THE 7TH INTERNATIONAL DEFENSE AND HOMELAND SECURITY SIMULATION WORKSHOP

SEPTEMBER 18 – 20, 2017
BARCELONA, SPAIN

EDITED BY

AGOSTINO BRUZZONE
ERDAL CAYIRCI
ROBERT SOTTILARE

PRINTED IN RENDE (CS), ITALY, SEPTEMBER 2017

ISBN 978-88-97999-98-0 (Paperback)
ISBN 978-88-97999-90-4 (PDF)

© 2017 DIME UNIVERSITÀ DI GENOVA

RESPONSIBILITY FOR THE ACCURACY OF ALL STATEMENTS IN EACH PAPER RESTS SOLELY WITH THE AUTHOR(S). STATEMENTS ARE NOT NECESSARILY REPRESENTATIVE OF NOR ENDORSED BY THE DIME, UNIVERSITY OF GENOA. PERMISSION IS GRANTED TO PHOTOCOPY PORTIONS OF THE PUBLICATION FOR PERSONAL USE AND FOR THE USE OF STUDENTS PROVIDING CREDIT IS GIVEN TO THE CONFERENCES AND PUBLICATION. PERMISSION DOES NOT EXTEND TO OTHER TYPES OF REPRODUCTION NOR TO COPYING FOR INCORPORATION INTO COMMERCIAL ADVERTISING NOR FOR ANY OTHER PROFIT – MAKING PURPOSE. OTHER PUBLICATIONS ARE ENCOURAGED TO INCLUDE 300 TO 500 WORD ABSTRACTS OR EXCERPTS FROM ANY PAPER CONTAINED IN THIS BOOK, PROVIDED CREDITS ARE GIVEN TO THE AUTHOR(S) AND THE WORKSHOP.

FOR PERMISSION TO PUBLISH A COMPLETE PAPER WRITE TO: DIME UNIVERSITY OF GENOA, PROF AGOSTINO BRUZZONE, VIA OPERA PIA 15, 16145 GENOVA, ITALY. ADDITIONAL COPIES OF THE PROCEEDINGS OF THE *DHSS* ARE AVAILABLE FROM DIME UNIVERSITY OF GENOA, PROF. AGOSTINO BRUZZONE, VIA OPERA PIA 15, 16145 GENOVA, ITALY.

ISBN 978-88-97999-98-0 (Paperback)
ISBN 978-88-97999-90-4 (PDF)

THE 7TH INTERNATIONAL DEFENSE AND HOMELAND SECURITY SIMULATION WORKSHOP, DHSS 2016

SEPTEMBER 18 – 20, 2017

BARCELONA, SPAIN

ORGANIZED BY

DIME – UNIVERSITY OF GENOA

LIOPHANT SIMULATION

SIMULATION TEAM

IMCS – INTERNATIONAL MEDITERRANEAN & LATIN AMERICAN COUNCIL OF SIMULATION

DIMEG, UNIVERSITY OF CALABRIA

MSC-LES, MODELING & SIMULATION CENTER, LABORATORY OF ENTERPRISE SOLUTIONS

AUTONOMOUS UNIVERSITY OF BARCELONA

MODELING AND SIMULATION CENTER OF EXCELLENCE (MSCOE)

LATVIAN SIMULATION CENTER – RIGA TECHNICAL UNIVERSITY

LOGISIM

LSIS – LABORATOIRE DES SCIENCES DE L'INFORMATION ET DES SYSTEMES

MIMOS – MOVIMENTO ITALIANO MODELLAZIONE E SIMULAZIONE

MITIM PERUGIA CENTER – UNIVERSITY OF PERUGIA

BRASILIAN SIMULATION CENTER, LAMCE-COPPE-UFRJ

MITIM - MCLEOD INSTITUTE OF TECHNOLOGY AND INTEROPERABLE MODELING AND SIMULATION – GENOA CENTER

M&SNET - MCLEOD MODELING AND SIMULATION NETWORK

LATVIAN SIMULATION SOCIETY

ECOLE SUPERIEURE D'INGENIERIE EN SCIENCES APPLIQUEES

FACULTAD DE CIENCIAS EXACTAS. INGEGNERIA Y AGRIMENSURA

UNIVERSITY OF LA LAGUNA

CIFASIS: CONICET-UNR-UPCAM

INSTICC - INSTITUTE FOR SYSTEMS AND TECHNOLOGIES OF INFORMATION, CONTROL AND COMMUNICATION

NATIONAL RUSSIAN SIMULATION SOCIETY

CEA - IFAC

UNIVERSITY OF BORDEAUX

UNIVERSITY OF CYPRUS

DUTCH BENELUX SIMULATION SOCIETY

I3M 2017 INDUSTRIAL SPONSORS

CAL-TEK SRL

LIOTECH LTD

MAST SRL

SIM-4-FUTURE

I3M 2017 Media Partners

Inderscience Publishers – International Journal of Simulation and Process Modeling

Inderscience Publishers – International Journal of Oil, Gas and Coal Technology

Inderscience Publishers – International Journal of Service and Computing Oriented Manufacturing

IGI Global – International Journal of Privacy and Health Information Management

Halldale Media Group: Military Simulation and Training Magazine

Halldale Media Group: The Journal for Healthcare Education, Simulation and Training

SAGE
Simulation Transaction of SCS

De Gruyter
International Journal of Food Engineering

EDITORS

AGOSTINO BRUZZONE
MITIM-DIME, UNIVERSITY OF GENOA, ITALY
agostino@itim.unige.it

ERDAL CAYIRCI
UNIVERSITY OF STAVANGER, NORWAY
erdal.cayirci@uis.no

ROBERT SOTTILARE
US ARMY RESEARCH LABORATORY, USA
Robert.Sottilare@us.army.mil

THE INTERNATIONAL MULTIDISCIPLINARY MODELING AND SIMULATION MULTICONFERENCE, I3M 2017

GENERAL CO-CHAIRS

AGOSTINO BRUZZONE, *MITIM DIME, UNIVERSITY OF GENOA, ITALY*
MIQUEL ANGEL PIERA, *AUTONOMOUS UNIVERSITY OF BARCELONA, SPAIN*

PROGRAM CO-CHAIRS

FRANCESCO LONGO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
YURY MERKURYEV, *RIGA TECHNICAL UNIVERSITY, LATVIA*

THE 7TH INTERNATIONAL DEFENSE AND HOMELAND SECURITY SIMULATION WORKSHOP, DHSS 2017

GENERAL CHAIR

AGOSTINO BRUZZONE, *MITIM-DIME, UNIVERSITY OF GENOA, ITALY*

PROGRAM CHAIR

ROBERT SOTTILARE, ***US*** *ARMY RESEARCH LABORATORY, USA*

DHSS 2017 International Program Committee

Carole Adam, LIG – UJF, France
Joni Amorim, University of Skövde, Sweden
Matteo Agresta, *University of Genoa, Italy*
Claudia Baisini, *Swedish National Defense College, Sweden*
Marco Biagini, NATO M&S COE, Italy
Joy Bruce, Vastpark, USA
Agostino Bruzzone, *University of Genoa, Italy*
Riccardo di Matteo, *University of Genoa, Italy*
Jean Guy Fontaine, *NATO STO CMRE*
Joaquin Garcia-Alfaro, *Telecom SudParis, France*
Peter Kemper, College of William and Mary, USA
Francesco Longo, *University of Calabria, Italy*
Edgar Alonso Lopez-Rojas, *Blekinge Institute of Technology, Sweden*
Gianluca Maglione, *University of Genoa, Italy*
Marina Massei, *University of Genoa, Italy*
Letizia Nicoletti, *Cal-Tek Srl, Italy*
Gabriele Oliva, *University Campus BioMedico of Rome, Italy*
Antonio Padovano, University of Calabria, Italy
Federica Pascucci, *University of Roma 3, Italy*
Alexandre Petrenko, Centre de Recherche Informatique de Montreal, Canada
Paolo Proietti, Mimos, Italy
Elaine Raybourn, *Sandia National Laboratories, USA*
Robert Sottilare, *Army Research Laboratory, USA*
Richard Zaluski, *CSCSS, UK*

Tracks and Workshop Chairs

Ports and Littoral Protection
Chair: Francesco Longo, University of Calabria, Italy

Serious Games: Toward Effective Training
Chairs: Paolo Proietti, MIMOS, Italy; Elaine Raybourn, Sandia National Labs/ADL, Usa

Urban Area Security
Chair: Agostino Bruzzone, Simulation Team, University of Genoa (Italy)

CyberSecurity
Chair: Richard Zaluski, CSCSS, USA

Adaptive and Predictive Computer-Based Tutoring
Chair: Robert Sottilare, Army Research Laboratory, USA

Next generation C4ISR Modeling & Simulation technology: Mixed Reality, Crowd sourcing and Interoperable C2-LVC architectures to support operations, training, and analysis
Chairs: *Marco Biagini, University of Genoa, Italy; Joy Bruce, Vastpark*

Simulations as Tools to Develop Intuitive Contextual Sensitivity, Re-Framing, and Mental Agility
Chair: Claudia Baisini, Swedish National Defence College, Sweden; Enrico Bocca, Simulation Team, Italy

Homeland Security: a Control System Point of View
Chairs: Gabriele Oliva, University Campus BioMedico of Rome, Italy; Federica Pascucci, University Roma Tre, Italy

Simulation for Autonomous & Intelligent Systems: Moving from Coordination to Collaboration and Competition
Chairs: Agostino G. Bruzzone, University of Genoa, Italy

Computer Simulation for Security
Chairs: Edgar Alonso Lopez-Rojas, Blekinge Institute of Technology, Sweden; Joni A. Amorim, University of Skövde, Sweden; Richard Zaluski, Center for Strategic Cyberspace + Security Science, Uk

Critical Systems Modeling and Simulation
Chairs: Joaquin Garcia-Alfaro, Telecom SudParis, France; Francesco Longo, University of Calabria, Italy

Co-Chairs' Message

Welcome To DHSS 2017!

As General and Program Chairs, we are pleased to welcome you to the International Defense and Homeland Security Simulation (DHSS 2017) Workshop that this year is held in the beautiful Barcelona, Spain!

The DHSS Workshop renovates once again the opportunity for simulation experts to meet and discuss the latest advances, the current state of the art and future potential of using M&S in the Defense and Homeland Security framework. DHSS includes contributions that document, in a rigorous and comprehensive manner, technical lessons and practical experimental results related to defense and homeland security systems modeling and simulation technology, methodology and theory. This year, the contributions range from surveys and state of the art reviews in the military operations domain to theoretical lessons and new knowledge created for the next generation of training systems, from advanced applications on real case studies to the design and development of new methodologies.

What emerges from this valuable researches is that Modeling & Simulation is a very powerful tool which can successfully support decision-making and training activities (leveraging on interoperability and serious games). Additionally, the focus for defense and security applications is increasingly moving away from traditional warfare towards the challenges of human, social, cultural, and behavioral factors.

In summary, all contributions focus on a particular facet that helps to understand the conceptual, technical, and organizational challenges we are currently facing.

We therefore hope you will enjoy your stay in Barcelona and find several ideas, old and new friends, collaborators, colleagues and sparks for new research topics that will advance the current state of the art in the field of Defense and Homeland Security Simulation.

Agostino G. Bruzzone
MITIM-DIME
University of Genoa, Italy

Robert A. Sottilare
U.S. Army Research
Laboratory, USA

ISBN 978-88-97999-90-4; Bruzzone, Cayirci and Sottilare Eds.

ACKNOWLEDGEMENTS

The DHSS 2017 International Program Committee (IPC) has selected the papers for the Conference among many submissions; therefore, based on this effort, a very successful event is expected. The DHSS 2017 IPC would like to thank all the authors as well as the reviewers for their invaluable work.
A special thank goes to Prof. Miquel Angel Piera from Autonomous University of Barcelona, as Local Organizer and to all the organizations, institutions and societies that have supported and technically sponsored the event.

I3M 2017 INTERNAL STAFF

MATTEO AGRESTA, *SIMULATION TEAM, ITALY*
MARGARITA BAGAMANOVA, *AUTONOMOUS UNIVERSITY OF BARCELONA, SPAIN*
LUIGI BRUNO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
AGOSTINO G. BRUZZONE, *DIME, UNIVERSITY OF GENOA, ITALY*
ALESSANDRO CHIURCO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
RICCARDO DI MATTEO, *SIMULATION TEAM, ITALY*
CATERINA FUSTO, *CAL-TEK SRL, ITALY*
THIMJO KOÇA, *AUTONOMOUS UNIVERSITY OF BARCELONA, SPAIN*
FRANCESCO LONGO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
GIANLUCA MAGLIONE, *SIMULATION TEAM, ITALY*
MARINA MASSEI, *DIME, UNIVERSITY OF GENOA, ITALY*
ROMUALDO MORENO, *AUTONOMOUS UNIVERSITY OF BARCELONA, SPAIN*
MERCEDES NARCISO, *AUTONOMOUS UNIVERSITY OF BARCELONA, SPAIN*
LETIZIA NICOLETTI, *CAL-TEK SRL, ITALY*
ANTONIO PADOVANO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
MIQUEL ANGEL PIERA, AUTONOMOUS UNIVERSITY OF BARCELONA, SPAIN
MARKO RADANOVIC, *AUTONOMOUS UNIVERSITY OF BARCELONA, SPAIN*
JUAN JOSE RAMOS, *AUTONOMOUS UNIVERSITY OF BARCELONA, SPAIN*
CATALDO RUSSO, *CAL-TEK SRL, ITALY*
NINA SCHEFERS, *AUTONOMOUS UNIVERSITY OF BARCELONA, SPAIN*
MARCO VETRANO, *CAL-TEK SRL, ITALY*

This International Workshop is part of the I3M Multiconference: the Congress leading **Simulation around the World and Along the Years**

ISBN 978-88-97999-90-4; Bruzzone, Cayirci and Sottilare Eds.

INDEX

JOINT MILITARY SPACE OPERATIONS SIMULATION: A SURVEY

Erdal Cayirci[(a)], Hakan Karapinar[(b)], Lutfu Ozcakir[(c)], Erdem Yazgan[(d)]

[(a)] University of Stavanger
[(b)]HAVELSAN
[(c)]Hacettepe University
[(d)]TED University

[(a)]erdal.cayirci@uis.no
[(b)]hakank@havelsan.com.tr
[(c)]lutfu.ozcakir@gmail.com
[(d)] erdem.yazgan@tedu.edu.tr

ABSTRACT
Space capabilities must be integrated into joint operations, which requires a common and clear understanding on space capabilities, their effects and the procedures for their employment by the joint force commanders and staff. Therefore, they need to be practiced during joint computer assisted exercises. In this paper, the military space mission areas and simulation tools that can be used for them are surveyed. The conclusion is that the joint military space operations simulation capabilities are limited and do not suffice to fulfil the requirement.

Keywords: modelling, simulation, space, joint military space operations

1. INTRODUCTION

Space capabilities are complex systems made up of various components including satellite production, checkout and storage facilities, launch facilities, user terminals, ground stations, manned or unmanned spacecrafts, payloads (e.g., sensors) and communication links (Rainey and Davis 2004). Many space capabilities require multiples of these components, such as tens of satellites and ground stations. The components of a space capability are typically procured from different commercial or governmental organizations. Recently, the vendors for space capabilities, including the space lift, are more and more frequently commercial companies.

Space technologies, capabilities and their components used to be controlled only by few nations. They are now available and affordable not only to state but also non-state entities. Moreover, it is not necessary to own all its components for having the space capability. It is possible to access the services by the space capabilities that the others own. Therefore, space is not a safe and secure place for the sophisticated intelligence, surveillance, reconnaissance (ISR), communications and navigation technologies for few nations anymore, but a challenging and integrated part of the joint military operations especially when defending against hybrid threats (Cayirci, Bruzzone, Longo and Gunneriusson 2016). Moreover, many terrestrial systems critical to military operations, such as navigation and communications, depend on space systems, although it is sometimes not easy to recognize this dependence. Therefore, military must have processes:

- To determine the space capability requirements for reaching the strategic and operational goals
- To contribute the development of new space technologies applicable for military
- To use the available space capabilities optimally
- To defend the components of the space capabilities
- To prevent the belligerents/adversaries from using space capabilities effectively

For these processes, and to realize the global advantages provided by space forces, all space capabilities and the means to protect them should be integrated into all kinds of military planning including defense, advance and response planning. Hence, the commanders and their staff *must understand the applications of space capabilities, have access to space-based support sufficient to accomplish their missions, use space systems to the degree needed for completing required tasks expeditiously, and make recommendations to deny or limit an adversary's access to space and use of space systems* (NATO 2009). Therefore, it is necessary to provide computer assisted exercises and wargames especially in the operational and higher levels with joint military space operations simulation (JMSOS) support. However, we can say that JMSOS support especially in the international collective training events do not suffice to fulfill the requirement. The main contributions of this paper are analyzing these shortcomings related to JMSOS support to operational level computer assisted exercises/wargames and making recommendations for eliminating them.

In Section 2, we elaborate on the need for military space operations simulation. We pay a special attention on collective training and exercises in operational level. In Section 3, we report the results from our literature survey on the available tools that address these requirements. Section 4 is about the potential future M&S approaches, such as modelling and simulation as a service (Cayirci 2013; NATO MSG-136 2017; Taylor et al. 2015; Zehe et al., 2015) applicable for JMSOS. We conclude our paper in Section 5.

2. MILITARY SPACE OPERATIONS SIMULATION REQUIREMENTS

We categorize the requirements for JMSOS into five classes of joint military processes in strategic and operational level as follows:

- JMSOS Support to Defense Planning Process: As depicted in Figure 1, defense planning process determines the military capability requirements for the potential scenarios given by the political level. Where and how the simulation support is needed within this process are explained in detail in (Cayirci and Ozcakir 2016). To the best of our knowledge, the capability for JMSOS support as explained in (Cayirci and Ozcakir 2016) is not available to any armed forces or alliance.

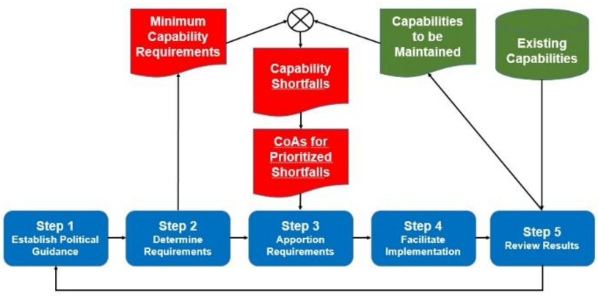


Figure 1: NATO Defense Planning Process.

- JMSOS Support to Capability Design and Acquisition Processes: The capability design and acquisition processes of the US DoD are illustrated in Figure 2 (DoD 2015). JMSOS support is needed for exploring the alternatives, analyzing the measurements, assessing the designs, evaluating the operating procedures, and predicting the performance during capability design and acquisition (Rainey and Davis 2004). Modelling and simulation (M&S) can be useful for the design of an overall space capability or solving complex engineering problems related to various topics, such as spacecraft, payload and orbit design. There are various tools commercially available for these purposes, and we can say that this is a mature field especially in civilian domain (ESA 2017; NASA 2010; Rainey and Davis 2004).
- JMSOS Support to Joint Operations Planning Process: Joint operations planning can be made advanced as standing defense plans and contingency plans or as a response to a crisis. Planning space capabilities must be an integrated part of the joint operations planning especially in operational level (SHAPE 2013). M&S tools are available for joint operations and logistics planning. However, JMSOS capabilities of these M&S tools are very limited.

