0-5-8-2-4-7-0	155.83

As described in section 2.3, the functional integration approach involves the simultaneous routing of empty container allocations and loaded container transport requests. The results of this approach for the first instance are shown in table 6. Due to the integration of the routing decisions, only three instead of four vehicles are needed. The third vehicle will even be used less than half the time available. Furthermore, variable transportation costs are reduced by €182.21 to €484.34.

Table 6: Vehicle Routes (FI)

Vehicle	Route	Cost
1	0-4-8-2-8-1-6-7-0	196.23
2	0-2-5-8-5-0	193.24
3	0-1-4-7-0	94.88

Finally, with the deep integration approach (DI), both allocation and routing decisions are fully integrated and modelled by a single model. Vehicle routes calculated by the integrated model are shown in table 7. As for the functional integration approach, only three vehicles are needed to satisfy all requests. Variable transportation costs for the deep integration approach are €453.31, which is respectively €213.25 and €31.04 lower than for the sequential planning and functional integration approaches.

For comparison purposes, empty container allocations made by the integrated model are shown in table 8. These container allocations differ from those proposed by the container allocation model (see table 2). Five container allocations are made, of which only two are street turns. Due to the simultaneous modelling of allocation and routing decisions, empty container allocations allowing for optimal vehicle routes are made.

Table 7: Vehicle Routes (DI)

Vehicle	Route	Cost
1	0-7-5-8-1-7-0	196.43
2	0-4-8-5-8-2-4-7-0	179.85
3	0-1-6-7-0	77.02

Table 8: Empty Container Allocations (DI)

Origin	Destination	# Containers
1	6	1
1	7	1
2	4	1
7	5	1
8	5	1

4.2. Other Results

An overview of results for all problem instances is presented in table 9. Column two, three and four show the total variable transportation costs for each integration approach. The last three columns show the number of vehicles required. Both variable routing costs and fleet size are always lower for functional and deep integration than for sequential planning.

In three cases, the functional and deep integration approach have the same results. In the seven other cases, deep integration causes lower variable transportation costs than functional integration. In these cases, empty container allocations made by the integrated model differ from those proposed by the container allocation model. Once the deep integration approach even results in a reduction of the number of vehicles required compared to the functional integration approach. From these results may be concluded that more integration leads to better results.

Table 9: Overview of Results

No.	Variable Costs			# Vehicles		
	SP	FI	DI	SP	FI	DI
1	666.56	484.34	453.31	4	3	3
2	565.79	440.57	440.57	4	3	3
3	574.09	462.53	447.75	4	3	3
4	687.82	413.91	389.10	4	3	2
5	542.24	363.47	363.47	4	2	2
6	626.05	533.42	471.84	4	3	3
7	655.66	423.96	423.15	4	3	3
8	577.84	472.96	472.96	4	3	3
9	719.85	562.89	510.84	5	3	3
10	496.95	365.52	357.68	4	2	2

5. CONCLUSIONS AND FUTURE RESEARCH

In order to satisfy demand for containers, empty containers have to be repositioned at a regional level between importers, exporters, intermodal terminals, inland depots and ports. Often, decisions on repositioning movements are based on a container allocation model, without taking into account vehicle routing.

This paper shows the advantage of integrating container allocation and vehicle routing decisions for the full-truckload problem with a truck capacity of a single container. Two approaches, functional integration and deep integration, may be used to integrate these decisions. First computational examples indicate that even for relatively small problem instances, both integration approaches result in smaller fleet size and significantly lower transportation costs. Best results are achieved with full integration of empty container allocation and vehicle routing decisions.

In future research an experimental design will be set up to determine which problem characteristics lead to the largest cost savings. Next, future research will focus on solution methods for larger problems and on the extension of the models presented in this paper. The container allocation model may be extended to a dynamic, multi-period model. This way repositioning movements in a certain period for fulfilling requests in a subsequent period may be modelled. Extra constraints, such as time windows for pickup and delivery of containers, may be imposed on the vehicle routing model. Finally, as node-based routing models are often computationally faster than arc-based routing models, future research will look at the opportunity of integrating the container allocation model with a node-based vehicle routing model.

REFERENCES

Abrache, J., Crainic, T.G., Gendreau, M., 1999. *A New Decomposition Algorithm for the Deterministic Dynamic Allocation of Empty Containers.* Publication CRT-99-49, Centre de recherche sur les transports, Université de Montréal.

Boile, M., Theofanis, S., Baveja, A., Mittal, N., 2008. Regional Repositioning of Empty Containers: a Case for Inland Depots. *Proceedings of 87th Transportation Research Board Annual Meeting*, pp. 31-40. January 13-17, Washington DC (USA).

Chang, H., Jula, H., Chassiakos, A., Ioannou, P., 2006. Empty Container Reuse in the Los Angeles/Long Beach Port Area. *Proceedings of National Urban Freight Conference*. February 1-3, Long Beach (California, USA).

Cheung, R.K., Chen, C.Y., 1998. A Two-Stage Stochastic Network Model and Solution Methods for the Dynamic Empty Container Allocation Problem. *Transportation Science* 32: 142-162.

Chu, Q., 1995. *Dynamic and stochastic models for container allocation*. Thesis (PhD). Department of Ocean Engineering. Massachusetts Institute of Technology (MIT). Available from: http://dspace.mit.edu/handle/1721.1/11742 [accessed 22 December 2008]

Cordeau, J.F., Laporte, G., Potvin, J.Y., Savelsbergh, M.W.P., 2007. Transportation on Demand. In: Barnhart, C., Laporte, G., eds. *Handbooks in Operations Research and Management Science. Volume 14: Transportation.* Amsterdam: North-Holland, 429-466.

Crainic, T.G., Gendreau, M., Dejax, P., 1993. Dynamic and Stochastic Models for the Allocation of Empty Containers. *Operations Research* 41: 102-126.

Dejax, P., Crainic, T.G., 1987. A Review of Empty Flows and Fleet Management Models in Freight Transportation. *Transportation Science* 21: 227-248.

Di Francesco, M., Manca, A., Zuddas, P., 2006. Optimal Management of Heterogeneous Fleets of Empty Containers. *Proceedings of International Conference on Information Systems, Logistics and Supply Chain*, pp. 922-931. May 14-17, Lyon (France).

Di Francesco, M., 2007. *New Optimization Models For Empty Container Management.* Thesis (Phd). Faculty of Engineering (Land Engineering). University of Cagliari.

Di Francesco, M., Crainic, T.G., Zuddas, P., 2009. The effect of multi-scenario policies on empty container repositioning. *Transportation Research Part E: Logistics and Transportation Review* 45: 758-770.

Dong, J., Song, D., 2009. Container fleet sizing and empty repositioning in liner shipping systems. *Transportation Research Part E: Logistics and Transportation Review* In Press.

Erera, A.L., Morales, J.C., Savelsbergh, M., 2005. Global intermodal tank container management for the chemical industry. *Transportation Research Part E: Logistics and Transportation Review* 41: 551-566.

Geoffrion, A.M., 1989. Integrated Modeling Systems. *Computational Economics* 2: 3-15.

Geoffrion, A.M., 1999. Structured modeling: survey and future research directions. *ITORMS: Interactive Transactions on Operations Research and Management Science* 1. Available from: http://www.anderson.ucla.edu/faculty/art.geoffrion/home/csts/index.htm [accessed 15 January 2009]

Huth, T., Mattfeld, D.C., 2009. Integration of vehicle routing and resource allocation in a dynamic logistics network. *Transportation Research Part C: Emerging Technologies* 17: 149-162.

Ioannou, P., Chassiakos, A., Jula, H., Unglaub, R., 2002. *Dynamic Optimization of Cargo Movement by Trucks in Metropolitan Areas with Adjacent Ports.* Metrans Technical Report, Center for Advanced Transportation Technologies, University of Southern California.

Jansen, B., Swinkels, P.C.J., Teeuwen, G.J.A., van Antwerpen de Fluiter, B., Fleuren, H.A., 2004. Operational planning of a large-scale multi-modal transportation system. *European Journal of Operational Research* 156: 41-53.

Jula, H., Chassiakos, A., Ioannou, P., 2003. *Methods for Modeling and Routing of Empty Containers in the Los Angeles and Long Beach Port Area.* Technical Report, Center for the Commercial Deployment of Transportation Technologies, California State University.

Lam, S.W., Lee, L.H., Tang, L.C., 2007. An approximate dynamic programming approach for the empty container allocation problem. *Transportation Research Part C*: *Emerging Technologies* 15: 265-277.

Lai, K.K., Lam, K., Chan, W.K., 1995. Shipping Container Logistics and Allocation. *Journal of the Operational Research Society* 46: 687-697.

Li, J.A., Leung, S.C.H., Wu, Y. Liu, K., 2007. Allocation of empty containers between multi-ports. *European Journal of Operational Research* 182: 400-412.

Mitrovic-Minic, S., 1998. *Pickup and Delivery Problem with Time Windows: A Survey*. Technical Report 1998-12, School of Computing Science, Simon Fraser University, Burnaby (BC, Canada).

Olivo, A., Zuddas, P., Di Francesco, M., Manca, A., 2005. An Operational Model for Empty Container Management. *Maritime Economics and Logistics* 7: 199-222.

Parragh, S., Doerner, K., Hartl, R., 2008. A survey on pickup and delivery problems - Part II: Transportation between pickup and delivery locations. *Journal für Betriebswirtschaft* 58: 81-117.

Savelsbergh, M.W.P., Sol, M., 1995. The General Pickup and Delivery Problem. *Transportation Science* 29: 17-29.

MULTI-AGENT SIMULATION FOR ANALYSING INLAND CONTAINER TERMINAL NETWORKS

Edith Schindlbacher[(a)], Manfred Gronalt[(b)], Hans Häuslmayer[(c)]

[(a), (b)] University of Natural Resources and Applied Life Sciences Vienna, Feistmantelstraße 4, 1180 Vienna, Austria
[(c)]h2 projekt.beratung KG, Obere Viaduktgasse 10/7, 1030 Vienna, Austria

[(a)]edith.schindlbacher@boku.ac.at, [(b)]manfred.gronalt@boku.ac.at [(c)]hh@h2pro.at

ABSTRACT
Using a multi-agent system, representing the nodes in the network of the Austrian inland container terminals, and using system dynamics to depict the terminals' internal structures and processes in an aggregate manner, we perform network flow analyses of intermodal load units in case of unforeseen disturbances. Comprehensive case studies of disturbances and irregularities in the flow of goods in Austrian container transport chains and transport systems are the basis of the definition of several risk scenarios, and are used in order to investigate the robustness of the network.

Keywords: multi-agent simulation, network flow analyses, disruption risks, inland container terminals

1. INTRODUCTION

Ever more frequent disturbances and irregularities in the flow of goods in transport chains and transport systems reflect the need to evaluate transport chain vulnerabilities and risk potential. These disturbances can stem from natural hazards, e.g. flooding or earthquakes, but can also be caused by acts of sabotage on transport chains or their infrastructure. Particularly in Austria, natural hazards like storms, heavy rain, avalanches and floods are responsible for such disruptions.

An intermodal network consists of multi-element transport chains, including pre-carriage of cargo to a terminal, the transshipment to another means of transport and the following actual change of location on that modus. The transportation process ends after a further transshipment and the on-carriage of the goods or containers. Container terminals are therefore essential infrastructure in intermodal networks.

Based on extensive field studies, comprising almost all Austrian container terminals, we have identified additional risk factors to these essential nodes, other than natural hazards or sabotage, for example risks originating partly from upstream and downstream terminals in the chain or from other parties in the transport system such as train operators or infrastructure operators. However, risks emanating from the daily operational business rank among the most prevalent sources of disruption in transport chains (Gronalt et al. 2008).

A longer persisting disruption of a certain stage - respectively a node - of a transport chain can cause massive damages which are hard to quantify. Such an incident can affect the whole supply chain, i.e. from suppliers, who can no longer deliver their goods, to the customers who do not receive them.

Immanent to networks are causal relationships amongst the participants, which include the emergence of negative cascading effects, i.e. one partner waiting for goods or load units (in the case of terminals) to be transferred by another partner, who is suffering a disruption, also experiences severe problems. Hence, these flow problems in the transport chain require information about rerouting possibilities. Based on the mentioned risk potential analysis and developed risk profiles for every Austrian container terminal, we present an agent-based simulation model of the terminal network to evaluate strategies for coping with disruptions.

In our model we consider a terminal as an individual unit in the network, which pursues its main objective of maintaining the maximum throughput. From this perspective, an analogy with software agents in computer science can be drawn. According to Jennings (2000), the crucial characteristic of an agent is its autonomous behaviour; the capacity to make independent decisions is the primary property of an agent. This implies the need for planning and active responses to the environment to achieve their particular objectives. Software agents are used in the areas of object-oriented programming and concurrent object-based systems. The agents are modelled to represent natural entities in the system under consideration, and therefore are applicable for representing network participants.

The operational activities in a container terminal are subject to dynamic causal processes, influenced by the present infrastructure and operational strategies. An aggregate view of these activities can be provided by the modeling approach of system dynamics, which represents real-world processes in terms of stocks, flows between them and the information that determines their value (see Schieritz and Milling 2003).

2. RELATED WORKS

Social processes are a common field of agent-based modeling applications. Even more important than modeling agent behavior, is modeling the agent interactions. The main issues hereby are the questions about which agent is connected to another, and of course, what the governing mechanisms of their interactions are. For social interactions a network interaction topology may provide a more accurate description of the agents' interaction patterns than a cellular one would. Moreover, agent-based modeling and simulation applications range from the areas of business and organizations, economics, crowds, society and culture and biology, to the subject of infrastructure (Macal and North 2008).

A survey of Davidsson et al. (2005) shows that agent technology has been applied to various problem areas within transport logistics, such as (transport) planning and scheduling, fleet management and traffic control and management. The papers reviewed cover the domains of transport, traffic and terminals, as well as the modes of transportation (air, rail, road, sea) and the topic of intermodal transport. The authors come to the conclusion that very little work has been done in the field of strategic decision-making. For intermodal terminals, Henesey et al. (2009) evaluate operational policies for the transshipment of containers, more precisely policies concerning the sequencing of ships, berth allocation and stacking rules.

Related questions to our investigation have been considered in a commodity flow network with arbitrary topology by Weiskircher et al. (2009). The solution of the profit maximization problem for the network with distributed control is implemented by using software-agents representing the network nodes.

In the field of supply chain modeling, numerous computer-based models deploy system dynamics. Größler and Schieritz (2005) demonstrate a combined approach of system dynamics and agent-based simulation to test the stability of supply chain structures under different levels of uncertainty. They judge a combination of both of these approaches which are helpful in the investigation of a supply network structure that emerges from the interaction of at least partly independent companies.

3. SIMULATION MODEL OF THE AUSTRIAN INLAND CONTAINER TERMINAL NETWORK

The aim of our developed simulation model is to emulate the flow of load units between the considered terminals and to analyse the network behavior in case of disturbances, covering the whole range from reduced transshipment performance to a total breakdown of a terminal. It is, therefore, of considerable interest, which other terminal in the network would be able to overtake a certain amount of load units. Furthermore, it is important to know how long it takes for the network to recover and regain initial conditions. In answering these questions, vulnerabilities of the Austrian terminal network can be identified and its robustness and resilience can be assessed. Therefore, the following performance measures for the network are applied:

1. The overall system performance is measured by the throughput of load units per week, since daily fluctuations should balance in this time.
2. The average utilisation of storage and lifting capacities per terminal, disruption and recovery period.
3. The number of days in total runtime, when storage and transshipment utilisation exceed a critical value of 75 percent.
4. The frequency of occurrence of a queue of load units waiting to be processed in a terminal.

3.1. Model structure and components

The main part of the model is formed by the *terminal agents*, which represent the single terminals in the network under consideration. They communicate with each other about the amounts of load units interchanged, and pass on and receive information from the *system dynamics models*, which perform terminal internal operations. An *administrator* manages the whole simulation cycle and structures, and controls the communication between the terminal agents. The scenarios of disturbance events are implemented through the so called *environment*, which emulates potential incidents of different origin in altering the regular flows of load units, or in affecting available storage and transshipment volumes. Figure 1 schematically shows the systems components and their interrelations.

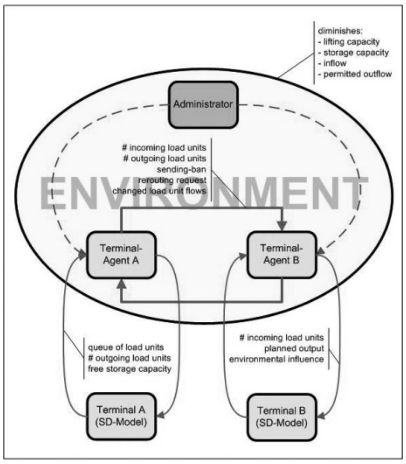


Figure 1: Model Components

The system dynamics models of terminal internal operations

The central elements of the system dynamics models are the storage and lifting capacities. The storage capacity is a stock variable, which is determined by the amount of the daily incoming and outgoing load units. The actual admission to storage results from the amount of the everyday incoming load units, the share of direct turnover to other means of transport without interim

ISBN 978-84-692-5416-5

storage, the available lifting capacities and of course the storage capacity. Likewise, the daily outcome derives from the planned amount of outgoing load units, the direct turnover, the lifting capacity and the existing stock. The theoretical lifting capacity of all the various transshipment equipment is dependent on the terminals' operating hours and reduced by the necessary reshuffling moves. Incoming load units which cannot be moved to the storage yard form a queue, and can be considered as prestow capacity.

The system dynamics models receive information about up-to-date incoming volumes and current planned outgoing amounts of load units from the associated terminal agent, as well as constraints on the storage and lifting capacities, passed on from the environment through the terminal agent. Vice versa, the information flow contains data about available storage capacities, the queue length and the actual output.

The terminal agents

The main function of the terminal agents is the coordination among themselves in case of deviations in the planned network flows by rail. Each terminal's interest is to fulfil its quotas and to provide for an alternative itinerary when it cannot complete its tasks. There are several scenarios which require terminal agents' action:

- The terminal could not process the planned number of outgoing load units e.g. in consequence of a lack of lifting capacity. So the other terminals in the network have to be informed about changed amounts of incoming load units, and furthermore, these quantities have to be processed additionally in the subsequent periods.
- Due to a deficiency in the storage capacity a queue of load units arose. In order to allow the queue to be processed, delivery bans have to be sent out for the next period.
- Incoming messages from the environment, which signify changes in the incoming or outgoing load unit flows or declines in lifting or storage capacities, will likely entail modifications in the network flows and therefore demand agents' actions.
- Given a ban from another terminal to send load units, the terminal agent tries, according to its focus on maintaining the maximum throughput, to reroute that quantities to another terminal in the network.
- As a reaction to the aforementioned situation, an agent has to accept these load units according to its capacities, which again alters the transshipment amount in the subsequent periods.

The environment

Based on previous work on risk analyses and vulnerability assessment of the Austrian inland container terminals, the environment emulates risk scenarios, which (negatively) influence the operational activities of a terminal. The terminal's turnover capability can thus be affected in terms of a reduced lifting or storage capacity, or changed conditions caused by altered amounts of incoming or outgoing load units.

The administrator

The administrator is responsible for the control of the whole simulation run, i.e. it determines the sequence of the terminal agents activities and the system dynamics components. To avoid problems due to concurrency issues, only one agent can be active at any given time, all other agents are inactive or in a waiting position. For this reason, the administrator additionally determines the order in which the terminals are allowed to conduct the different actions and coordinates their communication.

3.2. The communication of the model components

The communication structure of the agents is of considerable impact on the network performance. This is because the decision to reroute strongly influences the outcome of the simulation run, since these rerouting possibilities are then no longer available to another terminal, which might make a request later. Correspondingly, the rules for announcing bans on incoming load units, and selecting the terminal which is affected by changed outgoing amounts, can result in considerable discrepancies in the overall network performance.

The communication between all agents is implemented as follows. On the one hand, information is transferred by the type of message they send each other, and on the other hand by the carried content. The terminal agents and the system dynamics components are interconnected through interfaces in the form of variables. The different types of messages are:

- Activation of the environment by the administrator.
- Activation of a certain function of a terminal agent by the administrator.
- Announcement of a (partial) ban on incoming load units from one terminal agent to another.
- Request of rerouting possibilities and the corresponding answer among two terminal agents.
- Terminal agents give out information to the others about reductions in the flows of load units, and ask for operability in case of additional amounts and receive accordant reply.
- Confirmation of finished actions to the particular communication partner.

3.3. Activity scheduling

One total run of the models functionality represents the actions in a certain time period, so that it can be applied in a more or less aggregated way for a certain period of time, split into cycles representing data with different levels of aggregation. As mentioned before, the sequence in which the terminal agents get their turn for different actions in a cycle may change the outcome of the simulation, as the decisions of one agent can limit the options for the other agents later in the sequence. We choose their sequence randomly in each cycle.

At the beginning of every cycle the system dynamics components prepare the data basis for the terminal agents' actions. Next, the environment provides the risk driven influencing factors for the period in question, which are adjacently processed by the terminal agents. Afterwards, potential queues are checked, in case necessary bans for incoming load units are announced for the subsequent cycle. Obviously, in the random turn sequence, the reaction of all terminal agents to delivery bans is to demand rerouting opportunities. The sequence of asking the other terminal agents to take over the load units is again chosen at random. This situation is one of several, when one terminal agent has to wait for the response of another one, in order to continue its actions.

Next, the agents check whether modifications emerged in the planned number of load units to be sent or to be received. Deficiencies in the outgoing amounts only have to be announced. However, when additional transport is demanded, the intended recipient can refuse to accept them due to a lack of capacity.

The final action for the terminal agent is to calculate the actual values for the number of incoming and outgoing load units for the next cycle as input parameter for the terminal internal processes, realised in the system dynamics components.

3.4. Proposed model assumptions

The simulation is executed on a daily basis, so the flows of load units are the aggregated amounts of incoming and outgoing units per day. The depicted flows are the amounts to and from the other terminals in the considered network, as well as the cumulative volume of the inflows and outflows per truck and ship, and the other terminals which are not represented in the model, e.g. the flows to and from overseas.

The possibility of asking other terminals for the opportunity to reroute is limited by a certain probability, and their sequence taken by chance.

Basically, the acceptance of a rerouting demand is reliant on free storage capacities and actually accepted requests. Even provided sufficient storage capacities, a terminal is not obliged to accept a rerouting request. Therefore, a probability distribution is used to make this decision.

Every other terminal in the network not having announced a sending-ban, is an option for a rerouting request, as well as for processing additional volumes evoked through certain events in previous periods. Their order is also determined randomly.

No capacity constraints with regard to the links between the terminals are considered.

4. COMPUTATIONAL EXPERIMENTS

For our analyses a group of three representative Austrian terminals were chosen, later referred to as terminal A, B and C. Two of them hold an important strategic position for and in the Austrian terminal network, as they take over the function as a gateway and hub respectively, and belong to the biggest Austrian terminals measured in terms of their yearly turnover. The third terminal has been selected due to its strong interlinkage of transport volumes to the other ones, and, because of its geographical position, inland transport and also transit traffic along the arterial Austrian transport links are covered.

4.1. Data basis and model parameters

As already mentioned, the simulation and investigation of the network performance is done for aggregated daily flows, whereas empty container volumes are not taken into account, and neither are the corresponding capacities for storage and transshipment. The following assumptions are made regarding the model parameters:

- The daily amount of incoming and outgoing load units, given as number of load units, is an average value of the total transshipment volume per year, but also considers fluctuations in the different days of the week.
- Load unit interchange takes place between the modeled terminals, it is also implemented for the "terminal-environment" using aggregated volumes.
- The storage capacity is set for charged containers, measured in TEU (twenty-foot equivalent unit).
- The cumulated lifting capacities per simulation cycle result from the possible moves per hour and equipment, totalized for the terminals operating hours, and reduced by the moves necessary for empty container movement.
- Another limiting factor for the actual available lifting capacity is the share of unavoidable reshuffling moves, which depends on the utilisation of the storage capacity and the number of tiers.
- To convert the number of incoming and outgoing amounts of load units into the storage capacity measure TEU, the average length of stored load units and the share of not stackable load units are applied.
- Not every incoming load unit is put into (interim) storage. This fact is taken into account by implementing a certain share of direct moves per inflow.
- The initial filling degree of the storage yard is estimated at the rather low value of 65 % for two of the terminals; for the third, the actual utilisation of 56 % is taken.

4.2. Disruption Scenarios

The analysed disruption scenarios comprise diminished storage capacities, diminished lifting capacities and constraints on the planned amount of outgoing trains (in fact batches of load units) to the terminal-environment.

Capacity reductions are assumed with 50 and 100 percent respectively, the third type of disruptive event is defined to an extent of one and three trains for each terminal. All disruption scenarios were simulated with a time dimension of one, three and five days of suffering the disruption for every terminal. Every scenario is analysed under ceteris paribus conditions, so that no events occur contemporaneously. Consequentially, a total number of 55 scenarios (one reference scenario without incidence plus 54 disruption scenarios) were calculated. Table 1 provides an overview about the 18 scenarios applied for the first terminal.

Table 1: Disruption Scenarios using the Example of Terminal A

disruption scenario		*duration (days)*	*intensity (% / number of trains)*	*scenario number*
DISRUPTION TERMINAL A	reduction of storage capacity	1	50	1
			100	2
		3	50	3
			100	4
		5	50	5
			100	6
	reduction of lifting capacity	1	50	7
			100	8
		3	50	9
			100	10
		5	50	11
			100	12
	reduction in number of outgoing trains	1	1	13
			3	14
		3	1	15
			3	16
		5	1	17
			3	18

4.3. Simulation results and analysis

The simulation run was carried out for one year, presuming five operating days per week and 50 operating weeks per year. The performance measures for all scenarios were calculated by the average of 20 replications each. To ensure a justification period and to sufficiently observe disruptive events' consequences, the occurrence points in runtime were defined permissible starting four weeks after the start of the simulation, to allow for a system warm-up, until four weeks before the end of runtime, to allow for an observation period of the consequences. During this period, disruptions happen randomly according to a uniform distribution.

Every disruptions' consequences to the network are assessed for the time spans of 5, 10 and 15 days after occurrence, by means of averaging the already mentioned performance measures utilisation of storage capacity, utilisation of lifting capacity and the number of days exceeding their critical value, as well as the emerging of a queue of load units waiting to be processed. As a reference, the average values of a simulation run without a disruption is applied. In the following, we employ selected scenarios to show how certain disruptive events affect the performance of the single nodes, and accordingly disperse in the considered network.

Example 1: Diminishment of lifting capacity at Terminal B

The impact of a total breakdown of the transshipment equipment in Terminal B for a time period of five days on the average utilisation of the lifting capacities of all three terminals is shown in Figure 2. Apparently, the utilisation in the concerned Terminal B rises immediately. This effect lasts up to 10 days after the occurrence of the disruption with the same severity, until the recovery phase sets in. In the same period the effects on Terminal A and Terminal C start to become visible. Terminal A shows an increase in lifting capacities utilisation of about 10 percent and Terminal C shows an increment from about 11 percent to 14 percent. Several things can be shown to be responsible for this development. On the one hand, the other two terminals have to take over additional transport volumes from the system until the recovery of Terminal B has completed, on the other hand, the processing of the arisen queue of load units in Terminal B after its recovery, implies extra workload for the other network nodes. In the course of time, utilisation values drop again.

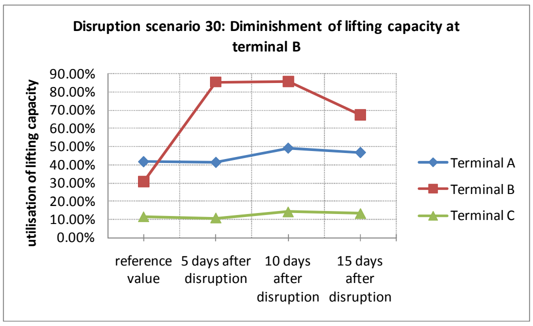


Figure 2: Diminishment of Lifting Capacity

Example 2: Reduction in number of outgoing trains at Terminal C

An incidence concerning the flow of load units to the external, not explicitly modeled, network emphasizes the role of Terminal C as an important gateway for the Austrian terminal network. As a real world example, this can be a situation in which maritime transports from Austria have to wait at an Austrian node located close to the border, because of capacity bottlenecks in the German railway network.

The infrastructure of Terminal C is of sufficient dimension to cope with such a situation, resulting in an extraordinary burden on storage capacities without markedly impairment of its lifting capacities. Effects on the other two terminals keep within limits, both due to the internal capacities of Terminal C and the fact that not every rerouting possibility is reasonable. As shown in Figure 3, efficiency curves of Terminal A and B remain unaffected.

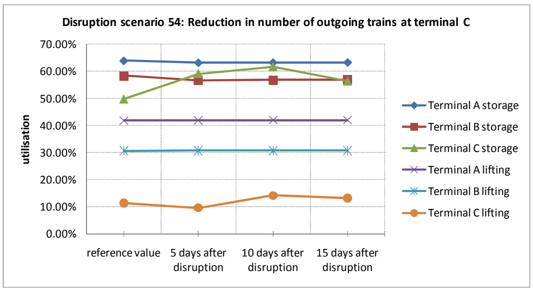


Figure 3: Reduction in Number of Outgoing Trains

Example 3: Comparison of consequences according to the network role of a terminal
According to the role of a terminal in the network, varying from a local node to a gateway or hub, distinctions can be made between the consequences of a disruption for the terminals that are only indirectly concerned. Figure 4 shows the effects of all events occurring at Terminal B and Terminal C on the other terminals, measured in the number of days their respective storage utilisation exceeds its critical value.

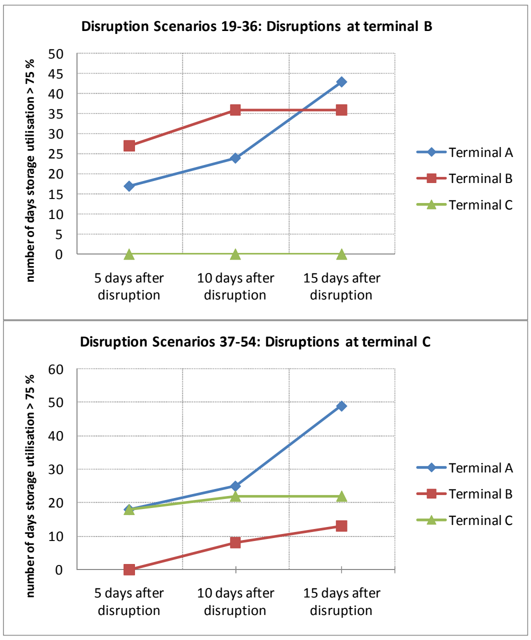


Figure 4: Comparison of Consequences on the Network

Given disruptive events at Terminal B, a considerable increase of the storage utilisation of Terminal A can be seen, whereas Terminal C is hardly affected. On the contrary, disruptions at Terminal C influence both other nodes, but Terminal A encounters stronger impacts than the others. Herefrom, conclusions about mutual interrelationships and dependencies can be drawn.

5. CONCLUSION

The implemented simulation experiments and findings that have been gained from it, conspicuously show that disruptive events in the (here considered) Austrian terminal network can be handled properly. The particular terminals have enough capacities in order to cope with disruptions of either terminal external or internal origin. Where this is not the case, there exist enough rerouting possibilities to geographically closer or otherwise connected terminals anyway.

Another, always feasible coping strategy with disruptions, is to use redundancies in transport routes and (re-)deflect transport from rail to road if necessary. However, limitations of the capacities of these links in the network have to be taken into account. These factors have not yet been considered in this study but are currently being considered in follow-up research works by the authors.

ACKNOWLEDGMENTS

This research was funded by a grant of the Austrian Security Research Programme KIRAS of the Federal Ministry for Transport, Innovation and Technology, managed by the Austrian Research Promotion Agency (FFG).

REFERENCES

Davidsson, P., Henesey, L., Ramstedt, L., Törnquist, J. and Wernstedt, F., 2005. An analysis of agent-based approaches to transport logistics. *Transportation Research Part C*, 13, 255-271.

Gronalt, M., Häuslmayer, H., Jammernegg, W., Schindlbacher, E. and Weishäupl, M., 2008. *A risk assessment approach for inland container terminals*. Proceedings of the 11th International Workshop on Harbor, Maritime & Multimodal Logistics Modeling & Simulation, September 17-19, Campora S. Giovanni, Italy.

Größler, A. and Schieritz, N., 2005. Of stocks, flows, agents and rules – "Strategic" simulations in supply chain research. In: Kotzab, H., Seuring, S., Müller, M. and Reiner, G., eds. Research Methodologies in Supply Chain Management. Heidelberg: Physica-Verlag, 445-460.

Henesey, L., Davidsson, P. and Persson, J.A., 2009. Agent based simulation architecture for evaluating operational policies in transshipping containers. *Autonomous Agents and Multi-Agent Systems*, 18, 220-238.

Jennings, N.R., 2000. On agent-based software engineering. *Artificial Intelligence*, 117, 277-296.

Macal, C.M. and North, M.J., 2008. *Agent-based modeling and simulation: ABMS examples*. Proceedings of the 2008 Winter Simulation Conference, Dec. 7-10, Miami, USA.

Schieritz, N. and Milling, P.M., 2003. *Modeling the forest or modeling the trees: A comparison of System Dynamics and Agent-Based Simulation.* Proceedings of the 21st System Dynamics Conference, July 20-24, New York City, USA.

Weiskircher, R., Kontoleon, N., Garcia-Flores, R. and Dunstall, S., 2009. Using agents for solving a multi-commodity-flow problem. *European Journal of Operational Research*, 194, 888-900.

NASS: SYSTEM SIMULATION OF INLAND WATERWAYS

Keith D. Hofseth[(a)], Cory M. Rogers[(b)]

[(a)] U.S. Army Corps of Engineers Institute for Water Resources, Alexandria VA, USA
[(b)] Camp Dresser & McKee Inc., Carbondale IL, USA

[(a)]keith.d.hofseth@usace.army.mil, [(b)]rogerscm@cdm.com

ABSTRACT
The Navigation System Simulation (NaSS) suite of tools has been developed by the Institute for Water Resources of the U.S. Army Corps of Engineers (Corps), as part of the Navigation Technologies Research Program. NaSS is composed of two primary applications, a Monte Carlo simulation model of vessel movements on an inland waterway system and a data processing and analysis tool for mining of historical data. The NaSS tools are used in concert for economic analyses of an inland waterway system. The NaSS suite was designed to answer potential questions of a waterway system under examination, for example; What is the overall system performance of a waterway network under different operating, demand load and reliability conditions?; How effective are alternative lockage polices at reducing delays and delay costs?; How does any single lock improvement project affect delays at other locks? This paper focuses on the Monte Carlo waterway simulation component called BasinSym.

Keywords: Monte Carlo simulation, waterway transportation, engineering reliability, queuing analysis.

1. INTRODUCTION

The United States economy is serviced by a vast network of 12,000 miles of navigable "inland waterways". These waterways take the form of rivers and canals that connect the inland States to each other and to international commerce moving through coastal ports. Approximately 69 billion gallons of petroleum, 20% of U.S. coal, and 60% of grain exports travel along the inland waterways on a fleet of over 4,000 tug and towboats and 27,000 barges. Should this network of rivers and waterways be inaccessible to commercial traffic, 60 semi-trucks or 15 rail cars would be needed to transport the cargo from one barge (American Mariner 2009). As over one billion tons of transported cargo was recorded on the U.S. inland waterway system in 2007 (Waterborne Commerce Statistics Center 2009), maintaining the inland waterways serviceability is critically important to the nation's economy, transportation system, and environmental health. Infrastructure capital investment and maintenance,, including locks, dams, and flood protection structures, along with regular channel dredging is necessary to ensure the smooth flow of vessels and commodities.

Improvements to the U.S. navigable waterway channels and infrastructure under the jurisdiction of the Corps must be justified based on analysis of the navigation benefits that the improvements will provide. Within the Corps, economic justification is done within a framework of comparing the "with-project" and "without-project" conditions, effectively discounting the benefits that would accrue if the Corps did nothing to improve the system. These benefits are dependent upon their affect on lock processing time or easing of congestion due to constraints within the system, for example, improving lockage times through rehabilitation or investment in locks, or altering the traffic flow through bend easings and relief of movement restrictions. These benefits are measured and compared to the cost of the particular improvements being analyzed. The BasinSym simulation model allows the estimate of transportation time and cost under varying future conditions.

2. OVERVIEW

System simulation of this problem at the microscopic level is highly complex. In order to simulate the individual movements of the vessels, a meaningful way to describe those movements is required. Tows and the barges they move must be represented as complex aggregations of individual vessels and the commodities they carry. While some tows operate in a shuttling capacity, moving barges from point A to point B and back again, others operate in a long haul capacity and still more are picking up empty barges and dropping them at economically advantageous destinations. Further complications arise, as tows change their barge configurations during their journey in order to effectively navigate the inland system. As well, an understanding of the motivations behind shipper response to congestion (i.e., when is a change in transportation mode likely?) must be gained in order to model the problem correctly.

The BasinSym model captures these complexities to assist with analysis of the problem. BasinSym is designed to simulate inland traffic movements on the waterway system and by making runs with differing configurations and fleets allows for comparative analysis of different conditions. Detailed analysis of potential improvements over a 50-year planning horizon requires determination of not only the existing condition but the operating characteristics of proposed lock improvements, the future servicing fleet and anticipated demand in commodities as well. BasinSym has been designed to provide the analyst with a set of tools to thoroughly investigate these problems in order to provide a justifiable basis for extrapolation of this information 50 years into the future.

3. SYSTEM REPRESENTATION

BasinSym utilizes a user-defined network to describe the physical characteristics of the waterway system being studied. The network consists of a series of reaches (channels), locks and docks as well as information about the physical characteristics and capacity of each.