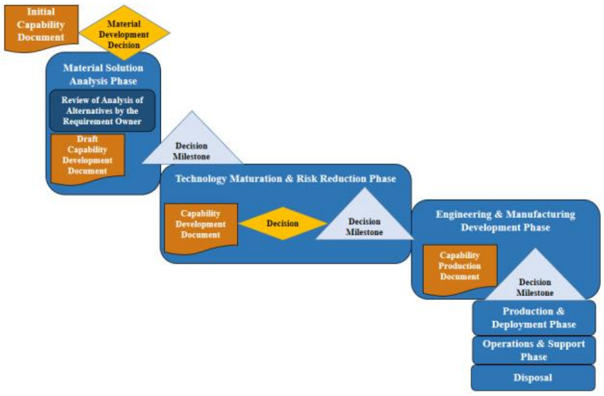


Figure 2: Capability Design and Acquisition Process.

- JMSOS Support to Joint Operations: Plans need to be adapted according to the evolving situation in current operations, which requires situational awareness and forecasts. Orbits and payloads of satellites, their sensing capabilities, their effects on operations need to be known and predicted. Similarly, space weather forecasts are important not only for the space assets but also many other military capabilities. Models and simulators are available for these purposes.
- JMSOS Support to Education and Training Processes: Models and simulators are available also for education and individual training. These include virtual simulators for spacecraft crew. However, constructive simulations in JMSOS domain do not suffice to fulfil the requirements for the collective training, exercises and wargames especially in joint operational level, which is the area that this paper focuses on.

For the collective training, exercises and wargaming purposes, JMSOS support should be provided in four space mission areas given in Table 1 (NATO 2009) and explained below:

- Space control operations (SCO) are conducted to attain and maintain the space superiority which involves the counter measures against the adversaries' space capabilities. These measures include actions by air, land, maritime, space and

special forces. SCO requires space situational awareness about space related conditions including space weather, constraints, capabilities and activities in, from, toward and through space. The details like orbits, payloads, frequencies are all of interest. SCO can be offensive or defensive. Offensive operations can be against not only the assets in space but also ground facilities and stations. Electronic control measures, such as, jamming an uplink or downlink are among the offensive SCO. Available combat models can be adapted to simulate a large subset of offensive and defensive SCO.
- Space force enhancement operations (SFEO) are to support the warfighter and to enhance the battlespace awareness. *There are five force enhancement functions: ISR; integrated tactical warning and attack assessment; environmental monitoring; communications; and position, velocity, time, and navigation* (NATO 2009). The combat models available for operational level exercises in NATO provide functions to simulate the results of a subset of SFEO. However, they are far from being sufficient.
- Space Support Operations (SSO) include space lift, satellite operations, reconstitution of space forces. Space lift delivers satellites, payloads and material to space. Satellite operations are conducted to maneuver, to configure and to sustain on-orbit forces and to activate on-orbit spares. Finally, reconstitution operations are for replenishing space forces when the existing forces degrade due to various reasons. SSO is seldom practiced in operational level exercises.
- Space Force Application Operations (SFAO) carried out by the weapon systems operating in or through space against terrestrial based targets. SFAO includes ballistic missile defense (BMD), theater ballistic missile defense (TBMD) and force projection. Please note that TBMD and BMD can be conducted also by means other than SFAO. This mission area is not practiced very often because there is not any known asset available in space for this purpose.

The following characteristics special to military space operations (NATO 2009) (Rainey Davis 2004) have to be taken into consideration in the design of the simulation services for these space mission areas:

- Global access and persistence: Satellites can fly over any location on Earth, and stay on orbit for extended period of time. However, except for geostationary satellites, they stay over a location on earth only a limited time.
- Coverage and propagation delay in communications: As the orbit altitude gets higher, the coverage area gets larger, nevertheless, the propagation delay in communications also gets longer. The return trip time for an electromagnetic signal from earth to a geostationary satellite is around 500 msec.
- Design life: Most satellites cannot be maintained or repaired. They also can have limited fuel on board to maintain the orbit or making changes in the orbit. Therefore, the lifetime of satellites is limited.
- Older technology: Although software defined technologies are changing this fact, typically the technology in a satellite is not the latest but the technology available before the launching day.
- Increasing affordability: New technologies introduce smaller and smaller satellites, such as, micro, nano, pico and femto satellites. More sophisticated satellites can be produced in less sophisticated production facilities and lifted into space more easily and less costly.
- Predictable orbits: Satellite orbits are predictable.
- Vulnerability: Ground to satellite links are susceptible to electronic counter measures and ground facilities and stations can be attacked.
- Resource considerations: Replacing or replenishing space forces need long lead times.
- Legal considerations: Numerous national and international laws must be considered during planning.
- Space treaties: Although currently there is no treaty that forbids the deployment of weapons other than weapons of mass destruction in space, many of them introduce constraints to the military use of space.

Table 1: Military Space Mission Areas.

Abbreviation	Name	Mission Types
SCO	Space Control Operations	-Space situational awareness -Offensive SCO -Defensive SCO
SFEO	Space Force Enhancement Operations	-Intelligence Surveillance Reconnaissance (ISR) -Tactical warning and attack assessment; -Environmental monitoring -Communications - Navigation
SSO	Space Support Operations	-Space lift -Satellite operations -Reconstitution of space forces
SFAO	Space Force Application Operations	-BMD/TBMD -Force projection

3. SIMULATION SYSTEMS FOR JOINT MILITARY SPACE OPERATIONS

In this section, we elaborate on the JMSOS capabilities of the constructive simulation systems available for joint operational and higher-level computer assisted exercises (CAX) and wargames. Although, our list is not exhaustive, it covers all the tools used in major coalition or alliance CAX known to us. Please note that, not only models and simulators but also the following tools are required for the successful employment of M&S including JMSOS during a CAX (Cayirci, 2006):

- Setting and scenario management tools
- Training objective management tools
- Main event list, main incident list management tools
- Command and control system data population and stimulation tools

There have been two main approaches followed for integrating JMSOS into joint operational CAX: federating a JMSOS system into a federation of simulators or incorporating JMSOS functionalities into a joint or service simulation system. Well known examples of both approaches are listed in Table 2 and elaborated on below:

Table 2: Examples for the Constructive Simulation Systems with the Models Related to the Space Mission Areas.

Tool	Type	Federation	Mission Areas	Status
PSM	Space Model	JTC and JLVC	SCO (partial), SFEO (partial), SFAO (partial)	Availability is unknown.
AWSIM	Service (Air) Model	JTC and JLVC	SCO (partial)	Available
JECEWSI	Electronic Warfare Model	JTC	SCO	Availability is unknown.
STORM	Air Heavy Joint (analysis)	HLA	SCO, SFEO, SSO (partial)	Available
FLAMES	Air	ITC, HLA	SCO (partial)	Available
ACE-IOS	Air	JLVC	SCO (partial), SFEO (partial)	Available
JTLS-GO	Joint	JLVC, NTF	SCO (partial), SFEO (partial)	Available
hTEC	Joint	MSaaS	SCO, SFEO, SSO, SFAO	Not available in 2017
ASCCE	Joint		SCO, SFEO, SSO (partial)	Available
MDST	Air, Missile Warning	-	SFAO	Available
Serious Games	Gaming	-	-	-
Communications / Network Simulators	Communications and Network System/Scheme Design	-	SCO (partial), SFEO (partial), SSO (partial)	Available

- **Portable Space Model** (PSM): One of the early attempts in integrating a space simulator into a federation is Joint Training Confederation (JTC) with Portable Space Model (Cayirci and Marincic 2009). JTC is an Aggregate Level Simulation Protocol (ALSP) federation, and one of the pioneering works for federating military constructive simulation systems. Joint Simulation System (JSIMS) and high level architecture (HLA) (IEEE 2010; Tolk 2012) made both JTC and ALSP obsolete before 2000. PSM is a US Space Command (SPACECOM) model. It is a discrete event simulation written in C that models satellite detection and early warning for tactical ballistic missiles. In other words, it partially covers SCO, SFEO (i.e., tactical warning) and SFAO (i.e., TBMD/BMD) domains.
- **Air Warfare Simulation** (AWSIM): AWSIM is the primary model of the Air Force Modeling and Simulation Training Toolkit (AFMSTT). It was a federate in JTC, and later became a federate also in Joint Live Virtual Constructive (JLVC) Federation, which was an HLA federation maintained by US Joint Forces Command (USJFCOM 2010). AWSIM was developed in the 1980s, and the favored simulation system of the US Air Force for conducting simulation exercises in Air Warfare and Space Operations (Tolk, 2012). AWSIM can support the SCO mission area.
- **Joint Electronic Combat Electronic Warfare Simulation (JECEWSI):** JECEWSI is an electronic warfare simulation system and was a federate in JTC (Cayirci and Marincic 2009). The JTC was composed of nine simulations or actors, and one of them was JECEWSI (Strickland 2011). JECEWSI focuses on electronic warfare and electronic combat environments in support of tactical air, electronic warfare defense and air defense operations (Cayirci and Marincic 2009; Tolk 2012). It can provide partial support for SCO and SFEO.
- **Synthetic Theater Operations Research Model** (STORM): STORM is the successor of

the air force comprehensive theater level analytical campaign simulator, called THUNDER. STORM is a campaign level simulator and supports in depth analysis of the contributions of air and space power to a campaign (Pugh 2000). It is designed as an aid to senior decision makers across the acquisition, policy and operations communities. The focus of the model is on the air-air and air-ground combat. However, since the models in STORM includes space objects and their interactions with the air and surface/ground objects, the relations between air/ground combat and space capabilities, and therefore SCO and SFEO can be simulated. Partial support for SSO can also be provided by STORM.

- **FLAMES**: FLAMES is an air simulation system, and a part of NATO Integrated Training Capability (ITC). ITC is used for training NATO combined air operations centers (CAOCs). Flames is also available in NATO Integrated Command and Control (ICC) system to provide a testing environment for air tasking orders (ATO). Its support to space mission areas is limited and indirect. FLAMES addresses many aspects of constructive simulation development and use, including customizable scenario creation, execution, visualization, and analysis, as well as interfaces to constructive, virtual, and live systems (Ternion 2009). It can partially support SCO.
- **Air and Space Constructive Environment Information Operations Suit** (ACE-IOS): The Air Constructive Environment-Information Operations Suite (ACE-IOS) provides the authoritative representation of Air Force information operations. ACE-IOS is comprised of models that support training and mission rehearsal for the Air Force, Joint Task Force commanders, and battle staffs during Joint and Service exercises and experimentations (DTIC 2013). ACE-IOS is the Joint Network Simulation (JNETS) tool for using to create the cyber range environment (Wells and Bryan, 2015).
- **Joint Theater Level Simulation Global Operations** (JTLS-GO): *The Joint Theater Level Simulation - Global Operations (JTLS-GO) is an interactive, Internet-enabled simulation that models multi-sided air, ground, and naval civil-military operations with logistical, Special Operation Force (SOF), and intelligence support. JTLS - GO development began in 1983 as a project funded by the U.S. Readiness Command, the U.S. Army Concepts Analysis Agency, and the U.S. Army War College. The simulation was originally designed as a tool for development and analysis of joint as well as combined (coalition) operations plans. Today, JTLS is primarily used as a training support model that is theater-independent and does not require a knowledge of programming to operate effectively* (Rolands and Associates 2017).
  The primary focus of the JTLS - GO system is conventional joint and combined operations at the Operational Level of War as defined by the Joint Staff's Universal Joint Task List. JTLS explicitly models air, land, sea, amphibious, and SOF operations. The Simulation supports limited nuclear and chemical effects, low-intensity conflict and pre-conflict operations, as well as support of Humanitarian Assistance and Disaster Relief (HA/DR) Operations (Rolands and Associates 2017).
  JTLS has been a federate to various federations including JLVC and NATO Training Federation (NTF)(NATO MSG-068 2016).
  JTLS-GO is the main simulation tool to support NATO's operational and higher level CAX. In the NATO's exercises, it is used to simulate/visualize satellite orbits, and ISR based on satellite payload. Therefore, it provides partial support for SCO (i.e., situational awareness) and SFEO (i.e., ISR). Please also note that space weather forecasts are injected as an incident content in NATO CAXs, but not simulated by JTLS-GO.
- **HAVELSAN Training and Experimentation Cloud** (hTEC): hTEC has been developed following modelling and simulation as a service architecture (Cayirci, Karapinar and Ozcakir 2017). At the time that this paper is written, it is being implemented on a testbed called BSigma (Cayirci, Karapinar and Ozcakir 2017). hTEC includes services that covers SCO, SFEO and SFAO completely and SSO partially.
- **The Air Space Cyber Constructive Environment** (ASCCE): *ASCCE simulations are the authoritative representation of air, space, and cyber power for U.S. ASCCE is used throughout the USAF for warfighter events. It is the air, space, and cyber representation for certifying Joint Force Air Component Commanders and their staff* (Deforest 2009). *It includes the Air Force Modeling and Simulation Training Toolkit (AFMSTT), which provides the representation of Air Force and Joint theater-level air and space power and is used to train Air and Space Operations Center (AOC) personnel and Combat Commanders (COCOM) staffs.* (DTIC 2013).
- **Missile Defense Space Tool** (MDST): MDST supports live or simulated exercises. *MDST can be used for simulating Military Space Operation by injecting messages into operational communication and simulation networks* (Rainey and Davis 2004). *It provides space simulation support, and space exercise by conducting National Missile Defense and*

Theater Air and Missile Defense activities. It provides partial support SFAO and SFEO.

- **Serious Games:** Gaming products may also provide solutions for the requirement. There are hundreds of space games. We surveyed tens of them, such as, Deliver the Moon, Limit Theory, Infinity Battlespace, The Mandate, Everspace and Dreadnought. We could not find any space game that can be useful for any of the space mission areas listed in Table 2. However, the game called Star Citizen (Star Citizen 2017) may evolve and become useful for the SSO and SFAO mission areas in the following years.
- **Communications / Network Simulators:** There are many tools used for simulating the communications/networking techniques, algorithms, schemes and systems in scientific and industrial research (Network 2017; Pan 2017), such as, Network Simulator (ns3) (NS3 2017), OPNET and OMNet++ (Pan 2017). Although these tools are designed for the technical evaluation of the schemes, protocols and systems, they may become useful also for the training in the SFEO and SSO mission areas.

Apart from the simulation systems listed in Table 2 which are mostly from the U.S., there are other military constructive simulation systems with air domain functionalities, such as the following (Cayirci and Dusan 2009):

- SCEPTRE, France
- Air land interactive conflict evaluation (ALICE), Germany
- Simulation Modell fur ubungen Operativer Fuhrung (SIMOF), Germany.
- Joint Operational Command and Staff Training System (JOCASTS), UK
- CATS-TYR, Sweden

However, to the best of our knowledge, their support to space mission areas are either very limited and indirect or not existing at all.

4. ALTERNATIVE APPROACHES FOR DEVELOPING JMSOS SYSTEMS

One of the following four approaches can be applied for developing a JMSOS capability:

- Models for military space mission areas can be included into the joint or air military simulation systems.
- An integrated JMSOS system that consists of functionalities for all the joint military space mission areas can be developed. This system can be federated with the other military simulation systems by using distributed simulation technologies, such as, HLA (IEEE 2010) or used standalone. They can populate and stimulate the command and control systems by using various C2 interoperability or military datalink protocols (Hura et al. 2000). Software as a service model can also be used for making this system available to the users (Cayirci 2013).
- Alternatively, multiple simulation systems can be developed for each of the space mission areas or even sub topics under the mission areas. They can be federated or used standalone.
- Finally service oriented cloud approach as described in (Cayirci Karapinar and Ozcakir 2017) can be followed.

For JMSOS, we follow the last approach in hTEC, i.e., service oriented cloud architecture, in which the following services are available as models ready to be integrated into a software as a service (Cayirci Karapinar and Ozcakir 2017):

- Spacecraft and orbit
- Space weather
- Weapon effects on space assets
- Space electronic warfare
- Space sensors
- Space ISR
- Space communications
- GPS
- Space weapon effects

These models are in the form of software libraries that can be composed into a simulator by the hTEC composition layer. The composed service can also be federated with the federates from the other domains by using HLA. Please note that, hTEC services (i.e., models) are not limited to only space mission areas. The hTEC architecture follows NATO recommendations on modelling and simulation as a service (MSaaS) (Cayirci Karapinar and Ozcakir 2017; NATO MSG-136 2017).

5. CONCLUSION

Although the domain of joint military space operations is critical for the success of joint operations and need to be integrated into joint operations planning and therefore practiced during CAXs and wargames, the JMSOS capability is typically not available or very limited for the exercises. We think that the main reasons for that are two folded:

- Operational space picture is very seldom of interest for the commanders and staff in operational and higher levels, although it must be.
- Changing the operational space picture during the execution phase of an exercise is typically not possible.

However, space mission areas are not limited to SCO and SSO but also SFEO, which has a major impact on joint operations and operational picture. Without a proper JMSOS capability, it is too difficult to portray a complete and realistic synthetic theater, and therefore it is important to employ the JMSOS tools especially in joint operational and higher level CAXs and wargames. Nevertheless, proper JMSOS tools are not available. Typically, a limited and indirect set of JMSOS functionalities are implemented in some military constructive simulation systems.

ACKNOWLEDGMENTS

This work is conducted as part of the HAVELSAN project titled as “Service Oriented Cloud Based Simulator/Simulation Development and Modelling Infrastructure” funded by TUBITAK 1511-BIT-BBIL-2015-2 program.

REFERENCES

Cayirci E. 2006. NATO Joint Warfare Center’s Perspective on CAX Support Tools and Requirements. *ITEC’2006*, London, U.K.

Cayirci, E. and D. Marincic. 2009. *Computer Assisted Exercises and Training: A Reference Guide*. New Jersey: John Wiley & Sons.

Cayirci, E. 2013. Modelling and Simulation as a Service: A Survey. *In Proceedings of the 2013 Winter Simulation Conference, edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl.* Washington DC, USA.

Cayirci E., A. Bruzzone, F. Longo and H Gunneriusson, 2016. A Model to Describe Hybrid Conflict Environments. I3M.

Cayirci, E., H. Karapinar, and L. Ozcakir. 2017. Joint Military Space Operations Simulation as a Service. *In Proceedings of the 2017 Winter Simulation Conference, edited by W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page.* Las Vegas, USA.

Cayirci E. and L. Ozcakir, 2016. ‘Modeling and Simulation Support to Defense Planning Process’. Journal of Defense Modeling and Simulation: Applications, Methodology, Technology, DOI: 10.1177/1548512916675433.

Deforest, R. 2009. Department of the Air Force, Orlando, Florida. *The Air Force Agency for Modeling and Simulation: Advancing Modeling and Simulation for the Warfighter*. ITEA Journal 2009; 30: 23–28.

DoD. 2015. Operation of the Defense Acquisition System. *The US Department of Defense Instruction 5002*. Issue Date: January 7, 2015.

DTIC, Defense Technical Information Center, 2013. http://www.dtic.mil/descriptivesum/Y2013/AirForce/stamped/0207601F_7_PB_2013.pdf last time accessed on July 20, 2017.

ESA. 2017. European Space Agency Modelling and Simulation Section. http://www.esa.int/TEC/Modelling_and_simulation/TEC89CNWTPE0.html. last time accessed on June 26, 2017.

Hura, M., G. McLeod, E. Larson, J. Schneider, and D. Gonzales. 2000. Interoperability: A Continuing Challenge in Coalition Air Operations. ISBN-10: 0833029126, ISBN-13: 978-0833029126.

IEEE. 2010. *1516-2010 - IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA)-- Framework and Rules*. IEEE Computer Society.

NASA. 2010. *National Aeronautics and Space Administration DRAFT Modeling, Simulation, Information Technology & Processing Roadmap*. Technical Report November 2010.

NATO MSG-068. 2016. NATO Education and Training Network. NATO Science and Technology Organization Technical Report RTO-TR-MSG-068, DOI: 10.14339.

NATO MSG 136. 2017. *Operational Concept Document for the Allied Framework for M&S as a Service*. Draft v0.5.0. NATO Science and Technology Organization Technical Report STO-TR-MSG-136-OCD.

NATO. 2009. *Allied Joint Doctrine for Air and Space Operations*. NATO AJP-3.3(A).

Network. 2017. Open Source Routing and Network Simulation. http://www.brianlinkletter.com/open-source-network-simulators/. Last time accessed on July 20, 2017.

NS3. 2017. https://www.nsnam.org/. Last time accessed on July 20, 2017.

Pan, J. 2017. A Survey of Network Simulation Tools: Current Status and Future Developments. http://www.cse.wustl.edu/~jain/cse567-08/ftp/simtools/index.html. Last time accessed on July 20, 2017.

Pugh, D.M., 2000. A Validation Assessment of the STORM Air-toAir Prototype Algorithm. Master Thesis, Air Force Institute of Technology, 20000613 068.

Rainey, L.B. and P.K. Davis. 2004. *Space Modeling and Simulation: Roles and Applications Throughout the System Life Cycle*. AIAA. ISBN: 1884989152, 9781884989155.

Rolands and Associates. 2017. https://www.rolands.com/jtls/j_over.php. last time accessed on June 26, 2017.

SHAPE. 2013. *Allied Command Operations Comprehensive Operations Planning Directive COPD Interim V2.0*. Issue Date: October 04, 2013.

Star Citizen. 2017. https://robertsspaceindustries.com/ last time accesses on July 20, 2017.

Strickland, J.S. 2011. *Using Math to Defeat the Enemy: Combat Modeling for Simulation.* Lulu Inc, ISBN 978-1-257-83225-5.

Taylor, S.J.E., A. Khan, K. L Morse, A. Tolk, L. Yilmaz, J. Zander, and P. J Mosterman. 2015. Grand challenges for modeling and simulation: simulation everywhere from cyberinfrastructure to clouds to citizens. *Simulation Transactions of the Society for Modelling and Simulation International, Vol. 91, Issue 7, pp. 648-665.*

Ternion, 2009. *FLAMES Simulation Framework Brochure*, http://www.ternion.com/flames-overview/, *last time accessed on July 20, 2017.*

Tolk, A. 2012. Engineering Principles of Combat Modeling and Distributed Simulation. John Wiley and Sons, New Jersey, USA. ISBN 978-0-470-87429-5.

USJFCOM. 2010. *Joint Live Virtual and Constructive (JLVC) Federation Integration Guide Version 3.1.* Issued on January 13, 2010.

Wells, D. and D. Bryan. 2015. *Cyber Operational Architecture Training System Cyber for All.* Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), Orlando, USA.

Zehe, D., W. Cai, A. Knoll, and H. Aydt. 2015. Tutorial on a *Modelling and Simulation Cloud Service. In Proceedings of the 2015 Winter Simulation Conference, edited by L.Yılmaz, W.K.V. Chan, I. Moon, T.M.K. Roeder, C. Macal, and M.D. Rossetti. Huttington Beach, USA.*

LESSONS LEARNED FOR AFFECTIVE DATA AND INTELLIGENT TUTORING SYSTEMS

Keith Brawner[(a)]

[(a)] United States Army Research Laboratory, Human Research and Engineering Directorate

[(a)]Keith.W.Brawner.Civ@mail.mil

ABSTRACT
Human-to-human tutoring is the most effective manner for learning. Human tutors, however, pay attention to the cognitive and affective states of the learners and use the knowledge to adjust instructional strategies. According to theory, data about the learner informs data about instruction informs instructional designs and impacts student learning. The crux of these type of operations is the ability to recognize affect in the learner, but deployed systems to do this have not had significant success with the "one size fits all" type of generalized models. An alternative to this approach is individualized models, which have shown limited success to date in other domains. This paper furthers the investigation of individualized models and validates the approach, showing that it can produce models of acceptable quality, but that doing so does not obviate the experimenter from creating quality generalized models prior to individualizing.

Keywords: Adaptive and Predictive Computer-based Training, Intelligent Tutoring Systems, Architectural Components, Emerging Standards

1. INTRODUCTION
Tutoring by an expert human tutor is extraordinarily effective. There is debate among the literature about how effective human tutors are, but it is commonly found to be between one and two standard deviations of improvement; between one and two letter grades (Bloom, 1984; VanLehn, 2011), with the highest reported gains in the range of four standard deviations (Fletcher & Morrison, 2012). Theory indicates that learner data inform learner states which inform instructional strategy selection which influences learning gains (Sottilare, Brawner, Goldberg, & Holden, 2012), shown in Fig. 1. As ITS research moves towards highly adaptable and individualized tutoring, the need to automatically assess the cognitive and affective states of the individual learner for instructional adjustment has been well documented (Department of the Army, 2011; Woolf, 2010), and extensive work has been performed to recognize the emotional state of a learner by incorporating sensors to monitor both behavioral and physiological markers and is discussed in the coming sections (AlZoubi, Calvo, & Stevens, 2009; Calvo & D'Mello, 2010; D'Mello et al., 2005; D'Mello, Taylor, & Graesser, 2007; McQuiggan, Lee, & Lester, 2007).

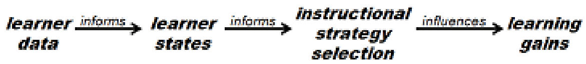


Fig. 1: Learning Effect Chain (Sottilare, 2012)

The traditional manner of incorporating these models is to take measurements of a group of classroom measurement, move the data offline, perform significant amounts of feature extraction, create models, and apply the models to the next set of classroom learners. However, these systems have not been shown to transition well into a field of use due to a variety of reasons including differences in classroom structure, sensor placement, time of day, season of the year,, individual differences, and other reasons (Brawner, 2013). An alternative to this practice are individualized modeling techniques (AlZoubi et al., 2009) using realtime analysis (Brawner, 2013).

However, realtime data presents unique problems for algorithmic modeling purposes. There are four main problems with realtime data, they are 1) the data can be of potentially infinite length, 2) concept detection, 3) concept drift, and 4) concept evolution. The combination of these issues present a problem for whichever type of algorithm is used to solve it. The realtime construction and use approach necessitates a stream model of the data, with the following assumptions, and corresponding design limitations, as (Beringer & Hüllermeier, 2006) outlines:

- Operations must be done on the data as they become available
    - Data ordering is outside of control
    - Data cannot be requested
    - Knowledge about prior points must be encoded, if they are to be related to each other
- The dataset is of infinite length; it is not possible to store or analyze all of the data
    - Data elements are not available for repeated request
- Data must either be saved or discarded

    - Practical processing and memory limits necessitate the discard of most data
- There are strict time constraints
    - Data must be processed in real time
- An approximate solution is an acceptable substitute for an ideal one (Considine, Li, Kollios, & Byers, 2004)

These problems are the starting point for an avenue of research into realtime modeling of student state at runtime, rather than *a priori*. However, there are lessons learned from these modeling techniques – the foremost of which is that the models must have feature extraction along the order of the offline models. This paper presents several algorithmic approaches towards dealing with the problem of affective computing at runtime. It discusses lessons learned across 4 algorithmic approaches, 3 labeling approaches, and 5 datasets. The algorithm and labeling approaches are discussed next, followed by a discussion of the various datasets, before performance and lessons learned are addressed.

2. ALGORITHMIC APPROACHES TO DEAL WITH REALTIME DATASTREAMS

2.1. Introduction

The goal of this research is to further investigate the approach of creating realtime models. It has previously shown promise for small sample sizes (Brawner, 2013; Brawner & Gonzalez, 2016), and the current work is an attempt to expand and validate the approach on a larger and more complete sample. In order to determine if the *approach* is valid, a number of disparate but representative machine learning methods were chosen, including an incremental clustering approach, a neural network approach, a linear regression approach, and a graphical model approach.

Further, it is useful to consider that, since the models are being made for individual learners as they begin, information *from* the learner is available during model construction time. Considering this, it is possible to ask the learner to selectively tag individual states when they are experienced. Each algorithm should take an approach that uses this information to tag the most pervasive (but currently unknown) category. In addition to an outline of each algorithm, modifications to adjust the algorithm for active learning are taken into account.

2.2. Clustering Approach

The clustering approach that was taken was an incremental algorithm described in pseudo-code in Fig. 2. The modifications made for semi-supervised and active learning are shown in Fig. 3.

Fig. 2:Pseudo-code description of clustering algorithm

```
For each new point, incrementally
Compare each point to all known centroids
If no cluster is within range of <vigilance>
this point is a new centroid
Else, move the matched cluster <delta> in new point direction
Merge closest centroids based on <vigilance>, if appropriate
Keep track of the number of points in these centroids, and the last
point which modified the centroid, label if possible
```

Fig. 3:Pseudo-code of active learning changes for clustering

```
When a label is requested
Find the largest size centroid which does not currently have a
label, Return the last seen datapoint which modified this centroid
```

2.3. Linear Regression Approach

A regression approach was additionally taken based on the superior performance of the logistic regression from the initial models. The Vowpal Wabbit (VW) software package is typically extraordinarily effective for the creation of regression models, and is capable of both one-pass (online) learning and supports the use of active learning by default (Langford, Karampatziakis, Hsu, & Hoffman, 2010). The authors would encourage the readers to see the original publications for a description of the active learning approach used, as their complexity is beyond the scope of this work (Beygelzimer, Hsu, Langford, & Zhang, 2010). A basic description of the algorithm used is described within Fig. 4.

Fig. 4:Pseudo-code of linear regression algorithm; VW

```
Start with for all i: wi= 0          Within the loop:
Get an example:        x ∈ (∞, ∞)
Make a prediction:     y = Σi  wixi
Learn the Truth:       y ∈ [0,1] with importance I
Update the weight:     wi = wi + 2η(y-yi)I
Repeat for specified number of passes or other criteria
```

2.4. Neural Network Approach

Adaptive Resonance Theory (ART) is a type of neural network architecture which classifies objects based on the activation of nodes in a structure that is build as new data is presented to it. It was developed to classify data in a one-pass learning environment (Carpenter and Grossberg 1995), and has historic performance roughly equivalent to neural networks, but with significantly reduced training time. In its most basic form, ART draws n-dimensional hypercubes around similar input patterns, where n is the dimension of the input data. Although sometimes viewed as a disadvantage, ART systems are capable of one-pass learning, which makes them appropriate for realtime classification problems. The basic algorithm is described within Fig. 5. Similar alterations to Fig. 3 were added for compatibility with the active learning process.

Fig. 5:Pseudo-code of ART algorithm

```
For each new data point
Compute each neurons' weighted activation to it ( yi = Σwij*xi )
Select the neuron with the highest activation
Test if this neuron in vigilance (xi fuzzyAnd wx < vigilance)
If it is, update the weights:
wi = learningRate*xi + (1-learningRate)*wi
Otherwise, create a new category with xi weights
```

2.5. Online Semi-Supervised Growing Neural Gas (OSSGNG)

Neural Gas is a robustly converging alternative to the k-means approach of clustering that finds optimal representations based on feature vectors. These feature vectors construct a topographical map overlaying the data. An example of such an overlay map is included in **Error! Reference source not found.**. This approach has its roots in Self Organizing Maps (SOMs) (Kohonen, 1982) and Neural Gas topologies (Martinetz & Schulten, 1991). Growing Neural Gas (GNG) is an incremental version of Neural Gas which is appropriate for datastream analysis (Holmstrom, 2002), and was initially proposed by Fritzke (Fritzke, 1995). Semi-Supervised GNGs are a further outgrowth of these methods to make use of unlabelled datapoints for classification (Zaki & Yin, 2008). Beyer and Cimiano have modified the initial algorithm to make it appropriate to realtime problems (Beyer & Cimiano, 2011). They present OSSGNG as a topographical mapping algorithm synthesized from the various contributing fields. They examine several metrics for determination of the establishment of clusters, and find that the minimum distance metric has the best performance on problems of interest, which is used in the current work. The algorithm, as well as the realtime modifications and active learning modifications are described in Fig. 6 and Fig. 7.

Fig. 6: Initial Pseudo-code description of OSSGNG algorithm

Present the set of labeled data (LD) to the network, train only on it, label accordingly
Present an input from unlabeled data set (UD), xj,with the previous distance metric
Label xj according to the winning node, remove it from UD, enter it into the LD' set
Loop until UD set is empty
Present LD and LD' to evaluate performance

Fig. 7: Realtime appropriate pseudo-code description of OSSGNG

Present a datapoint, finding the two closest items s1 and s2
If there is a missing label, assign a label based on the nearest item (unlabeled is possible)
Increment ages (detailed originally)
Proceed with GNG steps, do not loop to reevaluate

Active Learning Modifications:
When a label is requested, find the network of the largest unknown class
Compute the centroid of this node-created network
Find and request the label of the point closest to the centroid

2.6. Evaluation Approach

Given that this method of model creation is relatively new, it requires a novel method of assessment. To do this assessment, for each individual a model is created over time in each of a supervised, semi-supervised, and unsupervised fashion. At the time of assessment, each of the unlabeled groups is given a label according to the majority-class of the true labels contained within it. As such, an unsupervised approach which draws a single cluster over each datapoint will have a 100% accuracy score, despite a lack of utility. Configuration settings for each algorithm were taken to create a minimal number of categories for labeling.

As a byproduct of the evaluation algorithm, each of the models begins with 100% accuracy, as, in each algorithmic case a single datapoint generates a single cluster and the majority-class of the cluster is correctly labeled. Gradually, as more data about both the user and labels comes available, the overall accuracy of the model decreases. This is seen in the overall trend in each graph. The research question that this paper is trying to answer is whether the approach of creating realtime models on the individual level is useful. As such, it is useful to see the overall effect of the model, and how useful it *would* have been, on average, for a given unit of time. The algorithm which is used to assess the performance of each of the methods is described below in Fig. 6.

Fig. 8:Pseudo-code of assessment algorithm

For x from 10-100, in increments of 10
Feed x% of the data to the algorithm
For each class created by unlabeled class boundaries
Label this class the majority label of true set
Evaluate for AUC ROC accuracy through input of data for classification (next, previous, all seen)
Destroy model, loop

The results presented later were analyzed with the Receiver Operator Characteristic (ROC) benchmark (Hanley, 1989), which plots the proportion of correctly-classified observations from the positive class (true positive rate) against the incorrectly-classified observations (false positive rate). The Area Under the Curve (AUC) of this function was calculated. The AUC ROC is designed to compensate for the misleading figures of “percentage accuracy” for unbalanced data. The AUC ROC measurement allows an algorithm with lower overall error rates, either true positive or false negative, to score well (Hanley & McNeil, 1983), as the all of the categories of possible classification are weighted equally. A measure of 0.5 indicates a simple majority-class classifier. In general, AUC metrics of greater than 0.8 are considered good, while classifiers lower than 0.6 are considered poor; those scoring in the 0.2 range in between those values are considered to be fair. The AUC ROC is comparable to the A’ value traditional in machine learning literature.

3. DATASETS AND COLLECTION

3.1. Dataset #1, #2: Affective and Cognitive Dataset from High- and Low-Cost Sensors

The first two dataset were collected as part of an experiment to evaluate low-cost sensors. College-aged military learners experienced a breadth of learning-relevant emotions while watching videos or playing video games. They were measured by a suite of sensors. Cognitive states, such as distraction, are labeled with a high-cost sensor. Affective states, such as frustration, are labeled with a self-reporting tool called EmoPro (Kokini et al., 2012). There were 14 usable sets of cognitive data, with 19 usable sets of emotional data. The data was collected while users interacted with a training video game with validated scenarios to induce specific cognitive and emotional states, or while watching videos validated to produce cognitive and emotional states. Models developed under this effort are designed to replace the high-cost sensors measures. This dataset, and the experiment from which it was produced, is referred to as Dataset #1, when using cognitive labels, or as Dataset #2, when using affective labels. The features of this dataset are described in summary in Table 1. More information on the collection, population dynamics, and usability criteria of this dataset can be found in prior work (Brawner, 2013).

3.2. Dataset #3: Vigilance and Eyetracking Dataset

The experiment that produced the third dataset was part of a larger suite of experiments, each of which was targeted towards different objectives. The first of these was the objective to examine the relationship of workload and multi-tasking performance as part of a Mixed Initiative Experimental (MIX) testbed, which incorporates theory-driven tasks into a moderately high-fidelity military simulation designed for multi-tasking and physiological data capture (Reinerman-Jones, Barber, Lackey, & Nicholson, 2010). The most relevant experimental purpose is to *create generalized models of physiological response* to situations of changing workload in order to preemptively reduce workload in the future (Barber & Hudson, 2011). This dataset was collected as part of an experiment to evaluate physiological response to situations of changing workload, a cognitively relevant learning state. College students experienced simultaneous tasking on detecting changes and indentifying threats on a displayed monitor. Their cognitive state was monitored by a suite of sensors, with the data cognitively labeled with a high cost sensor. Models developed under this effort are intended to aid in classification of workload, with the intent of having a system compensate during times of high/low operator workload. This dataset, and the experiment that produced it, is referred to as Dataset #3. More information on the college and population dynamics of this dataset can be found in prior work (Brawner, 2013).

3.3. Dataset #4, #5: Medical Learning Kinect Dataset, 2013 and 2016

There are two datasets subject to analysis in this paper, one from 2013 and one from 2016 (DeFalco, et al., 2017). They were both collected from a class of United States Military Academy (USMA) cadets as they interacted with a Tactical Combat Casualty Care Simulation (TC3Sim). Participants interacted with the system for approximately an hour of total protocol, while approximately 25 minutes were spent within the simulation. The participants were monitored via within-system interactions as well as via Microsoft Kinect sensor. While the participants interacted with the system, the BROMP protocol (Ocumpaugh, Baker, & Rodrigo, 2012) was used in order to label the data of affective states of the learners as it was observed among the participants in order to label the "ground truth". There are advantages and disadvantages to different labeling schemes (DeFalco, et al., 2017), but in-field observations have been found to be relatively stable over time (Ocumpaugh et al., 2012).

The initial 2013 collection saw the development of various feature extraction methods, which were used in both studies to compare benchmark performance. The features and models from the 2013 study were used in 2016. The raw features of the Kinect data include the center shoulder position, head distance, and top of skull distance. Of the 91 vertices recorded by the Kinect sensor, only three are utilized for posture analysis: top_skull, head, and center_shoulder. These vertices were selected based on prior work investigating postural indicators of emotion with Kinect data (Grafsgaard, Wiggins, Boyer, Wiebe, & Lester, 2014). Derived statistical and windowed features are calculated over top of these items, including the minimum observed, maximum observed, median, variance; each of these features is additionally calculated for 5/10/20 second windows. Further and more specific information on the dataset can be found in other works (DeFalco, et al., 2017; Rowe, Mott, & Lester, 2014; Rowe, Lobene, & Sabourin, 2013).