The system is represented as a node-link network. Each link is a reach, representing a portion of the waterway (between nodes) on which vessels travel. Nodes may be referenced to a river mile indexing system, as well as by geodetic coordinates. Upstream and downstream nodes are defined based on direction of water flow.

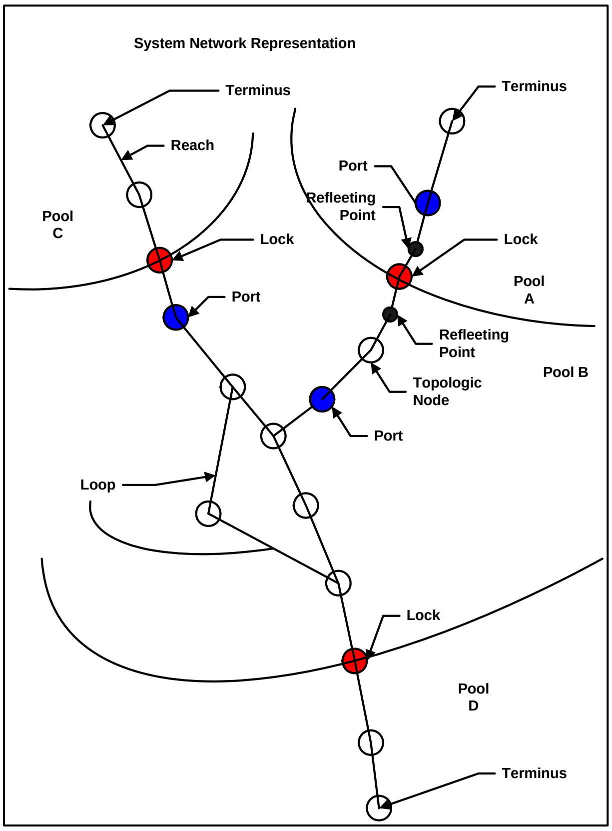


Figure 1: System Network Representation

The following node types are considered:

- Topologic nodes—serving only as start and end points of reaches;
- Port/dock nodes—serve as origin/destination of commodity and/or vessel movements;
- Terminus nodes—serve as entrance/exit points to the portion of the network under study and serve as sources and sinks.
- Re-fleeting Points—locations at which a tow can be reorganized into a different configuration with a different towboat. These can be co-located with a port node, to allow for representation of re-fleeting at a port.

All reaches are connections between a pair of nodes. Reaches may be either Open Channel or Lock Reaches. Open Channel reaches are segments where vessels can move subject only to transit rules while Lock Reaches consist of a lock and associated internal geometry. Reaches may have additional descriptive attributes assigned. In addition, vessel speed is associated with individual reaches, that is, vessels may travel at different speeds in different reaches, as specified by input data. Transit restrictions in a reach are defined by rules associated with the reach, specifically, no overtaking, no meeting, single vessel reach and one-way traffic.

Regions can also be defined by the user, e.g., pools, governmental jurisdictions, port hinterlands, etc., by associating nodes and reaches of the network with such regions. Allowing for user-defined regional definitions is useful for performing disaggregation on external data that is available on a spatial region basis or for aggregations on model-generated data.

4. FLEET DEFINITION

The calling fleet is described as a set of *vessel classes* that contain information on the characteristics of each type of tow and barge. A single classification table houses the data necessary for defining categories of vessels, whether they be barges, tugs, or even recreational vessels. The following data items are used to classify the vessels used during simulation:

- VesselClass - user defined name (e.g., Jumbo Hopper, 3000 HP, etc…)
- CostStaffed - total vessel cost
- CostMoored - cost at-port for barges
- ISPowered - indicates whether class contains powered vessels or barges
- Horsepower - triangular distribution specifying horsepower class
- LOA – triangular distribution specifying overall length for the class
- Beam - triangular distribution specifying beam width for the class

Unique flotillas (power vessels and barges) are stored separately in a table of unique barges and a table of unique powered vessels respectively. For barges, the following fields apply:

- LOA - overall length of barge

- Beam - beam width
- TPI – tons per inch indicating amount of draft added per ton of load
- LightDraft - draft of barge when empty
- MaxDraft – draft of barge when fully loaded

For power vessels, the following data items are used:

- Name - name of the vessel
- HP - horsepower of vessel
- LOA - overall length of vessel
- Beam - beam width of vessel
- TPI - tons per inch indicating amount of draft added per ton of load
- LightDraft - draft of vessel when empty
- IdleFuelBurn – daily number of gallons burned by vessel when waiting
- RatedHPFuelBurn - daily number of gallons burned per horsepower produced
- VesselCapability - integer indicator of what vessel capability level, valid values are: (1) vessel is capable of pushing barges, (2) vessel is capable of carrying cargo

Powered vessels can incur time at port nodes (for loading and unloading and for tow configuration), re-fleeting points (for re-configuration and change of tow) and at locks (for passage through the queue and lock). It is assumed that vessels do not incur any time when traversing a topologic node, unless they must wait before entering the next reach, due to transit rules.

4.1. Trip Specification

With this foundational information for a particular system in place, BasinSym employs a shipping manifest to provide the simulation with specific tow/barge movements that are subsequently routed through the system. This manifest, or shipment list, includes information on when the vessel departs on a journey, and commodities picked up and/or dropped off along the way. The specification of trips is accomplished through a structure consisting of a set of relational database tables which are slightly denormalized in the tables below for simplification purposes. In the real world, vessels are continuously traveling the waterway; for purposes of simulation an individual trip is defined as the set of reaches between a change in direction of the powering vessel. Even with this simplification, a complex data structure is needed to capture the information relevant to a trip.

Table 1: Trip Specifications

TripID	Trip Date	Power Vessel	Origin Node	Destination Node
1	Jan 11, 2009	1	1	6
2	Jan 12, 2009	8	4	10

The data in Table 1 defines each specific trip of an individual powered tow. This includes the begin date of the overall trip as well as the origin and ultimate destination of the trip overall. The individual node visits where commodities are loaded /unloaded or barges are added/removed along the entire route are specified in Table 2. The node visits are also given a specific order within the overall trip to provide for sequencing. A visit contains 1 or more transactions related to changes in power vessel, additions or removal of barges from the tow, or cargo related operations.

Table 2: Individual Node Visits

Visit ID	Trip ID	Visit Node	Visit Order
1	1	1	1
2	1	6	2
3	2	7	1
4	2	8	2
5	2	10	3

Barge removals/additions to the tow are handled through specification of an add or drop command for a specific number of unpowered barges at pertinent visit nodes as shown in Table 3. It is important to note that although powered vessels are stored and identified individually, this data structure does not indicate specific barges to be transacted but rather a count by class of vessel and by default they are assumed to be unloaded.

Table 3: Add/Drop Command

Action ID	Visit ID	Action	Vessel Class	Count
1	1	ADD	Jumbo Hopper	1
2	1	DROP	EMPTY	2
3	2	DROP	Jumbo Hopper	1
4	3	ADD	Hopper	2
5	4	ADD	Deck	2
6	4	ADD	Hopper	4
7	4	DROP	EMPTY	2
8	5	DROP	Hopper	4

Cargo transactions, displayed in Table 4, account for the final element of the trip specification. Cargo that is loaded or unloaded is described in terms of commodity, quantity in tons, and whether the cargo is to be loaded or unloaded. These transactions coupled with the ordered set of visits and actions within a power-trip fully define a trip specification.

Table 4: Cargo Transactions

Action ID	Commodity	Qty (tons)	Load/Unload
1	CORN	10	LOAD
3	CORN	10	UNLOAD
4	COAL	12	LOAD
5	MECH EQUIP	8	LOAD

6	COAL	30	LOAD
8	COAL	30	UNLOAD

5. ROUTING

Vessels travel on a designated route from origin port to destination port. Where alternative routes are possible (in a network with loops) it is necessary to determine which route is chosen. This is done adaptively as a vessel observes congestion and chooses an alternate route. Route selection in the simulation of barge traffic represents a balance between computational tractability and emulation of a real-world decision making process. Three major questions arise during assessment of the decision making process – When does a tow determine its route? When will a tow change its route and what triggers the route change? Will a tow always choose the least expensive route and what metric determines cost of a route? These questions and the assumptions made as to the answers form the foundation of routing in the BasinSym model.

First, it is necessary to understand how tow movements are specified in the simulation. As stated earlier, a tow is assigned a trip which consists of an ordered set of node visits. The order is assigned and cannot be changed as the order of cargo and barge transactions must be preserved to avoid the case in which barges are dropped off at a node before they are picked up. Thus, a trip is an ordered set of (potentially) smaller voyages or legs between nodes to be visited. BasinSym assumes that the initial (preferred) route for the entire trip is established prior to the first movement. Based on this assumption, prior to the first movement, a trip route consisting of the summation, with order preserved, of all leg routes, is created for the tow/trip.

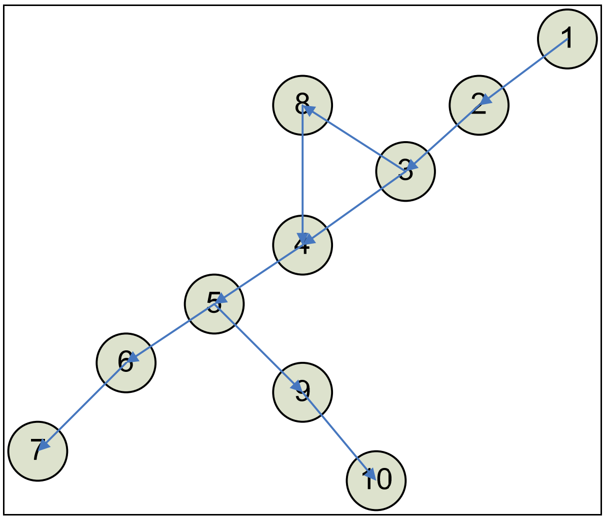


Figure 2: Trip-Visit Route

An example of a trip-visit route is shown in the figure above. A tow is given a trip starting at node 1 and ending at node 7. It has visits, in order of 1, 8, 9, and 7. Thus, we would determine routes between nodes 1 and 8, 8 and 9, and 9 and 7. Those routes would appear as follows:

1. (1 to 8): 1-2-3-8
2. (8 to 9): 8-4-5-9
3. (9 to 7): 9-5-6-7

The overall trip route would be 1-2-3-8-4-5-9-5-6-7. Tows store their visits in a FIFO queue structure and remove visits from the queue as they are completed, the leg routes are stored as FIFO queues as a subset of the visit storage structure.

5.1. Algorithm

BasinSym processes the shipment list by date and routes the movements on a least-cost path from origin to destination using the A* (A-Star) algorithm (Hart, Nilsson, Raphael 1968). A* utilizes a distance- plus cost heuristic for determination of the route. The cost portion of the heuristic within BasinSym is based on estimated transit time along the route to take into account any network congestion that may exist.

A* is both optimal and complete given that the evaluation function f(x) is admissible. The evaluation function in A* is a compound function consisting of g(x), the cost of the route from the origin to the current node, and h(x), the estimated cost of the remaining route from the current node to the destination. While this sounds simple, there are considerations. The heuristic, h(x), must be optimistic in that it must never over estimate the remaining travel cost. Poor construction of the heuristic will negatively impact the algorithm. In the worst case scenario, h(x) is set to a constant value of 0. This degrades the search from Best-First to a simple Depth-First search.

In the transportation modeling domain, the straight-line distance between two nodes is often used as a heuristic for optimistic shortest path given that the shortest distance between two points, in simple Euclidean geometry, is a straight line. The heuristic, h(x), implementation within BasinSym is the straight-line distance between the two points divided by the fastest speed limit on any reach for the vessel class of the power vessel of the tow. This requires that all nodes have valid geospatial coordinates assigned to them in a format that allows the determination of distance. The Great Circle Distance formula is applied to determine "straight-line" distance on the surface of the Earth. The distance obtained is divided by the top speed of the tow as determined by its power vessel class in order to provide a measure in hours as opposed to miles. The choice of the top speed is made as the evaluation function must be optimistic in order to be admissible and thus never overestimate the path cost.

The second component of the evaluation function, f(x), is the cost of the path from the origin or g(x). In that the evaluation function must be optimistic, each of its components must also be optimistic to include the cost of the path traversed. G(x) within BasinSym uses the vessel's top speed, governed by reach speed limits, over the length of the reaches traversed. This is implemented through use of a weighted graph structure to capture the nodes and reaches of the navigation network. Since simple weighting of the arcs without consideration to limiting factors is insufficient for effective route production, the vessel class of the tow's power vessel is used to set arc weights. This allow for establishment of extreme levels of impedance where a tow would be violating simple reach rules. The weights are stored in each arc as a dictionary of traversal times by vessel class. Cost to traverse a lock represents a slightly more complex situation. Given that the estimate must be optimistic, the cost is determined by the lower bound of any processing distributions for the fastest chamber at the lock given that there are no vessels in any queues. In order to simplify this, the lock objects contain a method called *Lock.EstimatedTraversal* based on vessel class. The result of this method is used to weight the inter-node arcs where the reach is a lock. Time to traverse the path is summed across the reaches to produce a consistent unit of measure – time to traverse.

5.2. Route Generation and Selection

Each tow retains the necessary information to describe its complete route. The route is stored as a set of intermediate destination nodes within the trip-visit structure of the tow. As each route is generated, it is stored in a dictionary of routes indexed by origin, destination, and class of vessel. These routes are determined upon creation of the tow and potentially added to when congestion or closure dictates a route alteration is desired. The predetermined routes are unchanging regardless of network condition or situation. These fixed routes are not purged under any circumstance for the duration of a simulation. This implementation avoids the resource cost of searching for a route that has already been found and provides a commensurate improvement in execution speed.

Periodically, a tow will need to alter the initial route due to congestion on the waterway or closures at a lock. Route selection is altered or recomputed when a high impact event such as a lock closure, an event that makes any route impassable, or an estimated lock traversal time that exceeds some user defined threshold. Upon receipt of a high impact event message, all tows in the system recalculate the best route from their current position to the completion of the current leg as well as any remaining voyage legs. When the simulation encounters a high impact event, the route dictionary is purged of all prior information with the exception of fixed routes. The model subsequently iterates through all tows that are active in the system. For each of those tows, routing information is invalidated and the tows regenerate routes from their current position through the end of the voyage.

6. LOCKS

BasinSym provides the user the ability to define all locks in the system to be as one of three types depending upon the level of detailed analysis required at a particular lock in the system. First tier locks are the simplest and utilize a user-defined distribution to define the amount of time it takes for a particular tow class to transit the lock. Processing distributions are fed to the model for upbound and downbound movements of tows respectively. An additional distribution set exists for the upbound and downbound movement of recreation vessels. All lockage tiers utilize a separate distribution to represent the time it takes to fill and empty the chamber.

Flotillas traversing the inland waterways of the US are often larger than the locks they must pass through. To pass through the lock the flotilla must be broken into "cuts". Second tier locks add the complexity of processing vessels using a set of distributions depending on the number of cuts required to completely lock the vessel through. For example, the first cut through a lock would utilize a different processing distribution than the second and/or third cut. Cut determination is handled using a packing algorithm which will be discussed in detail later in this document.

Tier three locks are the most detailed and move locks through using a set of processing distributions for tow/barge class and number of cuts but these distributions are broken out by lockage type on approach and exit as well.. For example, a particular distribution exists to represent the processing time for an upbound tow making a "fly" approach while another distribution exists for the same tow utilizing a "turnback" exit. Additionally, tier 3 locks have the added capability to model gate or approach area interference as well. Multiple lockage polices are available for use at the lock chamber level for tier three locks. While first and second tier locks use a First-In-First-Out (FIFO) queue, third tier locks chambers can be defined to utilize FIFO, shortest processing time and n-up/m-down lockage policies as well.

7. CUT COUNTING

Within the BasinSym, several of the lock simulation models require determination of the number of operations or 'cuts' that will be required to completely pass a tow in order to simulate the lockage, which is a function of the number and size of barges in a tow, and the manner in which the barbes can be packed into a lock chamber. The development team originally employed a min/max technique for simulation of cut-level lockages in which a pseudo-barge whose dimensions are determined based on greatest overall beam and LOA was tiled into the chamber to determine the cut count. While this method worked well when tows of uniform composition were processed, it would overestimate the cut count, potentially drastically, with tows of heterogeneous barge compositions. In order to reduce this estimation error, a new technique based on bin packing has been implemented.

7.1. Chamber Packing

The problem, put simply, is how to fill a rectangular space (the chamber) the fewest number of times given a required set of rectangles that must be included (the tow). Computer science lends us candidate solutions in the form of a family of spatial packing algorithms known as strip packing algorithms. Within the strip packing family of algorithms, there are two sub-families: online and offline. In online strip packing, a stream of rectangles is placed in the strip as they are presented to the algorithm. An example of this would be the loading of a truck with parcels from a conveyer belt. The offline sub-family of algorithms assume that the entire input stream is fully known to the processor prior to placement of the first rectangle. An example of this would be the packing of a truck where the loader is shown the pile of boxes to be loaded. Given that a tow operator has knowledge of locks that a tow will pass through, advanced knowledge of the tow configuration prior to lockage leads to the selection of an offline strip packing algorithm. The First Fit, Decreasing Height (FFDH) algorithm (Cardiff University 2008), was selected due to a balance between processing time and packing efficiency. In this algorithm, the barges are sorted based on length with the longest barge first in the list. The list of barges is processed in order (longest to shortest). A "strip" is created with its maximum length being that of the chamber. On that strip, "levels" are created. The maximum width of a level is the width of the chamber while the length of the level is determined by the first barge placed on that level.. The barge at the head of the processing list is placed in the first level that can accommodate its width. In the event that there is insufficient width in any of the existing levels, a new level is created if there is sufficient remaining length on the strip. If there is not sufficient remaining length, the strip is ended, cleared, and the cut count incremented. This is the basic functionality of the FFDH strip packing algorithm.

Within BasinSym, a distinction is made between a chamber with some form of assistance and a chamber without. In a chamber featuring assistance, the power vessel is included in only one cut while chambers without any form of assistance require the presence of the power vessel on each cut. For reasons of simplicity, the power vessel is processed in the first cut.

7.2. Chamber Packing Example

The step-wise execution that results from the specific implementation of the FFDH algorithm in BasinSym is as follows:

1. The barges are sorted based on decreasing length yielding a processing list of: A1, A2, A3, A4, B1, B2, B3. The power vessel is not included in the processing list.

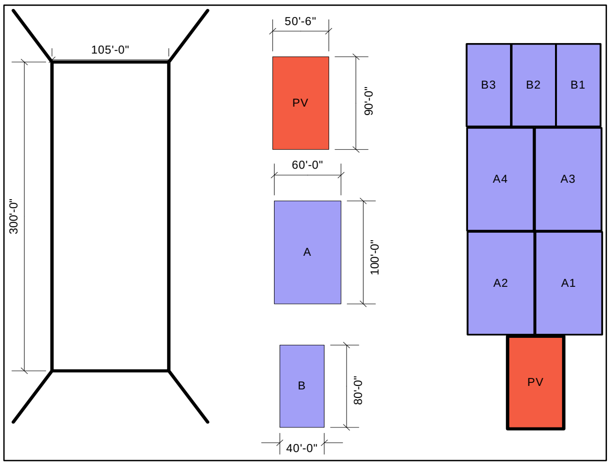


Figure 3: Chamber Parking – Step #1

2. The power vessel is placed in a new level on the strip – L1. L1 90' long and has an available width of 105'-50.6'.

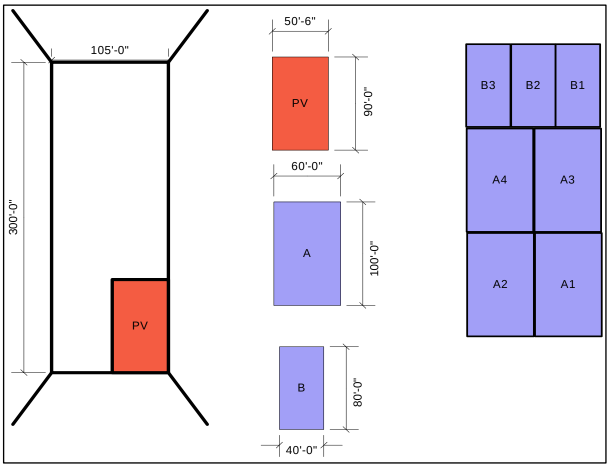


Figure 4: Chamber Packing – Step #2

3. Barge A1 is examined. It cannot fit on L1 due to length so a new level is created on the strip – L2. L2 is 100' long and has an available width of 105'-60'. The strip now has 300'-90'-100' of available length.

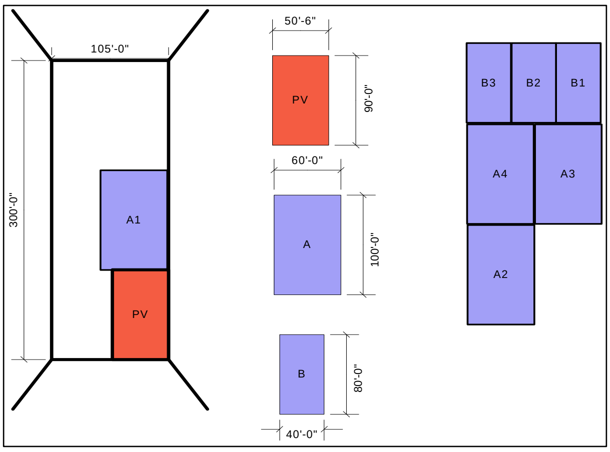


Figure 5: Chamber Packing – Step #3

4. Barge A2 is examined. It cannot fit on L1 due to length. It does not fit on L2 as the available width is insufficient. A new level L3 is created and A2 is placed on it. L3 is 100' long and has an available width of 105' – 60'. The strip now has 300' – 90' -100' -100' (10' available).

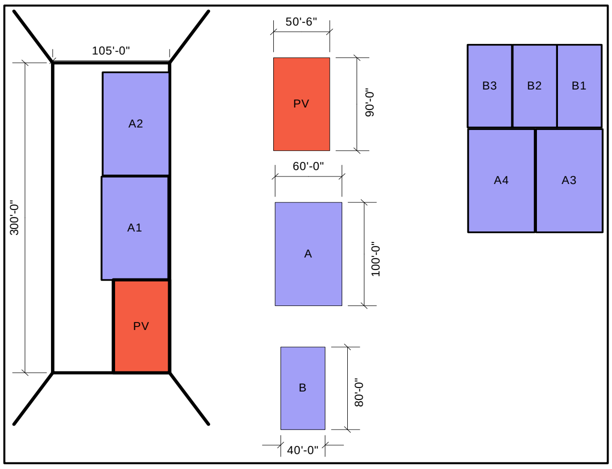


Figure 6: Chamber Packing – Step #4

5. Barge A3 is examined. It cannot fit on any of the existing levels and a new level would exceed the maximum strip length. The strip is cleared and the cut count incremented to 1. L1 is created on the strip with a length of 100' and an available width of 45'. The power vessel is not reintroduced as assistance does not require it.

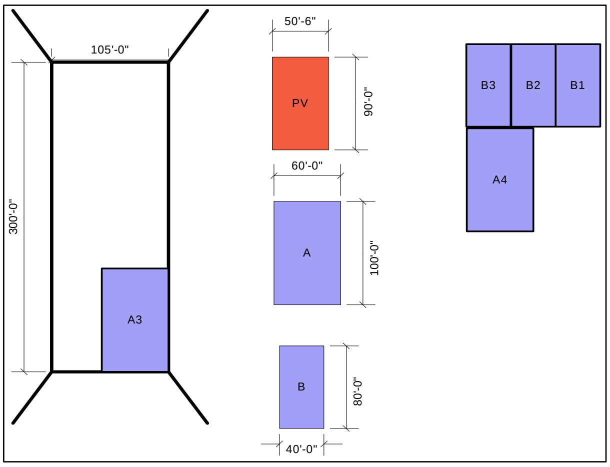


Figure 7: Chamber Packing – Step #5

6. Barge A4 is examined. It cannot fit on L1 as there is insufficient remaining width. L2 is created with a length of 100' and a remaining width of 45'. The strip now has 100' of available length.

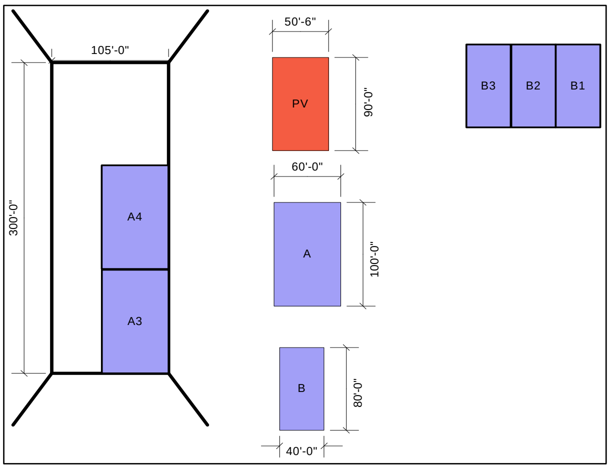


Figure 8: Chamber Packing – Step #6

7. Barge B1 is examined. It fits on L1 leaving L1 with 45' – 40' of available width.

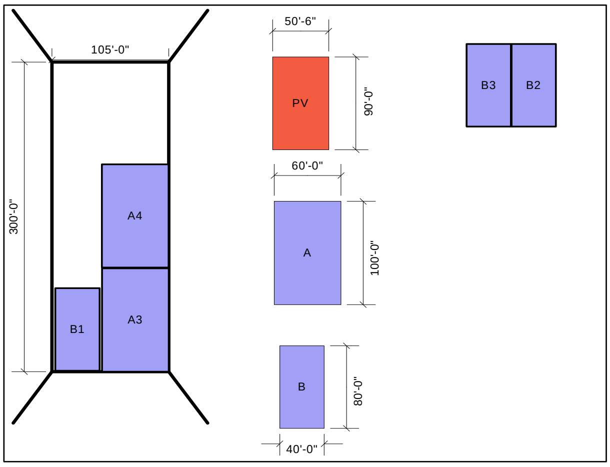


Figure 9: Chamber Packing – Step #7

8. Barge B2 is examined. It does not fit on L1 due to insufficient width but does fit on L2 leaving L2 with 5' of available width.

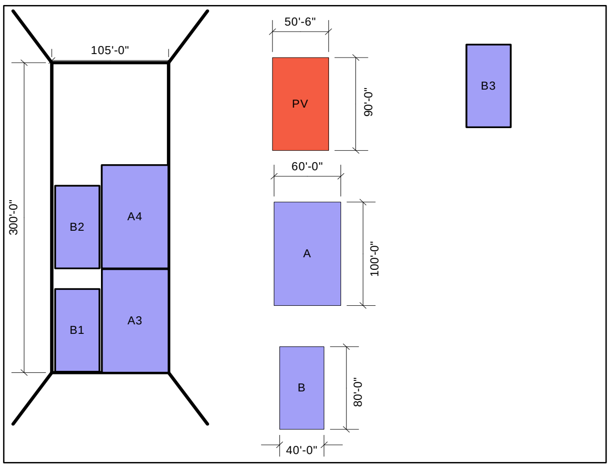


Figure 10: Chamber Packing – Step #8

ISBN 978-84-692-5416-5

9. Barge B3 is examined. It does not fit on either L1 or L2 due to insufficient width. L3 is created with a length of 80’ and an available width of 105’ – 40’. The strip has 300’ – 100’ -100’ -80’ of available length.

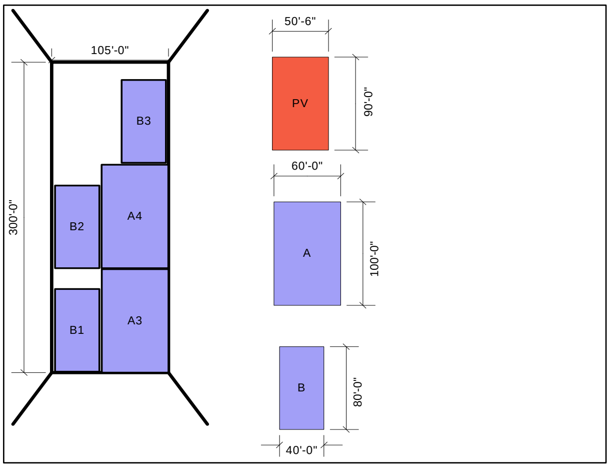


Figure 11: Chamber Packing – Step #9

The barge processing list is now empty. The cut count is incremented as the strip is not empty. This yields a cut count of 2 in our example. If assistance had not been available at the chamber, the power vessel would have been inserted each time a new strip (cut) had been created.

7.3. Implementation Details

The process of cut counting is not trivial in terms of computational resources. In a simulation consisting of hundreds of iterations over a 50 year horizon in which each year could have tens of thousands of lockages, this represents a tremendous amount of time expended. In consultation with USACE field experts, several assumptions have been made that allow for potentially huge savings in computational resources through the reuse of previous results. The first assumption is that the specific type of barge does not affect how it is packed in the chamber. This assumption can be reduced to “a barge that is X feet long by Y feet wide is a barge X feet long by Y feet wide”. The second simplifying assumption is that cut count does not vary by particular chamber or location for chambers of a given dimension.

The data structure employed for this is a tiered dictionary of cut counts based on a chamber dimension signature and a tow composition signature based on barge dimensions. This dictionary is persistent across simulations thus eliminating the need to count cuts unless a previously unseen combination of chamber dimensions and tow is encountered at which point the information is determined and never recalculated.

The chamber signature is a string used to identify chambers of equivalent dimension and assistance availability. In order to be more open to inspection, this string is quite plainly human readable. The components for the chamber signature are length, width, and assistance with each element separated by an ‘x’ and the entire triplet enclosed in parens. An example for a chamber that is 100.5 feet long by 56 feet wide without assistance would be: (100.5x56xNOASSIST). Thus, the signatures of dimensionally equivalent chambers are identical.

8. RELIABILITY

Reliability is an important issue, as many of the locks on US waterways are more than 70 years old, well beyond their initial design life. Reliability considerations are handled within BasinSym using the concepts of component states and state transitions. Under this approach, at minimum one component is assigned to each chamber. The component can be in one of a number of states, with transitions between the states defined probabilistically, based on component state transition functions (Males 2004).

A lock is represented as a hierarchy, consisting of a lock, composed of one or more chambers. Each chamber is composed of one or more components. There is no particular physical definition or behavior associated with a component—it is an abstract general concept. A component is simply something that can fail and whose current status participates in determining the overall performance of the lock. This concept allows the modeling effort to focus on the specific components of interest. Each component has an associated value of “age” and operating “cycles,” that is incremented as the simulation proceeds and that can be reset by a component repair or rehabilitation iteration.

In order to make data handling and modeling feasible, the concept of component states is used. Each component can occupy, at any given time, one of a set of user-defined states specific to that component. A miter gate at the entrance/exit of a lock chamber, for example, might be in one of three possible states—excellent, poor and non-operational. A guidewall might be in one of four possible states—very good, medium, poor and highly degraded. Each component can transition between its available states. The probability of moving to another state is a function of the current state. This is generally referred to as a Markov process.

The driving force for moving a component from one state to another state can be either the passage of time or the cycling of the lock due to the passage of a vessel. Thus, over time, components can fail due to age, stresses associated with repeated opening and closing, or a barge can collide with the gate, causing major or minor damage.

8.1. State Change Probabilities

In order to define probabilities associated with a change in state of a component, *state transition probability curves* are associated with the component for each of the event types and failure modes that are expected. These curves are typically referred to as PUP (probability of unacceptable performance) functions which may be defined as either age or cycle based. An example of a cycle based PUP function is shown in figure 12 below.

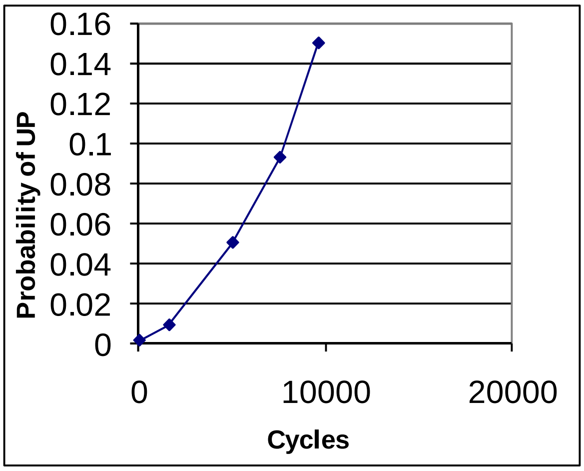


Figure 4: Cycle Based PUP Function

Recall that, for each component, the age and number of cycles are known (and continuously updated during the simulation). Thus, at any given time, when an event takes place that may cause a state transition, it is possible to determine, by curve lookup, the associated current probability of that state transition.

8.2. Failure Modes

For each component state, it is possible to have multiple failure modes, with different probabilities of occurrence. When an event that can trigger a state transition takes place, the current values of the point probabilities of each failure mode are determined from the associated lookups into the PUP functions (based on current component age or cycles). These probabilities are arrayed cumulatively on a probability line, a random number is generated, and the determination is made if there is a failure and, if so, which failure mode should be selected.

For each failure mode, a repair cost and duration are defined. The repair for a given failure can involve changes to more than the component that underwent failure. For example, if a gate fails, then the electrical system component can also be replaced. The user defines all the component repairs associated with each failure mode. Each such component repair involves setting the current state for that component and revising the current age and cycle.

If the particular failure mode is activated, then the chamber is out of service for the duration period and the associated cost is added to the economic analysis. At the end of the repair, the component states, ages and cycles are re-set to the user-defined values. The post-repair state can be the same as the current state or it can be different. For example, a minor repair can simply consume time and dollar resources, without making a significant enough performance change to move to a different component state.

The resetting of age and cycles associated with a repair can be either absolute or relative. If the values for post-repair age and cycle are positive, then the component's current values are set to these values. If, however, either of the values is negative, then the change is made relative to the current value. For example, if the current age of a component that is being repaired is 25 years, and the user entered value post-repair is 10, then the age is reset to 10 years. However, if the entered value is –5, then the post-repair age becomes 20 (25-5).

8.3. Scheduled Rehabilitation

A rehabilitation (rehab) is a scheduled outage at the end of which component state changes (or age/cycle changes) take place. Rehab events take place at user-defined times, with specified costs and outage durations and associated component rehabs, exactly analogous to component repairs in that a new post-rehab state, age and cycle are specified. Each set of rehab events is grouped into a rehab plan that can be activated when the simulation is run. Thus, multiple rehab plans (alternatives) consisting of different combinations of component rehabs can be stored and tested.

This approach is implemented for each lock/chamber, handling the level of detail through data. That is, at minimum each chamber at a lock is represented by a single component. Some locks, e.g., a lock at which a rehab study is being done, would be represented by more components, with more states and failure modes. A “simple” component for another lock might have two states, a single failure mode and a single associated repair.

8.4. Performance Penalties

The concept of a performance penalty is used to relate the state of the components to the performance (in terms of time spent serving vessels) of a chamber. At any given time, the total performance penalty that is added to the lockage time is calculated as the sum of the individual state-based performance penalties, given the state that each component is occupying. The performance penalty is probabilistic and associated with a particular stage of locking (e.g., approach, entry, chambering or exit). Thus, a component representing a set of valves would apply the performance penalty to the chambering or chamber turnback stage, as would the gate performance penalties.

9. DATA DEVELOPMENT

Data for development of the shipment lists comes primarily from the Corps Lock Performance Monitoring System (LPMS) and the Operations and Maintenance of Navigation Information (OMNI) system, corporate data warehouses that store information on every lock transit under Corps jurisdiction. As part of the NaSS suite of tools, a data extraction and analysis tool was developed for the LPMS / OMNI data stores and subsequently named the Data Analysis Pre-Processing (DAPP) module. The DAPP takes as input a database schema from the external data warehouses and provides a set of tools for developing shipment lists for input into the simulation model based on historical movements, statistics and commodity demands. There are three types of shipment lists available, historical – for calibrating the model; statistical – for using historical movements to create synthetic movements through distributions; and commodity driven – for developing a synthetic fleet to satisfy a projected demand of commodities at the dock level. Once a particular shipment list is exported from the DAPP it is ready for use as the simulation driver within BasinSym.

The DAPP includes utilities to produce distributions which feed the probabilistic data requirements of the BasinSym. The DAPP is able to quickly analyze a subset of traffic information for a set of locks and provide cumulative distribution functions to represent lock transit time based on a user defined set of parameters, such as vessel class, lockage type, number of cuts and date range.

10. CONCULSION

The NaSS suite of tools can be used to perform a basin-level analysis of inland waterway traffic and compare baseline conditions to a set of alternative projects. This information provides the basis of an economic analysis required by the USACE for evaluation of investment decisions. Detailed output logs are available once the simulation completes as well as .csv formatted outputs of transit times by reach, vessel class and lock.

The NaSS suite has been initially populated with historic vessel movement data from 2007 for the Ohio River basin, which includes over 2,500 specific shipments for the one year period. The NaSS suite was successfully used to analyze the potential benefits of conducting concurrent, as opposed to sequential, lock chamber closures on the Ohio River. The analysis showed concurrent lock closures would save about $1.3 million in delay cost at Byrd-Greenup Lock and Dam's compared to sequential closures at those locks. The analysis further showed concurrent closures at Myers-Smithland Lock and Dam's would save only about $126,000 due to the main and auxiliary chambers at Smithland being of identical size. The entire analysis was performed within a 2 week timeframe and the results were subsequently used to help set the 2009 lock and dam maintenance schedule on the Ohio River.

ACKNOWLEDGMENTS

The authors wish to express their sincerest gratitude to Dr. Richard Males of RMM Technical Services in Cincinnati, Ohio, USA who was responsible for the initial design document that served as the blueprint for this model. Mr. Mark Lisney of the US Army Corps of Engineers Institute of Water Resources serves as the project manager of the NASS suite and was an invaluable resource as a subject matter expert on inland waterway movements. Model development for NASS was carried out by Mr. William Woelbeling and Mr. Michael Dean of Camp, Dresser, McKee, Inc. in Carbondale, Illinois, USA.