3.4. Summary of Dataset Features

Due to paper length limits, a discussion of all of the features, meaning behind the features, and derived features of all of the sensors, across each of the five datasets, is not possible. Feature extraction was performed on dataset #1 and #2, with simple statistical measures (mean over a time period, difference from last observation, etc.). No effort for feature engineering was attempted on dataset #3. Extensive feature engineering was performed on dataset #4 and #5, and more information can be found in the publication which discusses the process (Paquette et al., 2016). The sensors, labels, and measurement across the number of datasets are presented in Table 2. Dataset #5 did not have any occurences of Boredom during observation, so models of Boredom are not considered in any of the models created.

Table 1 – Summary of sensors and ground truth across datasets (***bold italics*** indicates ground truth, * indicates feature extraction previously performed)

Dataset	Sensor	Measure
#1*	***ABM EEG***	***HighEngagement Distraction Workload***
#2*	***EmoPro***	***Anger Boredom Fear***
#12	Neurosky EEG	Alpha1, Alpha2, Gamma1, Gamma2, Delta, Beta1, Beta2, Theta, Attention, Meditation

#12	Zephyr HxM	Heart
#12	Motion (custom)	Motion
#12	Chair (custom)	Chair1-8
#12	Eye Tracking (custom)	LeftEyePupilDiameter
#3	***FaceLab 5***	***IndexofCognitveActivity***
#3	FaceLab 5	FixationDuration, PupilDiameter
#45*	***BROMP***	***Boredom, Confusion, Engaged Concentration, Frustration, Surprise***
#45	Microsoft® Kinect	Various, see text.

4. RESULTS

Given that a single user had ten datapoints, one for each of ten percentage points among the timeline, a summary presentation of data is required. Firstly, it is useful to note that performance starts high, as a single point represented in a single cluster has a 100% accurate classifier to categorize it.

Models with a performance benchmark below 0.6 are typically considered unusable. The above table shows that the best performance models using realtime methods for Dataset #1 and Dataset #3 are below any reasonable margin of usability. Among the usable models, the most common usable techniques consistently are the neural and grouping techniques represented by ART and K-means clustering.

Summary results can be presented in multiple manners – graphically, numerically in a number of single-model tables, or via summary table. The results are presented in this manner in Fig. 9, Fig. 10, and Table 2, respectively. Many results from many models and many studies are presented in this paper, with the goal of answering the "does this approach work?" question. Table 2 is best to answer this question, but it is based on the data shown in Fig. 9, Fig. 10, and many other similar figures, a substantial fraction of which have been published elsewhere (Brawner, Gonzalez, 2013).

Fig. 9: Unsupervised, Supervised, and Semi-Supervised performance of the ART algorithm over time, Dataset #5

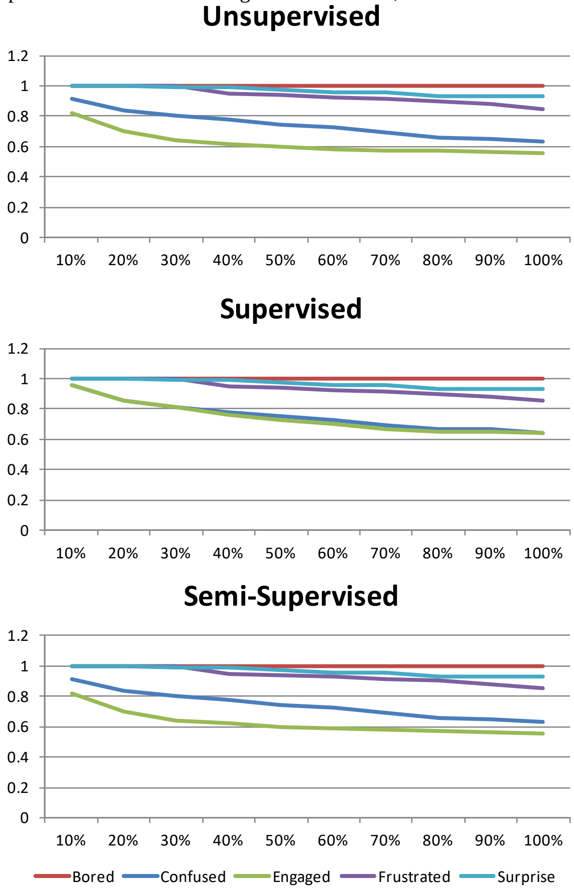


Fig. 10: Boredom model qualities with supervised ART algorithm, Dataset #2

User	20%	30%	40%	50%	60%	70%	80%	90%	100%	Avg
4134	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**1.000**
4133	0.96	0.92	0.92	0.77	0.70	0.71	0.68	0.72	0.58	**0.773**
4131	0.97	0.66	0.97	0.97	0.97	0.97	0.97	0.97	0.97	**0.939**
4127	0.63	0.67	0.60	0.60	0.53	0.51	0.62	0.51	0.65	**0.591**
4121	0.80	0.95	0.82	0.81	0.83	0.83	0.81	0.84	0.77	**0.829**
4111	1.00	1.00	1.00	0.75	0.73	0.79	0.83	0.74	0.79	**0.846**
4115	1.00	1.00	1.00	0.95	0.63	0.91	0.52	0.79	0.58	**0.821**
4135	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**1.000**
4136	1.00	1.00	1.00	0.60	0.60	0.60	0.60	0.60	0.60	**0.733**
4137	0.91	0.78	0.74	0.73	0.76	0.78	0.80	0.81	0.82	**0.792**
4101	0.66	0.66	0.66	0.64	0.65	0.74	0.72	0.64	0.63	**0.665**
4117	1.00	0.67	0.52	0.67	0.72	0.51	0.56	0.58	0.59	**0.648**
4102	0.85	0.79	0.78	0.85	0.80	0.67	0.71	0.76	0.81	**0.780**
4105	0.80	0.84	0.84	0.69	0.62	0.66	0.63	0.62	0.66	**0.708**
4104	1.00	1.00	0.75	0.87	0.80	0.74	0.66	0.73	0.75	**0.810**
4107	1.00	1.00	1.00	0.75	0.79	0.74	0.79	0.79	0.79	**0.848**
4106	0.65	0.75	0.67	0.75	0.70	0.64	0.70	0.70	0.72	**0.699**
4112	0.98	0.88	0.88	0.88	0.78	0.78	0.65	0.60	0.58	**0.779**
4132	1.00	1.00	0.79	0.89	0.84	0.85	0.87	0.86	0.75	**0.873**
Average	**0.906**	**0.872**	**0.839**	**0.800**	**0.760**	**0.760**	**0.743**	**0.750**	**0.739**	**0.796**
Total Usable (avg ROC >0.6):				18		Percent Usable:			95%	

5. DISCUSSION

Methods for realtime model creation are fundamentally handicapped when compared to traditional model creation methods. They are tasked to predict with a *fragment* of the total data, under *strict time constraints*. Given how unlikely the offline-created models are to transfer to a field of use, a practitioner should consider a small drop in overall accuracy to be acceptable. Large drops in accuracy, or models which poorly model the

ISBN 978-88-97999-90-4; Bruzzone, Cayirci and Sottilare Eds.

Table 2 – Summary of Model Performances. TIE indicates similar performance for ART and clustering unless otherwise noted. Instances where models are usable are bolded, usable instances (semi-supervised only) are italicized.

Dataset	Best Supervised Model / Quality	Best Semi-Supervised Model / Quality	Best Un-Supervised Model / Quality	Prior Benchmark
#1 (Distract)	TIE .552 Inck & GNG	GNG .538	Inck .537	.81
#1 (Engage)	TIE .560	TIE .555	TIE .555	.80
#1 (Workload)	Inck .529	VW .520	GNG .524	.82
#2 (Anger)	**ART .776**	***Inck .677***	**TIE .652**	<0.6
#2 (Boredom)	**ART .796**	***Inck .626***	**TIE .612**	.79
#2 (Fear)	**ART .841**	***Inck .810***	**TIE .805**	.83
#3 (ICA)	GNG .505	GNG .506	GNG .504	N/A
#4 (Boredom)	**Inck .886**	***Inck .888***	**Inck .891**	0.528
#4 (Confusion)	**Inck .820**	***Inck .820***	**Inck .831**	0.535
#4 (E. Conc.)	**Inck .949**	***Inck .765***	**Inck .952**	0.532
#4 (Frust.)	**ART .939**	***TIE .939***	**ART .941**	0.518
#4 (Surprise)	**ART .954**	***ART .954***	**ART .955**	0.493
#5 (Confusion)	**ART .630**	***Inck .640***	**Inck .750**	0.489
#5 (E. Conc.)	**Inck .595**	***Inck .595***	**Inck .647**	0.546
#5 (Frust.)	**TIE .851**	***ART .851***	**TIE .851**	0.331
#5 (Surprise)	**TIE .932**	***TIE .932***	**TIE .932**	0.51

underlying phenomenon should, of course, not be considered for application.

These methods failed on two of the above datasets (#1 and #3), experienced only moderate success on one of the above datasets (#2), and experienced significant success on two of the datasets (#4, #5). In a discussion, we should consider the things that both failure and success have in common. In short, these are the frequency of data, quality of labels, and quality of feature extraction.

5.1. Quality of Labels

Firstly, there should be a note about the variance in the quality of the labels. There are a variety of labeling techniques, each with its own advantages and disadvantages (Brawner & Boyce, 2017). Dataset #1 and #3 used an automated sensor system to assign label individual cognitive states. These cognitive state labels change multiple times per second. In contrast, Dataset #2 used self-report labels to label a state experienced during the total, but short, interaction. This creates a more stable label, but may not be an accurate representation of ground truth due to the time between labeling and the experience. In contrast, The BROMP protocol labels of Dataset #4 and #5 are relatively stable at approximately 1 minute resolution, but taken during live interactions without delay. The best overall performance was observed with the stable labeling techniques that are likely to be observed in real-world performance.

The datasets with the worst performance (#1, #3) used labeling techniques for *cognitive* states, while the datasets with the best performance (#2, #4, #5) used labeling techniques for *affective* states. D'Mello would categorize affect as either states, traits, moods, or emotions (D'Mello, Blanchard, Baker, Ocumpaugh, & Brawner, 2014), which are defined partly by the longevity of their experience. The "sweet spot" for ITS is in relatively short affective emotions (anger, fear) or affective-cognitive blends (confusion, concentration), and *not* long-lasting moods (depression) or ephemeral cognitive states (distraction). The methods described in this work function well inside the "sweet spot", but generally fail outside of it. Simply, the methods presented here work well for the types of occasions where it would be used – relatively short term classifications of relatively short term states.

5.2. Frequency of Data

The frequencies of the Datasets varied significantly, as shown in Table 3. Each of the method selected is capable of realtime performance, but unusually high frequency is one possible reason for poor model performance with Dataset #3, and may have served to compound errors with labeling quality.

Table 3 – Dataset Frequencies

Dataset	Frequency
#12	~1 Hz
#3	~3500 Hz
#45	~0.7 Hz

5.3. Quality of Feature Extraction

The best performing datasets (#4, #5) already had significant feature extraction performed in addition to useful-quality models based on the extracted features. Moderately performing datasets (#1, #2) had only statistical features (mean, difference, etc.). Poorly performing dataset (#3) had no manufactured features whatsoever. Generally speaking, the higher quality, as shown by their utility in offline-created models, of the features indicates that the realtime models will perform better. It is worthwhile to note that effort was spent in order to make an apples-to-apples comparison of model

– the realtime and offline models were always made with the exact same features. A data scientist making models for utility is not bound by these constraints and is free to manufacture features as they see fit.

6. CONCLUSIONS

Generally speaking, the technique of realtime modeling, as opposed to offline modeling, *works*. It should not be forgotten that the alternative to this type of approach is the traditional process involving online collection, offline creation, and online usage, which has yet to see results in the field of practice; it *doesn't work*. With that said, however, there are places where this approach shines and places where the approach fails.

The first area where this approach fails is when the learner state is highly fluid in comparison to the collection window. If the learner state changes at 30 Hz, but data is collected 1 Hz, it is difficult to algorithmically determine learner state. The experimenter should be aware of instances where the learner state is fluid. The second major area of failure is when proper feature extraction hasn't been performed. Feature extraction is a useful process to offline signals and online signals alike; failure to isolate proper features will produce unusable models.

Naturally, the opposite of the above failures indicates the successes – realtime models can be created if the state is changing slowly relative to the feature collection window and when proper features have been considered prior to interpretation by models. The goal of this paper is to set up provide a guide to creating realtime models and to set expectations for performance. The best realtime models observed thus far have had the following features:

- They attempt to model an affective-cognitive combined state, which is relatively stable on the order of minutes.
- They make use of offline experiments for validation of feature extraction techniques

An example experimental setup for affective state detection within an intelligent tutoring systems may have the following features:

- Hardware sensors of physiological state - ideally standoff sensors such as the Microsoft Band for GSR and heartrate measures
- Existing feature extraction shown useful in other contexts – such as a 300ms sliding window of signal power on a GSR signal
- Participant able to label affect states as they come available; a system able to request these items
- Use of one of more machine learning measures, such as ART or incremental clustering.

The intelligent tutoring system described above must also be able to perform the other functions of an intelligent tutoring system, such as changing the instruction to better suit the learner. Modeling the state of the learner is only the start of the process of customizing instruction and feedback towards individual needs.

ACKNOWLEDGMENTS

The author would like to thank Benjamin Goldberg for assistance in collection/provision of the initial dataset and numerous enlightening conversations, Lauren Reinerman and Julian Abich for the collection/provision of the second dataset, and Ryan Baker, Jonathan Rowe, and Jeanine DeFalco for aiding in the collection of the final two datasets. The research described herein has been sponsored by the U.S. Army Research Laboratory. The statements and opinions expressed in this article do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

REFERENCES

AlZoubi, O., Calvo, R., & Stevens, R. (2009). Classification of EEG for Affect Recognition: An Adaptive Approach. *AI 2009: Advances in Artificial Intelligence*, 52-61.

Barber, D., & Hudson, I. (2011). Distributed logging and synchronization of physiological and performance measures to support adaptive automation strategies. *Foundations of Augmented Cognition. Directing the Future of Adaptive Systems*, 559-566.

Beringer, J., & Hüllermeier, E. (2006). Online clustering of parallel data streams. *Data Knowl. Eng, 58*(2), 180-204.

Beyer, O., & Cimiano, P. (2011). *Online semi-supervised growing neural gas.* Paper presented at the Workshop New Challenges in Neural Computation 2011.

Beygelzimer, A., Hsu, D., Langford, J., & Zhang, T. (2010). Agnostic active learning without constraints. *arXiv preprint arXiv:1006.2588*.

Bloom, B. S. (1984). The 2-Sigma Problem: The search for methods of group instruction as effective as one-to-one tutoring,. *Educational Researcher, 13*(6), 4-16.

Brawner, K., & Boyce, M. (2017). *Establishing ground truth on pyschophysiological models for training machine learning algorithms: Options for ground truth proxies*. Paper presented at the International Conference on Augmented Cognition a part of the Human Computer and Intelligent Interaction (HCII) multi-conference.

Brawner, K. W. (2013). *Modeling Learner Mood In Realtime Through Biosensors For Intelligent Tutoring Improvements.* Doctor of Philosophy in EECS, University of Central Florida.

Brawner, K. W., & Gonzalez, A. J. (2016). Modelling a learner's affective state in real time to improve intelligent tutoring effectiveness. *Theoretical Issues in Ergonomics Science, 17*(2), 183-210.

Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *Affective Computing, IEEE Transactions on, 1*(1), 18-37.

Considine, J., Li, F., Kollios, G., & Byers, J. (2004). *Approximate aggregation techniques for sensor*

databases. Paper presented at the 20th IEEE International Conference on Data Engineering.
DeFalco, J., Rowe, J., Paquette, L., Georgoulas-Sherry, V., Brawner, K.W., Mott, B., Baker, R.S., Lester, J. C., (2017), Detecting and Addressing Frustration in a Serious Game for Military Training. International Journal of Artificial Intelligence and Education (IJAIED). Springer.
D'Mello, S., Blanchard, N., Baker, R., Ocumpaugh, J., & Brawner, K. (2014). I Feel Your Pain: A selective Review of Affect-Sensitive Instructional Strategies. In B. Sottilare, A. Graesser, X. Hu & B. Goldberg (Eds.), *Design Recommendations for Intelligent Tutoring Systems: Volume 2 - Instructional Management* (Vol. 2, pp. 35): Army Research Laboratory.
D'Mello, S. K., Craig, S. D., Gholson, B., Franklin, S., Picard, R. W., & Graesser, A. C. (2005). Integrating Affect Sensors in an Intelligent Tutoring System *Affective Interactions: The Computer in the Affective Loop Workshop at 2005 International Conference on Intelligent User Interfaces* (pp. 7-13). New York: AMC Press.
D'Mello, S. K., Taylor, R., & Graesser, A. C. (2007). Monitoring Affective Trajectories during Complex Learning. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 203-208). Austin, TX: Cognitive Science Society.
Department of the Army. (2011). The U.S. Army Learning Concept for 2015. TRADOC.
Fletcher, J., & Morrison, J. (2012). DARPA digital Tutor: Assessment data. *Retrieved from Institute for Defense Analyses: http://www. acuitus. com/web/pdf/D4686-DF. pdf*.
Fritzke, B. (1995). A growing neural gas network learns topologies. *Advances in neural information processing systems, 7*, 625-632.
Grafsgaard, J., Wiggins, J., Boyer, K. E., Wiebe, E., & Lester, J. (2014). *Predicting learning and affect from multimodal data streams in task-oriented tutorial dialogue*. Paper presented at the Educational Data Mining 2014.
Hanley, J. A. (1989). Receiver operating characteristic (ROC) methodology: the state of the art. *Critical reviews in diagnostic imaging, 29*(3), 307.
Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology, 148*(3), 839-843.
Holmstrom, J. (2002). *Growing neural gas.* Master's Thesis, Uppsala University.
Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological cybernetics, 43*(1), 59-69.
Kokini, C., Carroll, M., Ramirez-Padron, R., Hale, K., Sottilare, R., & Goldberg, B. (2012). *Quantification of trainee affective and cognitive state in real-time*. Paper presented at the The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC).
Langford, J., Karampatziakis, N., Hsu, D., & Hoffman, M. (2010). Vowpal Wabbit 5.0 *Neural Information Processing Systems 2010 Conference Tutorial*.
Martinetz, T., & Schulten, K. (1991). *A" neural-gas" network learns topologies*: University of Illinois at Urbana-Champaign.
McQuiggan, S., Lee, S., & Lester, J. (2007). Early prediction of student frustration. *Affective Computing and Intelligent Interaction*, 698-709.
Ocumpaugh, J., Baker, R. S. J. d., & Rodrigo, M. M. T. (2012). Baker-Rodrigo Observation Method Protocol (BROMP) 1.0. Training Manual version 1.0. *New York, NY: EdLab.*(Technical Report).
Paquette, L., Rowe, J., Baker, R., Mott, B., Lester, J., DeFalco, J., . . . Georgoulas, V. (2016). Sensor-Free or Sensor-Full: A Comparison of Data Modalities in Multi-Channel Affect Detection. *International Educational Data Mining Society*.
Reinerman-Jones, L., Barber, D., Lackey, S. J., & Nicholson, D. (2010, July 17-20). *Developing methods for utilizing physiological measures.* Paper presented at the Applied Human Factors and Ergonomics Society Conference 2010.
Rowe, J., Lobene, E. V., & Sabourin, J. (2013). *Run-Time Affect Modeling in a Serious Game with the Generalized Intelligent Framework for Tutoring*. Paper presented at the AIED 2013 Workshops Proceedings Volume 7.
Rowe, J. P., Mott, B. W., & Lester, J. C. (2014). *It's All About the Process: Building Sensor-Driven Emotion Detectors with GIFT*. Paper presented at the GIFTSym2, Pittsburgh, PA.
Sottilare, R. A. (2012, September 2012). *Considerations in the Development of an Ontology for a Generalized Intelligent Framework for Tutoring*. Paper presented at the International Defense & Homeland Security Simulation Workshop Vienna, Austria.
Sottilare, R. A., Brawner, K. W., Goldberg, B. S., & Holden, H. A. (2012). The Generalized Intelligent Framework for Tutoring (GIFT).
VanLehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. *Educational Psychologist, 46*(4), 197-221.
Woolf, B. P. (2010). A Roadmap for Education Technology: National Science Foundation.
Zaki, S. M., & Yin, H. (2008). A semi-supervised learning algorithm for growing neural gas in face recognition. *Journal of Mathematical Modelling and Algorithms, 7*(4), 425-435.