REFERENCES

American Mariner LLC. 2008. *Inland Waterways*. Available from: http://www.americanmariner.net/11_inland_waterways.php [accessed 11 April 2009]

Cardiff University, 2008. School of Computer Science, Level and Shelf Algorithms. Available from: http://users.cs.cf.ac.uk/C.L.Mumford/heidi/Approaches.html [accessed 8 March 2008]

Hart, P. E.; Nilsson, N. J.; Raphael, B. (1968). "A Formal Basis for the Heuristic Determination of Minimum Cost Paths". *IEEE Transactions on Systems Science and Cybernetics* SSC4 4 (2): 100–107.

Males, R., 2004. *LockSym Draft Documentation,* Available from US Army Corps Of Engineers Institute for Water Resources.

Waterborne Commerce Statistics Center, New Orleans, Louisiana, USA. 2009. *Final Waterborne Commerce Statistics For Calendar Year 2007.* Available from: http://www.iwr.usace.army.mil/ndc/wcsc/wcsc.htm [accessed 16 April 2009]

AUTHORS BIOGRAPHY

Keith D. Hofseth

Mr. Hofseth has served as an economist at the Corp of Engineers for over 24 years; conducting studies in flood control, deep draft navigation, inland navigation and hydropower. He co-authored the first guidance for major rehabilitation analysis and was a pioneer in bringing risk and uncertainty into the Corp's planning process. Mr. Hofseth started with the Corps in the Seattle district and later served as the Chief Economist for the Alaska District. He is currently the Technical Director of the Navigation Economic Technologies research program at the Institute for Water Resources and is a graduate of the Corps' Planners Associate Program. Mr. Hofseth holds a B.S. degree in Finance and Economics from the University of Montana and an M.A. degree in Applied Economics from Johns Hopkins University.

Cory M. Rogers

Mr. Rogers is a computer scientist currently serving as the National Practice Leader for Applications Development at Camp Dresser & McKee, Inc. He has 15 years of experience in the design and development of software applications to solve complex decision making problems. He has managed teams to develop successful software applications ranging from single-user desktop applications to web-enabled systems with a user base of thousands. Prior to joining Camp Dresser McKee, Inc, he served as the Director for Software Engineering for a small consulting firm and was a founding member of the Curriculum Advisory Board for the Information Systems Technology program at Southern Illinois University. He has worked on numerous decision support software models which are currently being used to perform crucial functions such as estimating flood damages caused by coastal storms, investigate the economic feasibility of improvements to the U.S. inland waterway and port infrastructure and manage financial risk at global food and beverage companies.

A STATE OF THE ART ON DISCRETE EVENT SIMULATION AND AN APPLICATION TO SUPPLY CHAIN INVENTORY MANAGEMENT

Francesco Longo, Giovanni Mirabelli

Modeling & Simulation Center, Laboratory of Enterprise Solutions (MSC-LES)
Mechanical Department, University of Calabria
Via P. Bucci, Cubo 44C, 87036 Rende, Italy

f.longo@unical.it, g.mirabelli@unical.it – URL www.msc-les.org

ABSTRACT
Nowadays there is a large availability of discrete event simulation software that can be easily used in different domains: from industry to supply chain, from healthcare to business management, from training to complex systems design. Simulation engines of commercial discrete event simulation software use specific rules and logics for simulation time and events management. They also provide the user with a number of libraries based on the traditional object oriented modeling approach. Difficulties and limitations come up when commercial discrete event simulation software are used for modeling complex real world-systems (i.e. supply chains, industrial plants). The objective of this paper is twofold: first a state of the art on commercial discrete event simulation software and an overview on discrete event simulation models development by using general purpose programming languages are presented; then a Supply Chain Order Performance Simulator (SCOPS, developed in C++) for investigating the inventory management problem along the supply chain under different supply chain scenarios is proposed to readers.

Keywords: Discrete Event Simulation, Simulation languages, Supply Chain, Inventory Management

1. INTRODUCTION

As reported in Banks (1998), discrete-event simulation software selection could be an exceeding difficult task especially for inexpert users. Simulation software selection problem was already known many years ago; Banks and Gibson (1997) propose a simulation buyer's guide identifying possible features to consider in simulation software selection. Among others Banks and Gibson (1997) include in their analysis considerations about Input, Processing, Output, Environment, Vendor and Costs. Hlupic (1999) presents a survey on users' requirements about discrete-event simulation software. The analysis shows that simulation software with good visualization/animation properties are easier to use but limited in case of complex and non-standard problems and. Further limitations include lack of software compatibility, output analysis tools, advanced programming languages.

Swain (2005) and Swain (2007) report functionalities and potentialities of different commercial discrete-event simulation software, in order to support users in software selection. In this case the author provides the reader with information about software vendor, primary software applications, hardware platform requirements, simulation animation, support, training and pricing.

Needless to say that Modeling & Simulation should be used when analytical approaches do not succeed in identifying proper solutions for analyzing complex systems (i.e. supply chains, industrial plants, etc.). For many of these systems, simulation models must be: (i) flexible and parametric (for supporting scenarios evaluation) (ii) time efficient (even in correspondence of very complex real-world systems) and (iii) repetitive in their architectures for scalability purposes (Longo and Mirabelli, 2008).

Let us consider the traditional modeling approach proposed by two commercial discrete event simulation software, Em-Plant by Siemens PLM Software solutions and Anylogic by Xj-Technologies. Both of them propose a typical object oriented modeling approach. Each discrete event simulation model is made up by system state variables, entities and attributes, lists processing, activities and delays. Usually complex systems involve high numbers of resources and entities flowing within the simulation model. The time required for executing a simulation run depends on the numbers of entities in the simulation model: the higher is the number of entities the higher is the time required for executing a simulation run. In addition, libraries objects, which should be used for modeling static entities, very often fall short of recreating the real system with satisfactory accuracy.

In other words, the traditional modeling approach (proposed by eM-Plant and Anylogic as well as by a number of discrete event simulation software), presents two problems: (i) difficulties in modeling complex scenarios; (ii) too many entities could cause computational heavy simulation models.

An alternative to commercial discrete event simulation software is to develop simulation models based on general purpose programming languages (i.e. C++, Java). The use of general purpose programming languages allows to develop ad-hoc simulation models with class-objects able to recreate carefully the behavior of the real world system.

The objective of this paper is twofold: first a state of the art on commercial discrete event simulation software and an overview on discrete event simulation models development by using general purpose programming languages are presented; then a Supply Chain Order Performance Simulator (SCOPS, developed in C++) for investigating the inventory management problem along the supply chain under different supply chain scenarios is proposed to readers.

Before getting into details of the work, in the sequel a brief overview of paper sections is reported. Section 2 provides the reader with a detailed description of different commercial discrete event simulation software. Section 3 presents a general overview of programming languages and describes the main steps to develop a simulation model based on general purpose programming languages. Finally, Section 4 presents a three stages supply chain simulation model (called SCOPS) used for investigating inventory problems along the supply chain (simulation experiments carried out by using the simulation model analyze the fill rate behavior under different demand intensity, demand variability and lead times levels). Finally the last section reports conclusions and research activities still on going.

2. DISCRETE EVENT SIMULATION SOFTWARE

Table 1 reports the results of a survey on the most widely used discrete event simulation software (conducted on 100 people working in the simulation field). The survey considers among others some critical aspects such as domains of application (specifically manufacturing and logistics), 3D and virtual reality potentialities, simulation languages, prices, etc. For each aspect and for each software the survey reports a score between 0 and 10. Table 1 help modelers in discrete event simulation software selection.

Moreover the following sections reports a brief description of some software of table 1 (Anylogic, Arena, Automod and EM-Plant) in terms of domains of applicability, types of libraries (i.e. modeling libraries, optimization libraries, etc.), input-output functionalities, animation functionalities, etc.

2.1. Anylogic

Anylogic is a Java based simulation software, by XJ Technologies (www.kjtek.com), used for forecasting and strategic planning, processes analysis and optimization, optimal operational management, processes visualization. It is widely used in logistics, supply chains, manufacturing, healthcare, consumer markets, project management, business processes and military.

Anylogic supports Agent Based, Discrete Event and System Dynamics modeling and simulation. The latest Anylogic version (Anylogic 6) has been released in 2007, it supports both graphical and flow-chart modeling and provides the user with Java code for simulation models extension.

For input data analysis, Anylogic provides the user with Stat-Fit (a simulation support software by Geer Mountain Software Corp.) for distributions fitting and statistics analysis.

Output analysis functionalities are provided by different types of datasets, charts and histograms (including export function to text files or excel spreadsheet). Finally simulation optimization is performed by using Optquest, an optimization tool integrated in Anylogic.

2.2. Arena

Arena is a simulation software by Rockwell Corporation (http://www.arenasimulation.com/) and it is used in different application domains: from manufacturing to supply chain (including logistics, warehousing and distribution) from customers service and strategies to internal business processes.

Arena (as Anylogic) provides the user with objects libraries for systems modeling and with a domain-specific simulation language (SIMAN). Simulation optimizations are carried out by using Optquest.

Arena includes three modules respectively called Arena Input Analyzer (for distributions fitting), Arena Output Analyzer (for simulation output analysis) and Arena Process Analyzer (for simulation experiments design).

Table 1 – Survey on most widely used Simulation software

	Arena	Witness	Promodel	AutoMod	Flexsim	Emplant	Anylogic
Logistic	7.5	7.5	6.5	7	7	7.2	6.5
Manufacturing	7.5	7.5	6.7	6.5	6.7	7.2	6.6
3D Virtual Reality	6.9	7	6.7	7.3	7.2	6.8	6.6
Simulation Engine	8	8	7	7.5	7.5	8	7
User Ability	8	8	9	6	7.5	7	7
User Community	9	8.5	7.5	6.7	6.6	6.5	6.2
Simulation Language	7	6.5	6.5	6.25	6.7	6.5	6.8
Runtime	7	7	7.5	6.5	6	6.5	7.5
Analysis tools	8	7.8	7.7	6.9	6	7.1	6.5
Internal Programming	7	6.5	6.2	6	7	7	7.2
Modular Construction	7	7	7.5	6	7	6.5	6.1
Price	6	6	7	5.6	5.7	5.8	7

Moreover Arena also provides the users animation at run time as well as it allows to import CAD drawings to enhance animation capabilities.

2.3. Automod

Automod is a discrete event simulation software, developed by Applied Materials Inc. (http://www.automod.com/) and it is based on the domain-specific simulation language Automod.

Typical domains of application are manufacturing, supply chain, warehousing and distribution, automotive, airports and semiconductor. It is strongly focused on transportation systems including objects such as conveyor, Path Mover, Power & Free, Kinematic, Train Conveyor, AS/RS, Bridge Crane, Tank & Pipe (each one customizable by the user). For input data analysis, experimental design and simulation output analysis, Automod provides the user with AutoStat. Moreover the software includes different modules such as AutoView devoted to support simulation animation with AVI formats.

2.4. Em-Plant

Em-plant is a Siemens PLM Software solutions (http://www.emplant.com/), developed for strategic production decisions. EM-Plant enables users to create well-structured, hierarchical models of production facilities, lines and processes. Em-Plant object-oriented architecture and modeling capabilities allow users to create and maintain complex systems, including advanced control mechanisms.

The Application Object Libraries, support the user in modeling complex scenarios in short time Furthermore EM-Plant provides the user with a number of mathematical analysis and statistics functions for input distribution fitting and single or multi-level factor analysis, histograms, charts, bottleneck analyzer and Gantt diagram. Experiments Design functionalities (with Experiments Manager) are also provided.
Simulation optimization is carried out by using Genetic Algorithms and Artificial Neural Networks.

3. GENERAL PURPOSE AND SPECIFIC SIMULATION PROGRAMMING LANGUAGES

There are many programming languages, general purpose or domain-specific simulation language (DSL) that can be used for simulation models development. General purpose languages are usually adopted when the programming logics cannot be easily expressed in GUI-based systems or when simulation results are more important than advanced animation/visualization (Babich and Bylev 1991). Simulation models can be developed both by using discrete-event simulation software and general purpose languages, such as C++ or Java (Pidd and Cassel, 2000).

As reported in Banks (1998) a simulation study requires a number of different steps; a simulation study starts with problem formulation and passes through different and iterative steps: conceptual model definition, data collection, simulation model implementation, verification, validation and accreditation, simulation experiments, simulation results analysis, documentation and reports.

Simulation model development by using general purpose programming languages (i.e. C++) requires a deep knowledge of the logical foundation of discrete event simulation. Among different aspects to be considered, Schriber and Brunner (1998) say that discrete event simulation model consist of entities, resources control elements and operations.

Dynamic entities flow in the simulation model (i.e. parts in a manufacturing system, products in a supply chain, etc.), static entities usually works as resources (a system part that provides services to dynamic entities). Control elements (such as variables, boolean expressions, specific programming code, etc.) support simulation model states control. Finally operations represent all the actions generated by the flow of dynamic entities within the simulation model.

During its life within the simulation model, an entity changes its state different times. According to Schriber and Brunner (1998) there are five different entity states: Ready state (the entity is ready to be processed), Active state (the entity is currently being processed), Time-delayed state (the entity is delayed until a predetermined simulation time), Condition-delayed state (the entity is delayed until a specific condition will be solved) and Dormant state (in this case the condition solution that frees the entity is managed by the modeler)

Entity management is supported by different lists, each one corresponding to an entity state: the CEL, (Current Event List for active state entity), the FEL (Future Event List for Time-delayed entities), the DL (Delay List for condition-delayed entities) and UML (User-Managed Lists for dormant entities).

In particular, Siman and GPSS/H call the CEL list CEC list (Current Events Chain), while ProModel language calls it AL (Action List). The FEL is called FEP (Future Events Heap) and FEC (Future Event Chain) respectively by Siman and GPSS/H.

After entities states definition and lists creation, the next step is the implementation of the phases of a simulation run: the Initialization Phase (IP), the Entity Movement Phases (EMP) and the Clock Update Phase (CUP). A detailed explanation of the simulation run anatomy is reported in Schriber and Brunner (1998).

4. A SUPPLY CHAIN SIMULATION MODEL DEVELOPED IN C++

According to the idea to implement simulation models based on general purpose programming languages, the authors propose a three stage supply chain simulation model implemented by using the Borland C++ Builder to compile the code. The acronym of the simulation model is SCOPS (Supply-Chain Order Performance Simulator). SCOPS investigates the inventory management problem along a three stages supply chain and allows the user to test different scenarios in terms of demand intensity, demand variability and lead times.

The supply chain conceptual model includes suppliers, distribution centers, stores and final customers. In the supply chain conceptual model a single network node can be considered as store, distribution center or supplier. A supply chain begins with one or more suppliers and ends with one or more stores. Usually stores satisfy final customers' demand, distribution centers satisfy stores demand and plants satisfy distribution centers demand. By using these three types of nodes we can model a general supply chain (also including more than three stages).

Suppliers, distribution centers and stores work 6 days per week, 8 hours per day. Stores receive orders from customers. An order can be completely or partially satisfied. At the end of each day, on the basis of an Order-Point, Order-Up-to-Level (s, S) inventory control policy, the stores decide whether place an order to the distribution centers or not. Similarly distribution centers place orders to suppliers according to the same inventory control policies. Distribution centers select suppliers according to their lead times (that includes production times and transportation times).

According to the Order-Point, Order-Up-to-Level policy (Silver et al., 1998), an order is emitted whenever the available quantity drops to the order point (s) or lower. A variable replenishment quantity is ordered to raise the available quantity to the order-up-to-level (S). For each item the order point s is the safety stock calculated as standard deviation of the lead-time demand, the order-up to level S is the maximum number of items that can be stored in the warehouse space assigned to the item type considered. For the *i-th* item, the evaluation of the replenishment quantity, $Q_i(t)$, has to take into consideration the quantity available (in terms of inventory position) and the order-up-to-level S. The inventory position (equation 1) is the on on-hand inventory, plus the quantity already on order, minus the quantity to be shipped. The calculation of $s_j(t)$ requires the evaluation of the demand over the lead time. The lead time demand of the *i-th* item (see equation 2), is evaluated by using the moving average methodology. Both at stores and distribution centers levels, managers know their peak and off-peak periods, and they usually use that knowledge to correct manually future estimates based on moving average methodology. They also correct their future estimates based on trucks capacity and suppliers quantity discounts. Finally equations 3 and 4 respectively express the order condition and calculate the replenishment quantity.

$$P_i(t) = Oh_i(t) + Or_i(t) - Sh_i(t) \quad (1)$$

$$Dlt_i(t) = \sum_{k=t+1}^{t+LT_i} Df_i(k) \quad (2)$$

$$P_i(t) < (s_i(t) = SS_i(t)) \quad (3)$$

$$Q_i(t) = S_i - P_i(t) \quad (4)$$

- $P_i(t)$, inventory position of the *i-th* item;
- $Oh_i(t)$, on-hand inventory of the *i-th* item;
- $Or_i(t)$, quantity already on order of the *i-th* item;
- $Sh_i(t)$, quantity to be shipped of the *i-th* item;
- $Dlt_i(t)$, lead time demand of the *i-th* item;
- $Df_i(t)$, demand forecast of the *i-th* item (evaluated by means of the moving average methodology);
- LT_i, lead time of the *i-th* item;
- $s_i(t)$, order point at time t of the *i-th* item;
- S_i, order-up-to-level of the *i-th* item;
- $SS_i(t)$, safety stock at time t of the *i-th* item;
- $Q_i(t)$, quantity to be ordered at time t of the *i-th* item;

4.1. Supply Chain Orders Perfomance Simulator

SCOPS translates the supply chain conceptual model recreating the complex and high stochastic environment of a real supply chain. For each type of product, customers' demand to stores is assumed to be Poisson with independent arrival processes (in relation to product types). Quantity required at stores is based on triangular distributions with different levels of intensity and variability. Partially satisfied orders are recorded at stores and distribution center levels for performance measures calculation.

In our application example fifty stores, three distribution center, ten suppliers and thirty different items define the supply chain scenario. Figure 1 shows the SCOPS user interface.

The SCOPS graphic interface provides the user with many commands as, for instance, simulation time length, start, stop and reset buttons, a check box for unique simulation experiments (that should be used for resetting the random number generator in order to compare different scenarios under the same conditions), supply chain configurations (number of items, stores, distribution centers, suppliers, input data, etc.). For each supply chain node a button allows to access the following information number of orders, arrival times, ordered quantities, received quantities, waiting times, fill rates. SCOPS graphic interface also allows the user to export simulation results on txt and excel files.

One of the most important features of SCOPS is the flexibility in terms of scenarios definition. The graphic interface gives to the user the possibility to carry out a number of different what-if analysis by changing supply chain configuration and input parameters (i.e. inventory policies, demand forecast methods, demand intensity and variability, lead times, inter-arrival times, number of items, number of stores, distribution centers and plants, number of supply chain echelons, etc.).

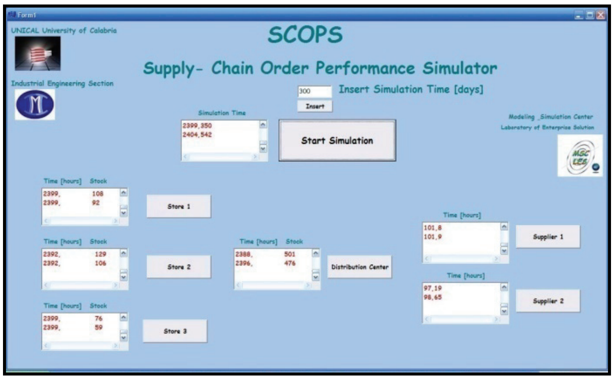


Figure 1: SCOPS User Interface

5. SUPPLY CHAIN CONFIGURATION AND DESIGN OF SIMULATION EXPERIMENTS

The authors propose as application example the investigation of 27 different supply chain scenarios. In particular simulation experiments take into account three different levels for demand intensity, demand variability and lead times (minimum, medium and maximum respectively indicated with "-", "0" and "+" signs). Table 1 reports (as example) factors and levels for one of the thirty items considered and table 3 reports scenarios description in terms of simulation experiments. Each simulation run has been replicated three times (totally 81 replications).

Table 2 – Factors and Levels

	Minimum	Medium	High
Demand Intensity [inter-arrival time]	3	5	8
Demand Variability [item]	[18,22]	[16,24]	[14,26]
Lead Time [days]	2	3	4

After the definition of factors levels and scenarios, the next step is the performance measures definition. SCOPS includes, among others, two fill rate performance measures defined as (i) the ratio between the number of satisfied Orders and the total number of orders; (ii) the ratio between the lost quantity and the total ordered quantity.

Simulation results, for each supply chain node and for each factors levels combination, are expressed in terms of average fill rate (intended as ratio between the number of satisfied Orders and the total number of orders).

Table 3 – Simulation experiments and supply chain scenarios

Run	Demand Intensity	Demand Variability	Lead Time
1	-	-	-
2	-	-	0
3	-	-	+
4	-	0	-
5	-	0	0
6	-	0	+
7	-	+	-
8	-	+	0
9	-	+	+
10	0	-	-
11	0	-	0
12	0	-	+
13	0	0	-
14	0	0	0
15	0	0	+
16	0	+	-
17	0	+	0
18	0	+	+
19	+	-	-
20	+	-	0
21	+	-	+
22	+	0	-
23	+	0	0
24	+	0	+
25	+	+	-
26	+	+	0
27	+	+	+

5.1. Supply Chain Scenarios analysis and comparison

The huge quantity of simulation results allows the analysis of a comprehensive set of supply chain operative scenarios. Let us consider the simulation results regarding the store #1; we have considered three different scenarios (low, medium and high lead times) and, within each scenario, the effects of demand variability and demand intensity are investigated.

Figure 2 shows the fill rate trend at store #1 in the case of low lead time.

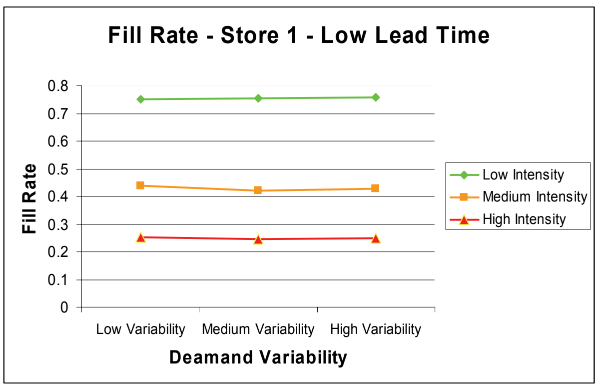


Figure 2: Fill rate at store #1, low lead time

The major effect is due to changes in demand intensity: as soon as the demand intensity increases there is a strong reduction of the fill rate. A similar trend can be observed in the case of medium and high lead time (respectively figure 3 and figure 4).

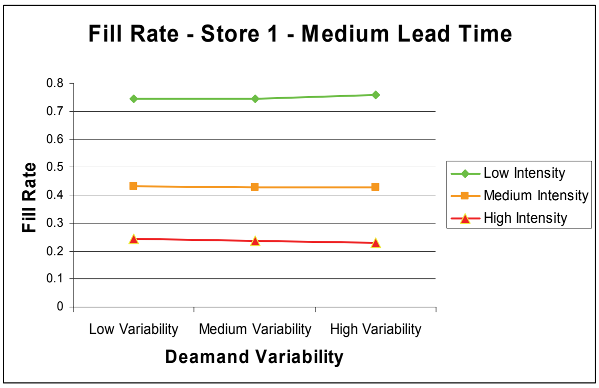


Figure 3: Fill Rate at store # 1, medium lead time

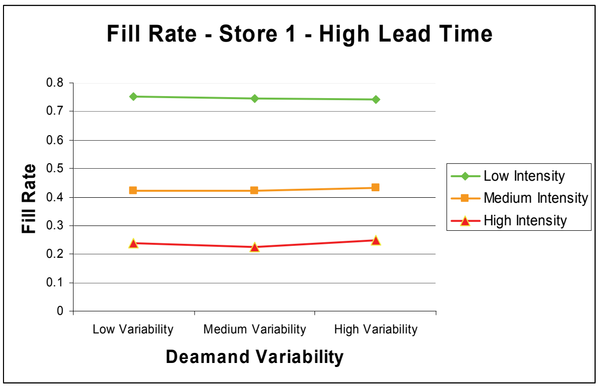


Figure 4: Fill Rate at store # 1, high lead time

The simultaneous comparison of figures 2, 3 and 4 shows the effect of different lead times on the average fill rate. The only minor issue is a small fill rate reduction passing from 2 days lead time to 3 and 4 days lead time.

As additional aspect (not shown in figures 2, 3, and 4), the higher is the demand intensity the higher is the average on hand inventory. Similarly the higher is the demand variability the higher is the average on hand inventory. In effect, the demand forecast usually overestimates the ordered quantity in case of high demand intensity and variability.

CONCLUSIONS

The paper first presents an overview on the most widely used discrete event simulation software in terms of domains of applicability, types of libraries (i.e. modeling libraries, optimization libraries, etc.), input-output functionalities, animation functionalities, etc. In the second part the paper proposes, as alternative to discrete event simulation software, the use of general purpose programming languages and provides the reader with a brief description about how a discrete event simulation model works.

As application example the authors propose a supply chain simulation model (SCOPS) developed in C++. SCOPS is a flexible simulator used for investigating different the inventory management problem along a three stages supply chain. SCOPS simulator is currently used for reverse logistics problems in the large scale retail supply chain.

REFERENCES

Anylogic by XjTech, www.xjtech.com

Arena by Rockwell Corporation, http://www.arenasimulation.com/

Automod by Applied Materials Inc., http://www.automod.com/

Babich V. P. and Bylev A. S. An approach to compiler construction for a general-purpose simulation language. Publisher on Cybernetics and Systems Analysis, Springer New York. ISSN: 1060-0396 (Print) 1573-8337 (Online), Issue: Volume 27, Number 5/ September, 1991.

Banks, J., 1998. *Handbook of simulation, Principles, Methodology, Advances, Application, and Practice.*

Banks, J., and R.G. Gibson , 1997. Simulation software buyer's guide, *IIE Solution,* May, pp 48-54.

Bantz, C.R., 1995. Social dimensions of software development. In: J. A. Anderson, ed. *Annual review of software management and development.* Newbury Park, CA: Sage, 502–510.

Bruzzone, A.G., Bocca, E., Longo, F., Massei, M., 2007. Training and Recruitment in Logistics Node Design by using Web Based Simulation, *International Journal of Internet Manufacturing and Services*, I(1), 2007, pp 32-50.

Carson, J.S., 1997. AutoStat: output statistical analysis for AutoMod users. *Proceedings of the 1997 Winter Simulation Conference,* New York, pp. 649-656.

Curcio, D., Longo, F., 2009. Inventory and Internal Logistics Management as Critical Factors Affecting the Supply Chain Performances. *International Journal of Simulation & Process Modelling*, Vol. 5(2), 127-137.

De Sensi G., Longo, F., Mirabelli, G., 2008. Inventory policies analysis under demand patterns and lead times constraints in a real supply chain, *International Journal of Production Research*, Vol. 46(24), 6997-7016.

Fishman George S. Discrete-Event Simulation: Modeling, Programming, and Analysis. Berlin: Springer-Verlag (2001).

Hhuente, D. J., 1987, Critique of SIMAN as a programming language, *ACM Annual Computer Science Conference* , p 385

Longo, F., Mirabelli, G., 2008. An Advanced supply chain management tool based on modeling and simulation. *Computer and Industrial Engineering*, 54/3: 570-588.

Longo, F., Massei, M., 2008. Advanced Supply Chain Protection & Integrated Decision Support System. *Proceedings of the Second Asia International*

Conference on Modeling & Simulation. Malesya. Kuala Lumpur, 13-15 May, pp. 716-721.

Longo F., Oren T., 2008. Supply Chain Vulnerability and Resilience: A state of the Art Overview. *Proceedings of the European Modeling & Simulation Symposium*. Campora S.Giovanni (CS), Italy, 17-19 September, pp. 527-533.

Pidd M. and Cassel R. A., April 2000. Using Java to Develop Discrete Event Simulations *The Journal of the Operational Research Society*, Vol. 51, No. 4, Progress in Simulation Research (Apr., 2000), pp. 405-412.

Vlatka Hlupic, 1999. Discrete-Event Simulation Software: What the Users Want. *SIMULATION*, Vol. 73, No. 6, 362-370

Schriber T.J., Brunner D.T., 1998. How discrete event simulation work. In: Banks J., Handbook of Simulation, pp. 765 – 811. Wiley Interscience.

Silver, Edward, David F. Pike, Rein Peterson (1998). "Inventory Management and Production Planning and Control", Third Edition, USA: John Wiley & Sons.

Swain J. J., October 2007. New Frontiers in Simulation, Biennial survey of discrete-event simulation software tools, *OR/MS Today*.

Swain J. J., 2005.Gaming Reality: Biennial survey of discrete-event simulation software tools. *OR/MS Today*, Vol. 32, No. 6, pp. 44-55.

Reisdorph Kent and Henderson Ken. Borland C++ Builder. Apogeo. ISBN: 8873033717 (2005).

AUTHORS BIOGRAPHY

FRANCESCO LONGO took the degree in Mechanical Engineering from University of Calabria (2002) and the PhD in Industrial Engineering (2005). He is currently scientific responsible of the MSC-LES (Modeling & Simulation Center – Laboratory of Enterprise Solutions). He is also Associate Editors of the international journal Simulation: Transaction of SCS. He supports the organization of the most important conferences on M&S (SCSC, EMSS, MAS, etc.). His research interests include modeling & simulation of manufacturing systems and supply chain management, supply chain vulnerability, resilience and security, DOE, ANOVA.

GIOVANNI MIRABELLI was born in Rende (Italy), on January the 24th, 1963. He took the degree in Industrial Technology Engineering from University of Calabria. He is currently researcher at the Mechanical Department (Industrial Engineering Section) of University of Calabria. His research interests include Modeling & Simulation for workstations effective design, work measurement and human reliability.

IMC CONTROL APPROACH FOR SUPPLY CHAIN MANAGEMENT

Carlos Andrés García [a], Ramón Vilanova [b], Carles Pedret[c]

[a)(b)(c] Department of Telecommunication and Systems Engineering Autonomous University Of Barcelona
Bellaterra, 08193, Spain

[a]CarlosAndres.Garcia@uab.cat , [b]Ramon.Vilanova .@uab.cat, [c]Carles.Pedret .@uab.cat

ABSTRACT
In this work, a discrete time series model of a supply chain system is derived using material balances and information flow. Transfer functions for each unit in the supply chain are obtained by z-transform. The entire chain can be modelled by combining these transfer functions into a close loop transfer function for the network. Controllers are designed using frequency analysis. The bullwhip effect, magnification of amplitudes of demand perturbations from the tail to upstream levels of the supply chain, is a very important phenomenon for supply chain systems. Furthermore, we show that by implementing a proportional integral or an IMC inventory control and properly synthesizing the controller parameters, we can effectively suppress the bullwhip effect. Moreover, the IMC control structure is superior in meeting customer demand due to its better tracking of long term trends of customer demand.

Keywords: supply chain, beer game, z-transform
bullwhip effect, PI control.

1. INTRODUCTION

A supply chain includes all the participants and processes involved in the satisfaction of customer demand: transportation, storages, wholesales, distributors and factories (Dejonckheere, Disney, Lambrecht and Towill 2002, Dejonckheere, Disney, Lambrecht and Towill 2004). A large number of participants, a variety of relations and processes, dynamics and the randomness in material and information flow prove that supply chains are complex systems in which coordination is one of the key elements of management.

A trouble and important phenomenon in supply chain management, known as the bullwhip effect, suggests that demand variability increases as one goes up in the supply chain (Hoberg, James, Bradley, Ulrich and Thonemann 2007, Marko, Rusjan 2008)

The causes of the bullwhip effect can be due to the forecasting demand, the lead times, order batching, supply shortages, or price fluctuations. We will mainly address the non-zero lead times and particularly the forecasting demand. The supply chain has attracted much attention among process system engineering researchers recently. There are many aspects in supply chain research.

One area is the analysis of the logistic problem of a supply chain using system control theory. In this paper the focus is on the analysis and control of the material balance and information flow among the system. In order to demonstrate the existence of bullwhip effect the beer game was created at the beginning of the sixties at the School of Management, Massachusetts Institute of Technology (MIT) (Ilhyung and Mark 2007). The game simulates a multi-echelon serial supply chain consisting of a Retailer (R), a Wholesaler (W), a Distributor (D) and a Factory/Manufacturer (M) We model the basic protocol of the ”beer distribution game”. The mathematical model used is the transfer function which represents the relation between the input and output of a linear time invariant system (LTI). In supply chain case we are dealing with discrete signals and system, therefore our theoretical analysis was performed by the z-transform. In control systems engineering, the transfer function of a system represents the relationship describing the dynamics of the system under consideration. It algebraically relates a systems output to its input and in this paper is defined as the ratio of the z-transform of the output variable to the z-transform of the input variable.

The Proportional-Integrative-Derivative (PID) controllers are without doubt the most extensive option that can be found on industrial control applications Ramon (2008). Their success is mainly due to their simple structure and meaning of the corresponding three parameters. This act makes the PID control easier to understand by the control engineer than most other advanced control techniques. An analysis of the effect of a P and PI controllers approach as an ordering strategy was made. Then, parameters values for stability and mitigate bullwhip was obtained. The derivate action is not analysed because this is not appropriate for noise systems. The internal model control structure has been introduced as an alternative to the feedback structure. Its main advantage is that closed-loop stability is assured simply by choosing a stable IMC controller. An analysis of the different controllers was made respect to its tuning simplicity, performance and mitigating of bullwhip effect.

2. BEER GAME MODEL

Let us consider a basic beer supply chain as shown in Figure (1). There are four logistic echelons: Retailer(R), wholesales (W), distributor (D), and Factory (F). We assume the following sequence of events: in each period t, the retailer first receives goods, then demand is observed and satisfied (if not backlogged), next, the retailer observes the new inventory level and finally places an order on the manufacturer. Let $I_i(k)$ denote the net stock inventory (the difference between on-hand

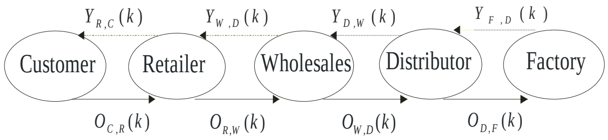


Figure 1. The Block Diagram Of Supply Chain

inventory and backorders) of a echelon of chain I(R,W,D, F) at any discrete time instant k. We also let $Y_{i,i-1}(k)$ indicate the amount of goods to be delivered to node $i-1$ by the upstream node i at the instant k, $O_{i-1,i}(k)$ indicates the demand received by node i from downstream $i-1$. A time delay of L due to transport is assumed for all delivery of goods so that goods dispatched for an upstream $i+1$ at time k will arrive at time $k+L$ at node i. However, due to the need for examination and administrative processing, this new delivery is only available to the node i at $k+L$. The orders placed in the upstream for the i node is denoted for $O_{i,i+1}(k)$

Coupling between two nodes through the ordering policy is eliminated when there is insufficient stock in any one node along the chain (Pin, David, Shan, Shi, Jang, Shyan and Zheng 2004). Therefore, propagation of demand fluctuations is only possible when every node has sufficient stock. Then in this work assume that the upstream $i+1$ supplier has sufficient inventory so that the orders of node i are always satisfied so that the amount of goods delivered by upstream $i+1$ in instant k. $Y_{i+1,i}(k)$ is equal to the order made for the downstream i in a previous time $O_{i,i+1}(k+L)$. It's also assumed the amount of goods delivered by upstream i at instant k. $Y_{i,i-1}(k)$ is equal to the order made for the downstream $i-1$ in a previous time $O_{i-1,i}(k)$.

The result of the integration of the difference between amount goods that entry from upstream node $i+1$ and amount goods that dispatched for downstream node $i-1$ is known as inventory balance, this has a role as a buffer to absorb the demand variability. In other words, the inventories should have stabilizing effect in material flow patterns. The equation for inventory balance at node i is given by:

$$I_i(k) = I_i(k-1) + O_{i,i+1}(k-L) - O_{i-1,i}(k) \qquad (1)$$

The manager aims at maintaining a certain inventory level without generating aggressive changes on orders, using the information signals. The control signal denoted by $O_{i,i+1}(k)$ can be the result of a feedback of any output and a proposed controller. For example, a inventory feedback and a simple P-control can be used. We assume that ordering information is communicated instantly. Hence, the order placed by the node i at the upstream $i+1$ is given by:

$$O_{i,i+1}(k) = C(k)(I_i^T(k) - I_i(k)) \qquad (2)$$

Where $C(k)$ denote the controller function and $I_I^T(k)$ is the inventory target at instant k.

The z-transform is a powerful operational method when one works with discrete control systems because the differential equation is converted to an algebraic problem. One important property of the Z-transform, which we use extensively, is the translation theorem Ogata (1996)

$$Z(x(k-n)) = z^{-n} X(z) \qquad (3)$$

Where z^{-n} is the operator of a time delay in space z and corresponds to a time delay of n time sampling periods T.

Using the translation theorem (3) , on equations (2) and(2), we obtained the z-transform of the above discrete time model, this is given by the equations (4) and (5).

$$I_i(z) - I_i(z)z^{-1} = O_{i,i+1}(z)z^{-L} - O_{i-1,i}(z) \qquad (4)$$

$$O_{i,i+1}(z) = C(z)(I_i^T(z) - I_i(z)) \qquad (5)$$

The resulting model in (4) is thus amenable to implement some controllers that exist in literature.

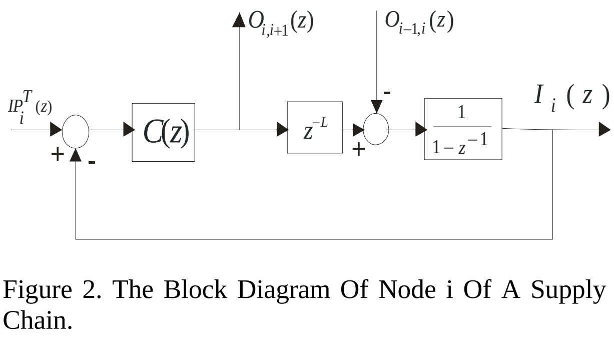


Figure 2. The Block Diagram Of Node i Of A Supply Chain.

The simplified block diagram is given in Figure (2). $C(z)$ represent the transfer function of anyone controller as replenishment policies. The above model seems simple. Nevertheless, it captures the basic dynamic feature of a supply chain system. A real supply chain usually has many customers, suppliers, and products. In a decentralized system, the inventory dynamics does not really depend on how many customers the node has, since all customer demands can be lumped into an aggregate demand. Obviously, if every node has sufficient inventory and has the same transportation delay, the distribution of order would not affect the dynamic behaviour of the system. We can assume that different suppliers have different inventory levels and different transportation delays.

2.1 Uncertainty description

The last model to describe an approximation of the dynamic of a supply chain, the real procces is nonlinearity and has uncertainty on the parameters.

To account for model uncertainty we will assume that the dynamic behaviour of a plant is described not by a single linear time invariant model but by a family $F(z)$ of LTI models. Algebraically, the family of plants is defined by:

$$F(z) = \tilde{P}(z)(1 + l_m(z)) :| l_m(z) | < | \bar{l}_m(z) \tag{6}$$

Where $\tilde{P}(z)$ denotes the nominal model, $l_m(z)$ is the relative (multiplicative) model error

$$\bar{l}_m(z) = \frac{P(z) - \tilde{P}(z)}{\tilde{P}(z)} \tag{7}$$

And $\bar{l}_m(z)$ is a frequency dependent weight that defines the family of plants by bounding the modelling error.

3. TRENSFER FUNCTIONS FOR A SUPPLY CHAIN MANAGEMENT

In order to derive the transfer function for a particular controller as replenishment inventory policy, let's refer to Figure (2). Here the control strategy $C(z)$ must be tuned in order to the system will be stable, keep the inventory $I_i(z)$ close to inventory target $I_i^T(z)$ and avoiding the bullwhip effect.

The transfers function that relates the physical inventory with the desired inventory were obtained

$$\frac{I_i(z)}{I_i^T(z)} = \frac{C(z)z^{-L}}{1 - z^{-1} + C(z)z^{-L}} \tag{8}$$

$$\frac{O_{i,i+1}(z)}{I_I^T(z)} = \frac{C(z)(z-1)}{1 - z^{-1} + C(z)z^{-L_i}} \tag{9}$$

Assuming that there is no change in the set point, the ratio of orders to successive nodes can be expressed as

$$\frac{O_{i,i+1}(z)}{O_{i-1,i}(z)} = \frac{C(z)}{1 - z^{-1} + C(z)z^{-L_i}} \tag{10}$$

The bullwhip effect can be represented as an amplification of demand fluctuations from downstream to upstream, it is translate to de amplification of amplitude of the relation (10) will be greater than 1.

All relations contain the same characteristic equation; hence it is possible to do stability analysis for both transfer functions using this equation. Then the characteristic equation is:

$$1 - z^{-1} + C(z)z^{-L_i} = 0 \tag{11}$$

With (8), (9), (10) and (11) the stability analysis and tuning for bullwhip mitigation can be realised for any controller.

3.1 Controller tuning criterion for bullwhip mitigation

Because of the widespread use of PID controllers, it is interesting to have simple but efficient methods for tuning the controller. In fact, since Ziegler Nichols proposed their first tuning rules Ramon (2008), an intensive research has been done. In this case an analytical tuning technique is used in order to mitigate the bullwhip effect. We assume that the customer demand is stochastic. However, the average of demand may be subjected to a low frequency disturbance such as a step change or seasonal cyclic changes. The objective of a simple inventory level controller is to maintain a given inventory position in the presence of such a low frequency disturbance. However, in addition to achieving the inventory position target, the objectives of a supply chain manager also include setting an inventory position target that is not too high (resulting inexcess inventory costs) or too low (resulting in customer dissatisfaction due to backorder) compared to the current average demand. Therefore, a manager should aim to create a fast response of the order to low frequency demand changes so that the inventory level can be maintained, but limit the ratio of order to demand to less than 1 at high frequency. The frequency response of $|O_{i,i+1}(z)| / |O_{i-1,i}(z)|$ of a closed loop supply chain node should be used for controller design. Standard textbooks (Pin, David, Shan, Shi, Jang, Shyan and Zheng 2004). suggest the following two factors to be considered:

- A wide bandwidth (The frequency at which the magnitude ratio is reduced to below 0.7.) indicates a faster response but poorer noise rejection capabilities. As we only discuss a discrete system; therefore, the highest frequency is at $\omega = \pi$. Therefore, a design ruler is choosing a controller than setting magnitude ratio less than 0.7 for this frequency.
- A higher resonance peak (RP) indicates a faster response but may be more oscillatory. Then the second rule is choosing suitability a controller which setting RP less than 1.

4 ANALISIS OF THE EFFECT OF PICONTROLLER AS REPLENISMENT INVENTORY

A control system is a combination of elements (components of the system) which enable us to control the dynamics of the selected process in a certain way. The PID (Proportional, Integral and Derivative) is the controller most commonly used in control engineering because of its flexibility and simplicity. Therefore there will by introductory analysis of this controller as inventory replenishment policy. In this section we show that by implementing a proportional integral inventory control and properly synthesizing the controller parameters, we can effectively suppress the bullwhip effect. We perform simulations with different values of parameters (k_p, k_i) and finally we compare the behaviour of the inventory and orders signals using PI actions. The derivate action is discarded because of its high noise sensitivity.

$$C(z) = k_p + \frac{k_i}{z-1} \tag{12}$$

4.1 Proportional Control

The main interest is management the of inventory level. Therefore we use a feedback-level inventory and a proportional controller. When there is sufficient supply and high stock, substituting the P-control into the transfer functions (8),(9) and (10) it is obtain the transfer functions of the system in close loop with the proportional action, then the stability analysis, the controller performance and the bullwhip effect can be analysed. Stability criterions and simulations have been used in previous works (Carlos , Pedro and Ramón 2008) by evaluate the stability and the behaviour of the orders $O_{i,i+1}(z)$ and the inventory level $I_i(z)$. With the proportional controller is possible to stabilize the system whit certain values of k_p but a phenomenon (offset) that is a mistake of steady state is inevitable because if the error is constant the control action is constant too.

One factor that bullwhip is usually attributed to aggressive ordering. We have demonstrated in last one work (Carlos, Pedro and Ramón 2008) that the system would become unstable when $k_p < 1$. When there is sufficient supply and high stock, substituting the P-control into the transfer Function (10) we get

$$\frac{O_{i,i+1}(z)}{O_{i-1,i}(z)} = \frac{k_p}{1 - z^{-1} + k_p z^{-L_i}} \tag{13}$$

Figure 3. gives the frequency response of an echelon with a proportional only controller to its inventory position. with several different controller gains. It can be shown that with a controller gain lower than one, the bullwhip effect of the echelon is suppressed. Thus $k_p = 0.2$ should be appropriate according to controller tuning criterion.

There is a large offset between the inventory position and set point (Carlos , Pedro and Ramón 2008). This offset will lead to accumulation of a large amount of backorder and low customer satisfaction. Since an offset cannot be avoided, customer dissatisfaction is inevitable for a P-only controller. To avoid this offset, a PI controller can be used.

4.2 Proportional Integral (PI) control

The integral action has some characteristics that mprove the response of the system in close loop.

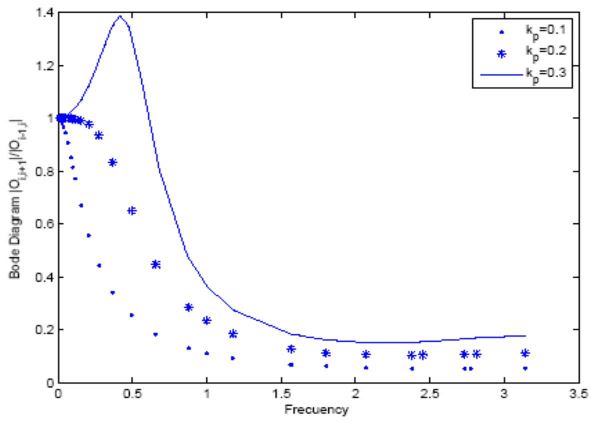


Figure 3. Frequency Responses Of $|O_{i,i+1}(z)|/|O_{i-1,i}(z)|$ with Different k_p Values Of The P Controller.

4.2 Proportional Integral (PI) control

The integral action has some characteristics that improve the response of the system in close loop. One of these properties is that it removes the offset because the control increases although the error remains constant (integrates the error), hence we analyze its behaviour in a supply chain. In this section we analyze the frequency response of the expression that related the orders an demand introducing this controller as replenishment inventory in order to tuning the PI parameters so that exist a trade off between inventory tracking and the mitigation of bullwhip effect. As well as we analyze the inventory and orders response. Substituting the PI-control give by (12) into the transfer function (10) we get the relation between the orders and the demand

The closed loop Bode plot of one supply Chain echelon with a PI controller is shown in Figure (4). It can see that for a lead time $L = 3$ and a given $k_p = 0.2$, the bullwhip effect still appears even the controller gain, k_i is close to zero. Hence there is a trade-off to be made between being responsive and following the demand changes very closely (large k_i-values) on the one hand and avoiding bullwhip (small k_i -values) on the other hand.Therefore we approach an alternative IMC scheme in order to improve the tracking and robust stability without generating bullwhip and following the demand changes.

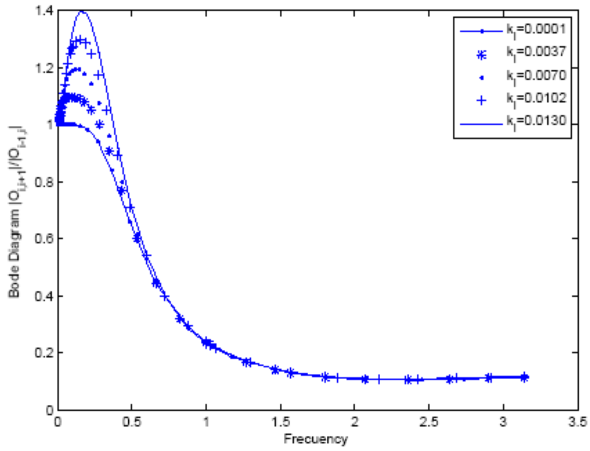


Figure 4. Frequency Responses Of $|O_{i,i+1}(z)|/|O_{i-1,i}(z)|$ with Different k_i Values Of The P Controller.

5 INTERNAL MODEL CONTROL BASED APPROACH

The internal model control structure has been introduced as an alternative to the feedback structure. Its main advantage is that closed-loop stability is assured simply by choosing a stable IMC controller (Manfred and Zafiriou 1989). Closed-loop performance characteristics (like settling time) are also, related directly to controller parameters which make on-line tuning of the IMC controller very convenient. A two-step design procedure was proposed. In the first step the controller is designed for optimal setpoint tracking (disturbance rejection) without regards for input saturation or model uncertainty. In the second step the controller is detuned for robust performance.

5.1 IMC structure for stable plants

The simplified block diagram of the IMC loop is shown in Figure (5). Where $P(z)$ denotes the plant and $\tilde{P}(z)$ is the nominal model. The controller $Q(z)$ determines the value of the input (manipulated variable) $u(z)$. The control objective is to keep $y(z)$ close to the reference (setpoint) $r(z)$.

If the model is exact ($\tilde{P}(z) = P(z)$) and there are no disturbances, then the output $\tilde{y}(z)$ and $y(z)$ are the same and the feedback signal $\tilde{d}(z)$ is zero. Thus, the control system is open-loop when there is no disturbance and no plant/model mismatch. The feedback signal $\tilde{d}(z)$ expresses the uncertainty about the process.

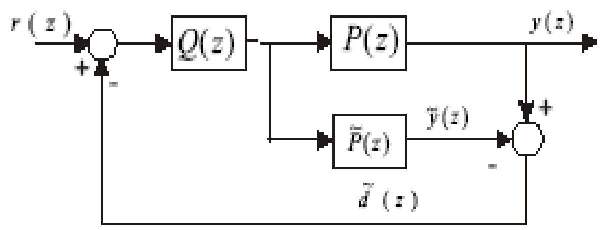


Figure 5. Internal Model Control Structure.

5.1.1 An overview of IMC design

The ultimate objective of control system design is clearly that the controller works well when implemented on the real plant. This goal can be best understood if decompose it into a series of subobjectives (nominal stability, Nominal performance, Robust Stability and Robust Performance) (Manfred and Zafiriou 1989) for this case, avoiding of bullwhip effect is introduced as the design criteria. The IMC has the property that allows us to discompose the design in steps in order to solve the objectives separately.

5.1.2 Stability conditions for IMC

Because it is assumed that the model is an approximate description of the true plant it is reasonable to require stability when the controller is applied to the plant model. Thus, the minimal requirement on the closed-loop system is nominal stability. Then, in order to test for internal stability the transfer functions between all possible system inputs and outputs are examined. From the block diagram in Figure (5). If there is no model error $\tilde{P}(z) = P(z)$, the inputs and outputs are related through the transfer matrix (15). Then the IMC system is internally stable if and only if both the plant $P(z)$ and the controller $Q(z)$ are stable

$$\begin{pmatrix} y(z) \\ u(z) \end{pmatrix} = (H) \begin{pmatrix} r(z) \\ u'(z) \end{pmatrix} \quad (14)$$

Where

$$H = \begin{pmatrix} \tilde{P}(z)Q(z) & P(z)(1-\tilde{P}(z)Q(z)) \\ Q(z) & 1-\tilde{P}(z)Q(z) \end{pmatrix} \quad (15)$$

On the other hand for good nominal performance and robust stability and performance the IMC uses a two-step approach which has no inherent optimality characteristic but should provide a good approximation to the optimal solution. It guarantees robustness but the performance is generally not optimal.

STEP1: For the Nominal performance the controller $\tilde{Q}(z)$ is selected to yield a good system response for the input(s) of interest, disregarding constrains and model uncertainty, in other words a weighted norm of the sensitivity function ε should be made small. Generally $\tilde{Q}(z)$ is chosen such that it is Integral Square Error or H_2 -optimal.

$$\min \| 1-\tilde{P}(z)\tilde{Q}(z) \|_2 \quad (16)$$

The optimal sensitivity function becomes

$$\varepsilon \doteq 1-\tilde{P}(z)\tilde{Q}(z) \quad (17)$$

And the optimal complementary sensitivity function

$$\eta \doteq \tilde{P}(z)\tilde{Q}(z) \quad (18)$$

The model inverse is an acceptable solution only for Minimum-Phase (MP) systems. For Nonminimum-Phase (NMP) systems the exact inverse model is unstable or noncausal then an approximate inverse of $\tilde{P}(z)$, must be found such that the weighted 2-norm of the sensitivity function is minimized. According to this, let's assume that $\tilde{P}(z)$ is stable.

Factor $\tilde{P}(z)$ into an allpass portion $\tilde{P}_A(z)$ and a MP portion $\tilde{P}_M(z)$

$$\tilde{P}(z) = \tilde{P}_M(z)\tilde{P}_A(z) \quad (19)$$

So that $\tilde{P}_A(z)$ includes all zeros outside the unite circle and delays and has the form

$$\tilde{P}_A(z) = z^N \prod_{j=1}^{h} \frac{(1-(\xi_j^H)^{-1})(z-\xi_j)}{(z-\xi_j)(1-(\xi_j^H)^{-1})} \quad (20)$$

Where the integer N is chosen such that $\tilde{P}_M(z)$ is semiproper and $\xi_j, j=1,....h$ are the unstable zeros. Therefore H_2 -optimal optimal controller $\tilde{Q}(z)$ is given by

$$\tilde{Q}(z) = \frac{1}{\tilde{P}_M(z)} \quad (21)$$

STEP2: For robust stability and performance at high frequencies when $l_m(z)$ exceeds unity $\eta(z)$ has to be rolled off. Therefore $\tilde{Q}(z)$ is augmented with low-pass filter $f(z)$.

$$Q(z) \doteq \tilde{Q}(z)f(z) \quad (22)$$

The order of $f(z)$ is such that $Q(z)$ is proper and its roll-off frequency low enough to satisfy the robust stability constrains. With the filter, $\varepsilon(z)$ and $\eta(z)$ become

$$\varepsilon \doteq 1-\tilde{P}(z)\tilde{Q}(z)f(z) \quad (23)$$

And the resulting complementary sensitivity function

$$\eta \doteq \tilde{P}(z)\tilde{Q}(z)f(z) \quad (24)$$

The function of f(z) is to detune the controller, to sacrifice performance (increase $|\varepsilon|$) for robustness (decrease $|\eta|$). In principle both the structure and parameters should be determined such that an optimal compromise between performance and robustness is reached. To simplify the design task the filter structure is fixed and search over a small number of filter parameters (just one) to obtain desired robustness characteristics. Typically the filter is chosen of the form given by

$$f(z) = \frac{1-\lambda}{1-\lambda z^{-1}} \quad (25)$$

Here the λ is an adjustable parameter.

5.1.3 Relationship with Classic Feedback

If we combine the two blocks $\tilde{Q}(z)$ and $\tilde{P}(z)$ in Figure (5), which are both part of the control system, into the one block $C(z)$ we obtain a single controller block.

$$C(z) = \frac{\tilde{Q}(z)}{1-\tilde{P}(z)\tilde{Q}(z)} \quad (26)$$

On the other hand its possible obtain $\tilde{Q}(z)$ from $C(z)$ used the following equation

$$\tilde{Q}(z) = \frac{C(z)}{1-\tilde{P}(z)C(z)} \quad (27)$$

5.1.4 Numerical Example

Consider the system given by

$$\tilde{P}(z) = 0.483\frac{z^2 + 1.01 + 0.0597}{z^3 - 0.116z^2 + 0.118z - 0.00315} \quad (28)$$

It has two zeros, both inside to unit circle, at $z = -0.95$ and $z = -0.95$. Therefore, the H_2 -optimal controller $\tilde{Q}(z)$ is determined by (21) as follow

$$\tilde{P}(z) = 1.001\frac{z^3 - 0.116z^2 + 0.118z - 0.00315}{z^3} \quad (29)$$

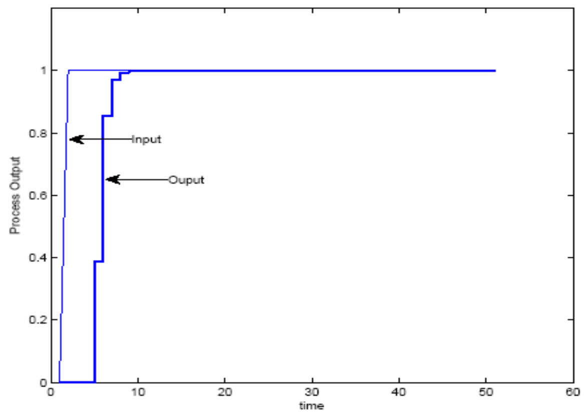


Figure 6. Output Response Of Close Loop System With A Control Based On IMC Design.

5.2 Modified IMC structure for unstable systems

We can see that the supply chain is an unstable processes with time delays, therefore the design of a conventional IMC is difficult, then a IMC modified structure (Wen, Horacio, Marquez and Tongwen 2003) can be an easy form to overcome the implementation problem as shown in Figure (7).

The procedure consists of first designing a compensator to stabilize the nominal plant, and then designing an IMC controller for the stabilized model (Wen, Horacio, Marquez and Tongwen 2003) with uncertainty. The block diagram of the modified IMC loop is shown in Figure (7). Here k is used to stabilize $\tilde{P}(z)$, the original (unstable) plant, ignoring the uncertainty in the time-delay and $Q(z)$ is designed to mitigate the bullwhip effect and robustly stabilize of new process.

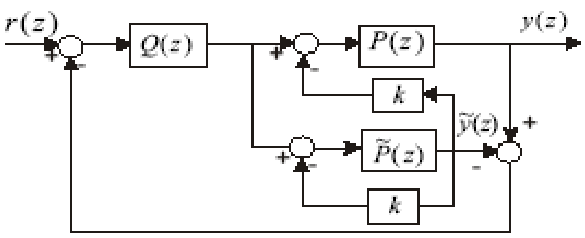


Figure 7. Modified Internal Model Control For Unstable Process.

5.2.1 Modified IMC design for an echelon of a Supply Chain

Let us consider the nominal model of the classical feedback ignoring uncertainty of an echelon as shown in Figure (8). It has an internal close-loop which is stabilized with a k controller. The transfer function is given by (30). Due to Lead time and the controller parameters the system poles can change, then is important to find the mathematical relation between the k controller and the lead time so that the system will be stable.

$$\tilde{P}(z) = \frac{I_i(z)}{I_i^T*(z)} = \frac{1}{1-z^{-1}+kz^{-L_i}} z^{-L_i} \qquad (30)$$

Any root or pole is a complex number then each of these can be represented as

$$z = re^{j\Omega} \qquad (31)$$

Our interest is to find a value for the controller so that the roots are into the unit circle $|r|<1$, then if we replace (32) on the characteristic equation we have

$$|k| = -r^{L_i} e^{jL_i\Omega} + r^{(L_i-1)} e^{j(L_i-1)\Omega} \qquad (32)$$

So the relationship between the lead time and the controller that guarantees the system is stable is given by

$$|k| < 2\cos\left(\frac{(L_i-1)\pi}{2L_i-1}\right) \qquad (33)$$

Using (33) we can chose a k value in order to stabilise the nominal plant. With the stabilized plant we can therefore apply the IMC design for stable plants.

STEP1: For the nominal performance the model inverse is an acceptable solution only for MP systems. Otherwise the exact inverse is unstable and/or no causal. Thus an approximate inverse of $\tilde{P}(z)$ must be found. Then, factor the nominal plant according to (19),(20) we have

$$\tilde{P}_A(z) = z^{-L_i} \qquad (34)$$

And

$$\tilde{P}_M(z) = \frac{1}{1-z^{-1}+kz^{-L_i}} \qquad (35)$$

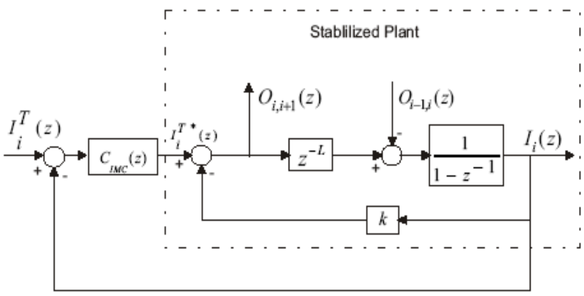


Figure 8. The Block Diagram Of Node i Stabilized With k Controller.

Thus we chose $\tilde{Q}(z)$ as shown bellow.

$$\tilde{Q}(z) = \frac{1}{\tilde{P}_M(z)} = 1 - z^{-1} + kz^{-L_i} \qquad (36)$$

STEP2: For robust stability and performance, $\tilde{Q}(z)$ is augmented by a low-pass filter $f(z)$ in this case we used the parameter for tuned so that the bullwhip is mitigate. Finally the structure of controller IMC would be

$$Q(z) = \frac{(1-\lambda)}{1-\lambda z^{-1}}(1 - z^{-1} + kz^{-L_i}) \qquad (37)$$

5.2.2 Tuning of for mitigate the bullwhip effect

$Q(z)$ is intended to be an IMC controller that mitigates the bullwhip effect, therefore it is necessary to obtain the expression that determines the behaviour of the orders vs demand with the IMC controller. The relationship between $|O_{i,i+1}(z)|/|O_{i-1,i}(z)|$ in close loop including the k controller and classical feedback controller $C(z)$ has been

$$\left|\frac{O_{i,i+1}(z)}{O_{i-1,i}(z)}\right| = \left|\frac{C(z)+k}{1-z^{-1}+(C(z)k\ \ )z^{-L_i}}\right| \le 1 \qquad (38)$$

Using the relation (26) we obtain the expression for the bullwhip effect with an IMC controller as

$$\left|\frac{Q(z)+k(1-Q(z)\tilde{P}(z))}{1-z^{-1}+k\ \ z^{-L_i}}\right| \le 1 \qquad (39)$$

This expression is so complicated to analyze, so we used the triangle inequality to separate it in two functions T_1 and T_2 as is shown bellow.

$$\overbrace{\left|Q(z)\frac{1-k\tilde{P}(z)}{1-z^{-1}+k\ \ z^{-L_i}}\right|}^{T_2} \le 1 - \overbrace{\left|\frac{k}{1-z^{-1}+k\ \ z^{-L_i}}\right|}^{T_1} \qquad (40)$$

In the light of the above equation we can see that once k is designed to stabilise the plant at hand using (33) T_1 gets fixed. Then, the λ parameter in $Q(z)$ can be used to satisfy (40) through T_2. Let's assume a nominal lead time $L = 3$, It has tuned the k controller in order to stabilize the plant and mitigate the bullwhip effect using the $T_1 < 1$ condition, then we tune the λ parameter of $Q(z)$ controller for improve the relation between tracking of inventory position and robust stability, the bullwhip effect is avoided if the relationship (40) is fulfilled, As shown on Figure (9), setting $\lambda = 0.212$ we have a wide bandwidth that indicates a faster response.

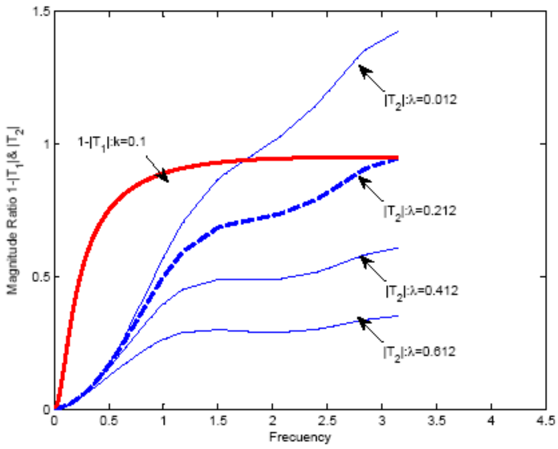


Figure 9. Magnitude Of $1-|T_1|$ and $|T_2|$ With Different Values Of λ.

Figure (10) shows that the IMC controller approach with the analytical design, improve the response with respect to PI controller. This implies that there is a faster replenishment inventory as well as Figure (11) shows that the bullwhip effect is avoided.

5.2.3 Demand forecasting

The forecasting demand will lead to bullwhip effect Pin, David, Shan, Shi, Jang, Shyan and Zheng 2004). We assume that the retailer is using a common method of simple exponential smoothing to estimate a demand forecast for the next period, that is

$$F(z) = \frac{\alpha}{1-\alpha z^{-1}} \qquad (41)$$

Then if we attempt to forecast the customer demand $F(z)$ with $\alpha = 1$ and set the inventory position target according to $L+2$, as shown in Figure (12), the closed loop responses of order to supplier $|O_{i,i+1}(z)|/|O_{i-1,i}(z)|$ become

$$\overbrace{\left|Q(z)\frac{F(z)(1-z^{-1})+1-k\tilde{P}(z)}{1-z^{-1}+k\ \ z^{-L_i}}\right|}^{T_2} \le 1 - \overbrace{\left|\frac{k}{1-z^{-1}+k\ \ z^{-L_i}}\right|}^{T_1} \qquad (42)$$

The forecasting demand function $F(z)$ does not affect the stability so the tuning criterion is the same. The Figure (13) shows the tuning for λ in order to improve the performance avoiding the bullwhip effect.

We simulated the orders for an echelon with the IMC scheme and forecasting demand as shown in Figure (14). IMC control scheme provides efficient control of the inventory position of a supply chain echelon avoiding the bullwhip effect.

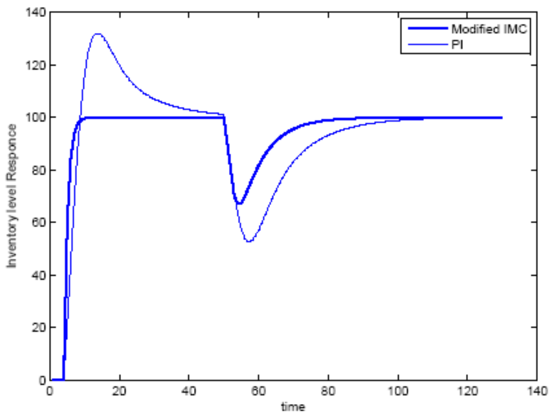


Figure 10. Inventory Level Response

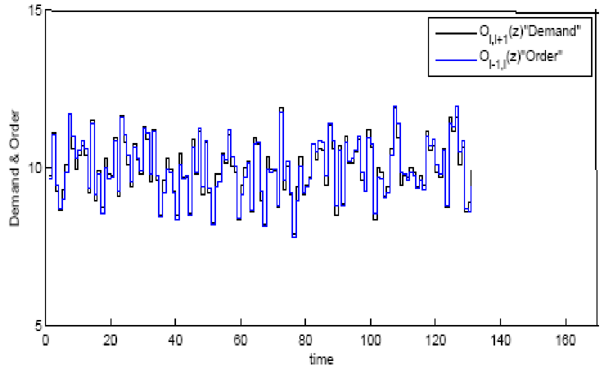


Figure 11. Simulation Results Of a Supply Chain Unit With A IMC ($\lambda = 0.212$) Controller.

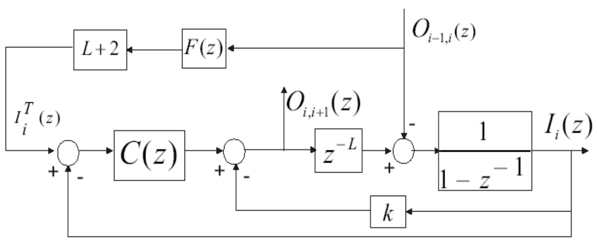


Figure 12. The Block Diagram Of Node i Of A Supply Chain With Forecasting Demand.

6 CONCLUSIONS

The discrete model for an echelon of the beer game has been derived using the z-transform. We obtain the characteristic equations of the closed loop. The bullwhip effect is also analyzed through frequency response. Some alternative ordering policies were formulated as P, PI and IMC control schemes and we can conclude that using P and PI controllers the bullwhip effect of a supply chain unit can be suppressed.

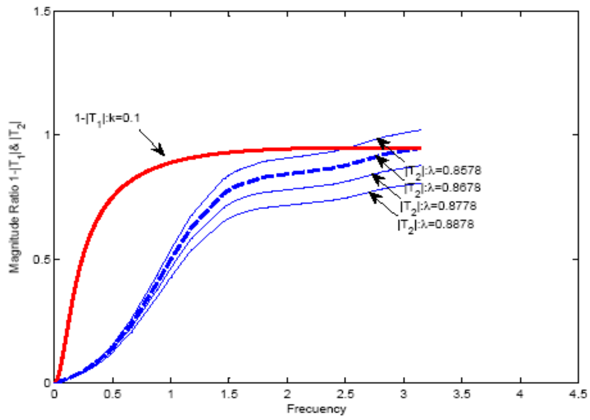


Figure 13. Magnitude Of $1-|T_1|$ and $|T_2|$ With Different Values Of λ

Moreover, the IMC control structure is superior in meeting customer demand due to its better tracking of long term trends of customer demand. We can therefore conclude that control theory is applicable to analysis of supply chains and that is would be possible to improve results using more efficient controllers as PI and IMC.

7 ACKNOWLEDGMENTS

This work has received financial support from the Spanish CICYT program under grants DPI2007-63356 and DPI2007-64570.

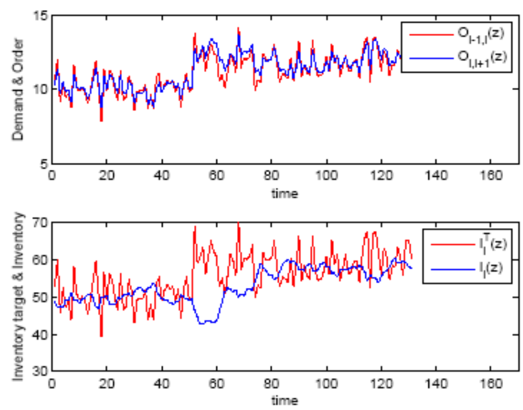


Figure 14. Simulation Results Of A Supply Chain Unit With A IMC Controller ($\lambda = 0.88$) And Stochastic Demand From Downstream.

REFERENCES

Carlos G, Pedro B, Ramon V. 2008. Supply Chain Management Through P and PI Controllers. *The 20TH European Modeling & Simulation Symposium.* University of La Calabria 265- 270

Dejonckheere, J., Disney, S.M., Lambrecht, M.R., Towill, D.R., 2002. Measuring and avoiding the bullwhip effect: A control theoretic approach.

European Journal of Operating Research, 147, 567–590.

Dejonckheere, J., Disney, S.M., Lambrecht, M.R., Towill, D.R., 2004. The impact of information enrichment on the bullwhip effect in supply chains: a control engineering perspective. European Journal of Operating Research, 153, 727–750

Hoberg, K, James, R., Bradley, Ulrich, W., and Thonemann. 2007. Analyzing the effect of the inventory policy on order and inventory variability with linear control theory. European Journal of Operational Research, 176, 1620–1642.

Ilhyung, K, Mark S, 2007, Measuring endogenous supply chain volatility: Beyond the bullwhip effect. European Journal of Operating Research 176.

Marko, J., Rusjan, B., 2008. The effect of replenishment policies on the bullwhip effect: a transfer function approach. European Journal of Operational Research, 184, 946–961.

Manfred, M, E. Zafiriou, Robust Process Control, 1989 Prentice Hall,Englewood Cliffs, NJ.

Ogata, K, (1996). Sistemas de Control en Tiempo Discreto, México: Prentice Hall.

Pin, H, David ,S, Shi- S, Shyan , S, Ji, Z , 2007 Controller design and reduction of bullwhip for a Model supply chain system using z-transform analysis. Journal of process control 14 487-499

Ramon, V 2008 IMC based Robust PID design: Tuning guidelines and automatic tuning. Journal Of Process Control 18 6170

Wen, T, Horacio J. Marquez, Tongwen,C 2003 IMC design for unstable processes with time delays. Journal of Process Control 13 203-213.

INTEGRATED SUPPLY CHAIN RISK MANAGEMENT PERFORMED BY LOGISTIC ASSISTANCE SYSTEMS

Gökhan Yüzgülec [(a)], Dr. Axel Wagenitz [(b)], Dr. Michael Toth [(c)]

[(a, b, c)] Fraunhofer-Institute for Material Flow and Logistics IML, Dortmund, Germany

[(a)] goekhan.yuezguelec@iml.fraunhofer.de, [(b)] axel.wagenitz@iml.fraunhofer.de, [(c)] michael.toth@iml.fraunhofer.de

ABSTRACT
The occurrence of risks usually leads to damages at relevant enterprise and moreover causes further risks at other partners. Consequently, more and more enterprises need to deal with risks. Especially, small and medium sized enterprises (SMEs) are forced to make statements concerning risks and intended actions, particularly regarding the dynamically fluctuating system load. There are no applicable tools and methods known to assess those risks and possible actions in order to control them profitable. The new SCM-Software generation Logistic Assistance Systems (LAS) opens new application fields in global supply chains. The novel LAS approach is currently being intended to provide a basis for an integrated Supply Chain Risk Management (SCRM) approach. With the aid of LAS, different risk key performance indicators (KPIs) can be obtained which can be used in order to assess risks and actions quantitatively and proactively; additionally to experience to what extend the risks can be minimised.

Key words: Supply Chain Risk Management, Risk Assessment, Logistic Assistance System, Simulation

1. INTRODUCTION
Consequences of the trends in recent years like the application of the lean production philosophy as well as the developing globalisation of the economy are efficient global supply chains, reduced inventory, high capacity utilisation and optimised lead times. Furthermore, increased end customer requirements regarding quality, price and availability as well as the increasing global competition lead many companies to a closer collaboration within their supply chains. Thus, these trends generally lead to significant efficiency improvements on the one hand; contrariwise by the optimisation, e.g. reduction of inventory, number of suppliers and delivery times, companies' interdependencies grow enormously and therefore lead to an increase of the vulnerability of the entire supply chain. Through worldwide interconnectedness, events in countries and continents faraway can affect adding value processes in other regions significantly (Kersten 2008). Ad hoc increases of customer demands, insolvencies of suppliers, malfunctions of information systems, strikes, fluctuating system loads, unexpected high declines in sales, natural catastrophes and other events are examples for risks, which can affect companies. The occurrence of such risks usually leads not only to damages at the relevant company but causes the entrance of further risks with according damages at other downstream or upstream supply chain partners. The understanding of supply chain risk is disruption, quantified with its occurrence probability; the risks' occurrence affects more than one company in the supply chain and the risks' origins are located within one company, its supply chain or within its environment (Kersten 2008).

The imminent loss by unforeseen events discloses the acuteness for companies to deal with their risks from an entire supply chain perspective in order to carry on being competitive, avoid turbulence costs along the entire supply chain and to generate a significant long-term advantage in competition. Though, risk management does not serve to avoid risks but to identify those and to provide a decision basis for strategic considerations.

To face these challenges it is essential for companies to proactively identify, evaluate, control and communicate the risks. Large-scale enterprises have already reacted to these challenges and operate a simple form of risk management in order to get a transparency of their risk situation. Nevertheless, in order to quantify the vulnerabilities in a supply chain context, companies have to consider not only the risks of their own operations but also risks, which are caused by the connections between the organisations; this means they need to consider their supply chain partners' risks.