Keith Brawner, Ph.D. is a researcher within the Learning in Intelligent Tutoring Environments (LITE) Lab within the U. S. Army Research Laboratory's Human Research & Engineering Directorate (ARL-HRED) in Orlando, Florida. He has 11 years of experience within U.S. Army and Navy acquisition, development, and research agencies. He holds a Masters and PhD degree in Computer Engineering with a focus on Intelligent Systems and Machine Learning from the University of Central Florida. His current research is in machine learning, active and semi-supervised learning, ITS architectures and cognitive architectures. He manages research in adaptive training, semi/fully automated user tools for adaptive training content, and architectural programs towards next-generation training.

UTILIZING INTELLIGENT TUTORING SYSTEMS AND THE GENERALIZED INTELLIGENT FRAMEWORK FOR TUTORING (GIFT) FOR RESEARCH

Anne M. Sinatra[(a)]

[(a)]U.S. Army Research Laboratory

[(a)]anne.m.sinatra.civ@mail.mil

ABSTRACT

This paper discusses using Intelligent Tutoring Systems (ITSs) and the Generalized Intelligent Framework for Tutoring (GIFT) to conduct research. The paper describes the functionalities and authoring tools in GIFT that can be utilized by an experimenter who wishes to create an experiment using GIFT. Lessons learned and suggestions are included. Further, there are examples and discussion of previous research that has used GIFT, and suggestions about improvements that could be made to use GIFT for research.

Keywords: intelligent tutoring systems, adaptive tutoring, research design, analysis

1. INTRODUCTION

Intelligent Tutoring Systems (ITSs) are adaptive computer based systems that can respond to an individual student with personalized materials based on performance, individual differences, and more. ITSs are generally designed by an instructional designer and/or course instructor who intends to use it as a supplement or primary means of providing information to students (Sottilare, Graesser, Hu, & Holden, 2013). While the utilization of adaptive tutoring has been shown to be as effective as a human tutor (VanLehn, 2011), it also has been shown to be a time-consuming and costly process (Sottilare, Brawner, Goldberg, & Holden, 2012). As a result, it is often difficult for an instructor or subject matter expert (SME) to sit down and construct their course without the help of individuals who have expertise with ITSs and computer programming. In fact, most ITSs are tightly coupled with the material that they are teaching, and are not necessarily editable or reusable in different domains. Domain-independent frameworks such as the Generalized Intelligent Framework for Tutoring (GIFT) aim to reduce the skill level and time required to author ITSs, and provide reusability (Sottilare, Brawner, Goldberg, & Holden, 2012). Specifically, one of the goals of GIFT is to provide an ability to conduct ITS research. While there have been a number of studies using ITSs in the classroom, and comparing ITS performance to non-ITS performance, there have not been as many studies about which characteristics or aspects an ITS should track and adapt on (Holden & Sinatra, 2014; Sinatra, Ososky, Sottilare, & Moss, 2017). One of the reasons for this is the lack of flexibility in the authoring of ITSs, and the lack of ability to easily manipulate the components of the ITS's learner, pedagogical, and domain models and modules. GIFT provides the ability to interchange and manipulate the elements of all of these modules, which allows for experiments to be conducted to determine what the most ideal elements of each module are for different educational domains.

Additionally, GIFT has been designed to be used to conduct research, export data, and assist in formatting the data for future analysis. The experiments that are run with GIFT do not necessarily have to be adaptive in nature, as multiple linear courses can be created as different experimental conditions in an experiment and data later exported and compared to each other (Sinatra, 2014; Sinatra, 2016). The current paper describes approaches that have been used for research in GIFT, provides suggestions on research that can be done using GIFT's authoring tools and setup, and provides future recommendations for features in GIFT that can assist in the research process.

2. CASE STUDY EXAMPLE OF A STUDY THAT HAVE BEEN CONDUCTED WITH GIFT, AND OTHER TYPES OF STUDIES

The following section describes a case study example of an experiment that was conducted with GIFT. Additionally, a general overview of other experiments and domains that have used GIFT for research are described.

2.1. Logic Grid Puzzle Tutorial Experiment

The Logic Grid Puzzle Tutorial experiment was one of the first set of studies that was conducted using the early versions of GIFT (GIFT version 2.5, and GIFT version 4.0). The Logic Grid Puzzle Tutorial study was a traditional psychology research experiment that was conducted using GIFT. The ultimate goals of the experiment were to examine the impact of different self-reference personalized instructional strategies on adaptive tutoring (Sinatra, Sottilare, & Sims, 2014; Sinatra, 2015). However, based on the features of GIFT at the time of the experiment, the adaptive manipulations occurred externally in PowerPoint using Visual Basic for Applications (VBA). The adaptive components to the course included feedback about the specific actions that the participant made during the PowerPoint based tutorial. However, the feedback did not change based on previous performance or on user characteristics.

GIFT was utilized to create an experimental flow that had three different conditions (self-reference, popular culture, and generic/baseline). The course included informational material, surveys/questionnaires, a tutorial that featured a manipulation during learning a skill (Logic Grid Puzzle performance), followed by assessments in the form of multiple choice questions, logic clue based questions, free recall, and performance of completing puzzles. The Q-sensor, which measures electrodermal activity (stress) was utilized and data was recorded using GIFT. The manipulation was in regard to the type of names that were present during the tutorial: names of the participants and people they knew, names of popular culture characters, and names that were generic. The study aimed to examine the impact of including names that are familiar or unfamiliar during the learning phase of a tutor to examine if they would assist in learning performance. See Figure 1 for a screen shot example of the popular culture name condition in which the names of characters were included in the learning materials, and Figure 2 for the generic names. Note that all of the content was identical except for the names that were present in both the puzzle grid, and the clues. The self-reference condition is not pictured, however, it asked participants to enter their own name, and names of friends, and they were incorporated into the puzzle. Individual difference characteristics such as Need for Cognition were collected, however, they were not adapted upon. One of the goals of the study was to examine the utility of the self-reference effect and different individual differences when teaching individuals a skill. Ultimately the results of this study, and other similar studies can be incorporated into GIFT in terms of types of adaptations that may occur based on relevant individual difference characteristics. For instance, if those with low need for cognition have better performance when popular culture names are used, that is a potential approach to adaptive content that can be given to a learner interacting with the system that has those characteristics.

One of the advantages of using GIFT for this study was the ease of data collection. The study was primarily linear with three different conditions. Participants were assigned to a condition when they arrived in person, and they were put in the corresponding version of the GIFT course. Everything about the courses except for the adaptive PowerPoint was identical. The software opened and closed surveys, PowerPoints, and materials as appropriate, which reduced the need for multiple research assistants. This experimental flow also ensured consistency, and allowed for having multiple participants engage with the study simultaneously on different computers.

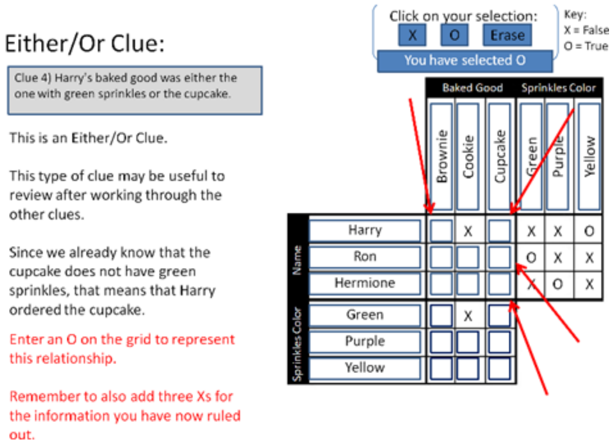


Figure 1. Example of the Popular Culture name adaptation in the Logic Grid Puzzle Tutorial Experiment.

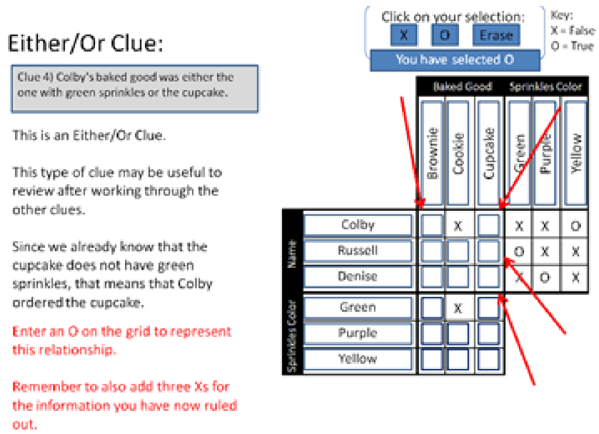


Figure 2. Example of the Generic Condition name adaptation in the Logic Grid Puzzle Tutorial Experiment. Notice that the names within the puzzle are different from those in Figure 1.

2.2. Other Studies that were conducted with GIFT

Other studies that have used GIFT include examining the location of feedback that comes from either inside or external to the environment (Goldberg & Cannon-Bowers, 2013). External programs such as VMedic (TC3 Sim) and VBS2 have been used to conduct affect research, as well as team based research (Paquette et al., 2016; Bonner et al., 2016). Specifically, a two person team tutor to evaluate surveillance performance was conducted using GIFT and VBS2 (Bonner et al., 2016). This study has demonstrated the ability to conduct team research using GIFT, as the two computers communicated with each other, and both team and individual feedback was provided. Additional work in the line of research is in progress that increases the number of team participants to three, and includes different roles/responsibilities for team members. GIFT has been used to conduct research in a psychomotor domain through using it for marksmanship training (Goldberg & Amburn, 2014). Additionally, GIFT has been used to assess the impact of using interactive (virtual reality sand table) vs. static means and displays to provide information to individuals (Boyce et al., 2016). Therefore, GIFT has been shown to be versatile in the types of domains that have been used for experiments. Further, it has been able to be incorporated in different virtual environments, as well as use sensors as an input for the system to adapt on. Many non-traditional educational domain and simulation based experiments can be conducted using GIFT.

GIFT has a gateway module that connects it with external training applications. The gateway module can be used to transmit data between GIFT and programs/sensors so that adaptations can occur. If an external program or sensor is of interest to a researcher, but not yet integrated with GIFT the developer's guide can be used to provide information about how to create new interop plugin to connect to GIFT. In this case there is the need for computer programming, however, if an already integrated program or sensor is used there is no need to engage in programming.

3. FEATURES OF GIFT THAT CAN BE UTILIZED TO CREATE RESEARCH EXPERIMENTS

GIFT has many tools that can be used to create experiments, and has a way of providing participants an opportunity to engage with an experiment online. When conducting a study there are a number of different course components or objects that a researcher may want to use. Among these are surveys/questionnaires, providing information to the participant, presenting slideshows, engaging in a dialogue with the participant, and engaging with an external computer-based environment. Additionally, it is important for the experimenter to be able to establish a controlled way in which to conduct research either in a laboratory or through an online approach.

The GIFT authoring tool has been carefully designed to not only provide the opportunity to create fully online ITSs, but also to be leveraged to conduct research. Among the researcher relevant items of the authoring tools are the ability to include actionable and non-actionable surveys that can be used to collect participant data, as well as conduct real-time assessments that can influence the flow of the GIFT course. Additionally, html, PDFs, as well as premade PowerPoint shows or Slide Shows can be incorporated into a GIFT course. Either a linear courseflow can be utilized for a traditional experiment, or an adaptive courseflow can be used to conduct research that involves testing different approaches to remediation in an ITS.

Once an experimenter has created their GIFT course that they would like to use for an experiment, they can create an experimental copy of it that will have an online link. The online link can be provided to participants who can then enter a participant number in a survey (or participate anonymously), and complete their interaction on their own computer. The data from the interaction will be recorded to GIFT's logs, and the researcher will have an opportunity to download and extract data from the logs after participation is complete. There are features in GIFT that also allow for the organization of data such that all data from an individual participant is exported onto one line so that it can easily be converted to excel or SPSS format for analysis.

3.1. Authoring Tools in GIFT

3.1.1. Development of the Authoring Tools over time

There have been many iterations of authoring tools in GIFT. While functionality is a high priority in regard to the development of GIFT, usability is also an extremely important element. In early versions of GIFT there were separate tools that were engaged with to create the courseflow (Course Authoring Tool), the learner configuration files, the surveys (Survey Authoring System), and more. The early tools emphasized functionality, and looked very similar to an .xml editor. The next step in the development of the authoring tools was to move toward a more straightforward and visual course authoring tool in order to design courseflow. Over time, the authoring tools not only became more visual and easier to populate (e.g., using drag and drop), but they also became more integrated with each other. In the current version of GIFT, the GIFT Authoring Tool includes the ability to not only create the courseflow, but to identify concepts, learner configuration, pedagogical configurations, adaptive courseflows, and create surveys. One of the main focuses in these developments is to make the process straightforward for a course author who is new to GIFT, as well as provide ease of use. See Figure 3 for an example of the early course authoring tools, and Figure 4 for the updated design. The same course (Logic Puzzle Tutorial) is being shown in both of the figures.

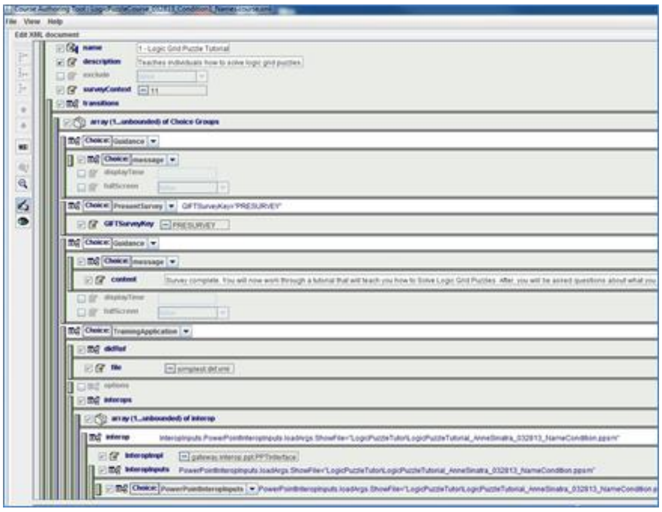


Figure 3. An example of an early version of the course authoring tool. Notice that while it is linear, it is not visual, and may look overwhelming to a new user.

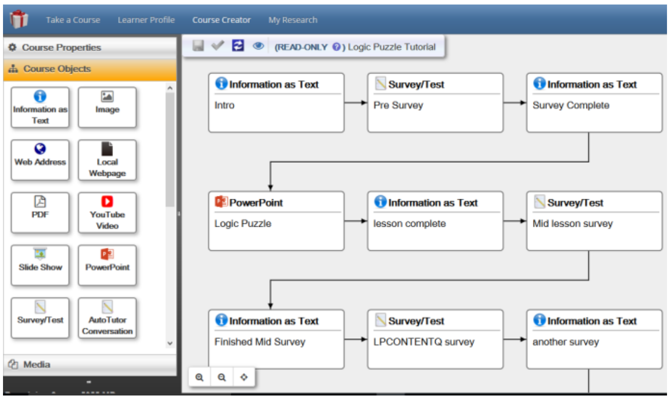


Figure 4. The current interface for GIFT Course Authoring. Note the visual interface as well as drag and drop functionality. The same course is loaded in this authoring tool as in Figure 3.

The changes and adjustments that have been made to the authoring tools have made it so that it is easier for an instructor or course author to engage with GIFT. One of the main goals for GIFT's design, is that a person does not have to have a background in computer programming to author in GIFT, which has become a reality. In the current state of the authoring tools a researcher in any domain (e.g., psychology, math, language), should be able to interact with the interface to create an experiment or course. Further, graduate students, or other students who wish to conduct their own traditional research experiments can utilize the tools without any external help needed in regard to programming.

3.1.2. An approach to getting started with the Authoring Tools

One recommendation for a user who is new to GIFT who wishes to author a course is to load one of the existing courses into the authoring tool and look at how it is constructed. Many of the course examples, including the Logic Grid Puzzle course which is available standard in GIFT, will demonstrate unique elements or flow to the courses that can help a new author gain an understanding of how GIFT works. Since GIFT is domain independent, if an author wanted to start by making a copy of an existing course and changing/replacing the content that is in it to be consistent with their desired topic, they could do so. It is also helpful to both engage with the original course by running it, then viewing the editor to see what both the original course looked like, and how it is different after changes have been made.

3.1.3. Functionality and distinction of the course objects in the Authoring Tools

As can be seen in Figure 4, the available course objects in the authoring tool include "Information as Text", "Image", "Web Address", "Local Webpage", "PDF", "YouTube Video", "Slide Show", "PowerPoint", "Survey/Test", as well as more, that are available through scrolling down the page. "Conversation Trees", "Question Bank", and "Adaptive Courseflow" objects are among the notable additional objects that can be included.

An experimenter who wishes to primarily engage in online surveys will need to make a decision between using the "Survey/Test" and "Question Bank" objects. As of GIFT 2017-1, these two objects are distinct and have different functionality. If a researcher wishes to use an established questionnaire or survey, then they would select the "Survey/Test" object. They would then make a selection between the type of survey (e.g.,non-actionable vs. actionable, etc.). For example, if demographics are simply being collected and looked at after the fact then a non-actionable survey may be appropriate. If a participant characteristic is being assessed, and the researcher wishes to adapt based on that characteristic later in the course, then one of the actionable types of surveys should be selected.

If a researcher wanted to generate a random set of questions from an overall question bank, then they would use the "Question Bank" course object. Using the "Question Bank" the researcher would add all of the questions that they wanted to be available, and they would tag them specifically with course concepts that they have defined within GIFT. For instance, if there were three different concepts that the questions covered this would be defined in course properties prior to authoring the course, and while authoring the question, the pre-defined concept can be selected for the question.

Once the question bank is established, the number of questions from each concept to present and the desired difficulty level is selected.

There are different functions and purposes the "Survey/Test", and "Question/Bank" tools can be used for. While the actual authoring interfaces are similar, it is important to understand the difference in the functionality between them and how they can be used in an experiment.

Another important distinction to make is between the "PowerPoint" and "Slide Show" objects. If the researcher desires to use VBA and adaptive functionality like in the Logic Grid Puzzle experiment example, then it is necessary for them to use the PowerPoint object. In the case of the PowerPoint object it will connect GIFT with the instance of PowerPoint that is on the participant's computer and will open and close it as needed. This can potentially result in a few challenges, as it requires the participant to have a license for PowerPoint, and in some cases their PowerPoint version may not be compatible with GIFT. In order for courses with this element to work in the Cloud version of GIFT, the gateway module of GIFT will need to connect to the instance of PowerPoint on the participant's computer, which requires a brief download process.

If an experimenter wants to present information to the user in a PowerPoint type format, but does not actually require interaction then they can use the "Slide Show" object. The "Slide Show" object will convert a PowerPoint show file (.pps) to images that are displayed with navigational arrows for the participant to use. One of the benefits of this is that it does not require the participant to have the external program installed on their computer, and it also does not require a brief download to connect GIFT and their computer before running the experiment. If the experimenter just wishes to present information to the participant then this may be the preferred method to use. Note, that GIFT requires, PowerPoint show files (.pps), which can be saved using PowerPoint.

Static images, PDFs, and other content can be inserted into the experimental courseflow if one wishes to display them to the participant. A local webpage can be used if the experimenter has constructed an .html file which they wish to display to the participant. In order to determine which might be most advantageous to use, the experimenter can look at previous courses to see how they were utilized, or use the course preview function in the Authoring Tool (it is a symbol that looks like an eye).

If one wishes to conduct adaptive training research, in which different adaptations are made based on learner characteristics and performance then the "Adaptive Courseflow" object can be used. This object is broken down into "Rules", "Examples", "Recall", and "Practice" based on Merrill's Component Display Theory (Merrill, 1983). The experimenter will enter instructional content for "Rules" and "Examples" in many possible forms (e.g., Slide Shows, images, web pages), and then they will select the number and types of questions they want to use from the question bank for the "Recall" component. The "Practice" component is optional, and currently requires the use of an external training application (e.g., VBS3). This adaptive courseflow may not be the preferred method for a researcher conducting a traditional linear experiment (e.g., psychology), however, it allows many possibilities for those who wish to conduct adaptive training research and examine the best pedagogical as well as individual differences strategies to use for adaptation.

There are many different features and approaches to displaying information to the participant in GIFT that may be of use to the researcher, but in some cases it may require examining the possible course objects and determining which would be the best approach to use for the goals.

4. EXAMPLES OF STUDIES THAT CAN BE CONDUCTED WITH GIFT

As mentioned earlier in the paper, there are a number of different approaches and types of studies that can be designed using GIFT. While traditional ITS studies tend to be in the area of Math and Physics, the majority of work that has been done to date with GIFT has not been in these areas. GIFT is domain-independent and very flexible. GIFT is not limited in regard to the types of questions that can be constructed, the methods of adaptation, or course objects that are used. GIFT has been designed such that different instructors or researchers that have different end goals can utilize it and the tools in order to meet their own goals. In particular this is one of the challenges in regard to the design of GIFT. It has to be powerful, but also needs to be constructed to be general, so that it can be used in different ways. In particular, this is one of the reasons that there are multiple approaches, as mentioned in the previous section, to presenting and collecting information. However, this flexibility is a great strength for the experimenter, as it allows for many different types of experiments to be conducted.

4.1. Linear Experiments

Linear experiments can easily be conducted with GIFT. The Logic Grid Puzzle experiment case study mentioned earlier in the paper was an example of a linear experiment. While the specific adaptive features of GIFT were not utilized, GIFT allowed for experimental control and facilitated data collection. Many times there can be human error when research assistants are required to open and close computer programs and interact with the participant. In the case of GIFT, the computer automatically leads the participant through the experiment. A research assistant can then be utilized mostly for startup and stopping of the experiment. Further, online experiments can be run that display information to participants, media (e.g., videos, images), and then collect survey responses. Traditional experiments, such as in the discipline of Psychology could use GIFT to streamline their data collection.

4.2. Adaptive Instruction and Intelligent Tutoring System Experiments

The main goal of GIFT is to facilitate the creation of ITSs. Therefore, GIFT can also be utilized to manipulate elements of an ITS to determine what types of adaptations have the best outcomes. In this case, the researcher may create competing types of media that they want to be presented to the participant and determine which is most effective in leading to the desired learning outcomes. In particular, different versions of a tutor can be created that contain different types of adaptations. For example, in the first version of the course, motivation level can be what the adaptation occurs based on. Whereas, in the second version of the course, it could be Need for Cognition. Or a potential third version of the course can examine the impact of adapting based on both Need for Cognition and Motivation. The participant's course outcomes can then be compared to each other. In this case, the research question would be in regard to the type of individual difference characteristics that should be taken into consideration. However, there are many other possibilities, as the learner module of GIFT and pedagogical module of GIFT are highly configurable and can be utilized in conducting ITS research. Therefore, GIFT can be used to conduct studies to determine the relevant components of the learner model, to determine adaptations based on individual differences, and for remediation based experiments. Conducting ITS research like these examples is generally difficult in domain-dependent systems as they are tied tightly to a concept or a way of adapting. With GIFT as long as the research question can be defined, the experimenter can utilize the tools in order to create the ideal configuration and experimental course.

5. RECOMMENDATIONS FOR FUTURE DIRECTIONS THAT COULD ASSIST WITH STUDY DEVELOPMENT

While the authoring tools for creating experiments in GIFT have been extensively worked on, there has not been as much focus on the data extraction tools. Early versions of GIFT used the Event Report Tool, which had boxes that could be checked by the experimenter after the fact to extract data from participant logs in the configuration that they desired. As GIFT moved toward the Cloud the My Research tab was created, and the ability to create an Experiment with GIFT became the ideal approach to data collection. These features are useful for experimenters, and over time they could be potentially improved as the authoring tool was.

As mentioned earlier in the paper, in current versions of GIFT, once a course is authored the experimenter clicks on the My Research tab and creates a new experiment. The experimenter will be prompted to enter an experiment name and description, then select the course that they want to make into an experiment. This makes a copy of the course as it currently is, and provides a web-link that the experimenter can provide to the participants so that they can engage with the experiment. If a change is made to the original course it will not populate into the experiment, a new experiment will need to be created.

Once the participant goes to the experimental link, they will encounter a "Start" button. After the button is pushed it will launch GIFT, and does not require a login. Therefore, if a participant number is desired to be collected, then there should be information and a question incorporated into a survey in the course to prompt the entry of the number.

Once data collection is completed, the experimenter finds their experiment in the My Research tab, and can build a report. By building a report the experiment is temporarily paused and no new data can be collected until it is started again. An interface pops up that allows the experimenter to select the information that they want to be included in the report, and how the data should be merged (e.g., all in one line). The experimenter can then save the report from all of their participants in a .csv file, which can be imported into Excel or SPSS for data analysis.

In the current version, there is no option for the experimenter to select specific logs that they want to extract, they are all extracted at the same time into the same file. This makes it difficult if there is an outlier or a data point that needs to be taken out of the data set because of an error. In the future it would be helpful to give the experimenter more control of the data logs that will be included in the output. However, it is possible to remove the data manually once it is extracted, and reconfigure it in Excel.

Further, it would be beneficial to have a participant number assigned at the beginning of the experiment, such that it is easier for the experimenter to manage and keep track of their data. This could be a potential option that is provided to the experimenter to include in the course, where a number is generated for the participant to continually use throughout the experiment.

Finally, while some of the features of the original desktop version of GIFT's extraction tools have been ported over to the new version, they are not all present. It would be beneficial to continue to add options for data extraction, while also being mindful of researcher needs and usability. The current version of the extraction tools do allow researchers to extract their data as needed, but small edits to them could potentially improve the functionality.

6. CONCUSIONS

GIFT is a powerful framework that is flexible and allows for researchers to conduct many different types of studies. In this paper a case study example was provided of a linear psychology experiment that was conducted using GIFT. Further, the authoring tools and different types of experiments that can be created with GIFT were discussed. GIFT is very flexible and can be utilized by experimenters from many backgrounds, with a variety of different research questions and goals. In the current form of GIFT, there are a few considerations to keep in mind, such as the most appropriate way to display information to the users, and collect data. Further, the experimenter should be mindful of the way that the My Research tab features work, and how to extract data. As with the authoring tools, the experimental process and support for data extraction tools will continue to improve based on user experiences and suggestions.

7. REFERENCES

Bonner, D., Gilbert, S., Dorneich, M. C., Winer, E., Sinatra, A. M., Slavina, A., ... & Holub, J. (2016, September). The Challenges of Building Intelligent Tutoring Systems for Teams. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 60, No. 1, pp. 1981-1985). SAGE Publications.

Boyce, M. W., Reyes, R. J., Cruz, D. E., Amburn, C. R., Goldberg, B., Moss, J. D., & Sottilare, R. A. (2016). Effect of Topography on Learning Military Tactics-Integration of Generalized Intelligent Framework for Tutoring (GIFT) and Augmented REality Sandtable (ARES) (No. ARL-TR-7792). ARMY RESEARCH LAB ABERDEEN PROVING GROUND MD ABERDEEN PROVING GROUND United States.

Goldberg, B., & Amburn, C. (2015, August). The application of GIFT in a psychomotor domain of instruction: a marksmanship use case. In *Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium (GIFTSym3)* (pp. 115-124).

Goldberg, B., & Cannon-Bowers, J. (2013, July). Experimentation with the generalized intelligent framework for tutoring (GIFT): a testbed use case. In *AIED 2013 Workshops Proceedings Volume* (Vol. 7, pp. 27-36).

Holden, H. K., Sinatra, A. M., (2013). The Need for Empirical Evaluation of Learner Model Elements. In Design Recommendations for Intelligent Tutoring Systems, Volume 1: Learner Modeling, 87.

Merrill, M. D. (1983). Component display theory. *Instructional-design theories and models: An overview of their current status*, *1*, 282-333.

Paquette, L., Rowe, J., Baker, R., Mott, B., Lester, J., DeFalco, J., ... & Georgoulas, V. (2016). Sensor-Free or Sensor-Full: A Comparison of data modalities in multi-channel affect detection. *International Educational Data Mining Society*.

Sinatra, A. M. (2014). The research psychologist's guide to GIFT. In *Proceedings of the Second Annual GIFT Users Symposium* (p. 86).

Sinatra, A. M. (2015). Developing an Experiment with GIFT: Then and Now. In *AIED Workshops*.

Sinatra, A. M. (2016). The Updated Research Psychologist's Guide to GIFT. In Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium (GIFTSym4) (p. 135).

Sinatra, A.M., Ososky, S., Sottilare, R., & Moss, J. (2017). Recommendations for use of Adaptive Tutoring Systems in the Classroom and in Educational Research. Accepted for publication in the Proceedings of the Human Computer Interaction International 2017 Conference, Vancouver, Canada, July 2017.

Sinatra, A. M., Sims, V. K., & Sottilare, R. A. (2014). *The Impact of Need for Cognition and Self-Reference on Tutoring a Deductive Reasoning Skill* (No. ARL-TR-6961). ARMY RESEARCH LAB ABERDEEN PROVING GROUND MD.

Sottilare, R. A., Brawner, K. W., Goldberg, B. S., & Holden, H. K. (2012). The generalized intelligent framework for tutoring (GIFT).

Sottilare, R.A., Graesser, A.C., Hu, X., & Holden, H. (2013). Preface. In Design Recommendations for Intelligent Tutoring Systems, Volume 1: Learner Modeling, i.

VanLehn, K .(2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197-221.

ACKNOWLEDGMENTS

The research described herein has been sponsored by the U.S. Army Research Laboratory. The statements and opinions expressed in this article do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred. Additionally, the author would like to acknowledge and thank the entire U.S. ARL Adaptive Training team, who have all been instrumental in the design and development of GIFT and the GIFT authoring tools.

THE COMPLIANCE OF PROJECT MANAGEMENT WITH CORPORATE'S RESEARCH AND DEVELOPMENT STRATEGIES: DTI'S VIRTUAL SHOOTING RANGE CASE STUDY

Chamnan Kumsap

Defense Technology Institute (Public Organization), Ministry of Defense,
Office of the Permanent Secretary of Defense, Ban Mai, Pak Kret, Nonthaburi, Thailand 11120

chamnan.k@dti.or.th

ABSTRACT

This original article was aimed to share with military simulation and training community how the project management of a military simulation and training project named '*Virtual Shooting Range*' or *VSR* was carried out in line with corporate's research and development strategies. The B.E. 2551 *Royal Decree of the Establishment of Defense Technology Institute (Public Organization)* or DTI was quoted to reflect the strategies that follow the objectives, roles and responsibilities defined in the Royal Decree. The strategies were elaborated each in turn to give an idea that research and development projects of DTI have to include activities in order to align with the strategies. The activities of the *VSR* project were discussed in terms of compliance with the explained strategies. It further contributes to the project management community by introducing a *Double Pyramid* approach for non-profit organizations where knowledge and innovations are highly expected from the public. One of the achievements of the project was measured by research grant conferred to *DTI* by the *National Research Council of Thailand* in order for *DTI* to extend products of the defense technology research to a commercialized stage. Byproducts of the project apart from deliverable systems were adopted as a key performance indicator to demonstrate how this defense research and development project was managed to build knowledge and innovation, develop collaborative and cooperative network, and develop a sustainable organization in line with the strategies. The paper illustrates the activities carried out with partners and sectors outside *DTI* that were central to the philosophy of *DTI*'s existence. It concludes that the *VSR*'s research and development activities were defined, designed, developed and delivered to comply with *DTI*'s research and development strategies that helped to partially improve the national competitiveness in building the potential and capabilities of hardware and software industry.

Keywords: virtual shooting range, corporate strategy, project management, defense research and development

1. INTRODUCTION

The scientific and technological contributions of defense and military research in Thailand have recently been directed towards a commercialization and production stage. The research and development institute such as *DTI* was established to realize and achieve the goal. In *DTI*, the research and development strategies were formulated to define missions for *DTI* researchers to utilize available resources to conduct research and development tasks. The *VSR* project was used as a case study to showcase its imperatives for project management including managing the explorative phase, managing the involvement of stakeholders in the project, and managing the project in relation to the strategizing process of the institute (Ben Mahmoud-Jouini et al, 2016). The project management was planned to contain not only within the *Iron Triangle* (Atkinson, 1999) but also collaborative and cooperative network of key stakeholders that Atkinson (1999) proposed for the framework to consider success criteria, the *Square Route*.

This article reports how the VSR project management was planned and conducted in a manner that aligned with institute strategies and had the public informed in various channels. It further contributes to the project management community by introducing the *Double Pyramid* approach for the planned research and development activities in response to public expectations provided that the activities were carried out within non-profit organizations where knowledge and innovations are highly expected from the public especially on the project of huge budget investment. Numerical investigation on the proposed equations, however, is needed further to validate the approach and equations. Due to the fact that the *VSR* project management created partnerships and research network that involved individuals, the descriptive and illustrative report of the article was limited to contents that were not to violate individual rights and not to expose corporate confidentiality agreement. It was assumed that the project management of research and development activities that complies with corporate strategies would benefit defense and military research projects, reinforce national defense and security measures, and improve the national competitiveness in a scientific and technological domain at an international stage.

2. DTI RESEARCH AND DEVELOPMENT STRATEGIES

2.1. Strategy 1: Research and Development of Defense Technology

The goal of this strategy is to improve the capability of the ministry of defense by means of research and development of advanced defense technologies to achieve the output of weaponry prototypes. It is envisioned that domestic industries will be able to support relevant logistics and indigenous human resources for the selected weaponry prototype. The research and development take into account the strategic context of the nation and the region. This strategy encompasses the research and development of core defense technologies and how to transfer knowledge and technologies to the industrial line of production.

2.1.1. Research and development of core defense technologies

DIT identified eight core technologies for its research and development. Four technologies including *Rocket & Missile, Military Simulation and Training and Information and Communication Technology* and *Unmanned Vehicle System* were approved by the *Defense Ministry Council* for research and development. Other four technologies wait for the approval from the council.

2.1.2 Transfer of knowledge and technologies to industrial line of production

The process to realize the research and development to defense industry is believed to take 10 steps as illustrated in Fig. 1. An initial phase includes Steps 1 to 4 where an analysis of user requirements, feasibility studies for master plan, council approval for the master plan, and proposal for project and budget approval are conducted. Steps 5 to 7 are research and development that involve the interface with the armed forces in a regular basis to keep users in the loop.

2.2. Strategy 2: Development of Knowledge and Innovation to the Public

The goal of this strategy is to develop, to store, and to acquire defense technology knowledge in a continuous and sustainable manner. The knowledge is to be transferred to the public for exploitation via academic, commercial, and industrial sectors.

2.2.1. Development of defense science and technology knowledge

There are two different approaches of developing defense science and technology knowledge. The first one is conducted in-house by DTI researchers in form of basic research and applied research. The second approach is achieved by either subsystem procurement from industrial partners or research engagement with academic sectors. The academic research engagement is managed by hiring professors in relevant fields of expertise to build subsystems.

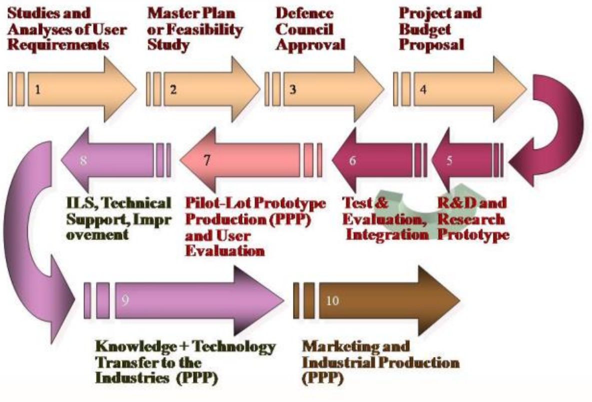


Figure 1: Steps from defense research and development toward defense industry

2.2.2. Human resource development of defense science and technology

Human resource development encompasses academic research engagement and direct human resource development in short training courses. Grants are provided for a research proposal in relation to DTI's core technologies. The master degree of defense engineering is also provided for individuals seeking higher education.

2.2.3. Management of defense science and technology to the public

The knowledge as the result of research and development is managed publically in two ways. The first one is in form of knowledge management prior to publication. Target users for this knowledge management are industrial and commercial sectors who have an agreement with DTI and intend to commercialize or invest in the project's line of production. The second way of knowledge management is in form of research papers for national or international journal publication, national or international conference proceedings, research and technical report, classified pieces of work procedures and instructions, and other writings for magazine or internet publications. National and international recognition and acceptance are built upon this knowledge management.

2.3. Strategy 3: Development of Collaborative and Cooperative Network

The goal of this strategy is for the *Ministry of Defense* and the country to manage and make utmost use of the knowledge and resources of defense technology from formally established network in an effective and efficient manner.

2.3.1. Development of common policy and vision

Agreement is in form of contract after fair and transparent bidding that can eventually lead to signed memorandum of understanding that defines in general the common policy and vision of the two parties toward the specific research and development project. Where domestic corporate and private sectors lack specific

ISBN 978-88-97999-90-4; Bruzzone, Cayirci and Sottilare Eds.

capacity for a given project, foreign partners are invited for technology demonstration via formal invitation or letters with the attention made to the *DTI* director. Non-disclosure agreement is basically reached before further agreement or commitment.

2.3.2. Development of cooperative and participative mechanism

A mechanism that key stakeholders within the ministry of defense can cooperate is developed at the initial stage of drafting the core technologies' master plan. That is to ensure the involvement of users from the *Royal Thai Armed Forces* to define projects for *DTI* researchers to carry out, to identify what military units to validate the product by means of test and evaluation, and to input military doctrine that is central to the design phase of the defense research and development.

2.4. Strategy 4: Development of Sustainable Organization

The goal of this strategy is for *DTI* administration to follow good governance policy of the government, for *DTI* researchers to perform efficient research and development and to attain excellence in core defense technologies, for *DTI*'s transparent logistics and support, and for *DTI*'s self-learning environment and sufficient research and development infrastructure.