That is why more and more big enterprises ask small and medium sized enterprises (SMEs) – usually as their key suppliers - to make statements concerning their risk situation regarding the satisfaction of requested demands. Malfunctions of SMEs can have significant effects on their partners. SMEs as partners in a network increase the risk for large-scale enterprises and endanger the stability of the whole supply chain because generally SMEs represent the most fragile partners within a supply chain. For instance, same risks can be serious risks for SMEs by being small or no risks for big enterprises. Consequently, these new challenges often cause trouble in SMEs, which were not confronted with such problems in the past; and therefore endanger the stability and performance not only in one but in all involved supply chains at the same time.

One of some major risks SMEs are exposed to extensively is the dynamically fluctuating system load. System load arises from the difference between supply and logistical performance object demands. It causes the flow of the logistical performance object through the processes (Jungmann, Kuhn, and Bandow 2007).

This means that demand requests from multiple customers, located in different supply chains, can change frequently in different ranges and affect SMEs enormously so that they may not perform and satisfy the requested transformation performance, for example due to limited capacities or due to missing implementation of risk minimisation actions. Requested transformation performance means that the requested demand is fully satisfied in terms of requested time, quantity and quality in such a manner that the performance of the supply chain is affected in no case. Consequently, the supply chain transformation performance cannot be assured by the SMEs and the fulfilment of requests can be at risk.

Over the past years Supply Chain Risk Management (SCRM) has been developed and integrated into the supply chain management (SCM) concept in order to meet the challenge mentioned above. This concept contains the identification, the assessment, the control as well as the monitoring of supply chain risks in order to improve the integration of companies in their supply chain (Kersten 2008). Several companies operate a simple type of risk management. However, the concept of SCRM is still in its infancy (Jüttner 2005). Furthermore, risk management not only serves for dealing with risks but to identify them and exhibit a basis of decision for strategic advisements.

Existing solutions in order to operate a simple form of risk management are predominantly applicable for big enterprises. The literature reveals that a lot of SMEs do not quantify and manage their risks adequately. Risk management should become a core issue in planning and management of every organisation (Finch 2004). An application-oriented methodology for identification, assessment, and control of risks from an intra-company perspective was developed by Ziegenbein (Ziegenbein 2007). For supply chains in practice neither tools nor an integrated application-orientated methodology with an entire supply chain perspective is known; that risk management shows substantial deficiencies regarding dissemination and implementation (Ziegenbein 2007).

For SMEs there is a lack of integrated supply chain risk management tools which systematically and proactively identify, assess and enable the communication of risks.

Regarding dynamically fluctuating system load which is one of the core risks for small and medium enterprises, SMEs are not able to make statements regarding their satisfaction risks of received purchase requisitions. Moreover, instruments for a decision-support are needed regarding which actions could minimise the satisfaction risks to which degree and how those purchase requisitions would influence the risks of existing purchase orders. Furthermore, there is a big need for key performance indicators (KPIs) which allow the measurement, the quantitative assessment, the transparency as well as the communication of the risks.

2. LOGISTIC ASSISTANCE SYSTEMS AS A NEW SCM SOFTWARE GENERATION

For the enhancements of an integrated SCRM concept and its successful widespread implementation in global supply chains by all involved partners there are no tools known which allow process owners and risk managers in enterprises a transparency and an exchange regarding existing risks and their extent. Basic requirements for SCRM instruments are the enabling of acting proactively in order to identify and assess risks, not only in the own company, in advance. Risk assessment needs to take place in a quantitative way, for instance with the aid of KPIs, especially in order to allow the communication of the risks. Furthermore, different possible risk minimisation actions need to be assessed proactively in order to determine the optimum way of risk control. Moreover, the tool handling complexity should not require special expertise by the user.

The new SCM-Software generation Logistic Assistance Systems (LAS) opens absolutely new application fields in the industry in global supply chains. Therefore, the novel LAS approach is being intended to provide likewise a basis for an integrated SCRM approach in global supply chains.

Assistance systems are computer-assisted tools which aim at delivering information and support for decision-makers and experts in order to assist them in terms of decision making and decision realisation in given planning situations (Schneider 2006). LAS offer transparency about all appropriate information and integrate specific decision support systems and planning approaches into one combined planning approach. Based on actual, high quality data LAS consist of additional APS functionality and decision support systems. The LAS concept focuses on relevant and consistent information along the supply chain, transparency about the current process status and planning functionality (Kuhn and Toth 2008). The basic idea of LAS is to plan and assess defined consideration scopes with a clearly specified planning scenario in supply chains, consisting out of several companies and service providers. In order to apply assistance systems in cross-company value added collaborations efficiently, an architecture which makes the technology globally accessible, by protecting the planning competence of the autonomic partners, is needed (Blutner et. al. 2007). The developed approach is based upon a decentralised cooperation of the value adding partners. All partners which are included in the process provide the required data. This data is process-oriented and flow into a central planning logistic, but can be locked for other partners if required.

Within the scope of an integrated SCRM, LAS are being intended to be applied by focusing on the planning approaches in order to allow quantitative analyses. The assistance system provides services for the available data and for the planning. A planning and simulation component is integrated in the LAS, enabling the analysis and assessment of the logistical network including the structures and processes (Wagenitz 2007; Hellingrath et. al. 2004). The simulation environment OTD-NET has been applied successfully in many projects in the automotive industry and therefore will be applied within the scope of the methodology for an integrated SCRM contribution.

OTD-NET was developed at the Fraunhofer-Institute for Material Flow and Logistics IML; OTD-NET allows very quick simulation runs as they are required for assistance systems. Process logical components are integrated in the simulator and therefore enable a simple modelling and parameterisation of different SCM strategies. Additionally, real demand data out of operational systems can be imported or generated. LAS extend a simulation model including all detail processes by providing a supply chain planning system and an integrated approach for operational and tactical planning of global supply chains.

Generally, every partner is able, according to his part in the planning process, to use the input data of other participants to accomplish his planning on real and current data. The planning components are provided as service adequate to the planning case. In case of disposition the planning component is a simulation tool, which displays the dynamic processes of the network, which has to be planned. Furthermore, it is able to simulate the supply chain's future performance of the based on program and flow data as well as on sales figures, by taking risks into account. Furthermore, risk minimisation actions can be simulated and different extent can be presented. The program planer of a distributor as well as the risk manager are able to make adjustments and to simulate in the planning domain (program data) which impact it has by given demand of the customer and reception of distributors. Here the demand is the input variable and the simulation model uses the actual plan data and also the process data (transports, inventory) to assure the validity of a demand adjustment within an available-to-promise-verification (Toth and Wagenitz 2009).

An important target of the planning component in an assistance system is the forecast of the supply chain's future performance. Besides the structure of the supply chain the considered parameters contain real time data with respect to demand situation, ship positions, inventory data, product capacity, etc. By the use of this new alternative of planning and decision assistance in a decentral-organised network, the planning quality as well as the risk transparency can be improved extraordinarily. Risk managers can simulate different system loads as well as the risks which affect them.

Based on the real data a simulation model is designed and integrates the actual status of the supply chain as well as the planed transports and the production volumes as load data which is the basis for a dynamic quantification of the future system development. This function establishes very accurate, adaptive plan scenarios for the expeditor as well as for the risk manager or process owner. The simulation structure, consisting out of modules is comparatively simple (without interference into the source code) and in most parts independent to diversify from the real data. This allows the user to diversify high complex simulation models with simple planning tables and to run simulation studies. By simulating the real data expeditors and risk managers can predict the future transformation performance and therefore the risk situation of the disposed supply chain considering all actual data. Furthermore, it is possible to manipulate single data systematically which offers possibilities for evaluating different scenarios (Deiseroth et. al. 2008).

In recent projects in cooperation with the Fraunhofer IML, LAS were successfully applied for several available-to-promise planning activities. It could be proved that LAS are absolutely capable for different planning challenges. Therefore, the LAS approach is being intended to act in the scope of SCRM.

3. INTEGRATED SUPPLY CHAIN RISK MANAGEMENT WITH THE AID OF LOGISTIC ASSISTANCE SYSTEMS

A novel application-orientated methodology approach for an integrated SCRM contribution is currently being developed at the Fraunhofer IML with the aid of the new SCM-Software generation LAS.

The novel LAS approach includes all specific supply chain structures, resources and processes which need to be considered and presents all relevant data information along the supply chain, delivering transparency about the current status. Besides those specific supply chain characteristics the simulation model contains all further information regarding the supply chain constraints.

The outcome of those given detailed information about the supply chain is that not only program planers of a distributor but also risk managers are able to make adjustments and to simulate in the planning domain in order to see the impact of changes. Risk managers are able to diversify high complex simulation models by implementing risk minimisation actions and to run simulation studies. By simulating the real data risk managers can predict the future performance as well as the risk situation of the disposed supply chain considering all actual data. The possibility to manipulate single data systematically allows an evaluation of different scenarios; therefore, for instance, implementations of different actions for system load fluctuations can be simulated evaluating the satisfaction risks for requested performances in advance.

In terms of system load fluctuations LAS are also able to deal with those risks which affect the system load. Multiple unexpected purchase requisition or increases of existing purchase orders (POs) by different individual customers can be evaluated regarding their fulfilment by taking into account implementing performance-enhancing actions; the risks and threats for non-fulfilment as well as possible implementation of actions can be evaluated and demonstrated in the LAS. Thus, alternatives for the control of critical risks can be shown by the LAS while demonstrating the impact of the potential delays or cancellations. The developed simulation component is capable to handle the complexity of multiple customer demands and multiple risks considering their occurrence probability and mutual impact, to forecast the supply chain behaviour in case of occurrence. Complex simulations can be run without any need for special IT expertise. Simple handling of the simulation runs which are based on actual, high quality data, allows users to create different risk scenarios which can be assessed for different possible actions. The extensiveness of the proved information covers current data of the supply chain as well as an intelligent concept in order to forecast future dynamic behaviour and is not limited to individual enterprises. Highly important supply chain partners' safety requirements are ensured by role based access limitations to critical data (Deiseroth et. al. 2008).

Applying LAS within the scope of integrated SCRM allows companies, especially SMEs, a common understanding and transparency by assessing risks in their supply chain. Optimum actions for dealing with risks can be identified. Finally, supply chain partners obtain a decision support for making strategic decisions, either in order to take actions proactively for the purpose of reducing risks and to avoid disruptions or in order to take no actions by knowing the potential extent of damage in advance.

In order to meet especially the requirements of SMEs, usually representing the most fragile partners within a supply chain, it has to be taken into account that due to limited resources for used tools the modelling duration and effort is to be kept as low as possible, which is an essential requirement and challenge. The duration of simulation studies varies by the level of complexity, the data availability and the assigned resources. Generally, for SMEs ideally durations more than a few weeks should not be extended for modelling efforts since quick and cheap results are required, which may not to be as detailed as for large-scale enterprises. In order to decrease the modelling effort, the creation of application-oriented libraries, predefined modules, is very profitable to minimise the implementation and modelling effort, compared to conventional simulation studies. Input and edit masks are available immediately and have only to be integrated into the new model. LAS integrate the following main characteristics and functions: information transparency, collaborative planning, decision support, risk management, process orientation and software flexibility (figure 1).

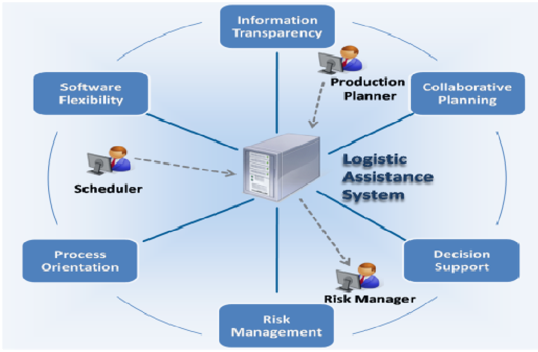


Figure 1: Logistic Assistance Systems

Currently, at the Fraunhofer IML LAS are being developed further in such a way that especially demands mentioned above, regarding dynamically fluctuating system load relevant to SMEs, can be met. First developments of an integrated SCRM approach with LAS have been approached and show promising results.

4. KEY PERFORMANCE INDICATORS FOR INTEGRATED SUPPLY CHAIN RISK MANAGEMENT

In the context of a SCRM approach, focusing on affected system load, LAS need to deliver different key performance indicators in order to enable the assessment and the communication of risks intra-company as well as with supply chain partners. Recent industry projects, managed by the Fraunhofer IML, show the benefits and the potentials of KPIs, which were delivered by LAS. By using components of the tool OTD-NET, presented in chapter 2, a simulation model of a production network was generated; due to the detailing, multifarious questions could be answered quantitatively with high quality (Motta et. al. 2008). The aim of using risk KPIs is to manage and adjust the risks in such a manner that companies keep being sufficiently robust to handle fluctuating system loads by performing requested transformation performances and by keeping costs as low as possible at the same time. This means, risks need to be opposed to costs which would become due by implementing actions or which would become due by taking risks without implementing any actions. Core questions which need to be answered by the aid of the risk KPIs are:

- Which extent do risks have regarding the company transformation performance due to system load fluctuations?
- Which impact do interactions between different supply chain demands have on company's transformation performance?
- Is it efficient to implement any actions in order to decrease the occurrence probability of the risks? If yes, which actions need to be implemented?
- What is the decrease of risks' occurrence probability and the residual risk extent by implementing actions?

In order to answer the questions mentioned above a detailed modelling of the network is essential. The products with variants and bill of material have to be part of the model like demand development within the given period as well as transport relations and supplier on multiple network levels. Besides the network structure the processes are essential. Rules for the planning of orders in factories, the selection of alternative means of extra transportation (e.g. plane) in case of shortages but also the selection of suppliers in case of multiple-sourcing have to be defined. Apart from that, combined used resources and stochastic influences e.g. fluctuation in transport times, have to be considered. Variant part changes have powerful impact on the network because advised variant parts and devices cannot be used anymore on the one hand, and on the other hand a doubling in demand of already produced obsolete variant parts occurs for the supplier.

Analyses showed that for a holistic risk management and risk assessment of risks affecting the system load following risk KPIs are required and can be gained by using Logistic Assistance Systems:

Dependability of delivery reliability in terms of time: This risk KPI gives information regarding the feasibility of the agreed delivery time, for each existing PO. A variance to the agreed delivery time results directly in noncompliance with the requested transformation performance.

Dependability of delivery reliability in terms of quantity: This risk KPI gives information regarding the feasibility of the agreed delivery quantity, for each existing PO. A variance to the agreed delivery quantity results directly in noncompliance with the requested transformation performance.

Dependability of lead times: This risk KPI gives information regarding the feasibility of the planned lead times, for each existing PO. A variance to planned lead times does not necessarily result directly in noncompliance with the requested transformation performance.

Dependability of supply availability: This risk KPI gives information regarding the supply availability, for each existing PO. A supply disability results directly in noncompliance with the requested transformation performance.

Availability of parts: This risk KPI gives information regarding the availability of the parts, for each existing PO. An unavailability of the parts does not necessarily result directly in noncompliance with the requested transformation performance.

Number of existing POs, affected by delayed deliveries: This risk KPI gives information regarding the number of existing POs which have been accepted or are already in progress and are affected by delivery delays due to incoming purchase requisitions. The occurrence of delayed POs results directly in noncompliance with the requested transformation performance.

Number of existing POs, which are affected positively: This risk KPI gives information regarding the number of existing POs which have been accepted or are already in progress and could be affected positively by taking actions. The occurrence of positively affected POs results directly in an increase of compliance with the requested transformation performance.

Degree of risk minimisation: This risk KPI gives information regarding the degree of risk minimisation due to taking actions. The occurrence of a degree of risk minimisation results directly in an increase of compliance with the requested transformation performance.

Costs by taking actions: This risk KPI gives information regarding the costs which would turn up for taking actions in order to minimise risks and increase the compliance with the requested transformation performance. This KPI always needs to be opposed to the risk KPIs “potential costs” and “degree of risk minimisation”. The actual quantitative risk assessment can only be meaningful when this comparison has taken place.

Potential costs: This risk KPI gives information regarding the costs which would turn up by not performing the required transformation performance, either by taking or not taking any actions in order to minimise the risks. This KPI always needs to be opposed to the risk KPIs “costs by taking actions” and “degree of risk minimisation”. The actual quantitative risk assessment can only be meaningful when these comparisons have taken place.

In terms of an integrated SCRM the risk KPIs listed above can be used as a communication medium in order to get a transparency and expose the actual situation. Furthermore, the communication of risk KPIs with downstream supply chain partners can be highly useful as a starting point for the partners’ risk management and risk assessment.

Missing risk KPIs, which cannot be communicated among the supply chain partners, there will be always a falsification of the supply chain’s actual state. The basic assumption for further risk assessment would draw wrong consequences. The usage of risk KPIs within the SCM context, obtained by LAS, leads to a new dimension of detailing in terms of risk assessment.

Usually, risks can be minimised by taking actions. Nevertheless, actions may not always be profitable. Apart from the risk KPIs mentioned above, recent projects showed that cost KPIs can be gained by using LAS. Thus, the risk KPI “potential costs” need to include cost KPIs obtained by LAS. The actual quantitative risk assessment can only be meaningful when this consideration is not neglected in order to have an accurate decision support in terms of taking actions.

Currently, at the Fraunhofer IML the risk KPIs as well as their implementation into the LAS is being investigated. The implementation of the interdependency between supply chain partners’ risks and costs is one of the challenges which are being met.

5. CONCLUSIONS & OUTLOOK

The new SCM-Software generation LAS allows large-scale enterprises as well as SMEs to get a transparency regarding their actual risk situation, especially for risks which are affecting the system load. Though, risk management does not serve to avoid risks but to identify those and to provide a decision basis for strategic considerations. LAS are able to deliver highly useful KPIs. For an integrated SCRM contribution a number of risk KPIs are introduced in this paper. On the on hand enterprises can use these presented KPIs as a communication basis in order to exchange risks. Benefits of the exchange of risk KPIs may be, for instance, the simple identification of potential win-win-situations. Enterprises can share costs for taking actions commonly to increase the compliance with requested transformation performance. On the other hand enterprises can use KPIs intra-company in order to analyse their actual situation and consequently to take actions purposeful.

For future research work it is imaginable to investigate the extension of the presented KPIs and to implement those into LAS which may act for further risk areas and SCM concepts.

REFERENCES

Blutner, D., Cramer, S., Kause, S., Mönks, T., Nagel, L., Reinholz, A., Witthaut, M., 2007. Assistenzsysteme für die Entscheidungsunterstützung. In: Buchholz, P., Clausen, C., ed. *Große Netze der Logistik. Die Ergebnsise des Sonderforschungsbereichs 559.*Berlin, Heidelberg: Springer

Deiseroth, J., Weibels, D., Toth, M., Wagenitz, A., 2008. Simulationsbasiertes Assistenzsystem für die Disposition von globalen Lieferketten. Advances in Simulation for Production and Logistics Applications. *Advances in Simulation for Production and Logistics Application*, pp. 41-51. October 1-2, Berlin.

Finch, P., 2004. Supply Chain Risk Management. *Supply Chain Management: an International Journal*, 2 (9), 183-196.

Hellingrath, B., Toth, M., Wagenitz, A., Motta, M., 2004. Simuliere und herrsche. *Beschaffung Aktuell*, 10, 32-33.

Jungmann, T., Kuhn, A., Bandow, G., 2007. Analyse der Dynamik der Systemlast eines Intralogistiksystems. *Forderungsgerechte Auslegung von intralogistischen Systemen-Logistics on Demand,* pp. 163-178. October 10, Dortmund.

Jüttner, U., 2005. Supply chain risk management. Understanding the business requirements, from a practitioner perspective. *The International Journal of Logistics Management*, 16 (1), 120-141.

Kersten, W., 2008. Risikomanagement in Wertschöpfungsnetzwerken – Status Quo und aktuelle Herausforderungen. *Wirtschaft und Management* ,8, 7-21.

Kuhn, A., Toth, M., 2008: Assistenzsysteme für die effektive Planung logistischer Netzwerke. In: Scholz-Reiter, B., ed. *Technologiegetriebene Veränderungen der Arbeitswelt*. Berlin: GITO-Verlag.

Motta, M., Wagenitz, A., Hellingrath, B., Weller, R., 2008. Gestaltung logistischer Netwerke-ein Praxisbericht. Design of Logistic Networks-A Case Study. *Advances in Simulation for Production and Logistics Application*, pp. 21-31. October 1-2, Berlin.

Schneider, M. 2006. *Assistenzsystem zur Strategiefestlegung in der Anlaufplanung - dargestellt am Beispiel des Pumpen- und Kompressorenbaus.* Thesis (PhD). TU University Dortmund.

Toth, M., Wagenitz, A., 2009: Neue Wege für die effektive Planung logistischer Netzwerke – Dynamische Verfügbarkeitsplanung mit Hilfe von Assistenzsystemen. *Industrie Management*, 25(2), 55-58.

Wagenitz, A., 2007. *Modellierungsmethode zur Auftragsabwicklung in der Automobilindustrie.* Thesis (PhD). TU University Dortmund.

Ziegenbein, A., 2007. *Supply Chain Risiken - Identifikation, Bewertung und Steuerung.* Thesis (PhD). ETH Zurich.

AUTHORS BIOGRAPHY

Dipl.-Ing. Gökhan Yüzgülec is consultant and scientist at Fraunhofer-Institute for Material Flow and Logistics IML in Dortmund, Germany. He is responsible for industry projects as well as application-oriented research projects, especially in the areas of supply chain risk management for small and medium enterprises.

Dr.-Ing. Axel Wagenitz is head of the department ”supply chain management” for the Fraunhofer-Institute for Material Flow and Logistics in Dortmund, Germany. His research interests are supply chain design, supply chain simulation and improvement of supply chain processes by means of model based approaches.

Dr.-Ing. Michael Toth is consultant and scientist at Fraunhofer-Institute for Material Flow and Logistics IML and team leader of the LAS-Department (logistic assistance systems). He has been working for the IML since 2003 and is responsible for industry projects as well as application-oriented research projects, especially in the areas automotive and plant engineering.

LOGISTIC ASSISTANCE SYSTEMS FOR DECISION SUPPORT IN COLLABORATIVE SUPPLY CHAIN PLANNING AND CONTROL

Felix Bockholt[(a)], Wolfgang Raabe[(b)], Dr. Michael Toth[(c)]

[(a)]Volkswagen Commercial Vehicles, Hanover, Germany
[(b)]Volkswagen AG, Wolfsburg, Germany
[(c)]Fraunhofer Institute for Material Flow and Logistics IML, Dortmund, Germany

[(a)]felix.bockholt@volkswagen.de, [(b)]wolfgang.raabe@volkswagen.de, [(c)]michael.toth@iml.fraunhofer.de

ABSTRACT
Global market changes and new business strategies have changed logistic processes. Today complex supply chains contain numerous manufacturing levels with long transportation distances. Effective planning and control have to follow various objectives like cost reduction, flexibility and ecological target values. Several supply chain management (SCM) concepts have been developed to support these challenging collaborative planning processes. For a successful implementation they all require software support. As will be shown in this paper existing solutions cannot offer a sufficient collaborative focus on supply networks and at the same time features for supply chain planning. This paper introduces logistic assistance systems (LAS) for decision support in collaborative supply chain planning and control. Designed as lean software systems, LAS focus on specific process parts and integrate selected methods for data management, information processing and supply chain planning. First achievements with LAS at Volkswagen AG are now followed by two new promising projects.

Keywords: logistic assistance system, supply chain collaboration, supply chain simulation, decision support system

1. INTRODUCTION

A progressing globalisation is challenging logistics and companies that cooperate as partners in a supply chain. Purchase parts are sourced globally and new production facilities on different continents have to be supplied. This results in global multimodal supply chains with long transport times, increased handling, numerous customs checks and the involvement of multiple service providers for transportation. Due to large geographical distances, a high number of process steps and numerous involved actors, supply chains sustain extended lead times and reduced flexibility to respond to process changes like new demand situations. Demand variation induces either supply shortfalls or high inventory levels, both causing additional costs (Deiseroth, Weibels and Toth 2008). Nevertheless global sourcing and the relocation of production sites can only be efficient as long as the logistic costs do not exceed the expected savings through lower production costs.

In addition to the mentioned attributes of multimodal supply chains external risks like supplier insolvencies, transport damage or transport delays due to customs checks, strikes, bad weather conditions or accidents are also responsible for an increased complexity and difficulty of managing global supply networks. (Wyk and Baerwaldt 2005)

These internal and external characteristics conflict with logistic targets and important competitive factors like short service times and high delivery reliability. To reach high delivery reliability, avoid supply shortfalls, utilise the left flexibility and enable short service times in those complex supply networks, methods and processes for an effective joint planning and control are required. To gain benefits for all, supply chain partners have to act in collaborative decision processes. (Kuhn, Hellingrath and Hinrichs 2008)

Collaboration in supply chains focuses on joint planning, coordination and process integration between organisational departments, customers, suppliers and other partners like logistic service providers. In order to optimise an entire supply network and not just to create local and often competing optima the involved actors have to communicate and jointly make supply and demand decisions. This way they can reach sustainable benefits like cost reductions, increased reliability and responsiveness to market needs (McLaren, Head and Yuan 2002). As a result collaboration becomes a necessary requirement for a comprehensive and successful SCM.

To work together the different actors have to define and follow joint objectives, share proprietary data and process information, and trust each other. Often a lack of trust between the different actors avoids successful supply chain collaboration (Barratt 2004; Ireland and Bruce 2000). To face this problem it is necessary to involve all relevant actors along the supply chain for defining a collaborative process and show the benefits for all partners (e. g. Dudek and Stadtler 2004). To implement a manageable process information technology is an essential enabler for a collaborative relationship across the supply chain. (Mentzer, Foggin and Golicic 2000)

For logistics and SCM the developments towards complex global multimodal supply chains also result in a higher importance of ecological efficiency which has become a new logistic target (Meißner 2008; Fleischmann 2008; Souren 2000). Growing geographic distances and a rising number of manufacturing levels result in higher demands of transportation resources. A rising number of transports with truckload, sea freight an air cargo leads to higher greenhouse gas emissions. Recently these impacts have been a main topic in political, cultural, economical and especially ecological discussions. This growing importance is mainly driven by the increasing deterioration of the environment (Srivastava 2007). Defined thresholds for specific types of transportation vehicles and a certificate market for carbon dioxide emissions (CO_2 emissions) are just two results of those discussions. For this reason ecological key performance indicators (KPIs) like CO_2 emissions emerge as relevant objectives for decision processes in supply chain planning and control.

To support collaborative planning, control and decision processes in complex global multimodal supply networks considering economical and ecological logistic targets, network partners have to apply logistic concepts supported by IT systems. Those software applications have to provide all relevant information along the supply chain and present it to the responsible experts and planners. The setup of accountabilities in the network and the way information is processed can only be determined individually for each case, according to the requirements of the given process and respective logistic concept. This also applies to the selection of methods and IT technologies used in the software system. As will be shown later, existing software solutions cannot offer all features needed for an effective supply chain planning and control.

Logistic Assistance Systems (LAS) can provide those features. They are designed as lean software systems and focus on specific process parts and integrate selected methods for data management, information processing and supply chain planning. How these LAS can be used and designed in practice and which requirements they have to meet is a current research question. (Kuhn, Hellingrath and Hinrichs 2008)

This article contributes to this topic and starts with summarising and analysing requirements of balancing between logistic targets, typical planning situations in global multimodal supply networks and existing SCM concepts (section 2). In section 3 the focus is set on existing software approaches for SCM, showing that they can not fulfil all requirements of collaborative planning and control in global multimodal supply networks. In section 4 the concept of LAS is outlined, in section 5 a prototypical implementation of a LAS application is described. A conclusion and prospects are given in section 6.

2. REQUIREMENTS OF SCM TO INFORMATION TECHNOLOGY

This section is structured into three parts. First it will be shown that in today's corporations the main targets of logistics can only be balanced with effective usage of information technology (2.1). Then typical planning challenges in global supply chains are outlined (2.2) and major SCM concepts are presented, focussing on their requirements to IT support (2.3).

2.1. Balancing Logistic Targets

According to the mentioned logistic focus to supply a demand with the right product and the right quantity and type, at the right time, in the right place with the right costs LAS focus on all parts of the order process of an enterprise (Kuhn, Hellingrath and Hinrichs 2008; Plowman 1964). The logistic targets in this context can be divided into economic, performance and ecologic objectives (Fleischmann 2008). The economic objectives aim for targets like high machine and transport unit utilisation and a low capital commitment. The performance objectives aim for targets like high supply flexibility, high adherence to delivery dates and high service standards (Kuhn, Hellingrath and Hinrichs 2008). As shown before recent socio-political and industrial ambitions are the reason for a rising number of 'green logistics' concepts which aim for ecologic values like reducing carbon dioxide emissions or resource consumptions (Meißner 2008; Souren 2000).

Nickel and Vogel state that an effective IT system is needed to position an enterprise in between its conflicting logistic objectives (Nickel and Vogel 2006). Deiseroth, Weibels and Toth show that for an optimal decision process current and future process data is needed which can only be offered by an IT based software system integrating methods for forecasting, e. g. simulation (Deiseroth, Weibels and Toth 2008). Kuhn, Hellingrath and Hinrichs point out that a balance between those concurring objectives can only be reached, when decision processes in global multimodal supply chain planning and control are either partly or totally supported by LAS (Kuhn, Hellingrath and Hinrichs 2008).

2.2. Planning Situations in Global SCM

Typical planning situations in global multimodal supply networks arise from short conditions like short-term fluctuating customer demands, short-term order changes or supply shortfalls.

Fluctuating customer demand results in unsteady demand for parts that are purchased from suppliers. With every additional manufacturing level uncertainty for demand planning rises. Global multimodal supply networks often have multiple manufacturing levels which lead to extended lead times, great geographic distances and difficult planning situations at the last level of supply chains. Extended lead times also mean extended times for customer orders and a reduced flexibility to respond to fluctuations. This results in long-term planning periods for all levels of the supply chain. Information about the customer demands has to be shared over all levels as fast as possible to reach the best possible planning situation. If information is not running fast or only by the reaction of one manufacturing level to the next by showing the current demand, the 'bull-whip' effect occurs. That means that demand variation rises along the supply chain and with it inventory levels rise or drop and cause a high capital commitment or supply shortfalls (Fransoo and Wouters 2000). A collaborative planning based on information transparency in a supply network can avoid these impacts.

Short-term order changes by end customers have large impact on a supply network. This especially applies for products with a great variability. Here, order changes can result in an unchanged production volume on the one side, but in short-term changes for the demand of production parts on the other side. For some parts demand turns out to be much higher and for others a lot lower. To manage these order changes businesses can provide higher safety stock levels for the affected product parts, reduce supply chain lead times and try to reach high supply chain flexibility. As has been shown global multimodal supply networks are characterised by external lead times and reduced flexibility. Since high safety stock levels conflict with economic logistic targets businesses try to use the remaining supply chain flexibility to answer capable-to-promise (CTP) requests. With those the question if a certain demand of a product or product part is available at a certain time can be answered. In contrast to available-to-promise (ATP) requests not only the bills of materials, local production capacities and inventory levels are considered (Zhao and Ball 2005) but also supplier production capacities and their inventory levels. This way information about local and supplier production capacities, transport times and inventory levels of all levels of the supply chain including all parts in transit can be used for an effective collaborative planning to reach delivery reliability and answer CTP requests. In coordination and synchronisation processes supported by calculations all possibilities and scenarios to avoid supply shortfalls can be considered.

Supply shortfalls can result from low demand forecasting accuracy, quality problems, transport damage, loss, incorrect data about inventory levels, or external effects. For early identification of upcoming supply shortfalls forecasts on production demand, inventory levels, supplier production capacities and transport times are needed. Using this information an early identification of supply shortfalls can be reached by applying calculations and defined thresholds. This enhanced collaborative planning depends on consistent and up-to-date supply network information. If a supply shortfall cannot be avoided by an improved planning process with integrated methods for answering CTP requests, business partners still can initiate activities like additional transports, extra working shifts or using additional machines. For an effective planning of all of these activities it is necessary to exactly know how many additional product parts can reach the location of demand in which time and for which expenses. With sufficient relevant information about the supply network different scenarios can be calculated or forecasted.

The described planning situations show that in every case specific consistent and up-to-date information from all relevant business partners are needed. They were described as single planning situations but in practice they can occur in the same processes at the same time and this way can have interdependencies. For all described planning situations information transparency about the current process status and methods for calculations and planning are required. The information and results have to be presented to the decision makers and experts.

For an effective planning in SCM various concepts and methods have been developed and published in literature as will be shown below.

2.3. Concepts and Methods for SCM

Since its appearance business logistics has developed from principles about efficient supply into a very comprehensive approach. SCM and subordinate concepts have emerged, addressing all levels from strategy to operations, as illustrated in figure 1 (Baumgarten 2008).

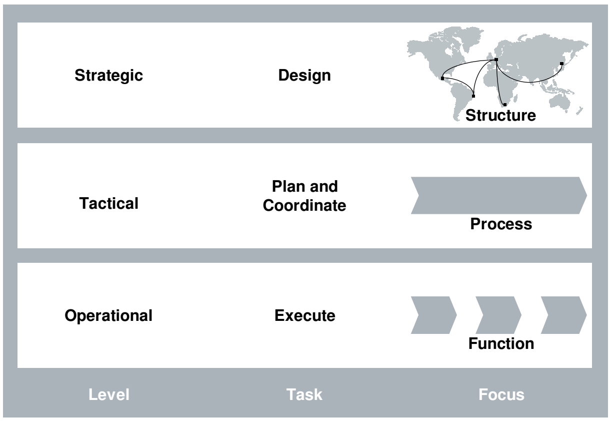


Figure 1: Scope of SCM (Baumgarten 2008)

Focusing on the customer and synchronising supply and demand, market needs should be met flexibly while stock levels are reduced (Kuhn and Hellingrath 2002). Among other things SCM includes the coordination and collaboration of the supply chain partners (CSCMP 2009). For a successful implementation of these concepts information and communication technology is essential.

One comprehensive approach which addresses the strategic, tactical and operational level is called collaborative planning, forecasting and replenishment (CPFR), a concept that identifies certain supply chain processes as being collaborative tasks, which means that they should be planned together by the supply chain partners. Having relevant information available is a precondition for effective collaboration, therefore a CPFR implementation should be supported by information technology. This can either be realised by one single CPFR application with centralised data management or through individual applications which can communicate via standardised interfaces like XML or EDI. (VICS 2004)

Visibility over the current state of the supply chain is an essential requirement for planning tasks on the tactical and operational level. The concept of supply chain monitoring (SCMo) supports supply chain partners by exchanging information about inventory and demand. This transparency should prevent stock-out and oversupply situations and optimise allocation in case of a bottleneck. For establishing visibility SCMo applications need data from the companies' back-end systems. (Odette 2003)

Access to relevant information is a very important precondition for collaborative demand and capacity planning (DCP) as well. Aligning the partners' production capacities with demand, DCP aims to avoid both capacity shortfalls and under-utilisation of capacities. A DCP process regularly checks if capacity is in line with demand, therefore DCP applications need data from the partners' back-end systems as well. (Odette 2004)

On the operational level the concept of supply chain event management (SCEM) aims to support the execution of agreed plans by automatically identifying unacceptable deviations and suggesting alternative solutions. A SCEM system should offer five basic functionalities: Monitoring, alerting, simulate, exception handling and measure. Monitoring and alerting are the core functions that a SCEM application has to offer. Therefore it needs status information from enterprise resource planning (ERP) and tracking systems, which then are processed and visualised. The other functionalities can be provided by other tools or components. (Hegmanns, Hellingrath and Toth 2008)

All concepts mentioned above have in common that reliable information about the supply chain is needed. It has to be collected, processed and presented by IT systems to support collaborative decision processes. Depending on individual requirements of specific planning situations in different supply chains, some elements of SCM concepts, like the ones mentioned above, are needed. In the following, existing software solutions are analysed to what extend they can support decision-making in collaborative supply chain planning.

3. EXISTING SOFTWARE SOLUTIONS FOR SCM DECISION SUPPORT

As previously shown an effective supply chain planning and control can only be done in collaborative decision processes and all mentioned concepts need software support that fulfils a number of assumptions. To support numerous business partners in a supply network with relevant and consistent data for collaborative planning and control, information about the current process like inventory levels, supplier capacities and production demand, have to be provided. Thus information asymmetries can be avoided. As shown above global and multimodal supply chains are often very complex and underlie fast and dynamic changes. Those characteristics determine that adequate software applications for decision support in SCM have to be very flexible to give sufficient possibility to readjust them to a changed process of the supply network. In addition to that smaller process changes (e. g. altered part numbers due to validities) should be integrated automatically by processing data from preceding operational systems. Further categories for the evaluation of supply chain decision support systems are: user interface capabilities; desired analytical tools like optimisation and simulation; methods for information presentation like reports, dashboards and tables; hardware and software requirements; compatibility for integration with existing systems (Simchi-Levi, Kaminsky and Simchi-Levi 2000).

In the past a great number of decision support systems focusing logistics and SCM concepts have been developed and vary from flexible mainframe systems, over isolated PC tools and client/server processes to high-performance and extensible enterprise decision-support applications (Simchi-Levi, Kaminsky and Simchi-Levi 2000).

In the following business intelligence (BI) technology for information analysis and processing and three major types of relevant existing software solutions for decision support in SCM will be analysed: advanced planning and scheduling (APS) systems; supply chain simulation (SCS) systems and SCMo systems as one example for a solution that focuses on a specific SCM concept. BI systems focus on consolidation, preparation and analysis of information for decision support of experts and executives. APS systems are optimisation-based modular extensions of ERP systems that focus on planning in SCM. SCS systems enable enterprises for strategic planning by using simulation studies and scenarios. As follows all concepts integrate some excellent methods for decision support, information processing and advanced planning but cannot fulfil the specific assumptions for the collaborative management and planning of high dynamic and complex supply networks.