2.4.1. Development of financial affairs

In the *Royal Decree of DTI's Establishment*, the Article 2 Code 8 (6) states that *DTI* has roles and responsibilities for charges from fee, maintenance fee, remuneration or services under regulated criteria and rate. *DTI* has ordered the regulation to support incomes of the charges without having to return them to the Ministry of Finance under the *Article 2 Code 10*.

2.4.2. Responsive development to user satisfaction

Apart from the initial stage of the research and development that user requirements are included in conceptual and detailed design. Logistics and schedules for test and evaluation are attached to the service plan to keep track on functionality and performance of the delivered system. Contact data of suppliers responsible for integrated subsystems is stored for later logistical services and maintenances.

2.4.3. Good governance for internal affairs

DTI is subject to regular performance assessment from a lawfully registered body that is deemed expert in the field. The result of the assessment is liable to reveal evident facts in terms of the effectiveness and achievements of corporate development and others specified by *DTI* board of directors.

2.4.4. Development through learning

The *Royal Decree Article 1 Codes 7 (3), (4),* and *(5)* require that *DTI* cooperate with other state agencies, academic sectors and domestic and international private sectors in defense technology. In addition, *DTI* is required to support trainings, to encourage research, to develop human resources in defense technology, to act as the center for the service of defense technology data and information, and to facilitate academic activities for the purpose of the dissemination of defense technology knowledge.

3. THE COMPLIANCE OF *VIRTUAL SHOOTING RANGE*

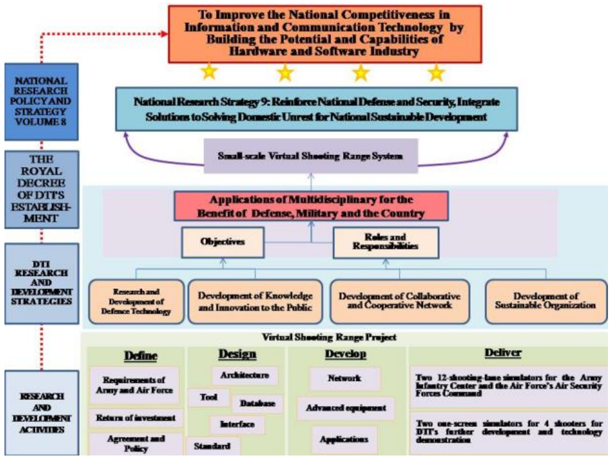


Figure 2: The compliance of *DTI's VSR* and *National Research Policy and Strategy*

3.1. Overview of the Project

The *VSR* project was the first one under the *Military Simulation and Training Master Plan* (see Fig. 2). The research and development activities were planned and managed to correspond to the institute strategies that follow closely the objectives, roles and responsibilities of *DTI*. Within two fiscal year *time*frame and a little over 1 million US dollar *cost*, the project was required to arrive at four training simulators that were ready for *delivery* to the *Army's Infantry Center* and the *Air Force's Air Security Forces Command*. Further validation from both units was needed to complete the prototypes with desired *quality* and acceptable standard of test and evaluation. The project's two fiscal year duration spanned from October 2013 to September 2015. The output was expected to be economically viable due to the rationale that the prototypes were to be commercialized, in line of production, and partially compensating foreign import.

3.2. Research and Development Activities

The research methodology was viewed from a 4-D (*define*, *design*, *develop*, *deliver*) management perspective. Workshops and seminars were held to collect users' requirements prior to the launch of the project. Fig. 3 above shows one of the workshops under the title '*Virtual World and Training Scenarios: The Essence of Newly Advanced Technology*'. It was held at the Ramar Garden Hotel, Bangkok, Thailand during 17-18 September 2013, as a forum that facilitated exchanges and sharing of research and development activities from the *Air Force's Air Security Forces Command*, *DTI* researchers, and Thailand's university academia. The activity responded to all the strategies described in Sections 2.1, 2.2, 2.3, and 2.4.

3.2.1. Defining user requirements

Figure 3: The workshop that corresponded to all *DTI's research and development strategies*

3.2.2. Designing system architecture

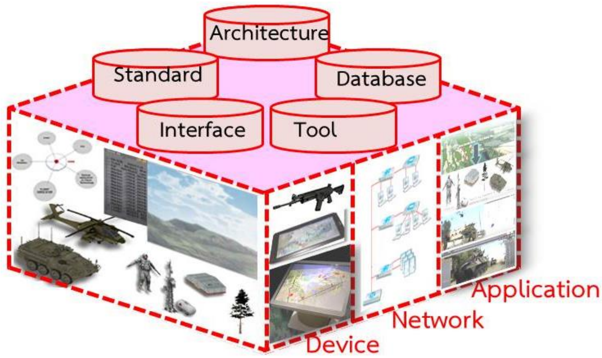


Figure 4: Five pillars of the projects of *DTI's Military Simulation and Training* Master Plan

The *Military Simulation and Training Master Plan* requires projects under it share the same pillars including *Architecture, Standard, Database, Interface, and Tool* (Kumsap and Meegla, 2013). The ultimate goal of this concept was to save cost of investment whereby each project will be able to share the existing infrastructure as shown in Fig. 4. Advanced and developed devices are the result of research and development activities that will house applications that match users' requirements, and to be networked with other existing simulators through shared and integrated *Architecture* (Kumsap et al, (2013) and Chieochan et al (2015)).

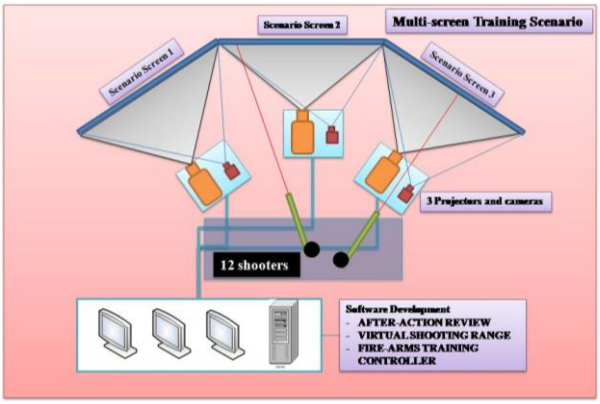


Figure 5: Conceptual Design of *VSR*

The *VSR* was conceptually designed as shown on Fig. 5 to interface main hardware devices including three scenario screens, three sets of cameras and projectors, twelve BB Guns for twelve shooters, three image generators, and one computer for instructor operating system. The software development included *After-Action Review, Virtual Shooting Range, Firearms Training Controller*. The activity mainly responded to the strategy described in Sections 2.1 that focused on the research and development of core defense technology.

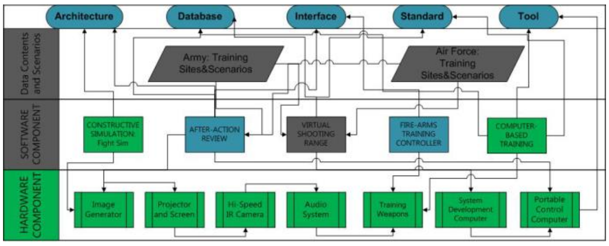


Figure 6: Detailed Design of *VSR*

The project's activities detailed in Fig. 6 are in three colors and geometric shapes. The green color represents collaboration in terms of purchase from and co-research with private sectors. The dark grey represents close cooperation with the *Army's Infantry Center* and the *Air Force's Air Security Forces Command*. The blue ones represent in-house research and development within *DTI*. The blue rounded rectangles are used for representing the research and development of military simulation and training technology since they resulted in establishing the foundation of the master plan. The green predefined process boxes ae used for representing the development of collaborative and cooperative network with private and public sectors. The green rectangles are used for representing the development of knowledge and innovation to the public. The dark grey parallelogram goes deeper to the development of cooperative and participative mechanism in which the *Army* and *Air Force* armed units that were central to the activities.

3.2.3. Developing system

The five pillars of the *DTI's Military Simulation and Training* Master Plan were utilized as selection criteria when choosing hardware devices for a system prototype. The standard of all the devices were strictly followed to ensure a seamless interface during system integration stage, see Fig. 7. A *Gun Recoil System* was vital to realistic firearms training, thereby being initiated as another research and development project in *DTI*, and more importantly responding to the Strategies of Section 2.1 and 2.2.

In order to network distributed simulators as *VPN Spoke #1* at the *Air Force's Air Security Forces Command* in *Don Mueang*, Bangkok (see Fig. 8) and as VPN Spoke #2 at the *Army's Infantry Center* in *Prachuab Khirikhan* province, there was a requirement to establish seamlessly networked communications. That was to be a good infrastructure for distributed simulation intended by the *Master Plan*. For data communication security,

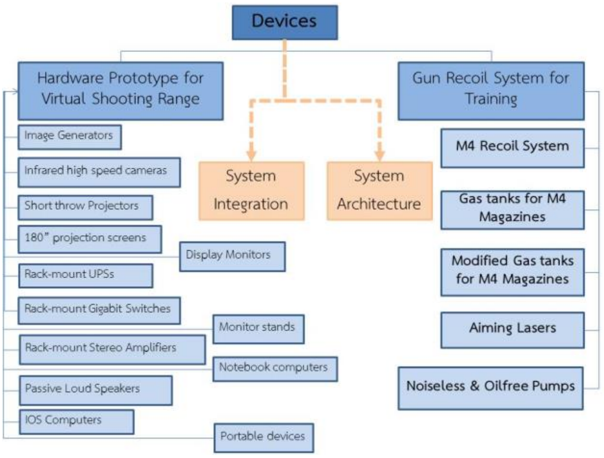


Figure 7: The importance of standard of devices for seamless system integration

the *Virtual Private Network* (*VPN*) was adopted as a central network under standard network-level encryption. The *VPN Hub* was at DTI with access to *Internet Cloud* and server for system monitoring. The VPN Spoke *#1 and #2* mentioned earlier were installed with access to *Internet Cloud* via 3G/4G.

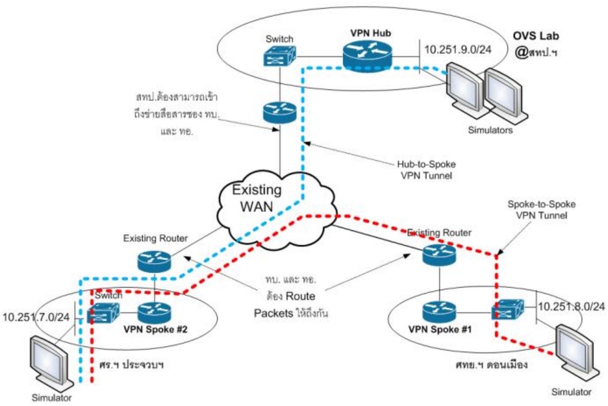


Figure 8: Dynamic multipoint *Virtual Private Network* (*VPN*) for distributed simulation over the Internet

Applications that run on the network include *training scenarios, after action review and computer – based training* illustrated in Fig. 9. The *instructor operation system* was each installed at the three nodes. The *training scenario development* answered to the Strategy in Section 2.1 on the research and development of *Military Simulation and Training* technology with the development of collaborative and cooperative network among DTI, academic sectors, the *Army's Infantry Center* and the *Air Force's Air Security Forces Command*. The *instructor operation system* and *after action review* component were complete in-house capability built upon the standard purchased tools with some input from the users. The *computer – based training* component was a collaborative work between DTI and Thailand's public university that signed the MoU with DTI. Military officers from the *Army* and the *Air Force* were source of system validation and user satisfaction.

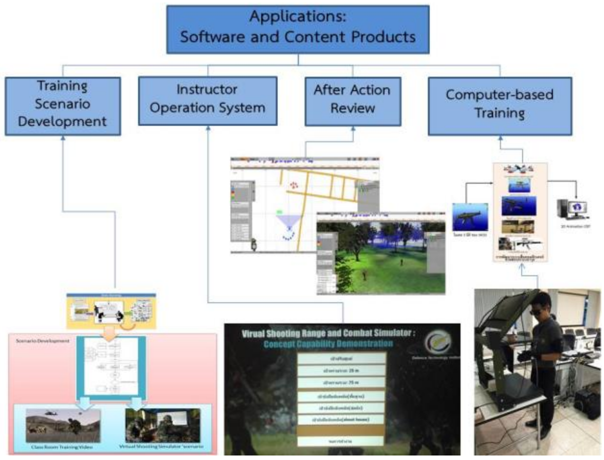


Figure 9: Applications of the *VSR* on the *Virtual Private Network* (*VPN*)

3.2.4. Delivering project output

Figure 10: *VSR* installed for test, experiment and use in the *Army's Infantry Center* (upper left and middle right) and the *Air Force's Air Security Forces Command* (middle left and lower)

The complete system of software installation, hardware implementation and integration as conceptually designed in Fig. 10 was delivered to the *Army* and the *Air Force* at the end of fiscal year 2015.

For the validation purpose, the delivered systems were validated in three categories namely system test of evaluation, software validation at the installation sites and *ISO/IEC 29110* at *DTI* (see Fig. 12). The first two were aimed at testing system's seamless integration, functionalities and response to users' requirements in terms of training effectiveness and representation realism and consequently responded to Section 2.4.2. The complete manual of the project validation of the *VSR* was reported in Haddawy (2016). The third one was to ensure that the research and development team in *DTI* complied with the *ISO/IEC 29110* in project management software implementation and the project team has recently filed for the *ISO/IEC 29110* certification.

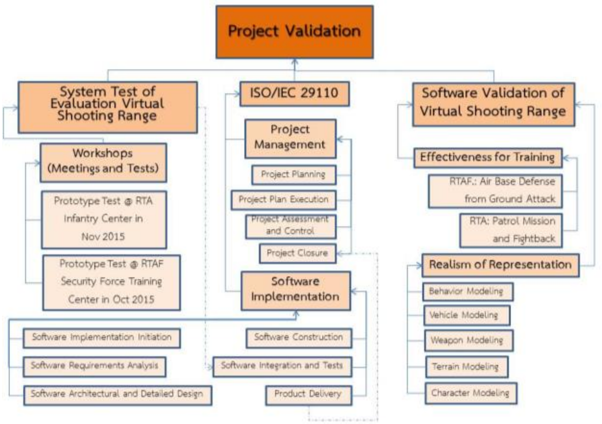


Figure 11: *VSR* project validation

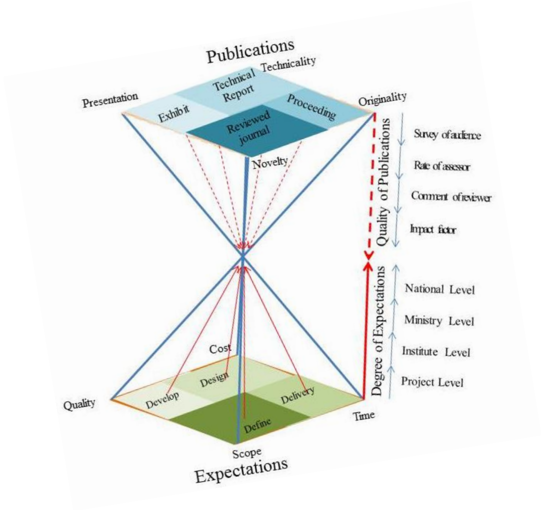


Figure 12: The *Double Pyramid* approach to cope with expectations for non-profit organizations

Two one-screen simulators for four shooters were installed with one system in *DTI* for further research and development and the other one was used for technology demonstration in related defense exhibitions (Fig. 13). This simplified and economical system architecture received research grant from the *National Research Council of Thailand* (*NRCT*) in order for *DTI* to extend products of the defense technology research to a commercialized stage. The architecture was designed to get a concise, yet, effective portable system that made it easier for only one person to carry, set up and operate and then called the *economy scale architecture* of the *VSR*. This research grant proved that the *VSR* complied with the strategy in Section 2.4 whereby the grant helped *DTI* to develop financial affairs of generating incomes and to continue developing the *VSR* according to an objective of the grant.

4. PROMOTING PUBLICATIONS TO COPE WITH EXPECTATIONS

Since the cost, time and scope of the *VSR* were fixed by allocated government budget, allowed fiscal years and signed agreements with the *Army* and the *Air Force*, the project quality was managed to keep the diamond in shape. The project was defined to respond to requirements of the *Army* and the *Air Force*, designed to stay within the budget allocation, developed with technologies that brought innovations, and delivered in time. The project needed to live up with strategically gradual expectations towards the tip of the *Double Pyramid* (see Fig. 12) where the improvement of national competitiveness (top of Fig. 2) in science and technology was expected by the public. To keep the equidistant diamond shape, the volumetric *Expectations* should be:

$$Exp = \frac{(Define+Design+Develop+Deliver)x\ DoE}{3} \qquad (1)$$

where *Exp* = Volumetric *Expectations* at 33.33 maximum while performed *Define*, *Design*, *Develop*, and *Deliver* are in percentage of success as the project ends, and *DoE* = Degree of Expectations from four levels which is $\leq \frac{1}{(national+ministry+institute+project)}$, whereby *national* = *planned_national/performed_national*, *ministry* = *planned_ministry/performed_ministry*, *institute* = *planned_institute/performed_institute*, *project* = *planned_ project/performed_ project*, and each level can reach the maximum value of 1 where planned equals performed.

It can be seen from Fig. 12 that (1) can arrive at 33.33 units when each variable to the right hand side reaches their maximum value. The upside-down pyramid of Fig. 12 illustrates various forms of *Publications* that include *Exhibition*, *Technical Report*, *Conference Proceeding*, and *Reviewed Journal* shaped respectively by *Presentation*, *Technicality*, *Originality*, and *Novelty*. The project could live up to *Expectations* of the public and complied with the *Strategy 2* provided that the *Publications* were made equal to *Expectations.* Each publication must include in their content *Presentation*, *Technicality*, *Novelty*, and *Originality* to achieve high *Quality of Publications*. Just like the equation (1), the volumetric *Publications* should be:

$$Pub = \frac{(Exhibition+TechReport+Proceeding+Journal)\ x\ QoP}{3} \qquad (2)$$

Where *Pub* = Volumetric *Publications* also at 33.33 maximum while measured *Exhibition*, *Technical Report*, *Proceeding*, and *Journal* are in percentage of achievement as the project ends, and *QoP* = *Quality o Publications* which is $\leq \frac{1}{(survey+rate+comment+impact)}$, whereby each *Publications* result can reach the maximum value of 1.

5. DISCUSSIONS

5.1. The Royal Decree of *DTI*'s Establishment

The compliance of the *VSR* with the *DTI*'s research and development strategies was evidently illustrated in Section 3. Objectives of the DTI establishment were fulfilled by the delivery of the *VSR* systems to the *Army*, *Air Force* and *DTI* and other collaborative and cooperative network built from within the project management. The project answered to roles and responsibilities of DTI in which it generated the MoU between DTI and the *Land Development Department* of the *Ministry of Agriculture and Cooperatives*, the MoU between DTI and the Department of Special Case Investigation (DSI) of the *Ministry of Justice*, and the MoU between DTI and *Mahidol University* of *Ministry of University Affairs*.