3.1. BI Systems

The first concepts of BI systems have been developed in the 1960s, starting with management information systems and were used for decision support for executives. With progresses in IT development and growing areas of applications further concepts followed (e. g. decision support systems and executive information systems). Today they are all known as management information systems and usually integrate graphical user interfaces, functions of exception reporting and various standard interfaces to operational standard software systems. Today they support executives and experts in decision processes (Chamoni and Gluchowski 2006). Typical exception reporting functions are visual effects (e. g. markings or traffic lights for exceeding specific thresholds), instant messages (e. g. by e-mail, pop-up window or short message service) and drill-up and -down for the aggregation and detailing of data (Zwerenz 2006). Systems with integrated concepts like data warehousing, online analytical processing (OLAP) and data mining are called BI systems. A data warehouse is a central and consistent database which obtains comprehensive data from other operational databases. It is used for data integration and analysis (Navrade 2008; Chamoni and Gluchowski 2006). Systems that integrate OLAP are usually based on multi-dimensional data storage and enable complex analytical and ad hoc queries with rapid execution times (Kemper, Mehanna and Unger 2006). Data mining makes it possible to identify unknown data structures and performances by data analysis (Chamoni and Gluchowski 2006; Gluchowski, Gabriel and Dittmar 2006). The additional integration of pivot tables gives users the possibility to manually analyse table data by creating cross tables or by automated sorting, counting or totalling (Zwerenz 2006).

Today BI systems are standard software systems which usually will be customised for the implementation in single enterprises of all business sectors. Since they usually do not integrate methods for forecasting and collaboration the solutions are not set up for planning and control in global supply networks integrating multiple network partners. Since they do not focus on planning they also lack the functionality to integrate expert knowledge and master data for a model-based representation of the supply network.

3.2. APS Systems

The first ERP systems were released in the 1990s (Al-Mashari 2002). Today they integrate almost all operational core functions of any type of enterprise (e. g. accounting or material planning). As modular based and customised standard software systems they support and contain inter-departmental processes and functions (Betge 2006). Usually ERP systems integrate a single and central database that contains the data of all software modules. APS systems are additional modules of ERP systems for forecasting and supply chain planning and were published in the mid 1990's for the first time. Typically APS systems cover the areas demand planning, supply planning and manufacturing and scheduling (Simchi-Levi, Kaminsky and Simchi-Levi 2000). Usually they are based on optimisation procedures or industry-sector-specific planning methods based on mathematical algorithms (Stadtler 2004).

The successful integration of an ERP system with an additional APS module requires complex process analysis and adjustments and high expenses for license fees and customising. The customising can only be done within the limited flexibility of the ERP system structure. In addition to that ERP systems with APS modules often have high hardware requirements especially for servers and fast connections for data transfers, both causing further expenses. Especially an inter-organisational linking-up and collaboration are main challenges to APS systems (Betge 2006). Betge shows that a production stage planning for the whole or parts of the order to delivery process with the integration of multiple network partners cannot be supported by APS systems sufficiently (Betge 2006). The needed inter-organisational supply chain integration for supply chain planning and control cannot be accomplished with APS systems.

3.3. SCS Systems

For strategic decision support enterprises often use SCS systems. With modelling the entire or just specific parts of the supply network and processing simulation studies they can reach valuable support for strategic decisions. Usually SCS systems are used prior to the execution of a plan or in irregular intervals to verify a given or planned situation. This way various benefits can be reached: Understanding of supply chain processes and characteristics; capturing system dynamics (using probability distributions gives the possibility to model events and analyse the impact of those) and testing alternatives with simulation scenarios (Chang and Makatsoris 2001). For daily operative use network partners would need a simulation-based real-time system to monitor the network status and make decisions. Therefore an organisation needs to offer a number of capabilities: interfaces to legacy databases to obtain information; hardware and software for short-time simulation runs; interfaces to operative systems to assign tasks and receive feedback (Chang and Makatsoris 2001). Available SCS systems and technologies can offer the flexibility to automatically integrate smaller process changes by using data from operative systems.

Since SCS systems usually do not provide graphical user interfaces with comprehensive methods for information and presentation simulation results of those dynamic systems usually have to be integrated into secondary systems to be analysed and presented for decision support. The same applies for the execution of decision results which is usually done in additional operative systems. Missing concepts for collecting, evaluating and presenting data, SCS systems also lack functionalities to provide information transparency to experts and decision makers.

3.4. SCMo Systems

As mentioned above SCMo is a collaborative multi-level SCM concept that needs software support for processing information between network partners (Odette 2003). These software applications are known as SCMo systems. Their basic function is the exchange of production demands and inventory levels between business partners in supply networks to reach information transparency and avoid time lags in information flow. This data will be processed by comparison between present and future production demands and inventory levels. Depending on the results of the comparison and defined thresholds, automatic alerts (e. g. by sending a message or traffic lights on a dashboard view) will inform experts if inventory levels are out of range. By cumulating the inventory levels over the whole supply chain and considering the demands of the last business partner in a supply chain demand solutions for the 1^{st} to n^{th} tier can be calculated (Odette 2003). With this advanced functionality the bull-whip effect can be avoided and a smooth and secure supply chain with low inventory levels is supported. For SCMo applications interoperability, scalability and technical robustness are the main technical requirements. Full interoperability can be reached when every business partner in a supply network can choose one application, depending on their specific needs. Standard interfaces can provide the flexibility for information exchange between different systems (Odette 2003). Recent developments in SCMo applications integrate Supply Chain KPI and assortment of graphical tools (e. g. predefined charts, cockpits or dashboards) for information presentation (Bäck and Gössler 2006). Typical KPIs that can be found in SCMo applications are forecasting accuracy and range of material (Hegmanns, Hellingrath and Toth 2008).

SCMo systems provide all functionalities for monitoring the current process status of a supply chain or network but they miss methods for forecasting and planning like optimisation and simulation.

3.5. Insufficient Existing Software Solutions

As shown above the analysed system technologies and methods cannot integrate the entire functionality of providing information transparency for supply chain control, forecasting for supply chain planning and executing for decision processing. The high complexity and fast dynamics of supply networks, the process specific supply chain structures and the integration of global network partners for collaboration require lean and affordable software systems that can offer comprehensive and process specific software functionality. The potential methods of BI technology, APS systems, SCS systems and SCMo systems should be chosen and applied as required in given process. Considering the mentioned planning situation an optimal software solution has to combine parts of the described IT concepts: The data collection and evaluation can be supported by methods of BI systems, an optimal information presentation for supply chains can be provided by concepts of SCMo systems and for forecasting and planning methods of APS or SCS systems can be used.

4. LOGISTIC ASSISTANCE SYSTEMS AS A NEW SCM SOFTWARE STANDARD

This section introduces LAS as a paradigm for the development of flexible, lean and specified software systems that can be applied to all parts of the order process of a manufacturing business and have a special focus on decision support in supply chain planning and control (Kuhn, Hellingrath and Hinrichs 2008). Focusing on global multimodal supply chains they can give multiple benefits to all involved network partners.

Subsection 4.1 presents the basic idea of LAS, illustrates the main characteristics and functions and shows which concepts and methods can be used for supply chain planning and control and ends with a concept definition. After that, the different possible areas of application are described (4.2) and potentials and limits of LAS are shown (4.3).

4.1. LAS as Specified Software Solutions

The basic idea of LAS is to design specified software solutions for given planning situations in global multimodal supply chains which integrate and present all relevant information for decision support. The concept focuses on enhanced collaboration, consistent information over the whole supply chain, transparency about the current process status, planning functionality and fast system development and implementation.

Taking a closer look LAS integrate the following main characteristics and functions: information transparency, information processing, decision support, collaborative planning, process orientation and software flexibility.

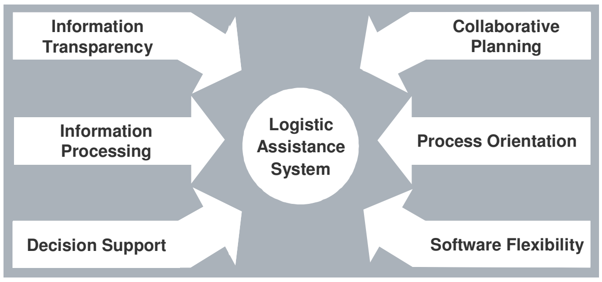


Figure 2: Characteristics and functions of LAS

'Information transparency' in LAS means that all relevant information about a planning task can be provided to decision makers and experts. On the one hand the software system reveals the current process status including information like inventory levels, production demands, positions of transport units or supplier production capacities and on the other hand planning information about forecasted future developments.

By 'information processing' data from experts and various back-end systems of all relevant and integrated network partners can be consolidated, evaluated and presented in the software system. The evaluation of the consolidated information can be processed by calculations and defined thresholds, rules for data developments and other data exceptions. The presentation of the consolidated and evaluated data to decision makers and experts can be done by using various forms of illustration like tables, text or diagrams.

'Decision support' in LAS means that depending on the given process and the integrated functionality the software system can support single or all parts of the human decision process from the decision preparation over option selection to decision execution and control (Kuhn, Hellingrath and Hinrichs 2008). In the decision preparation for identifying a solution LAS help to reveal all needed information for a correct alternative selection. After choosing an alternative the user or expert is able to directly execute the decision in LAS. The last part the decision control can enable the user to control the chosen decision execution and its impact on the subsequent process (Blutner, Cramer and Krause 2007; Toth and Wagenitz 2009).

To support 'collaborative planning' LAS combine data from all relevant network partners. This way information about all considered levels of the supply network can be used for information processing and decision support. Supporting the planning process, manual changes in data can be tracked and documented by users. In addition to that it is possible to integrate individual user views by an implemented role-based user management. This way proprietary information can be revealed to selected user groups, network partners or organisational departments. To support an enhanced communication process for collaborative planning workflows and message systems can be applied.

'Process orientation' follows the basic idea of developing lean and specified software systems that provide all needed information and functionality for a given planning situation. That means that only relevant functionality and information are selected and integrated in the software system. In addition to that user views can be designed according to the focused process. This way, users reach a valuable decision support for a given planning situation, transparency about the interdependencies in the supply network and the possible impacts of planning scenarios and executed decisions.

'Software flexibility' in LAS means on the one hand that the software system is flexible to be adjusted to process changes in dynamic supply networks and on the other hand that it is possible to integrate data from various back-end system by using standard interfaces and selected IT methods for information processing and collaborative planning. In addition to that a fast system development and implementation at affordable expenses is followed.

Based on the specific requirements of a given process and the involved users and experts, an optimal selection of IT methods and concepts has to be found for a reliable and optimal integration of LAS. Usually it is necessary to consolidate data from various back-end systems and data bases for data analysis and information transparency. As shown above BI systems offer reliable methods for this use: data warehouse, OLAP, data mining and exception reporting. For forecasting and planning APS systems offer optimisation procedures and specific planning methods based on mathematical algorithms. In addition to that or as an alternative it is possible to use dynamic SCS systems for supply chain planning and forecasting. Under consideration of the assumptions of the selected methods of the shown IT concepts it is possible to combine them in a lean LAS specified to the given process. Since it will be used by multiple organisational departments and/or network partners an additional central user management system combined with a web-based graphical user interface can reduce the administrational efforts. At the same time it enables for the integration of role based and specified user windows and data access. The presentation of the data can be done with text, tables, diagrams or KPIs supported with methods of exception reporting, pivot tables and drill-up and drill-down functionalities. Dashboard views are another way for presentation like they are often used in SCMo systems. With an integrated message system (e. g. news, e-mail, SMS or pop-up window) an optimal communication between all experts and users can be supported. If there are standardised processes with successive tasks and decisions for different experts it is also an option to integrate them as semi-automatic workflows that will be processed in a previously defined and standardised order.

As an example considering the planning situation of short-term order changes described in section 2.2 a LAS can provide all needed functions for planning and control: To reveal the current process status of the supply network data from various back-end systems of all relevant network partners can be integrated in a database (e. g. data warehouse). For the given planning situation information about supplier production capacities, inventory levels of all levels of the supply chain and the current production demand are most important. In addition master data about the supply chain can be integrated and continuously updated from different IT systems or expert knowledge. Especially the supply chain structure including transit times and the bills of materials are important for planning calculations. To show a current and consistent process status to all involved network partners and organisational departments a graphical user interface (e. g. web-based browser views) can give an overview. Using technology for forecasting and planning (e. g. SCS systems) information about the current process status the production demands, bills of materials, inventory levels over the whole supply chain including parts in transit and transit times can be processed to identify future process developments. By using safety stock levels or defined thresholds the current system status can be evaluated by calculations if stock levels are too high or too low. An effective interpretation of the evaluated information can be achieved by using some of the described presentation concepts (e. g. diagrams or dashboards) and exception reporting methods. ATP or CTP requests for short-term order changes can be answered by manually changing production demand data and run simulation scenarios. If supplier capacities or inventory levels are too low the evaluation of the data can help to identify bottlenecks and present it to decision makers.

It is a great potential of LAS that they can offer the possibility of integrating large amounts of data and complex procedures which overtop the cognitive skills of humans (Blutner, Cramer and Krause 2007). LAS can integrate data, information and expert knowledge from various back-end systems, complex process structures, individually selected algorithms and methods for data processing and forecasting. Nevertheless they support processes in supply networks that are too complex and have too many intern and extern influences for total automatic and autonomous software based processing. At least the final option selection still has to be left to the involved experts. To them the LAS can provide all relevant information for an effective decision process.

After all, LAS can be defined as lean, flexible and specified software systems for decision support in supply chain planning and control and their central functions are collecting, evaluating and presenting all relevant information about the current process status and providing planning support. Depending on the specific area of application they integrate information about certain process steps and relations, about up-to-date data and analysis from preceding operational systems, mathematical methods and algorithms and concepts for forecasting and planning.

4.2. Area of Application

One way of classifying the area of application for LAS focuses on the type of decision process, if it regards operational control or tactical and strategic planning. The other one concentrates on the involved organisational units of the order process.

LAS can be used for decision processes with varying scope. They can regard multiple stages of the supply network and their impact can be different concerning time, geographic range, monetary influences and process changes. Within the operative logistic control especially tasks like active order planning, transport monitoring and container management can be found (Blutner, Cramer and Krause 2007). In transport monitoring LAS provide the possibility to consider ecologic objectives like carbon dioxide emissions. The tactical logistic planning regards operations and decision processes of the sales planning, the production planning, the transport planning and the purchase planning. The strategic logistic planning includes the design of supply networks, location planning and the design of stock and production sites (Kuhn and Hellingrath 2002).

The areas of application of LAS concerning the different organisational units are various. For the mentioned tasks of the strategic planning specified software systems can support in the logistic planning department. In the material control department an efficient and collaborative material control can be reached by providing information transparency over the whole supply network with history, current and future data. Using forecasted (e. g. by simulation) planning data for multiple stages of the supply network the department for demand and capacity management is able to reduce supply shortfalls by increasing supplier capacities with short or long term arrangements. By integrating simulation technology it is also possible to analyse the inventory levels of the whole supply chain for alternative production programs to achieve flexibility in the order process. With data preparation, aggregation, analysis and the integration of KPIs for supply chain planning and control the supply chain controlling reaches transparency over the current process and its history or future development. The KPIs can consider all mentioned logistic targets and focus on economic, performance and ecologic objectives.

4.3. Potentials and Limits

After all it can be determined that LAS offer great potentials for providing experts with relevant information and specified functionality for decision processing in global multimodal supply chain planning and control.

A high transparency of the stages of the supply network can help to identify future supply shortfalls or too high inventory levels. Both can help to avoid logistic costs due to production stops, additional transports or working shifts. Deviations to defined thresholds and other events can be noticed early and enable for quick response. Forecasts for inventory levels and multiple stages of the supply network and a collaborative planning may avoid the bull-whip effect and help to regulate safety stock levels concerning production flexibility and capital lockup. By the integration of transport routes and current information on transport unit positions (e. g. by satellite technology for trucks or ships) timetables can be checked and adjusted. A shared consistent and comprehensive data base (e. g. data warehouse) supports an effective collaborative planning. In using LAS it is possible to enhance decision and communication quality for collaborative decision processes. In addition to that experts and users may reach learning effects because they cognise their decisions impact on the supply chain.

To achieve these great potentials there are some challenges that have to be managed. Similar to all collaborative concepts a common future process and business targets have to be defined with all involved partners of the supply network. Responsibilities have to be assigned and unequal advantages have to be agreed on or balanced out. The main necessary requirement is that no involved business partner will reach advantages that come out as disadvantages for another one (Hegmanns, Hellingrath and Toth 2008). The implementation of collaborative processes means for all involved network partners that they have to trust each other and share data and process information. Often this means that proprietary data has to be shared as well. As mentioned before there are solutions in literature how to synchronise plans (e. g. Dudek and Stadtler 2004).

Since global multimodal supply chains integrate network partners from different countries this also means that they integrate different cultures which have varying communication processes, different technology standards with different computing performances and interfaces and different standards for data management (e. g. manual or automatic data collection). This results in a great challenge for continuously and automatically collecting all relevant, current and correct data from various back-end systems. IT-standards for interfaces like EDI or XML enormously help to solve those problems.

The LAS has to be constructed as a lean and flexible software system with acceptable computational performance requirements. This way it is possible to run the system in countries with low technology standards and adjust it to their local standards. Especially by using web-based technology it is possible to keep computational performance requirements and administration efforts low. In addition to that system flexibility is needed to be able to adjust the software system to process changes in the dynamic supply network immediately. The way of constructing a system only with the relevant and selected functions keeps the system development process short and affordable.

5. LAS FOR COLLABORATIVE SUPPLY CHAIN PLANNING IN THE AUTOMOTIVE INDUSTRY (VOLKSWAGEN AG)

The benefits from using LAS will now be shown on three examples from the automotive industry. Major original equipment manufacturers (OEM) produce at worldwide locations to enter growing markets and to utilise lower production costs. For cost reasons supplier parts are used in different vehicle models. Most suppliers are located in the countries of the established OEM sites. Due to a lack of local suppliers OEM production sites in growing markets have to be supplied over long distances. This is organised as multimodal transports with container vessels for main carriage. Using logistics service providers for consolidation scale effects can be realised. A typical global supply chain is shown in figure 3.

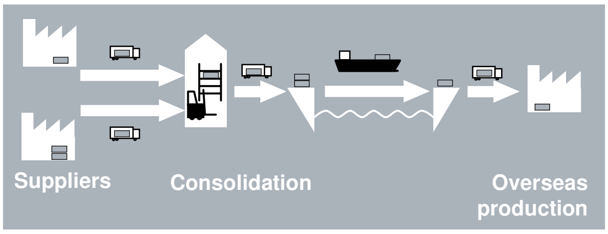


Figure 3: Supply chain in the automotive industry

Since 2006 Volkswagen AG corporate IT, Volkswagen Commercial Vehicles and Fraunhofer Institute for Material Flow and Logistics (IML) have been developing and testing LAS in the field of global SCM. A framework is used that can be easily customised to create individual LAS for specific applications. The supply network is represented by an object-relational database with predefined objects, classes and inheritance for the order-to-delivery process (OTD). Supply chain models can be created and filled with data via an XML interface.

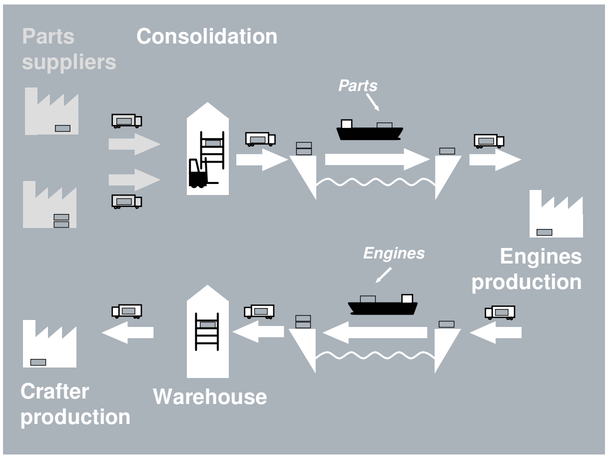


Figure 4: Crafter engines and parts supply chain

A first prototype was created to gain more flexibility in a challenging supply process: Parts built into the engines of the Volkswagen Crafter are delivered from Europe to South Africa, where the engines are produced. These engines are then shipped to Germany where the Crafter is built. Shipping alone takes several weeks. Since the safety stock has to be low it does not cover the entire time span between call-off and delivery. Production orders are released late in the process, so that call-offs are based on forecasted quantities. Varying customer demand can only be met as long as there is enough stock in the pipeline, consisting of safety stock and other stock, which has been built up due to batch sizes or earlier varying demand. For utilising this stock and gaining more flexibility, there has to be full visibility over demand, lead times, inventory levels and their interaction. Figure 4 shows a simplified illustration of the supply chain.

The LAS prototype is being used for controlling this supply chain of engine parts and assembled engines since 2007. It is based on a client-server-architecture, providing access to all relevant information on inventory and demand, which are extracted from different data sources and kept in a central supply chain database. The data also includes information about the shipping parts and engines and their estimated arrival. Having this data available is a precondition for answering capable-to-promise requests. For this functionality the OTD-NET simulation engine has been integrated into the LAS. OTD-NET has been developed by Fraunhofer IML and Volkswagen for simulations of the OTD process and has been used for numerous studies. Considering dynamic parameters, future inventory levels can be forecasted on each client. The results are presented in customised views (figure 5); a potential running out of stock is shown using traffic light colours. Flexibility is gained and emergency transports by airfreight can be avoided. (Deiseroth, Weibels and Toth 2008; Wagenitz 2007)

Figure 5: Planning view of the “Crafter Tool”

Recent projects extend these features adapting it to new requirements. In the automotive industry in-house logistics service providers centrally supply production sites in new markets with parts. Serving several customers within the cooperation, these divisions have to avoid shortage situations while demand is varying and delivery time is long. A new web-based application replaces the clients, offering features and views for significantly higher data volume. An integrated user management regulates the access to information and provides data security. Input data is extracted from different data warehouses and operative systems. Creating visibility over all integrated supply chains and simulating future scenarios can stabilise planning and reduce cost for additional transports. A prototype is currently developed by Fraunhofer IML and Volkswagen AG.

A third project deals with the control of supply chains from different logistics service providers to a production plant in South America. Besides the features already mentioned above it is planned to consider production capacities from suppliers as further input data. In a future step real customer orders, calculation of dependent demand requirements and cross-checking with inventory levels and supplier capacities will be integrated into a simulation model. In addition CO_2 emissions of transports will be calculated. They will be displayed historically, for current and for future transports and cover the standard process as well as additional future transports like airfreight. The prototype of the LAS is expected to be released in the fourth quarter of 2009.

6. CONCLUSION AND PROSPECTS

It has been shown that all mentioned SCM concepts need software systems that can be implemented for collaboration of multiple partners of a supply network. Existing software systems offer various concepts and methods for preparation, aggregation and analysis of data from multiple interfaces to support expert decision processes. They also provide forecasting functionalities for planning but they all cannot offer the flexibility to be immediately implemented to specific supply chain parts and for multiple network partners.

Existing software solutions cannot fulfil all requirements of collecting, evaluating, forecasting and presenting information for supply chain control and planning in global multimodal supply networks. With the integration of selected methods of these IT systems LAS can be developed as lean and specialised software systems that provide all relevant functionality and information for a given situation in supply chain planning and control. To realise the great potentials of LAS, they have to be specified and customised to the given process and involved experts and be implemented in a short time. Another assumption is the willingness of all involved network partners for collaboration. Both can only be reached with a preceding and common planning process. As shown above under consideration of these requirements of collaborative concepts first applications in industry show convincing results and follow-ups have been started already.

REFERENCES

Al-Mashari, M., 2002. Enterprise resource planning (ERP) systems: a research agenda. In: *Industrial Management & Data Systems*, 102 (3), pp. 165–170.

Bäck, S., Gössler, G., 2006. SCM-Kompetenz-Management – Focus: Planungs- und Dispositionssysteme. In: Corinna, E.-N., 2006. *Ausbildung in der Logistik*. 1st edition, Wiesbaden: Deutscher Universitäts-Verlag, pp. 155–179.

Betge, D., 2006. *Koordination in Advanced Planning and Scheduling-Systemen,* Wiesbaden: Deutscher Universitäts-Verlag.

Barratt, M., 2004. Understanding the meaning of collaboration in the supply chain. *Supply Chain Management: An International Journal*, 9 (1), pp. 30–42.

Baumgarten, H., 2008. Das Beste in der Logistik – Auf dem Weg zu logistischer Exzellenz. In: Baumgarten H., ed. *Das Beste der Logistik – Innovationen, Strategien, Umsetzungen.* Berlin: Springer-Verlag, pp. 12–19.

Blutner, D., Cramer, S., Krause, S., Möncks, T., Nagel, L., Reinholz, A., Witthaut, M., 2007. *Assistenzsysteme für die Entscheidungsunterstützung. Endbericht der Arbeitsgruppe 5. Technical Report SFB 559 (Modellierung großer Netze der Logistik),* Universität Dortmund.

Chang, Y., Makatsoris, H., 2001. Supply Chain Modelling Using Simulation. *International Journal of Simulation*, 2 (1), pp. 24–30.

Chamoni, P., Gluchowski, P., 2006. Analytische Informationssysteme. Einordnung und Überblick. In: *Analytische Informationssysteme. Business Intelligence-Technologien und -Anwendungen*, 3rd edition, Berlin: Springer-Verlag, pp. 1–22.

Council of Supply Chain Management Professionals. *CSCMP Supply Chain Management Definitions.* Available from: http://cscmp.org/aboutcscmp/definitions.asp [accessed 17 June 2009].

Deiseroth, J., Weibels, D., Toth, M., Wagenitz, A., 2008. Simulationsbasiertes Assistenzsystem für die Disposition von globalen Lieferketten. In: Markus Rabe ed. *Advances in Simulation for Production and Logistics Applications*. Stuttgart: Fraunhofer IRB-Verlag, pp. 41–50.

Dudek, G., Stadtler, H., 2005. Negotiation-based collaborative planning between supply chain partners, *European Journal of Operations Research*, 163 (3), pp. 668–687.

Fleischmann, B., 2008. Grundlagen: Begriff der Logistik, logistische Systeme und Prozesse. In: Arnold, D.. Isermann, H., Kuhn, A., Tempelmeier, H., Furmans, K., ed. *Handbuch Logistik*. 3rd edition, Berlin: Springer-Verlag, pp. 458–484.

Fransoo, J., Wouters, M., 2000. Measuring the bullwhip-effect in the supply chain. In: *Supply Chain Management: An International Journal*, 5 (2), pp. 78–89.

Gluchowski, P., Gabriel, R., Dittmar, C., 2008. *Management Support Systeme und Business Intelligence. Computergestützte Informationssysteme für Fach- und Führungskräfte.* 2nd edition, Berlin: Springer-Verlag.

Hegmanns, T., Hellingrath, B., Toth, M., Maaß, J.-C., 2008. Prozesse in Logistiknetzwerken. Supply Chain Management. In: Arnold, D., Isermann, H., Kuhn, A., Tempelmeier, H., Furmans, K., ed. *Handbuch Logistik*. 3rd edition, Berlin: Springer-Verlag, pp. 458–486.

Ireland, R., Bruce, R., 2000. CPFR Only the Beginning of Collaboration. *Supply Chain Management Review*, 4 (4), pp. 80–88.

Kuhn, A. and Hellingrath, B., 2002. *Supply Chain Management. Optimierte Zusammenarbeit in der Wertschöpfungskette.* Berlin: Springer.

Kuhn, A., Hellingrath, B., Hinrichs, J., 2008. Logistische Assistenzsysteme. In: *Software in der Logistik. Weltweit sichere Supply Chains*, München: huss-Verlag, pp. 20–26.

Kemper, H.-G., Mehanna, W., Unger, C., 2006. *Business Intelligence. Grundlagen und praktische Anwendungen.* 2nd edition, Wiesbaden: Vieweg.

McLaren, T., Head, T., Yuan, Y., 2002. Supply chain collaboration alternatives: understanding the expected costs and benefits. In: *Internet Research: Electronic Networking applications and Policy*, 12 (4), pp. 348–364.

Meißner, M., 2008. Alle reden vom Klima. In: *Software in der Logistik. Weltweit sichere Supply Chains*, München: huss-Verlag, pp. 27–32.

Mentzer, J., Foggin, J., Golicic, S., 2000. Collaboration – The Enablers, Impediments and Benefits. In: *Supply Chain Management Review*, 4 (4), pp. 52–58.

Navrade, F., 2008. *Strategische Planung mit Data-Warehouse-Systemen,* 1st edition, Wiesbaden: Gabler.

Nickel, R., Vogel, M., 2006. Entwicklung eines Assistenzsystems für das Produktionscontrolling, In: *Industrie Management,* 22 (4), Berlin: GITO Verlag, pp. 61–64.

Odette International, 2003. *Supply Chain Monitoring* Version 1.0. Available from: http://www.odette.org/ [accessed 17 June 2009]

Odette International, 2004. *Demand Capacity Planning*. Version 1.1. Available from: http://www.odette.org/ [accessed 17 June 2009]

Plowman, E., 1964. *Lectures on Elements of Business Logistics*, Stanford: Stanford University – Graduate School of Business.

Simchi-Levi, D., Kaminsky, P., Simchi-Levi, E., 2000. *Designing and Managing the Supply Chain – Concepts, Strategies, and Case Studies.* Singapore: McGraw-Hill.

Souren, R., 2000. Umweltorientierte Logistik. In: Dyckhoff, H. 2000. *Umweltmanagement. Zehn Lektionen in umweltorientierter Unternehmensführung*. Berlin: Springer-Verlag, pp. 151–168.

Srivastava, S., 2007. Green supply-chain management: A state-of-the-art literature review. In: *International Journal of Management Reviews*, 9 (1), pp. 53–80.

Stadtler, H., 2004. Supply Chain Management – An Overview. In: Stadtler, H.; Kliger, C., 2004, *Supply Chain Management and Advanced Planning – Concepts, Models, Software and Case Studies,* 3rd edition, Berlin: Springer-Verlag, pp. 9–33.

Toth, M., Wagenitz, A., 2009. Neue Wege für die effektive Planung logistischer Netzwerke. Dynamische Verfügbarkeitsplanung mit Hilfe von Assistenzsystemen. In: *Industrie Management*, 25 (2), Berlin: GITO Verlag, pp. 55–58.

Voluntary Interindustry Commerce Solutions Association (VICS), 2004. *CPFR. An Overview.* 18 May 2004. Available from: http://www.vics.org/ [accessed 17 June 2009].

Wagenitz, A., 2007. *Modellierungsmethode zur Auftragsabwicklung in der Automobilindustrie.* Dissertation Universität Dortmund.

Wyk, J. van, Baerwaldt, W., 2005. External Risks and the Global Supply Chain in the Chemical Industry. In: *Supply Chain Forum: An International Journal*, 6 (1), pp. 2–15.

Zhao, Z., Ball, M., 2005. Optimization-Based Available-To-Promise with Multi-Stage Resource Availability. In: *Annals of Operations Research*, 135, pp. 65–85.

Zwerenz, K., 2006. *Statistik – Datenanalyse mit EXCEL und SPSS*, 3rd edition, München: Oldenbourg.

AUTHORS BIOGRAPHY

Felix Bockholt is doctoral student at the faculty of mechanical engineering of the Dortmund University of Technology. At the same time he leads a project for the implementation of a logistic assistance system in the Demand Capacity Management of Volkswagen Commercial Vehicles in Hanover.

Wolfgang Raabe works as a doctoral student in the corporate IT department of Volkswagen. In a staff unit he is involved in the research on the order-to-delivery process. With a scientific focus on global supply chains he leads two projects for the prototypical implementation of logistic assistance systems.

Dr.-Ing. Michael Toth is consultant and scientist at Fraunhofer Institute for Material Flow and Logistics IML and head of the LAS-Department (logistic assistance system). He has been working for the IML since 2003 and is responsible for industry projects as well as application-oriented research projects, especially in the areas automotive and plant engineering.

MULTI-OBJECTIVE GENETIC LOCAL SEARCH ALGORITHM FOR SUPPLY CHAIN SIMULATION OPTIMISATION

Galina Merkuryeva[(a)], **Liana Napalkova**[(b)]

[(a) (b)] Department of Modelling and Simulation, Riga Technical University, 1 Kalku Street, Riga LV-1658, Latvia

[(a)]Galina.Merkurjeva@rtu.lv, [(b)]Liana.Napalkova@rtu.lv

ABSTRACT
This paper presents a hybrid simulation optimisation algorithm that integrates a multi-objective genetic algorithm and response surface-based metamodelling techniques. The optimisation problem involves a search in a high dimensional space with different ranges for decision variable scales, multiple stochastic objective functions and problem specific constraints. A case study demonstrates the application of a hybrid simulation optimisation algorithm to optimal cyclic planning for a generic supply chain network.

Keywords: simulation optimisation, genetic algorithm, response surface-based metamodelling, supply chain cyclic planning

1. INTRODUCTION

Pareto-based evolutionary algorithms are powerful algorithms for solving complex multi-objective problems. They present the class of direct search methods that apply the concepts of Pareto-optimality and dominance relation. They evolve multiple parallel solutions that allow generating a set of non-dominated solutions and preserving a diverse set of the candidate solutions. In addition, they are able to perform a search in a high dimensional space with different ranges of decision variables and could incorporate constraint handling techniques, such as rejection of unfeasible solutions, penalty function, etc. The obvious difference between various types of evolutionary algorithms is in encoding mechanism of candidate solutions, i.e. strings over a finite alphabet in genetic algorithms (GAs); real-valued vectors in evolution strategies; finite state automata in evolutionary programming and trees in genetic evolutionary programming. In particular, GA's encoding mechanism allows easy implementation of specific constraints in supply chain planning problems.

The need for hybridization of Pareto-based search algorithms with other methods and techniques is widely discussed in literature (Hiroyasu et al 1999). In this paper, hybridisation of Pareto-based genetic algorithms with response surface-based metamodelling techniques applied to simulation optimisation is investigated.

Response surface methodology (RSM) is a collection of statistical and mathematical techniques for optimisation of stochastic functions. RSM is developed in order to analyse experimental data and to build empirical models based on observations of the stochastic function. In case of a computer simulation model, simulation output dependence on its input variables could be interpreted by a response surface function. The main advantage of RSM methodology (Merkuryeva 2005) is its applicability for a small number of observations in conditions of costly and time-consuming simulation experiments. RSM-based optimization is based on local approximation of the simulation response surface by a regression type metamodel in a small region of independent factors. Exploration and optimisation of the resulting response surface function approximations provide effective tools for iterative optimisation of the simulation response function applicable for a small number of simulation experiments.

For the last years, there has been an increasing attention placed on the performance, design, and analysis of multi-echelon supply chains (Merkuryev et al 2008). Optimisation of multi-echelon cyclic plans in supply chains that is performed in the case study refers to the class of multi-objective optimisation problems, which are usually characterised by a large search space of decision variables, conflicting and stochastic objectives etc. While there is no a single optimal solution for a number of conflicting objectives, the development of an algorithm, which gives a large number of alternative solutions lying near the Pareto-optimal front and tackles the variations of a response generated from the uncertainties in the environment variables, is of great practical value.

The case study is aimed to find the optimal parameters of a multi-echelon cyclic plan for each of supply chain nodes in order to minimise the supply chain total cost and maximise end-customers fill rate taking into account cyclic planning constraints and assumptions of stochastic demand and backordering.

A network simulation model (Merkuryeva and Napalkova 2009) is built as a process-oriented model with a one-directional flow of goods. It is represented by two types of nodes: stock points and processes. The stock points correspond to stock keeping units, while the processes denote transformation activities (assembly, transportation and packaging operations).

2. OPTIMISATION PROBLEM

The stochastic optimisation problem is formulated as follows:

$$\text{Min E}[\mathbf{f}(\mathbf{x})] = \text{E}[f_1(\mathbf{x}),\ldots, f_M(\mathbf{x})], \quad (1)$$
$$\text{subject to: } \mathbf{g}(\mathbf{x}) = \text{E}[\mathbf{r}(\mathbf{x})] \leq 0 \text{ and } \mathbf{h}(\mathbf{x}) \leq 0,$$

where E[·] is a mathematical expectation; $\mathbf{x} = (x_1,\ldots x_K) \in X$, $\mathbf{f} = (f_1,\ldots,f_M) \in Z$; K is the number of decision variables; M is the number of objective functions; X is the decision space; Z is the objective space; $\mathbf{x}$ is a vector of decision variables; $\mathbf{f}$ is a vector of objective functions; $\mathbf{g}$ is a vector of stochastic constraints; $\mathbf{h}$ is a vector of deterministic constraints on the decision variables; $\mathbf{r}$ is a random vector that represents several responses of the simulation model for a given $\mathbf{x}$. Decision variables could be of different types (i.e. discrete, continuous) and have metrics with different ranges of possible values. Proceeding from (1), the solution of multi-objective simulation-based optimisation problem is interpreted as a vector of decision variables $\mathbf{x}$ that satisfies all feasible constraints and provides the best trade-off between multiple objectives.

To describe the objective vector function, one could use traditional methods of aggregating multiple objectives into a single one, or optimising the most important objective while treating others as constraints. The main strength of these techniques is their computational efficiency and simple implementation. The weakness is the difficulty to determine a priori information about the objectives, such as weight coefficients that reflect a relative importance of each criterion or their ranks. Moreover, this approach can produce only one optimal solution during a single experiment, which may not be the best trade-off. However, in multi-objective optimisation problems, each objective function could have individual optimal solution and none of them can be considered to be better than any other with respect to all objective functions. Therefore, in order to find trade-off solutions this paper applies the principles of Pareto optimality. The dominance relation can be formulated in the following way:

A trade-off solution $\mathbf{x}^* \in X$ *is said to dominate a solution* $\mathbf{x} \in X$ *iff* $\forall i \in \{1,\ldots,M\}: f_i(\mathbf{x}^*) \leq f_i(\mathbf{x})$ *and* $\exists j \in \{1,\ldots,M\}: f_j(\mathbf{x}^*) < f_j(\mathbf{x})$.

As it follows from the definition, a solution $\mathbf{x}^*$ is Pareto-optimal if it is not worse than any other solutions for all criteria and is better for at least one criterion. For simplification, it is assumed here that all objective functions are minimised.

In the case study, two objective functions that define the quality of a multi-echelon cyclic plan are introduced. The first one is aimed at minimising the average total cost of the supply chain, which includes the sum of production, inventory and reordering costs. The second objective function is aimed to maximise end-customer service requirements specified by the average order fill rate. The parameters of a multi-echelon cyclic plan identify decision variables. They are: replenishment of cycles Cy_i and order-up-to-levels S_i defined at each stock point i on the network. These variables determine the reorder period and quantity to be ordered or produced for each mature product and are interpreted as discrete and continuous type variables, correspondingly. A large number of decision variables in practice make conducting simulation experiments difficult. Specific constraints are introduced that define cycles by the power-of-two policy, in which cycles are integers and multiples of two.

3. ANALYSIS OF THE PROBLEM SOLVING METHODS

In order to solve the aforedescribed simulation-based optimisation problem (1), an appropriate method should be applied. Based on analysing the properties of the objective functions and decision search space, the following requirements are imposed on the problem solving method: (i) it should converge to the approximate Pareto-optimal front while keeping its diversity, (ii) it should be able to guide the search toward a near-optimal direction using only the numeric values of the multiple stochastic objective functions and constraints, (iii) it should incorporate some techniques for generating statistically significant candidate solutions and (iv) it should manage a search process in such a way that the total number of simulation experiments and, in consequence, the total computational time would be decreased.

The first and second requirements restrict the choice to the class of *direct search methods* that apply the concepts of Pareto-optimality and dominance relation. Pareto-based evolutionary algorithms (EAs) refer to the most efficient representatives of this class. The tremendous advantage of EAs over others is that they evolve multiple parallel solutions instead of a single one that allows generating a set of non-dominated solutions at each iteration. On the other hand, EAs are able to preserve a diverse set of non-dominated solutions using specific mechanisms. In addition, EAs are able to perform a search in a high dimensional space with different ranges for decision variables. Moreover, EAs have proved to be independent on strong problem structure, such as, for example, convexity and discontinuity of the objective function. Also, they allow one to incorporate different constraint handling techniques, such as rejection of unfeasible solutions, penalty function, etc.

Although Pareto-based EAs are powerful algorithms for solving complex multi-objective problems, they are unable to fulfil all of the above-formulated problem requirements. This fact clearly illustrates the need for hybridisation of Pareto-based EAs with others methods and techniques.

Typically, the hybridisation is performed following some predefined scheme. In literature, it is possible to outline three hybridisation schemes, such as *parallel hybridisation*, *sequential hybridisation* and *built-in hybridisation*. Parallel hybridisation requires that the search space is investigated independently by multiple optimisation methods. For instance, a population can be divided into sub-populations called islands, which are associated with particular objective functions or certain ranges of the Pareto-optimal front. Another example includes dividing the control of genetic operators between computer processors. The parallel hybridisation scheme is implemented in a divided range multi-objective genetic algorithm (DRMOGA) (Hiroyasu et al 1999), parallel strength Pareto multi-objective evolutionary algorithm (PSPMEA) and parallel multi-objective evolutionary algorithm with a hypergraph represented population structure (pMOHypEA).

In using sequential hybridisation, separate methods are sequentially combined based on predefined rules. According to the most widespread implementation of this scheme, Pareto-based EAs are combined with local search based methods. The reason is that EAs have overall global perspective, while the local search based methods have good convergence properties to a local optimal solution and can be used to extensively explore the search space around EA solutions. On this way, the simple multi-objective genetic local search (S-MOGLS) algorithm probabilistically applies the local search to candidate solutions found by the fast elitist non-dominated sorting genetic algorithm (NSGA-II) (Deb et al 2000). In using the local search, multiple objective functions are aggregated based on randomly generated weight coefficients. Another example of implementing the sequential hybridisation scheme is to apply the local search based method after running Pareto-based EA. The weight coefficients used are computed for each EA solution based on its location in the Pareto-optimal front.

Built-in hybridisation concerns with introducing some features into mechanisms of Pareto-based EAs. Recent developments in this domain include the application of fuzzy logic to (i) the dynamical adjustment of the crossover and mutation rates in the NSGA-II algorithm (Deb et al 2000), (ii) selection of more preferable solutions from the Pareto-optimal set based on their degrees of fuzzy optimality and (iii) incorporation of fuzzy ranking scheme, in which dominance degrees are measured by using membership functions.

The aforementioned hybridisation schemes have been mainly tested on deterministic analytical models. However, the simple combination of Pareto-based EAs with the simulation model may not provide efficient results because of time consuming simulation experiments and the simulation noise, which influence the objective function estimates and performance of EA operators. Thus, the ongoing section is dedicated to a hybridised approach to multi-objective simulation-based optimisation, which can be mentioned as a very promising and at the same time poorly investigated field of research.

4. OPTIMISATION ALGORITHM

The proposed simulation optimisation algorithm (Figure 1) is based on integration of the multi-objective genetic algorithm (GA) and RSM-based linear search algorithm (Merkuryeva and Napalkova 2008). While a GA is well suited to solve combinatorial problems and is used to guide the search towards the Pareto-optimal front, the RSM-based linear search is appropriate to improve GA solutions based on the local search.

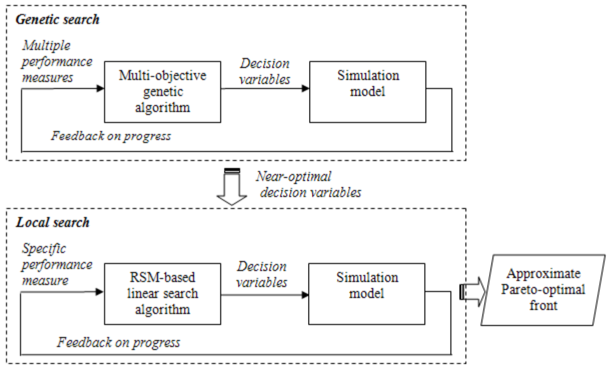


Figure 1: General scheme of the hybrid simulation optimisation algorithm

The multi-objective genetic algorithm starts with generating an initial population of decision variables values (Napalkova and Merkuryeva 2008). In order to smoothly cover the investigated search space, uniform distribution is applied. Decision variables such as cycles are encoded using a modified binary encoding procedure, which satisfy power-of-two synchronisation policy in supply chains. Afterwards, fitness values are defined based on multiple objectives, here by the average total cost and average fill rate that are obtained through simulation experiments. To estimate fitness values of chromosomes, a ranking-based fitness assignment is applied. It concerns the use of a dominance depth that is connected with dividing a population into several fronts in order to represent a front of a certain solution. In order to obtain solutions uniformly distributed over the Pareto-optimal front, the diversity preserving mechanism based on a crowding distance metric is implemented. The crowding distance is an estimate of the density of solutions surrounding the current solution. The larger a crowding distance value becomes, the less crowded an area around the solution is. As a result, every chromosome in the population has the following two attributes: (1) domination depth and (2) crowding distance.

Then, the penalty function is applied to decrease the survival probability of solutions, which provide the average fill rate lower than the pre-defined threshold. In order to choose chromosomes from the current population for breeding purposes, the algorithm applies a crowded two-tournament selection. The main idea of this selection strategy is that a crowded comparison operator is used to compare pairs of chromosomes. From two candidate solutions the one with the lower domination depth is preferable. If both solutions have the same depth, then the solution with larger crowding distance is selected. The crowded comparison operator ($\geq$) is defined as follows:

$$a \geq b \text{ if } (r_a < r_b) \text{ or } ((r_a = r_b) \text{ and } (\delta_a > \delta_b)), \qquad (2)$$

where r_a and r_b are domination depths, δ_a and δ_b are crowding distances for chromosomes *a* and *b*.

After applying the crossover and mutation operators, the new population is replaced by the union of the best parents and offspring to avoid the loss of non-dominated solutions during the evolution process. Domination depths of chromosomes in the combined population are updated. The first *N* solutions are gathered for the next generation, where *N* is a population size. This elitist strategy is often called ($\mu + \lambda$) – selection, where μ and λ assign parents and offspring, respectively. The multi-objective GA is automatically terminated, when the number of generations with stagnant non-dominated set is equal to the predefined value (usually set to 3).

In the local search, the RSM-based linear search iterative algorithm is used to improve decisions solutions of the genetic algorithm by adjusting specific decision variables, e.g. order-up-to levels in supply chains. The algorithm is based on local approximation of the simulation response surface by a regression type meta-model in a small region of independent factors; it integrates linear search techniques for optimising stock points' order-up-to levels. Finally, the approximate Pareto-optimal front initially generated by the GA is updated including solutions found by the response surface-based linear search algorithm.

5. CASE STUDY

The case study is aimed to find an optimal cyclic plan of a chemical product, i.e. liquid based raisin, in order to minimise production, ordering and inventory holding costs, and maximise end-customers fill rate. As a test bed, the chemical manufacturing supply chain is used. The main operations that occur in the supply chain network are the following. In the plant CH, the raw material is converted to the liquid based raisin. It is then either sourced to direct customers or shipped to the plant DE, where other components are added to make different products. From that plant, the end-products are shipped to different types of customers.

The Service Model-based simulation model of the above-described supply chain network is automatically generated in optimisation environment developed in (Merkuryeva and Napalkova 2007). The end-customer demand is normally distributed and cycles are defined according to the power-of-two policy. Cycles are represented in weeks as follows: 7, 14, 28, 56, where 56 days is the maximal cycle that corresponds to one full turn of a "planning wheel". In this business case, specific policies such as nested or inverted-nested ones are not analysed. Order-up-to levels are calculated using analytical formulas, where the cycle service level is set to 95%. Initial stocks are equal to order-up-to levels plus average demand multiplied by cycle delays. Stock point 1 has infinite on hand stock and is not controlled by any policy. Backorders are delivered in full.

Simulation run length is equal to 224 periods. This allows modelling of four full turns of the planning wheel, i.e. 4*56 periods. Number of simulation replications is equal to 5. The GA is executed with the following parameters: the population size is 40; crossover and mutation probabilities are 0.5 and 0.1, correspondingly; a tournament size is equal to 2. The GA works with 66 decision variables (i.e. cycles and order-up-to levels assigned to network stock points). Initial values of order-up-to levels are calculated analytically. When the number of generations with a stagnant non-domination set is equal to 3, the GA is terminated. Figure 2 shows solutions received from the final population.

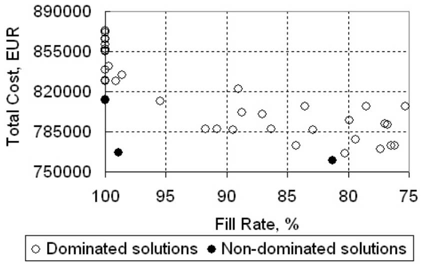


Figure 2: Final GA population

Figures 3 and 4 illustrate the execution of the GA. The average total cost and fill rate of parent chromosomes are plotted against the generation step. The GA makes quick progress at the beginning of the evolutionary process that is typical for genetic algorithms. Then, there are phases when it hits the local optimum before mutations further improve its performance. Finally, the GA finds three non-dominated solutions with the following performance average measures: 1) *total cost* = €787,431, *fill rate* = 100.00%; 2) *total cost* = €766,669, *fill rate* = 98.88%; and 3) *total cost* = €752,300, *fill rate* = 93.76%.

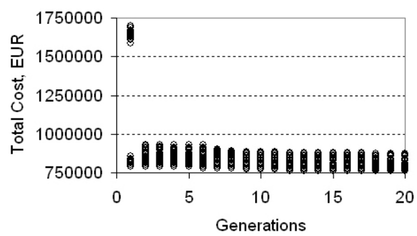


Figure 3: The GA convergence subject to total cost

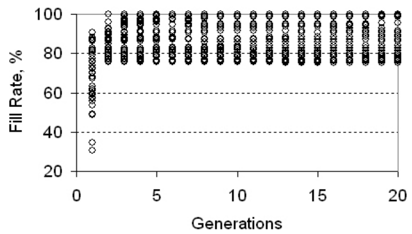


Figure 4: The GA convergence subject to fill rate

The response surface-based linear search algorithm is used to adjust order-up-to levels of three non-dominated solutions received with the GA while fixing stock point cycles. Finally, the average total cost and average fill rate of the second solution are equal to €756,178 and 98.88%, respectively. The updated Pareto-optimal front is given in Figure 5. There are three non-dominated solutions found by the GA, where the second solution is improved by the RSM-based linear search algorithm.

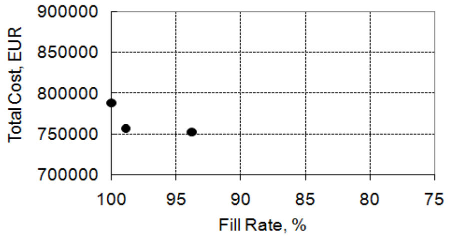


Figure 5: The approximate Pareto-optimal front

CONCLUSIONS

The paper has presented the hybrid simulation optimisation algorithm that integrates the multi-objective genetic algorithm and response surface-based linear search algorithm. Although genetic algorithms are widely applied at solving different real world multi-objective problems, they are often unable to ensure both the convergence to the Pareto-optimal front and its diversity. In this paper, genetic algorithm allows covering a broad region of the search space at each generation, while RSM-based linear search algorithm provides careful investigation of small portions of the search space and improves the current solution by moving to a better "neighbour" solution. The results of the case study have demonstrated performance efficiency of the proposed hybrid algorithm.

REFERENCES

Campbell, G.M. and Mabert, V.A., 1991. Cyclical Schedules for Capacitated Lot Sizing with Dynamic Demands. *Management Science*, 409 – 427.

Deb, K., Agrawal, S., Meyarivan, T., 2000. A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimisation: NSGA-II. *Proceedings of the Parallel Problem Solving from Nature VI Conference*, pp. 849 - 858.

Hiroyasu, T., Miki, M. and Watanabe, S. (1999). Divided range genetic algorithms in multiobjective optimization problems. *Proceedings of International Workshop on Emergent Synthesis (IWES'99)*, pp.57-66.

Merkuryev, Y., Merkuryeva, G., Desmet, B., Jacquet-Lagrèze, E., 2007. Integrating Analytical and Simulation Techniques in Multi-Echelon Cyclic Planning. *Proceedings of the First Asia International Conference on Modelling and Simulation (AMS 2007)*, March 27 – 30, Phuket, Thailand, pp.460 – 464.

Merkuryeva, G., 2005. Response surface-based simulation metamodelling methods with applications to optimisation problems. Chapter 15. In: Dolgui, A., Soldek, J. and Zaikin, O., eds. *Supply chain optimisation Product / Process Design, Facility Location and Flow control*. Springer, 205–215.

Merkuryeva, G., Merkuryev, J. and Napalkova, L., 2007. Simulation-based environment for multi-echelon cyclic planning and optimisation. *Proceedings of the 19th European Modeling and Simulation Symposium*, pp.318-325. October 4-6, Bergeggi, Italy.

Merkuryeva, G. and Napalkova, L., 2009. Supply Chain Cyclic Planning and Optimisation. In: Merkuryev Y., Merkuryeva G., Pierra M.A., Guash A., eds. *Simulation-Based studies in Logistics: Education and Applied Research*. UK: Springer-Verlag, 89.-111.

Merkuryeva, G. and Napalkova, L., 2009. Two-phase simulation optimization algorithm with applications to multi-echelon cyclic planning. *International Journal of Simulation and Process Modelling* (in press).

Merkuryeva, G. and Napalkova, L., 2007. Development of simulation-based environment for multi-echelon cyclic planning and optimisation. *Proceedings of the 6th EUROSIM Congress on Modelling and Simulation*, paper ID 452.

Napalkova, L. and Merkuryeva, G., 2008. Theoretical Framework of Multi-Objective Simulation-Based Genetic Algorithm for Supply Chain Cyclic Planning and Optimisation. *Proceedings of the 10th International Conference of Computer Modelling and Simulation*, pp.467-474. April 1-3, Cambridge, UK.

AUTHORS BIOGRAPHY

GALINA MERKURYEVA is Professor at the Department of Modelling and Simulation, Riga Technical University. She holds Doctor of Engineering (Dr.sc.ing) from Institute of Electronics and Computer Techniques of Latvian Academy of Sciences and Doctor of Sciences (DSc.) from Institute of Control Sciences of Russian Academy of Sciences. G. Merkuryeva has professional interests and experiences in discrete-event simulation, simulation metamodelling and optimisation, decision support systems, supply chain simulation and simulation-based training.

LIANA NAPALKOVA holds her MSc degree in Computer Science from Riga Technical University (2006). Currently, she is a PhD student at RTU Department of Modelling and Simulation, and participates in research projects in the logistics field. Her interests focus on the use of simulation-based optimisation techniques for providing competitive advantages in multi-echelon supply chain cyclic planning. Liana Napalkova is a member of the Latvian Simulation Society.

SIMULATION VERSUS ANALYTICAL MODELLING FOR SUPPLY CHAIN DYNAMICS ANALYSIS

Julija Petuhova[(a)], Yuri Merkuryev[(b)], Maris Buikis[(c)]

[(a), (b)]Department of Modelling and Simulation, Riga Technical University, 1 Kalku Street, Riga, LV-1658, Latvia
[(c)]Department of Mathematical Statistics, Riga Technical University, 1 Kalku Street, Riga, LV-1658, Latvia

[(a)]julija@itl.rtu.lv, [(b)]merkur@itl.rtu.lv, [(c)]mbuikis@latnet.lv

ABSTRACT
Simulation has become an important information technology tool for analysis and improvement of entire supply chain operation. One of the most important supply chain operation stability measures is a bullwhip effect value. The bullwhip effect can lead to holding an excessive inventory, insufficient capacities and high transportation costs. It is important to investigate the magnitude of this effect by quantifying it. There are a variety of methods, which address bullwhip effect modelling. Nevertheless, there is a lack of methods for its numerical evaluation in a supply chain. This paper proposes statistics and probability theory based analytical methods which allow quantification of the bullwhip effect value in a supply chain that operates under uncertain market demand. Simulation technique is used to validate results obtained from the analytical model. Based on validation results, the logic of the analytical model is examined, and some specifications of the analytical model are analysed and described.

Keywords: simulation, analytical modelling, the bullwhip effect, supply chain under uncertain demand

1. INTRODUCTION

A precise representation of dynamic, time-dependent changes in a system operation is important for analysis and evaluation of the bullwhip effect in a supply chain. The most common distinction between dynamic models is drawn between the so-called continuous and discrete models (Crosbie 2000). A continuous model is usually based on ordinary or partial differential equations with time as an independent variable. The state of the system is assumed to change in a continuous fashion so that at any particular instant in time the state of the system will be uniquely defined. A discrete model assumes that the state of the system changes only at specific times, often referred to as events, and the state of the system is unchanged between these times.

Both differential equation (more typical for technical system models), and discrete-event simulation (more typical for business system models) could be used to evaluate the bullwhip effect value in a supply chain.

The bullwhip effect evaluation methods are classified as follows:

1. Continuous system modelling (analysis of an order flow):
    - control theory based (Laplace transformation, z-transformation);
    - system dynamics.
2. Discrete-event system/Analytical modelling (analysis of an individual order):
    - discrete-event system simulation;
    - analytical simulation.

Simon (1952) was one of the first to use classical Laplace transform techniques to analyse simple production-inventory systems (Riddalls et al. 2000). This move was quickly translated into the newly favoured discrete z-domain by the Operations Research community (Dejonckheere et al. 2003). Contributions that utilise the Laplace transform are more numerous than those utilising the z-transform. This is probably due to more tractable algebraic manipulation required when using the Laplace transform. However, as the z-transform is a special case of the Laplace transform, many tools, techniques and best practices developed for the Laplace transform are readily exploited in the z-domain – usually after a small change in notation.

Disney and Towil (2002) developed a z-transform model from which an analytical expression for the bullwhip effect is derived that is directly equivalent to the common statistical measure, the Coefficient of Variation (COV) often used in simulation, statistical and empirical studies to quantify the bullwhip effect (Disney and Towill 2002). Dejonckheere et al. (2003) developed a control engineering insights based methodology (the transfer function, the frequency response plot) for the analysis of the bullwhip effect value using different demand forecasting methods (Dejonckheere et al. 2003). Forrester (1961) started to analyse a supply chain phenomena – demand amplification using a continuous time model (Forrester 1961). This work facilitates the development of the systems dynamics in the production and inventory control fields. Sterman (1989) characterised causes of unsuccessful decisions in supply chain management and realised a number of experiments using system dynamics principles (Sterman 1989). Barlas and Aksogan (1997) developed apparel industry supply chain model, using system dynamics simulation and analyse effectiveness of inventory and production policies in supply chain management (Barlas and Aksogan 1996). Hennet (2005) proposed a system dynamics based method to analyse multi-stage supply chain taking into consideration random variation of demand for final product under classical distributed production and ordering policies (Hennet 2005).

Zhang and Zhang (2007) used discrete-event simulation to evaluate the bullwhip effect value in three-stage supply chain with different information sharing strategies (Zhang and Zhang 2007). They also studied possibilities of decreasing the lead time uncertainty and the effect of this on the bullwhip effect using simulation (Zhang et al. 2006).

Ingals et al. (2005) analysed the impact of a control-based forecasting method on the bullwhip effect value by performing simulation-based experimental study (Ingals et al. 2005). Chandra et al. (2001) investigated information sharing impact on the demand forecast accuracy and the bullwhip effect value in a supply chain through discrete-event simulation (Chandra et al. 2001).

The developed analytical models usually support an analysis of different factors impact on the bullwhip effect value, but not evaluation of the value itself. For example, Simchi-Levi et al. (2002) explained that the increase in demand variability with the necessity for each supply chain stage makes orders based on the forecasted demand of the previous stage (Simchi-Levi 2002). Since variability in placed orders is significantly higher than that in customer demand, the supply chain stage is forced to carry more safety stock in order to meet the same service level. The proposed quantifying of the magnitude of increase in variability between two neighbour supply chain stages is expressed as a function of a lead time between the orders receiving and the number of demand observation on which forecast is made:

$$\frac{Var(Q)}{Var(D)} \geq 1 + \frac{2L}{p} + \frac{2L^2}{p^2}, \qquad (1)$$

where

Var(Q) – the variance of the orders placed by the supply chain stage;

Var(D) – the variance of the demand seen by this supply chain stage;

L – lead time between the orders receiving;

p – number of observation on which further demand forecast is based.

As a result, the bullwhip effect is magnified with increasing the lead time and decreasing the observations number.

Luong and Phien (2007) evaluated the bullwhip effect value in two echelon supply chain using AR(2) autoregression model. However, it is necessary to determine autoregression coefficients and for providing a precise evaluation of the bullwhip effect value it is recommended to use a higher degree autoregression model that leads to complication of the bullwhip effect value calculation (Luong Huynh Trung et al. 2007). Kelle and Milne (1999) suggested to evaluate a variance of placed orders (bullwhip effect) in inventory systems that implement the *S-s* inventory control policy using approximations of the quantitative model, developed in accordance with asymptotic renewal theory (Kelle and Milne 1999).

The performed analysis of the bullwhip effect evaluation methods elicits that the developed analytical models allow analysis of the impact of different factors on the bullwhip effect value, but not evaluation of the value itself. Researches in this area are still developing and are considered perspective in supply chain management.

This paper develops statistics and probability theory based methods for analytical evaluation of the bullwhip effect value based on the numerical measures of the customer demand distribution.

The rest of the paper is organised as follows. In the next section, the importance and impact of the stochastic factors on the management of a supply chain is discussed. This is followed by the description of the supply chain analysed. An explanation of the elaborated statistics and probability theory based methods for the evaluation of the bullwhip effect value is given. A numerical example and a simulation-based validation of the results of the analytical solution are discussed. Finally, conclusions are provided.

2. STOCHASTIC FACTORS IN A SUPPLY CHAIN MANAGEMENT

Stochastic factors have a major influence on the behaviour of a supply chain and its management. Stochastic nature of the customer demand is established as one of the most critical factors in decision-making, since many of the uncertainty sources can be handled adequately only at the tactical level (Van Landeghem and Vanmaele 2002), and planning cycle of supply chain processes is coordinated with material flow, which is characterised by demand volume. That's why, planning and controlling problems of a supply chain at the tactical level under the stochastic customer demand are analysed in the paper. Operation of any supply chain depends on the customer demand and its fluctuation. However internal demand between supply chain stages also plays an important role due to it considerably changes the information about the required product amount. An important phenomenon in the supply chain management is the increase in variability of the demand as it moves through the supply chain in the direction from customer to supplier. This phenomenon's name is the bullwhip effect (Simchi-Levi et al. 2002) because even small disturbances in demand at the customer level cause the demand amplification for the next supply chain member (see Fig. 1.).

A measure of the demand variability is the standard deviation of demand, σ. An increase in this value at each supply chain stage directly indicates the existence of the bullwhip effect. The bullwhip effect is considered to be an important characteristic of supply chain operation stability.

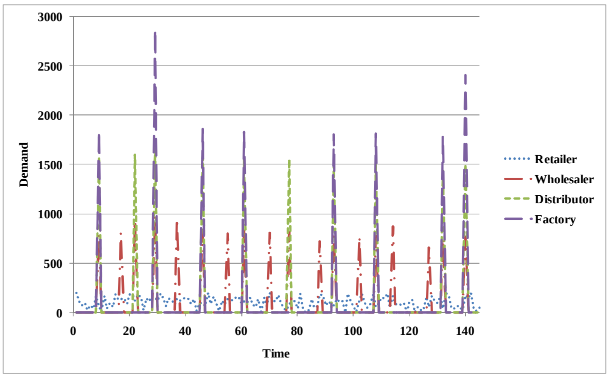


Figure 1: The Bullwhip Effect in a Supply Chain

The main consequence of the bullwhip effect appearance is holding an excessive inventory or/and inventory allocation in inappropriate supply chain stages.

3. SUPPLY CHAIN ANALYSED

The main objective of the research is to develop analytical methods for evaluation of the bullwhip effect value based on the numerical measures (mean and standard deviation) of the customer demand distribution in a supply chain.

The supply chain is analysed from the inventory management point of view, when it is represented as a serially connected inventory management systems chain. The considered supply chain consists of the end customer, retailer and supplier. The retailer supplies the customer single-item products according to the demand received, and replenishes its inventory by placing orders to the supplier. Customer demand is stochastic and stationary. For managing the inventory, the *s-S* inventory control strategy (see Fig. 2.) is used.

The magnitude of increase in variability of the placed order size with regard to variability of received demand value characterises the bullwhip effect. Its value can be expressed analytically, taking into consideration the numerical measures of the customer demand distribution – an expected value *E(X)* and variance *D(X)*.

It is assumed that the demand X_1, X_2, …, X_i is a discrete random sample observed from some population. Equivalently, these data are independent and identically distributed (IID) observations on some underlying random variable *X* whose distribution governs the population. Values that numerically characterise the population/distribution, such as an expected value *E(X)* and a variance *D(X)* of the discrete random variable *X* are given.

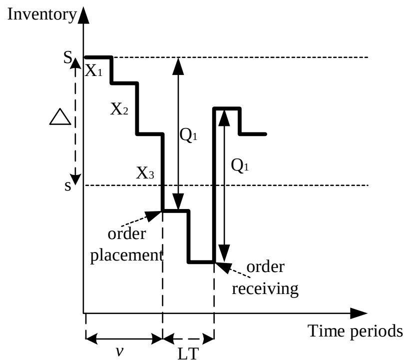


Figure 2: Inventory Control Strategy in the Retailer Stage

The inventory level to which inventory is allowed to drop before a replacement order is placed (reorder point level) is found by a formula:

$$s = E(X) * LT + STD(X) * \sqrt{LT} * z, \qquad (2)$$

where

LT – constant lead time between replenishments;

$STD(X) = \sqrt{D(X)}$ - standard deviation of the mean demand;

z - the safety stock factor, based on a defined in-stock probability during the lead time.

The total requirement for the stock amount or target stock level *S* is calculated as a sum of the reorder point level and a demand during the lead time value:

$$S = s + E(X) * LT \qquad (3)$$

The order size Q_i is demanded when the on-hand inventory drops below the reorder point; it is equal to the sum of the demand values between the order placements:

$$Q_i = X_1 + X_i + \ldots + X_v, \qquad (4)$$

where

v – random variable, the number of period in which order is placed.

Regular or cyclical in nature inventories with additional safety stock are considered. A strategy to control such inventories assumes that the conditions of demand level, its variability and lead time are known and involves the following main steps:

1. find the current on-hand inventories at the stocking point;
2. establish the stock availability level at the stocking point after the demand satisfaction;
3. calculate total requirements that is the amount of cycle stock plus additional quantities needed to cover the uncertainty in demand;
4. determine an order size as the difference between the total requirements and the quantity on hand in case if the on-hand inventory drops

below the allowed level when a replacement order should be placed.

4. ANALYTICAL METHODS FOR THE BULLWHIP EFFECT EVALUATION

4.1. Statistics-based method

Provided that the demand X is uncertain and the aforementioned inventory control strategy is employed, the placed order size Q is expected to be a random variable that depends on the demand values. The expected value $E(Q)$ and variance $D(Q)$ of the function $Q = \varphi(X)$ are estimated using the following formulas proposed by Feller (1967):

$$E(Q) = E(X) * E(v), \tag{5}$$

and

$$D(Q) = E(v) * D(X) + D(v) * [E(X)]^2, \tag{6}$$

where

$E(v)$ – expected value of a number of periods between placed orders;

$D(v)$ – variance of a number of periods between placed orders.

The number of periods between placed orders v characterises the frequency of order placements but its probabilistic behaviour is estimated by numerical characteristics: an expected value and a variance.

The multi-experimental realisation of the following algorithm allows one to collect statistics of v values (v_i, $i = \overline{1,n}$) and evaluate their probabilities p_i by relative frequencies $\hat{p}_i$ of their occurrence, which in its turn allows one to estimate v expected value and variance:

if $X_1 > \Delta$ THEN v=1 AND STOP
ELSE generate X_2
if $X_1 < \Delta$ and $X_1 + X_2 > \Delta$ THEN v=2 AND STOP
ELSE generate X_3
…
if $X_1 + X_2 + \ldots + X_{n-1} < \Delta$ and $X_1 + X_2 + \ldots + X_n > \Delta$ THEN v=n
STOP

The performance of the supply chain is evaluated under various factors such as end customer mean demand $E(X)$ and its standard deviation $STD(X)$, safety stock factor z and a lead time LT. It is assumed that end customer demands arrive with fixed time-intervals, and their value is variable and is derived from a normal distribution. A constant lead time between replenishment is considered. No order processing delay is taken into account, so all demand events are treated immediately by the inventory system. We will also assume no capacity constraints for supplier of the inventory system. In this case, stockouts will lead to backorders, not lost sales.

One of simulation techniques usage assignments in modelling and analysing inventory systems is analytical model validation and verification (Banks and Malave 1984). It is the most important part of a simulation study that enables determining whether an analytical model of an inventory management system performs as intended and is an accurate representation of the real-world system under study. The simulation model was developed using the ARENA 12.0 simulation modelling environment. Simulation is used to analyse and evaluate the increase in variability of placed orders in the described supply chain and to validate the results of the analytical solution (Petuhova and Merkuryev 2007).

The model was run for 1 replication. Replication length is defined as 2000 time periods. The warm-up period is avoided by setting the initial inventory level equal to the lower control limit called the reorder point s. Numerical results of comparison of the statistics and simulation based methods are given in Table 1.

Table 1: Validation Results of the Analytical, Statistics-Based Method for the Evaluation of the Bullwhip Effect Value

$STD(Q)_{an}$	$STD(Q)_{sim}$	χ^2_{fact}	$\left[\chi^2_{1-0.95,389}; \chi^2_{0.95,389}\right]$	result
29.89	17.93	233.28	[344.29; 435.99]	disagree
32.01	17.71	215.23	[344.29; 435.99]	disagree
35.13	17.89	198.08	[344.29; 435.99]	disagree
39.79	18.63	182.15	[344.29; 435.99]	disagree
47.36	20.74	170.35	[344.29; 435.99]	disagree

To compare results given by analytical and simulation models the χ^2 test is used. The validation indicates that results given by the analytical model proved to be in disagreement with those given by the simulation model with a probability of 95%. The variance of placed orders calculated by analytical model is much greater than actual variance of placed orders derived from the simulation model (see Table 1). The reason for the inadequate bullwhip effect quantification by the analytical model is an existing dependence between a number of periods when an order is placed v and realisations of the end customer demand X_i. In other words, the proposed formulae (5) and (6) assume independence of v and X, but in the described inventory management system they are dependent in the way of conditional probability of v occurrence $p_v = P(X_1 + X_2 + \ldots + X_v > \Delta / X_1 + X_2 + \ldots + X_{v-1} < \Delta)$.

The proposed statistics-based method could be used for the evaluation of the bullwhip effect value based on the numerical measures of the customer demand distribution in the case, when a period of order placement is independent of the value of the received demand. For example, if period of the order placement v is an independent random variable with known theoretical distribution, e.g., normal distribution with mean equal to 4 and standard deviation equal to 1, then performed experimental study confirms that analytically obtained results agree with those given by simulation (see Table 2).

Table 2: Validation Results of the Analytical, Statistics-Based Method for the Evaluation of the Bullwhip Effect Value in Case of Order Time Period Independency

$STD(Q)_{an}$	$STD(Q)_{sim}$	χ^2_{fact}	$\left[\chi^2_{1-0.95,245};\ \chi^2_{0.95,245}\right]$	result
53.85	58.19	264.72	[209.76;282.51]	agree
75.39	81.46	264.70	[209.76;282.51]	agree
96.93	104.68	264.58	[209.76;282.51]	agree
118.47	127.95	264.59	[209.76;282.51]	agree
140.01	151.20	264.58	[209.76;282.51]	agree

The validation indicates that results given by the analytical model agree with those given by the simulation model with a probability of 95%.

4.2. Probability theory based method

To consider a correlation between demand value X_i and time period of order placement v, a probability density function of the order size Q should be defined. It is obtained by integration of a total probability formula of the sum of received demand values ($S_Q=X_1+X_2+...+X_v$). The distribution function of the random variable S_Q with regard to the total probability formula is:

$$\begin{aligned} F_Q(x) &= P(S_Q < x) = P(X_1 + X_2 + ... + X_v < x) = \\ &= P(S_Q < x/v = 1)\cdot P(v = 1) + P(S_Q < x/v = 2)\cdot P(v = 2) + \\ &+ P(S_Q < x/v = 3)\cdot P(v = 3) + ... = P(X_1 < x)\cdot P(v = 1) + \\ &+ P(X_1 + X_2 < x)\cdot P(v = 2) + P(X_1 + X_2 + X_3 < x)\cdot P(v = 3) + ... \end{aligned} \qquad (7)$$

where

$P(v=i),\ i=1\div\infty$ – probability that the order will be placed in the i^{th} time period;

$P(S_Q<x/v=i)$ – probability that the order size will be less than x in time period v.

Analytically estimating the probability of the period when order is placed $P(v=1, 2, 3, ...)$ and probability when the sum of the demand values reaches the Δ level $P(X_1+X_2+X_{3+...}<x)$ it is possible to define the following distribution function of the order size $F_Q(x)$:

$$\begin{aligned} F_Q(x) &= \int_0^x f(x_1)dx_1 \cdot \left(1 - \int_0^\Delta f(x)dx\right) + \\ &+ \int_0^x f(x_1) \int_0^{x-x_1} f(x_2)dx_2 dx_1 \cdot \int_0^\Delta f(x_1) \int_{\Delta-x_1}^\infty f(x_2)dx_2 dx_1 + \\ &+ \int_0^x f(x_1) \int_0^{x-x_1} f(x_2) \int_0^{x-x_1-x_2} f(x_3)dx_3 dx_2 dx_1 \cdot \int_0^\Delta f(x_1) \int_0^{\Delta-x_1} f(x_2)\cdot \\ &\cdot \int_{\Delta-x_1-x_2}^\infty f(x_3)dx_3 dx_2 dx_1 + \\ &+ \int_0^x f(x_1) \int_0^{x-x_1} f(x_2) ... \int_0^{x-x_1-...-x_{i-1}} f(x_i)dx_i dx_{i-1}...dx_1 \cdot \\ &\cdot \int_0^\Delta f(x_1) \int_0^{\Delta-x_1} f(x_2) ... \int_{\Delta-x_1-x_2-x_{i-1}}^\infty f(x_i)dx_i dx_{i-1}...dx_1 \end{aligned} \qquad (8)$$

where

$f(x_i)$ – probability density function of the customer demand distribution;

Δ – the difference between target inventory level S and reorder point s;

i – the number of random variable values; the more addends are available, the more precisely the distribution describes the random variable Q behaviour.

To define a probability density function of the order size $f_Q(x)$, its distribution function $F_Q(x)$ (8) should be derived:

$$\begin{aligned} f_Q(x) &= [F_Q(x)]' = f(x)\cdot\left(1 - \int_0^\Delta f(x)dx\right) + \\ &+ \int_0^x f(x_1)f(x-x_1)dx_1 \cdot \int_0^\Delta f(x_1) \int_{\Delta-x_1}^\infty f(x_2)dx_2 dx_1 + \\ &... + \int_0^x f(x_1) \int_0^{x-x_1} f(x_2) \int_0^{x-x_1-...-x_{i-2}} f(x-x_1-...-x_{i-1})dx_{i-1}dx_{i-2}...dx_1 \cdot \\ &\cdot \int_0^\Delta f(x_1) \int_0^{\Delta-x_1} f(x_2) ... \int_{\Delta-x_1-x_2-x_{i-1}}^\infty f(x_i)dx_i dx_{i-1}...dx_1 \end{aligned} \quad . \qquad (9)$$

Knowing the random variable probability density function (9) it is possible to define analytically its numerical characteristics: expected value (10) and variance (11):

$$E(Q) = \int_\Delta^\infty x f_Q(x)dx\,, \qquad (10)$$

$$D(Q) = \int_\Delta^\infty x^2 f_Q(x)dx - [E(Q)]^2\,. \qquad (11)$$

where $f_Q(x)$ – probability density function of the order size.

The variance of order size Q characterises the bullwhip effect value in the supply chain. Knowing distribution function of the customer demand and its numerical characteristics it is possible to estimate the variance of the order size Q by applying the developed probability theory based method.

Practical application and validation of the proposed probability theory based methods is a subject of future research.

5. CONCLUSIONS

Two methods for the analytical evaluation of the bullwhip effect value in the supply chain are discussed in this paper. The statistics-based method could be used for evaluating the order size variance based on the distribution parameter values of the customer demand if the period when order is placed v does not depend on the customer demand value X. In its turn, the probability theory based method could be used when the dependence does exist.

Obviously, the simulation supports evaluation of the bullwhip effect in all supply chain configurations and it could be taken as a general technology for supply chain operation analysis and planning. Analytical approaches to analysing the supply chain operation do not need any special software; however mathematical methods are not always able to describe complicated dynamical and stochastic systems. The research performed allows one to conclude about the simulation technology advantages for analysing dynamical systems that operate in a stochastic environment.

With the aid of the developed statistics and probability theory based methods it is possible to conclude about supply chain operation stability and to verify the adequacy of the created supply chain simulation model. There are developed a number of simulation models during the research, and various experiments are performed with them in order to demonstrate practical applications of the developed statistics-based method in supply chain management area. The practical application of probability theory based methods and its validation by simulation is a subject for further research.

ACKNOWLEDGEMENTS

This work has been partly supported by the European Social Fund within the National Programme "Support for the carrying out doctoral study programmes and post-doctoral researches" project "Support for the development of doctoral studies at Riga Technical University".

REFERENCES

Banks J., Malave C.O. The simulation of inventory systems: An overview// Simulation Councils, Inc. – 1984. - P. 283-290.

Barlas, Y., Aksogan A. Product Diversification and Quick Response Order Strategies in Supply Chain Management// Proceedings of the International System Dynamics Conference. – Boston: 1996. – P. 51-54.

Chandra C., Chilov N., Smirnov A. Analysis of supply chain bullwhip effect using simulation modeling// CONSA Research Report Series "Simulation: Applications, Research and Education in the Baltic Area". - Linkoping: Presentations at the Special CONSA Conference, 2001. – 10 p.

Crosbie R. Modeling and simulation in supply chain management// Proceedings of Harbor Simulation Conference. Presentation – Portofino: 2000. – 9 p.

Dejonckheere J., Disney S.M., Lambrecht M.R. et al. Measuring and avoiding the bullwhip effect: A control theoretic approach// European Journal of Operational Research. – 2003. – No. 147. - P. 567–590.

Disney S.M., Towill, D.R. A robust and stable analytical solution to the production and inventory control problem via a z transform approach// Proceedings of the Twelfth International Working Conference on Production Economics. – Igls: Vol. 1, Austria, 2002. – P. 37-47.

Feller W. An introduction to probability theory and its applications. – New York: John Wiley&Sons, Inc., London: Chapman&Hall, Limited, 1967. – 498 p.

Forrester J.W. Industrial Dynamics. – New York: MIT Press and Wiley, 1961. - 479 p.

Hennet J.-C. Load and inventory fluctuations in supply chains// Proceedings of the 17th European Simulation Symposium and Exhibition "Simulation in industry". – Marseille: 2005. – P. 217-222.

Ingals R.G., Foote B.L., Krishnamoorthy A. Reducing the bullwhip effect in supply chains with control-based forecasting// International Journal of Simulation & Process Modelling. 2005. – No. 1/2. - P. 90-110.

Kelle P., Milne A. The effect of (s, S) ordering policy on the supply chain// International Journal of Production Economics. – 1999. – Volume 59. – P. 113-122.

Luong Huynh Trung, Phien Nguyen Huu. Measure of bullwhip effect in supply chains: The case of high order autoregressive demand process// European Journal of Operational Research. – 2007. – No. 183. – P. 197–209.

Petuhova J., Merkuryev Y. Combining analytical and simulation approaches to quantification of the bullwhip effect// International Journal of Simulation: Systems, Science & Technology (IJS^3T), United Kingdom Simulation Society, The Nottingham Trent University. – 2007. – Volume. 8, No. 1. – P. 16-24.

Riddalls C.E., Bennett S., Tipi N.S. Modelling the dynamic of supply chain// International Journal of System Science. – 2000. - Volume 31 (8). - P. 969-976.

Simchi-Levi D., Kaminsky P., Simchi-Levi E. Designing and Managing the Supply Chain. – USA: McGraw-Hill/Irwin, 2^{nd} ed., 2002. – 384 p.

Simon H.A. On the application of servomechanism theory in the study of production control// Econometrica. - 1952. – No.20. - P. 247-268.

Sterman J.D. Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment// Management Science. – 1989. – No. 35 (3). - P. 321-339.

Van Landeghem H., Vanmaele H. Robust planning: a new paradigm for demand chain planning// Journal of Operations Management. – 2002. – No. 20. - P. 69-783.

Zhang C, Tan G.W., Robb D.J. et al. Sharing shipping quantity information in the supply chain// Omega: International Journal of Management Science. – 2006. – No. 34 (5). – P. 427-438.

Zhang C., Zhang C. Design and simulation of demand information sharing in a supply chain// Simulation Modelling Practice and Theory.–2007.– Volume 15, No.1.– P. 32-46.

SUSTAINABLE CAR INDUSTRY BASED ON GREEN LOGISTICS SIMULATION MODELS

Agostino G. Bruzzone[(a)], Marina Massei[(a)], Federico Tarone[(b)], Francesco Longo[(c)]

[(a)]MISS DIPTEM University of Genoa
[(b)]Simulation Team
[(c)]MSC-LES University of Calabria

[(a)]{agostino, massei}@itim.unige.it
[(b)]federico.tarone@simulationteam.com
[(c)]f.longo@unical.it

ABSTRACT
Sustainability is major issue, that emerged during recent crisis as an important aspect in designing industrial processes (Taplin 2001); in this paper it is presented a model to analyze the supply chain in term of cost efficiency and environmental impact (Daly and Cobb 1989); this analysis is based on simulation; in fact the paper GreenLog, a simulation model developed by the authors for studying these problems; the experimental analysis provide an opportunity for validating and verifying the simulator

Keywords: Sustainability, Environmental Impact

1. INTRODUCTION

The definition of sustainability as result of the application of methods of harvesting or of using resource so that the resource is not depleted or permanently damaged (Denton 1998); in fact today a strong emphasis is attributed to sustainability in social context corresponding to human activities and lifestyle involving the use of sustainable methods even in references to social and human resources (Hardin 1968).

The current crisis highlighted that target functions need to include environmental, social and economic sustainability in addition to time payback period and profit estimations for next four or height trimesters.

These issues have been extensively approached in the past (Daly et al. 1989); however it results that the time horizon represent a critical issue, that need to be consistent with user perceptions (Velasco 2008).

However recent event highlight the necessity to move from qualitative analysis and economic models considering generic and simplified hypotheses to quantitative models based on simulation to reproduce realistic scenarios (Amico Vince et al. 2000).

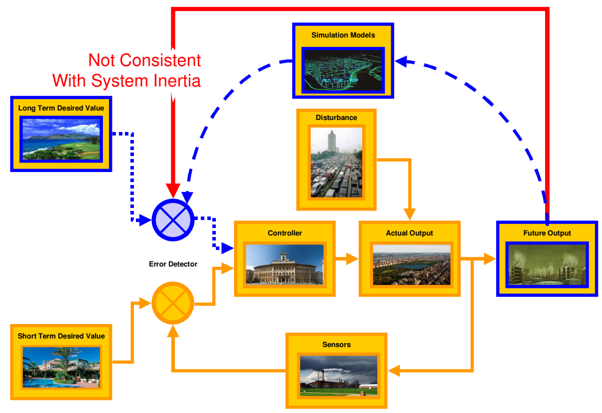


Figure 1 - Time Horizon and Sustainability

In fact a major problem, arising from the recent experience, corresponds to the short term horizon in decision making analysis; even large investments today are focusing on very short period pay back and usually long term considerations are depleted by contingency and short term targets.

In addition to analyze sustainability requires to consider additional factors that usually are regulated by pretty complex relations; so traditional quantitative techniques are often pretty inefficient in approaching these aspects and usually qualitative methods and/or expert feelings are the only approach taken in consideration.

Therefore the implications of our actions and their impact on the whole system, today are much stronger and complex than in the past, partially due to technology advances, partially due to population growth, without forgetting to take into consideration the "time horizon " issues (Miller 1998)

Due to these considerations, simulation (Spedding et al. 1999) represent an ideal approach for considering complex system, with many components interacting each other.

This paper propose a simulation model developed to investigate supply chain from point of view of environmental issues (Christensen 1999; Seamann 1992) in this case it is proposed a classical question, therefore the answer proposed is based on scientific models and quantitative hypotheses related to scenario configuration (Bruzzone et al. 2007). Obviously the result proposed are strongly related to the specific hypotheses and the specific configuration of the proposed case study (Melnyk and Handfield 1996), however the simulator allows to estimate the interactions and different factors; so while the results have just a local validity and the consideration can't be extended as general criteria due to the strongly not-linear nature of the problem, the model allows to test the sensitivity of the problem to different hypotheses and provides a methodological example in analyzing sustainability problems.

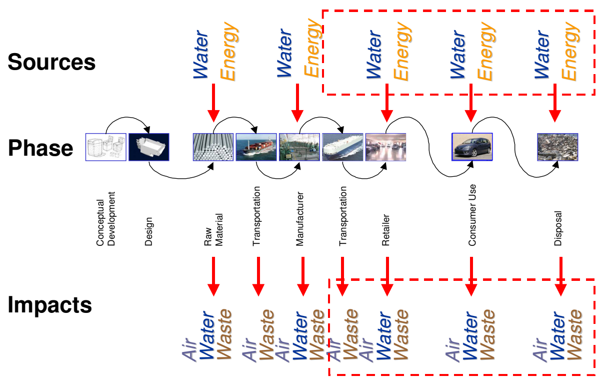


Figure 2 - Source and Impact of Car Supply Chain

2. SIMULATION FOR SUSTAINABILITY

It was already mentioned the critical issue in managing sustainability (Hibbert 1998) due to time horizon limitation; in fact it is interesting to analyze closed loop applied to this area as proposed in figure 1.

Therefore lets assume that the decision makers are driven by both short and long term goals (Ottman 1999); our real system react to the decisions taken by controllers in order to achieve the expected goals; usually short term target functions are emphasizing profit and social stability, while long term are paying more attention to quality of life and environmental issues (European Commission Dg Environment 2000).

From this point of view it is interesting to note that the long term target functions are often affected by significant inertia, so the time response results inconsistent respect the controller actions; due to this situation it becomes necessary to use simulation for estimating future state of the system by conducting experimental analysis on the models; this support the development of a predictive approach that is able to activate in time required actions for keeping the system under control.

In fact the long term impact of actual actions, due to the inertia of the real system, requires to adopt simulation as forecasting methodology in order to develop a predictive control.

Based on this approach it is critical to define performance indexes able to provide a guideline in guarantee sustainability in long term, without forgetting short term goals.

As always in analysis of complex systems it is critical to define target functions able to map completely the user expectations; usually in business activities it is important to include target functions covering at least: economic profits (Hick 2000), activities volumes, quality of the service provided; even in the case of complex system sustainability it is important to define KPI (key performance indexes) covering all the spectrum of short and long term (as well as medium) objectives (Stead and Stead 2000); this should include among the others:

- Profitability
- Value of Deliverables
- Quality of Deliverables
- Environmental Impact
- Social Stability

3. SUSTAINABILITY IN CAR INDUSTRY

Today an important aspect in manufacturing (Street 1986) and logistics is related to the sustainability and to the environmental impacts (Bruzzone 2004a). In fact traditional economic and financial aspects requires to be keep in consideration concurrently with social and environmental issues (Flachsbart 1999); therefore to proceed in development of interoperable models covering these aspects it is a big challenge, so the authors are proposing in this paper a specific case study as example to demonstrate these concepts and methodologies. It was decided to focus on car industry, currently very strongly affected by the financial crisis; however even in this sector an extensive analysis should result in complex scenarios affected by many uncertain variables (Quinn 2001) (i.e. market evolution, advertising effectiveness, factors affecting both value engineering and quality function deployment) (Bruzzone and Kerckhoffs 1996).

So the authors focused on a very specific case related to an on-going questions: "hybrid cars produced in far east and delivered in western countries are more sustainable than new generation traditional cars produced locally?" (Merkuryev et al. 2003; Sarkis 2006)

This context is still pretty large, and in order to be solved requires to define properly the scenario (type of car, production sites, recycling policies, mode of use, etc.).

Table I: Scenario Parameters				
	Value		Scenario	Phase
Road Car Carrier Cost	1.2	[Euro/km]	A	FE Factory-FE Port
Road Car Carrier Emission	0.00035	[tCO2/km]	A	FE Factory-FE Port
Road Car Carrier Capacity	9	[cars]	A	FE Factory-FE Port
Road Car Carrier Saturation	100%		A	FE Factory-FE Port
Impact on Road Car Carrier Transp.	100%		A	FE Factory-FE Port
Road Car Carrier Average Distance	12	[km]	A	FE Factory-FE Port
Road Car Carrier Empty Return	100%		A	FE Factory-FE Port
Ship Car Carrier Variable Cost	30,000	[Euro/day]	A	FE Port-SE Port
Ship Car Carrier Extra Cost (Suez)	200,000	[Euro/travel]	A	FE Port-SE Port
Ship Car Carrier Emission	0.1053	[tCO2/km]	A	FE Port-SE Port
Ship Car Carrier Capacity	6000	[cars]	A	FE Port-SE Port
Ship Car Carrier Saturation	9%		A	FE Port-SE Port
Impact on Ship Car Carrier Transp.	9%		A	FE Factory-FE Port
Ship Car Carrier Average Distance	19.21	[days]	A	FE Port-SE Port
Ship Car Carrier Average Distance	17,500.00	[km]	A	FE Port-SE Port
Ship Car Carrier Empty Return	0%		A	FE Port-SE Port
Road Car Carrier Cost	1.2	[Euro/km]	A	SE Port - SE Distributor
Road Car Carrier Emission	0.00035	[tCO2/km]	A	SE Port - SE Distributor
Road Car Carrier Capacity	9	cars	A	SE Port - SE Distributor
Road Car Carrier Saturation	100%		A	SE Port - SE Distributor
Impact on Road Car Carrier Transp.	100%		A	FE Factory-FE Port
Road Car Carrier Average Distance	200	[km]	A	SE Port - SE Distributor
Road Car Carrier Empty Return	100%		A	SE Port - SE Distributor
Hybrid Car Consumption	26.32	km/litre	A	Consumer Use
Hybrid Car Emission	104	[gCO2/km]	A	Consumer Use
Hybrid Car Acquistion Cost	28,000	[Euro/car]	A	Consumer Use
Hybrid Car Service Cost	750	[Euro/year]	A	Consumer Use
Hybrid Car Extra Disposal Cost (NiMH)	0.012	[Euro/car]	A	Dismissal
Hybrid Car Extra Emissions (NiMH)	480	[gCO2/car]	A	Dismissal
Road Car Carrier Cost	1.2	[Euro/km]	B	SE Factory-SE Distributor
Road Car Carrier Emission	0.00035	[tCO2/km]	B	SE Factory-SE Distributor
Road Car Carrier Capacity	9	[cars]	B	SE Factory-SE Distributor
Road Car Carrier Saturation	100%		B	SE Factory-SE Distributor
Impact on Road Car Carrier Transp.	100%		B	SE Factory-SE Distributor
Road Car Carrier Average Distance	800	[km]	B	SE Factory-SE Distributor
Conventional Car Consumption	16.39	km/litre	B	Consumer Use
Conventional Car Emission	155	[gCO2/km]	B	Consumer Use
Convential Car Acquistion Cost	14,000	[Euro/car]	A	Consumer Use
Convential Car Service Cost	500	[Euro/year]	A	Consumer Use
Notes: FE Far East / SE South Europe				

In addition the model reproducing these case is expected to be strongly not linear, so it is important to state that the obtained results are specific of the configuration of author's input and assumptions.

The case in fact focuses on comparing two different scenarios:

A] New Conventional Family Cars (unleaded gasoline) produced in South Europe and transferred within 1000 km to a regional distributor

B] New Hybrid Family Car (conventional combustion engine, electric engine and NiMH battery) produced in Far East and transferred in South Europe to a regional distributor

It is assumed that the cars are expected to be used for about 15'000 km/year and to have equivalent life cycle duration (i.e. a potential life cycle of 5 years); in table I it is proposed a synthesis of several parameters.

In figure 2 it is proposed a general scheme for analyzing the impact on environmental issues of car production (LockWood et al. 2007), in our case it was decided to consider equivalent the impact on environment and social aspect for manufacturing the car both in solution A and B; so the difference are computed just after the cars exit from production site parking; at the same time the consumer use profile is supposed to be equivalent on the both scenarios; service costs for the two cars have been hypothesized. The parameters adopted in these scenarios are fictional, even if they have been tunes on realistic values; any similarity to any car existing is merely coincidental. In figure 3 it is proposed the comparison among the different components of the A and B life cycle.

The sustainability it is measured respect CO_2 emission (Colvile et al. 2001) and total social cost; therefore the GreenLog simulator used for completing the analysis is even able to measure others factors such as different environmental impacts (i.e. tire disposal, oil waste), but in this case, due to the fact that it is supposed hybrid and conventional car have equivalent behavior it is not consider; the only difference evaluated it is related to service activities (additional electrical engine, controls and battery for case A) and extra disposal activities for batteries (in this case the hypotheses are focusing on NiMH batteries) (Krishnamurthi and Shanthi 2003).

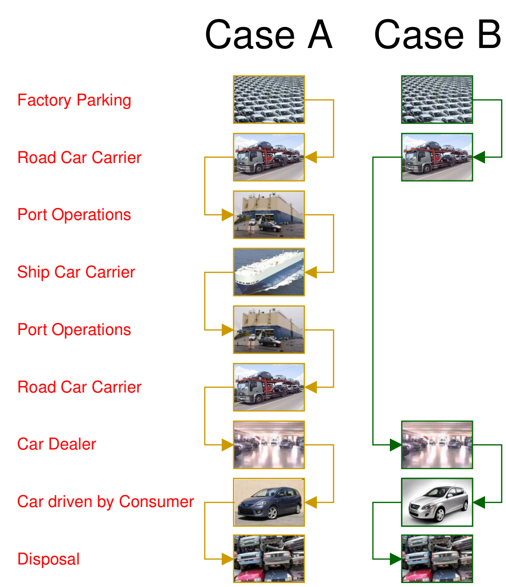


Figure 3 - Life Cycle of the two case studies

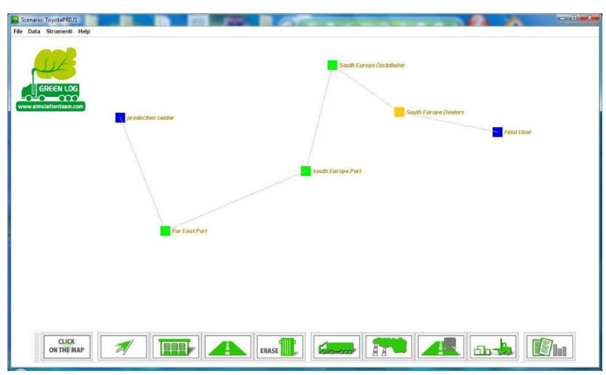


Figure 5 - Scenario A in Greenlog Simulator

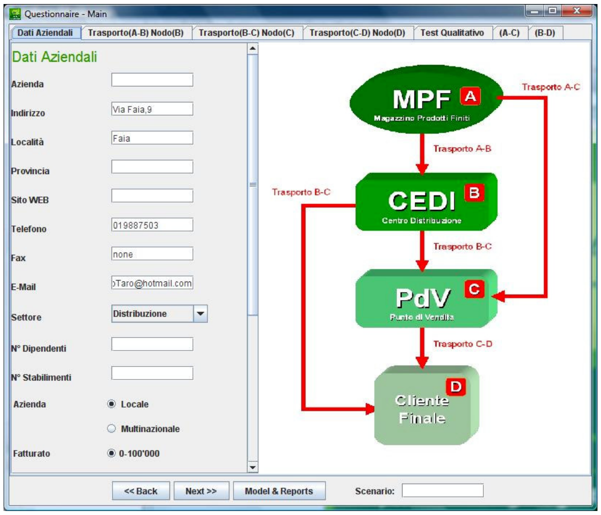


Figure 4 - Questionnaire for Company Self Evaluation

4. GREENLOG SIMULATOR

The authors developed an object oriented simulators for analyzing complex supply chain scenarios in term of overall efficiency and sustainability (Arnold 2006). This simulator it is named GreenLog (Green Logistics simulator) and it is based on web technologies for guaranteeing easy access and distributed use (Bruzzone et al. 1999) (Bruzzone et al. 2001) In fact GreenLog is part of a project for supporting environmental impact mitigation in Italian Logistics; this initiative involve Agencies (i.e. Assologistica Cultura e Formazione), Networks (i.e. Simulation Team), Institutions (i.e. MISS DIPTEM Genoa University) and Companies (i.e.Campari, CRAI, MARS, Sony, etc.); Green Logistics initiative (Gifford 1997) it is lead by MISS DIPTEM from technical point of view and it is based on a web portal (st.itim.unige.it/greenlogistics) where it is possible to access to several services:

- Company Qualitative and Quantitative Questionnaire (see figure 4)
- Self Measure of the Company Logistics Green Level based on automated configuration of the specific simulation model
- Supply Chain Simulation for measuring impacts and performances.

The model is based on the following main objects devoted to reproduce infrastructures, processes and performances; among the others it is possible to list the following objects:

- Logistics Node
- Logistics Link
- Vector
- Logistics Flows
- Environmental Impact

It could be interesting to dedicate just a special mention for the last object, the environmental impact (EI); EIs can be attributed to each of the other object and correspond to emissions, disposals, consumption etc. In fact each EI is related to a specific variable connected with the object where they are attributed, and by this relation it is possible to module the EI based on the dynamic evolution of the simulation (Bruzzone and Williams 2006) the relations are involving different variables including:

- Constant
- Distance
- Time
- Flow Volumes
- Flow Masses
- Costs
- Flow Types

Greenlog simulator is implemented in Java and the portal is currently active on Genoa University servers. Due to the user nature, the authors designed Greenlog in order to allow graphical construction of the conceptual model as proposed in figure 5 (Barros et al. 2005) The simulator provides easy GUI for all object creation and configuration (see figure 6) as well as automated reporting capabilities in figures (fig.7); another important feature it is the capability to present the results graphically issues directly over the logical network in order to identify critical component, for instance in figure 8, the Scenario A supply chain is related to the CO_2 environmental impact (Bruzzone 2002a). It is important to note that when environmental impacts are evaluated often, just the impacts related to the "on going" state are taken into consideration; sometime this is wrong approach because it is necessary to consider the whole life cycle to avoid mistakes (Cox 1999). So it's very important to consider all the boundary conditions and relations in order to evaluate the logistical impacts.

The nature of these phenomena and the uncertainty on major aspects are strongly affecting the reliability of the results: i.e. comparing environmental impact of different compound emissions and estimation of their social costs. Due to these reasons it is critical to proceed since the beginning in Verification, Validation process with extensive test; this paper in fact represents exactly one good example of this procedure and the authors used this case for teaching in master classes and in decision makers executive course; based on the above mentioned consideration it should be emphasized the accreditation process to guarantee that stakeholders and decision makers will get benefits in term of sustainability from valuable simulators through their direct involvement in the VV&A (verification, validation and accreditation) activities.

5. EXPERIMENTAL CASE STUDIES

In fact the proposed example related to cars represents a good application of the proposed approach (Bruzzone et al. 2006a).

If a car is manufactured in Far East and sold in South Europe, it results characterized by a different life cycle respect another one directly manufactured and sold in South Europe.

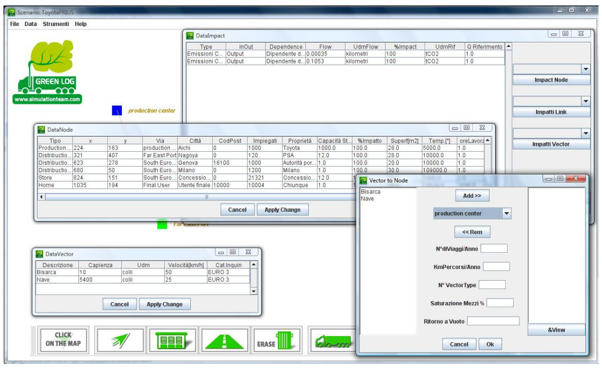


Figure 6 - Setting up the Object Parameters in GreenLog

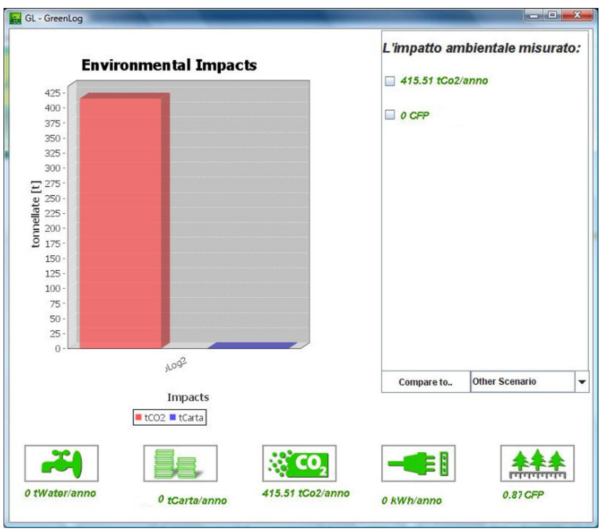


Figure 7 - Example of Greenlog Synthetic Output

Figure 8 - Greenlog highlight most critical element of the supply chain

It is interesting to evaluate the difference between the two life cycles considering to have different models as proposed in case A (Hybrid Car produced in Far East for South Europe) and B (Conventional Car locally produced and used).

This analysis is based on GreenLog Simulator adopting the setting above descripted proposed.

GreenLog simulator estimate the average emissions of CO2 for case A and B and based on their comparison respect car emissions it is possible to synthesize the results; for instance case B break even point in term of emission is obtained at 5'000 km, so the additional impact due to the long transportation by ship from Far East to South Europe is quickly balanced; viceversa based on current hypotheses on duty and insurance cost the B solution is more convenient in term of costs as proposed by figure 9.

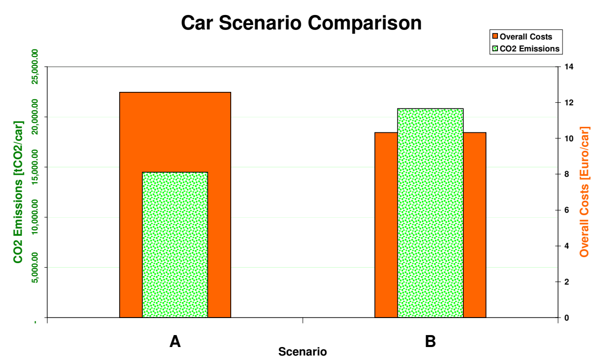


Figure 9 - Scenario A and B Comparison based on Greenlog Simulator Results

6. CONCLUSIONS

This paper present an application of Modeling and Simulation to sustainability issues (Bruzzone et al. 2004b) (Bruzzone et al. 2006b); the case used for the analysis represent a good example to highlight the complexity of this framework and the variety of variable and hypotheses required even to configure simple scenarios (Matthias 1999).

The approach for solving this problem it is based on the use of GreenLog simulator, developed by the authors to measure the performance of a supply chain for logistics and environmental point of view; to apply this model to the case study proposed resulted very easy and flexibile; in addition the authors used Greenlog, thanks to its web technology, as a support in classes and courses to involve student and decision makers on these subject with very good results (Bruzzone 2002b).

Currently the authors are working in developing new models to consider with major detail social aspects as well as long term effects; the development of scenarios on real case studies it is another important issue to diffuse knowledge on these methodologies for facing challenges related to sustainability and to complete validation, verification and accreditation on these models, working together with decision makers (Khoo et al. 2001; Shultz and Holbrook 1999

REFERENCES

Amico Vince, Guha R., Bruzzone A.G. (2000) "Critical Issues in Simulation", Proceedings of Summer Computer Simulation Conference, Vancouver, July

Arnold O. (2006) "Efficiency of Battery Channel", Technical Report French Ministry of Environment, Parisa

Barros F., Bruzzone A.G., Frydman C., Giambiasi N. (2005) "International Mediterranean Modelling Multiconfernece - Conceptual Modelling & Simulation Conference", LSIS Press, ISBN 2-9520712-2-5 (pp 232)

Bruzzone A.G. (2002) "Supply Chain Management", Simulation, Volume 78, No.5, May, 2002 pp 283-337 ISSN 0037-5497

Bruzzone A.G. (2002) "Web Integrated Logistics and Artificial Intelligence Application for Creating Smart Logistics Networks", Proceeding of SCI2002, Orlando, July

Bruzzone A.G. (2004) "Introduction to Modeling and Simulation Methodologies for Logistics and Manufacturing Optimization", Vol 80, No.3 pg 119-120 ISSN 0037-5497

Bruzzone A.G., Bocca E., Briano E., Poggis S. (2007) "Supply Chain Management and Vulnerability", Proceedings of Summer Computer Simulation Conference 2007, San Diego, July

Bruzzone A.G., Cunha G.G., Landau L., Saetta S. (2006) "Applied Modelling & Simulation", LAMCE Press Rio de Janerio, ISBN 85-285-0089-6 (240 pp)

Bruzzone A.G., E.Page, A.Uhrmacher (1999) "Web-based Modelling & Simulation", SCS International, San Francisco,ISBN 1-56555-156-7

Bruzzone A.G., Frydman C., Giambiasi N., Mosca R. (2004) "International Mediterranean Modelling Multiconfernece", DIPTEM Press, ISBN 88-900732-4-1 Vol I e II (884 pp)

Bruzzone A.G., Guasch A., Piera M.A., Rozenblit J. (2006) "International Mediterranean Modelling Multiconference", Logsim, ISBN 84-690-0726-2 (768 pp)

Bruzzone A.G., Kerckhoffs (1996) "Simulation in Industry ", Genoa, Italy, October, Vol. I & II, ISBN 1-56555-099-4

Bruzzone A.G., Mosca R., Revetria R., Coppa A. (2001) "Simulation Application in Distributed Logistics", Proceedings of 6th ISL2001, Salzsburg, July 8-10

Bruzzone A.G., Williams E. (2006) "Summer Computer Simulation Conference", SCS International, ISBN 1-56555-307-1 (504 pp)

Christensen J. (1999) "Seeing a Forest to Save the Trees" New York Times.

Colvile, R.N., Hutchinson, E.J., Mindell, J.S. and Warren, R.F. (2001) "The transport sector as a source of a air pollution", Atmospheric Environment, 35(9), 1537-65.

Cox, A.(1999) "Power, value and supply chain management". International Journal of Supply Chain Management, 4(4), 167-75.

Daly H.E. and Cobb J.B.Jr., (1989) "For the Common Good: Redirecting the Economy Toward Community, the Environment and a Sustainable Future", Beacon Press, Boston (1989).

Denton, T. (1998) "Sustainable development at the next level" Chemical Market reporter, 253(7) 3-4

EDCG (2000) "A study on the economic valuation of environmental externalities from landfill disposal and incineration of waste", Technical Final Report European Commission DG Environment , Bruxelles, October

Flachsbart, P.G. (1999) "Human exposure to carbon monoxide from mobile sources" Chemoshere Global Change Science, 1(1-3), 301-29

Gifford, D. (1997 "The value of going Green" Harvard Business Review, 75, 11-2

Hardin, G. (1968) "Tragedy of the Commons" Science, 162,(1293), 48

Hibbert, L.(1998) "Sustainable activity" Professional Engineering, 11, 32 -3

Hick, S (2000) "Morals maketh the money" Australian CPA, 70, 72-3

Khoo, H.H., Spedding, T.A., Tobin, L., Taplin, D.(2001) "Integrated simulation and modelling approach to Decision Making and Environmental Protection" Environment, Development and sustainability, 3(2), 93-108

Krishnamurthi, Shanthi. (2003) "Batteries in Mobile Applications Which Way Would The Penny Fall?", Technical Report Frost & Sullivan

LockWood C., Reihardt F., Sells B., Stern A. J. (2007) "Harvard Business Review on Green Business Strategy", Harvard Business School Press, Cambridge MA

Matthias, R. (1999) "Strategies for promoting a sustainable industrial ecology, Environmental Science & Thecnology, 33(13), 280-82

Melnyk, S. and Handfield, R. (1996) "Greenspeack" Purchasing Today, July, 32-6

Merkuryev Y., Bruzzone A.G., Merkuryeva G., Novitsky L., Williams E. (2003) "Harbour Maritime and Multimodal Logistics Modelling & Simulation 2003", DIPTEM Press, Riga, ISBN 9984-32-547-4 (400pp)

Miller, W.H (1998) "Citizenship: a competitive asset" Industry Week, 247(15), 104-8

Ottman, J.A. (1999) "How to develop really new, new products" Marketing News 33(3), 5-7

Quinn, B. (2001) "Manufacturers squeeze out more energy efficiency" Pollution Engineering, 33(1), 23 - 4

Sarkis J.(2006) "Greening the supply Chain", Springer Verlag, London

Seamann, R.(1992) "The environment and the need for new technology : empowerment and ethical values" Columbia Journal of World Business, 27, 186-93

Shultz II, C.J. and Holbrook, M.B. (1999) "Marketing and tragedy of the commons: a synthesis, commendary, and analysis for action" Journal of Public Policy and Marketing, 18(2), 218-29

Spedding, T.A., H.H., Taplin, D. (1999) "An integrated simulation approach for teaching industrial ecology" Proceedings of the HKK Conference" Waterloo: Canada

Stead, J.G. and Stead, E. (2000), Eco-enterprise strategy: standing for sustainability" Journal of Business Ethics, 24(4), 313-29

Street, A.C. (1986) "The Diecasting Book" 2nd Ediction, Portcullis Press

Taplin, D.M.R., Spedding, T.A., Khoo, H.H. (2001) "Environmental security: simulation and moddeling of Industrial Process Ecology" Paper presented at the Strasbourg Forum, Council of Europe, Strasbourg: France

Velasco H. (2008) "Sustainability: The matter of time horizon and semantic closure", Ecological Economics Volume 65, Issue 1, 15 March 2008, Pages 167-176

AUTHORS BIOGRAPHY

Agostino G. Bruzzone since 1991, he has taught "Theories and Techniques of Automatic Control" and in 1992 he has become a member of the industrial simulation work group at the ITIM University of Genoa; currently he is Full Professor in DIPTEM. He has utilized extensively simulation techniques in harbor terminals, maritime trading and sailboat racing sectors. He has been actively involved in the scientific community from several years and served as Director of the McLeod Institute of Simulation Science (MISS), Associate Vice-President and Member of the Board of the SCS (Society for Modelling & Simulation international), President of the Liophant Simulation, VicePresident of MIMOS (Movimento Italiano di Simulazione) and Italian Point of Contact for the ISAG (International Simulation Advisory Grouyp) and Sim-Serv. e has written more than 150 scientific papers in addition to technical and professional reports in partnerships with major companies (i.e. IBM, Fiat Group, Contship, Solvay) and agencies (i.e. Italian Navy, NASA, National Center for Simulation, US Army). He teaches "Project Management" and "Industrial Logistics" at the University for students in the Mechanical Engineering (4th year), Management Engineering (4th year) and Logistics & Production Engineering (3rd year) degree courses.

Marina Massei is currently enrolled in DIPTEM University of Genoa as member of the Simulation Team of Prof.Agostino Bruzzone. She is involved in the organization of International Scientific Events (i.e. Summer Computer Simulation Conference 2003-Montreal, 2004-San Jose', 2005-Philadelphia) She served as Chair for Applications in Management, Planning & Forecasting Workshop in SummerSim 2004, 2005, 2006. She is Associate Director of the MISS Center at the University of Perugia She was Finance Control & Administration Director of MAST.

Federico Tarone is Phd student at DIPTEM, University of Genoa. He is currently involved in some Simulation projects in the field of Logistics, Inventory Management and Web Applications. He is member of the Liophant Simulation. In 2008 he attended the I3M International Conference as speaker.

Francesco Longo was born in Crotone on February the 8th 1979. He took his degree in Mechanical Engineering, summa cum Laude, in October 2002 from the University of Calabria. He received his Ph.D. in Mechanical Engineering from the Department of Mechanical Engineering at University of Calabria in January 2006. He is currently researcher in the same Department. His research interests include Modeling & Simulation for production systems design and supply chain management. He was General Chairman of the 20th European Modeling & Simulation Symposium He was Program Co-Chair of the 4th International Modelling Mediterranean Multiconference and Program Co-Chair of the 19th European Modeling & Simulation Symposium. He belongs to the International Program Committee (IPC) of the Summer Computer Simulation Conference and Track Chair of the Emergency Simulation session in the same conference. He belongs to the IPC of the European Conference on Modeling & Simulation. He works as reviewer for the most important conferences on Modeling & Simulation such as the Summer Computer Simulation Conference, the European Conference on Modeling & Simulation, the Winter Simulation Conference, the Workshop on Applied Simulation, etc. He works also as reviewer for the international journal Computer & Industrial Engineering. He is Responsible of the Modeling & Simulation Center – Laboratory of Enterprise Solutions (MSC-LES). He teaches "Design of Production Systems" and "Laboratory of Industrial Plants" respectively for students of the 2nd and 5th year in Management Engineering. He is also tutor on the courses "Industrial Plants" and "Production Systems Management" respectively for students of the 3rd and 5th year in Mechanical Engineering.