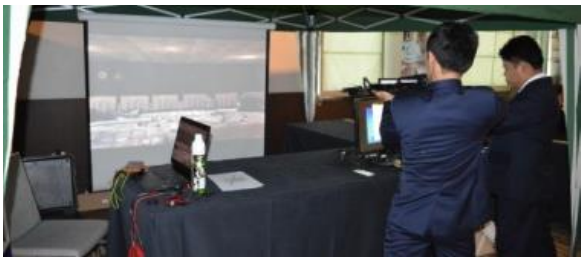


Figure 13: One-screen *VSR* simulator in technology demonstration

5.2. National Research Policy and Strategy

The *VSR* has mirrored Thailand's National research policy and strategy Volume 8, Strategy 9 that it reinforced national defense and security whereby providing the *fire-arms training simulator* that placed the users at the center of the research and development. It also integrated solutions as training scenarios (see the *Software Validation* part in *Project Validation* of Fig. 10) to solving domestic unrest such as *air base defense from ground attack* and *patrol mission and fightback* scenarios. The scenarios were an add-on feature of the *small scale architecture* or one-screen of the *VSR*.

5.3. Living up to Expectations

The author proposed the *Double Pyramid* approach that needs further numerical validation using the formulated (1) and (2). In (1), the variables planned_institute, performed_institute, planned_ project and performed_ project are easier to measure while the rest are cumbersome but achievable with focus group or questionnaire. All the variables of (2) are measurable and within the planned and achieved Publications. At the time of writing this report, the author is investigating more on the proposed approach.

5.4. Improvement of National Competitiveness

The *VSR* was successful with the improvement of national competitiveness when its extended version to the *small scale VSR* was presented at the *MilSim Asia* during 17-18 January 2017. *DTI* had gained international reputation since *DTI* researchers of this project submitted their research and development concepts and findings leading to and resulted from this project in several international conferences (Kumsap et al (2013), Chalainanont et al (2013), Kumsap et al (2014), Chieochan et al (2015), Tanvilaipong et al (2015), and Tepkhunchorn et al (2015)). Most importantly, the work by Robert et al (2016) with the use of *DTI*'s *tool* and *database* received the *Best Paper Award* from *International Defense and Homeland Security Simulation Workshop, September 26-28 2016, Cyprus*.

In addition, one of the research byproducts as fully explained in Chalainanont et al (2013) was developed under the realm of information and communication technology that was used in field survey and ground truth missions in the *VSR* project was used in the *DSI Map Extended* under the MoU between *DTI* and *DSI*. The researchers managed to build potential software from its extension to the capability as discussed by Kingkangwan et al (2015). *Computer - based Training* (see Fig. 9) and *Virtual Gun Assembly* (lower right inset of Fig. 9) were also identified by the *NRCT* as having the hardware potential and software capabilities to improve Thailand's hardware and software industry and advised to report for research grant from the *NRCT*.

6. CONCLUSIONS AND RECOMMENDATIONS

This original article shares with military simulation and training community the *VSR* project of *DTI*'s *military simulation and training Master Plan* as the case study to reflect its compliance with the explained strategies. The research and development activities were managed in order for the project's researchers to *define* the user requirements, *design* the system architecture, *develop* the system, and *deliver* the project output. It further contributed to the project management community by introducing the *Double Pyramid* approach that values Publications against Expectations. Numerical investigation on the proposed equations, however, is needed further to validate the approach and equations. Additionally, it discussed that the *VSR* had fulfilled the objectives of the *DTI Establishment* by the delivery of the systems to the users, by the creation of research partnerships through MoU, by the reinforcement of national defense and security with the training simulator, by the integration of hardware and software solutions to solving the domestic unrest, and by the improvement of Thailand's scientific and technological recognition in the international showcases.

DTI has missions, vision and core values to define its strategies that rule the objectives, roles and responsibilities entitled by the *Royal Decree of DTI's Establishment*. This article shows how the *VSR* project case study was managed to comply with the policy of the institute up to the national level. However, the *Royal Decree Article 1 Codes 7 (1)* also requires *DTI* to execute missions relevant or related to defense technology development. Research and development prototypes should be financially supported for the same actions as the research grant provide by the *NRCT* so that the potential prototype can be commercialized and put in line of production. Within the introduced *Double Pyramid* framework, the author is validating the viability of the approach and expecting is publication soon. However, it is recommended that further research on commercialization and manufacturing topics in the context of Thailand's defense technology research prototype should be carried out. Evidence from the *VSR* project conveys the message that the research and development project that is widely accepted either by national or international domains is feasible for upgrading to the commercialization and manufacturing phase. Last but not least, byproducts of the project can be extended and enhanced as a research project by the agreement of the signed memorandum of understanding.

REFERENCES

Atkinson, R., (1999) Project management: cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria, International Journal of Project Management, Volume 17, Issue 6, pp. 337–342.

Ben Mahmoud-Jouini, S., Midler, C. and Silberzahn, P. (2016), Contributions of Design Thinking to Project Management in an Innovation Context. Project Management Journal, 47: 144–156.

Chalainanont, N., Kumsap, C., Boonprasert, P. (2013) An Android-based Interface of Mobile Network for Real-time Field Survey and Ground Truth Missions, *Proceedings of the 34th Asian Conference on Remote Sensing 2013*, pp. 4193 – 4200.

Chieochan, S., Pratoomma, P., and Kumsap, C., (2015), Distributed Virtual Environments for Military Training Applications: Trends and Challenges, Proceedings of The First Asian Conference on Defense Technology, ACDT2015, April 23 - 25, 2015, Hua Hin, Thailand, pp. 27-32.

Defense Technology Road Map 2010 - 2024, Defense Technology Institute (Public Organization), 24 p. [*In Thai*]

Haddawy, P., (2016), Manual of Virtual Shooting Range Validation for the Virtual Shooting Range project, Information and Communication Technology, Mahidol University, 23 August 2016, 57 p.

Kingkangwan, K., Chalainanont, N. and Kumsap, C., (2015), Triple-Mode Location-Aware Andrioid Portable System for Field Operations, Proceedings of The First Asian Conference on Defense Technology, ACDT2015, April 23 - 25, 2015, Hua Hin, Thailand, pp. 63-66.

Kumsap, C., Chalainanont, N., Sontayamal, T. (2013) Using High Level Architecture for an Information Exchange: A DTI's Integrated Approach of Modeling and Simulation, Proceedings of the 34th Asian Conference on Remote Sensing 2013, pp. 4278 – 4284.

Kumsap, C., Chalainanont, N., Meekhla, W., (2014), An Approach of Geo-computation for Automatic Placement of 3D Building Models, Unpaginated CD-ROM, 10th Asia GIS 2014 International Conference, Chiang Mai, Thailand, 16-17 June 2014.

Kumsap, C., Meegla, W. (2013) Basic Components of DTI's Simulation and Training, *National Defense Studies Institute Journal*, pp. 64 – 70. (In Thai)

Kumsap, c. (2017), Speaker Interview for the topic of development and application of NRCT's Economy Virtual Shooting Range. http://www.milsimasia.com/page/interview-with-gp-capt-dr-chamnan-kumsap/ [Accessed 19 February 2017].

Memorandum of Understanding between Defense Technology Institute (Public Organization) and Land Development Department for the terms and understanding between the two parties to utilize digital map and mapping data of Land Development Department on defense technology signed on 18 May 2013. [*In Thai*]

Memorandum of Understanding between Defense Technology Institute (Public Organization) and Department of Special Case Investigation for the terms and understanding between the two parties to research and develop defense technologies for special case investigation signed on 23 July 2015. [*In Thai*]

Memorandum of Understanding between Defense Technology Institute (Public Organization) and Mahidol University for the terms and understanding between the two parties to academic collaboration and research knowledge and human resources development on defense science and technologies signed on 8 October 2014. [*In Thai*]

National Research Policy and Strategy volume 8 (2012 - 2016), the National Research Council of Thailand, 90 p.

Robert, O. P., Kumsap, C. and Janpengpen, A., (2016) Modeling drug trafficking routes using potential surface analysis for the simulation of military training scenarios, Proceedings of The 6th International Defense and Homeland Security Simulation Workshop, September 26-28 2016, Cyprus, pp. 35-43.

Royal Decree of the Establishment of Defense Technology Institute (Public Organization) on 27 December 2008, 73 p. [*In Thai*]

Tanvilaipong, N., Chalainanont, N., and Kumsap, C., (2015), Database Management to Support Automatic Placement of 3D Building Models to A Simulated Scene, Proceedings of The First Asian Conference on Defense Technology, ACDT2015, April 23 - 25, 2015, Hua Hin, Thailand, pp. 172-176.

Tepkhunchorn, K., Khantong, R., and Kumsap, C., (2015), Modeling and Simulation as an approach to create Awareness of Thailand Aerospace Technology Ambition, Accepted for Publication in the 2015 International Conference on Modelling, Simulation and Applied Mathematics (MSAM2015), August 23-24, 2015, Phuket, Thailand.

AUTHOR'S BIOGRAPHY

Chamnan Kumsap used to work in the Royal Thai Air Force until the rank of Group Captain and recently worked as a project director of the Virtual Shooting Range project at *Defense Technology Institute (Public Organization)*. He received the Ph.D. degree in Remote Sensing and *GIS* in 2005. His research interests include military simulation and training, modeling and simulation, *GIS*, terrain modeling, *UAV*-based terrain modeling. His e-mail address is: chamnan.k@dti.or.th.

TASK ALLOCATION AND AUTOMATION LEVEL IN THE DESIGN OF INTELLIGENT TUTORING SYSTEM AUTHORING TOOLS

Scott Ososky[(a)], Robert A. Sottilare[(b)], & Keith Brawner[(c)]

[(a),(b), (c)] U.S. Army Research Laboratory
[(a)] Oak Ridge Associated Universities

[(a)] scott.j.ososky.ctr@mail.mil, [(b)] robert.a.sottilare.civ@mail.mil, [(c)] keith.w.brawner.civ@mail.mil

ABSTRACT
Intelligent tutoring system (ITS) *authoring tools* are productivity applications that aid humans in developing adaptive tutors. Authoring tools have the potential to reduce the skill, resources, and effort required to produce a tutor, thereby enabling ITS authoring for individuals without programming or instructional design experience. However ITSs continue to evolve and authoring tool designs iterate to keep pace. Authoring tool design has been previously framed as a tradeoff between usability, depth, and flexibility. This current paper, however, takes a different look at the design of authoring tools, through the lens of function allocation and levels of automation. The evolution of authoring tools within the Generalized Intelligent Framework for Tutoring (GIFT) will be used as an example of how tasks can be allocated (and automated) between humans and software. The paper will discuss lessons learned, implications for system design, and considerations regarding human mental models of adaptive tutor authoring processes.

Keywords: intelligent tutoring systems, authoring, automation, function allocation

1. INTRODUCTION

Intelligent tutoring systems (ITSs), or adaptive tutors, are learning systems that can collect data about a learner through assessments, reports, and sensors, in order to select and present optimal instructional content based on a unique learner profile. ITSs have been demonstrated to be more effective than one-to-many instruction (e.g., classroom instruction), approaching the effectiveness of one-to-one human tutoring (VanLehn, 2011). However, ITSs have not been widely adopted in educational settings (Bill & Melinda Gates Foundation, 2015; Murray, 2004) or military training environments (Sottilare, Graesser, Hu, & Holden, 2013). There are many factors that inhibit the adoption of ITSs within these contexts. A lack of usable and accessible authoring tools for ITSs is among those factors.

Much has been written on the topic of theory - and engineering-based efforts to develop authoring tools that provide to potential authors the functions necessary to create tutors without computer science or instructional design knowledge (Aleven & Sewall, 2010; Mitrovic et al., 2009; Olsen, Belenky, Aleven, & Rummel, 2013; Suraweera, Mitrovic, & Martin, 2010). Murray, specifically, has published numerous works on authoring tools over the past two decades. His work includes analysis of the problem (or opportunity) in which design tradeoffs are made in authoring tools between usability, depth, and flexibility (Murray, 1996, 2014). In summary, increasing the power of the authoring tools (i.e., depth), the applicability of the tools to different domains and problem spaces (i.e., flexibility), or the usability of the tools themselves (i.e., learnability, productivity), comes at a cost to one or both of the other two (Murray, 2004).

This paper does not endeavor to duplicate the effort of those prior works, rather to explore the design of ITS authoring tools in a slightly different way. Specifically, this paper will delve into the following three questions: a) how should ITS authoring tasks be delegated between humans and software? b) For tasks that are delegated to software, what level of task automation is currently appropriate given the average user's current understanding of ITSs? c) How might automated tutor authoring evolve over the long term? There may not be a clear answers to these questions. Adaptive tutor authoring is a relatively new productivity paradigm. It would be useful to be able to educate current and potential authors about aspects of ITSs and the authoring process in order to build robust mental models of the authoring process and ITSs, in general. Simultaneously, the relative benefit of tutor authoring tools in terms of effectiveness, efficiency, and cost must also be considered. As such, specific efforts in automating parts of the ITS authoring process will also be presented. The current paper will use the authoring tools associated with the Generalized Intelligent Framework for Tutoring (GIFT) as a case study in balancing knowledge-building with automation.

2. BACKGROUND

2.1. Mental Models

Murray (2014) explained that authoring tools should help users build accurate mental models of the ITS building blocks, configurations, and workflow afforded by the authoring tool. This is inherently difficult, because ITSs are evolving, and each ITS will differ in some ways from others. However, mental model theory can provide guidance to approaching this interaction problem en route to an *accurate* mental model for the ITS authoring process. Rouse and Morris (1986) described mental models as "mechanisms whereby humans are able to generate descriptions of purpose and form, explanations of system functioning and observed system states, and predictions of future states" (p. 7). Mental models influence users' expectations regarding a system's functionality and guide user interaction behavior (Ososky, 2013). Human mental models do not need to be complete or even accurate to be applied to a specific system interaction (Norman, 1986). It is important that human authors understand what ITSs are capable of in order to fully realize the potential of adaptive tutors. Thus, it may be necessary to delegate certain ITS authoring tasks to humans in service of cultivating accurate mental models, even if the computer could do the task more efficiently. Therefore, it would be appropriate to dynamically allocate tasks along various levels of automation, supporting human authors of various skill levels and experiences.

2.2. Task allocation

Task (or function) allocation is the practice of determining which tasks should be delegated to a human, a machine / computer, or some combination of the two. Generally, decisions in task allocation are informed by the principle that humans are better at some tasks than computers, and vice versa (Fitts, 1951). In reality, there are a variety of factors that are considered in task allocation including error rates, fatigue, safety, technology limitations, human values, and human desire to work or learn (see section 2.1). Specifically, in ITS authoring, there is a strong need to build mental models of ITS authoring by humans, specifically instructors, instructional designers, and subject matter experts, because the output of authoring (i.e., tutors) will have far reaching implications in the education and training of learners. Further, since the individual(s) designing the instruction may not be the same individual(s) that administering it, knowledge regarding the functioning of the tutor will need to be communicated. Thus the allocation of authoring tasks between humans and automation (e.g., the authoring tools) should seek to be complementary of one another (Grote, Weik, Wäfler, & Zölch, 1995), supporting general ITS goals, like developing pedagogically sound instruction, providing individualized tutoring, delivering time/cost savings to stakeholders, and so on.

Another aspect of task allocation to consider is the ongoing development of ITS authoring tools. That is the notion of *balanced work,* as described by Hollangel and Bye (2000), resulting from "an adjustment by the working system to the performance demands" (p. 255). Resources are derived from humans, technology, and the organization under which the work occurs. In that paper, performance demands included safety and efficiency in the context of nuclear power production. For ITS authoring, the demand may be better described as efficiency and *user acceptance*. Authors may become more familiar with ITS authoring over time, thus making the process more efficient. However, ITS and their corresponding tools continue to evolve with new features and technologies, which disturb this balance. New features place new burdens on authors, leading to stress and fatigue. Thus, an inefficient (read: frustrating) system may turn authors away from ITSs entirely in favor of easier or "good enough" alternatives to adaptive tutoring (Ososky, 2016). "Good enough" alternatives include non-adaptive learning systems, analog alternatives (e.g., books, flashcards) and/or human-to-human tutoring.

2.3. Automation

Automation, then, is the use of a machine to perform a function or task. Regarding automation of a function for particular system, Parasuraman, Sheridan, & Wickens (2000) suggested the application of a model for types and levels of automation along a 10 point scale ranging from a fully manual to a fully automated task (Table 1). The middle automation layers are further differentiated by the complementary four-stage model of information processing which includes sensory processing, working memory, decision making, and response selection.

Table 1: Levels of Automation of decision and action selection (Parasuraman et al., 2000, p. 287)

10.	HIGH. The computer decides everything, acts autonomously, ignoring the human.
9.	Informs the human only if it, the computer, decides to
8.	Informs the human only if asked, or
7.	Executes automatically, the necessarily informs the human, and
6.	Allows the human a restricted time to veto before automatic execution, or
5.	Executes the suggestion if the human approves
4.	Suggests one alternative
3.	Narrows the selection down to a few, or
2.	The computer offers a complete set of decision/action alternatives, or
1.	LOW. The computer offers no assistance; human must take all decisions and actions.

That model is operationalized within an ITS authoring tool typically in support of efficiency/usability: auto-fill, intelligent defaults, advanced configuration options, templates, etc. With that comes the same concern that may manifest in other automated systems, automation misuse or disuse. Misuse describes failures resulting from an overreliance on automation capability, while the underutilization of automation is known as disuse (Parasuraman & Riley, 1997). Thus, if too much of the authoring system is automated, humans will have little idea of what it is doing (i.e., poor mental model) and will not be able to review/make changes to the tutor once it is produced. Alternatively, if the automation is poor, cumbersome, or not transparent, an author may lose trust in the ability of the tool to build a valid tutor, and discontinue use in favor of some other instructional alternative. Again, a dynamic approach to automation may provide a pathway to a useful and reliable authoring system for potential authors of different skill levels, roles, and/or experiences.

3. CASE STUDY

The widespread adoption of ITSs depends upon the ability for relatively low-skill users to be able to create tutoring systems. These tutoring systems give each user a customized/tailored learning experience, which depends runtime content availability. One of the primary goals for the Generalized Intelligent Framework for Tutoring (GIFT; Sottilare, Brawner, Goldberg, & Holden, 2012), an adaptive tutoring architecture, is to lower the entry skill and reduce the time to author ITSs. Pursuant to this goal, two principle methods are being investigated simultaneously: 1) improving the usability of authoring tools and 2) automating parts of the authoring process to reduce the author's workload. In order to support the former with the latter, the discussion now turns toward the authoring tools contained within GIFT.

3.1. Authoring tools in GIFT

The Generalized Intelligent Framework for Tutoring (GIFT) is "an empirically-based, service-oriented framework of tools, methods, and standards to make it easier to author computer-based tutoring systems (CBTS), manage instruction, and assess the effect of CBTS, components and methodologies" (Sottilare et al., 2013). GIFT is currently under development and includes a number of technologies, features, tools, and methods intended to support a variety of users including instructional designers, authors, instructors, researchers, and learners. The GIFT Authoring Tool (GAT) has undergone a number of major revisions since the project's origin. Each new version of the GAT has endeavored, in part, to improve authoring usability and efficiency, concurrent with the notion of *balanced work.*

3.2. ITS Authoring tasks for humans

High-level tasks associated with tutor authoring and management include: 1.) Defining objectives, 2.) Content curation 3.) Sequencing content and overall "course flow" 4.) Creating discrete-time assessments, and 5.) Creating real-time assessments, 6.) Generating learner reports and, 7.) Analyzing ITS data for purposes of tutor refinement. Each of those tasks will be described in the following sections, through the lens of task automation.

3.2.1. Define learning objectives

GIFT represents concepts to be learned as either a flat list (i.e., no hierarchical relationship) or a hierarchy. The lowest level in the hierarchy of concepts require an assessment to determine if the learner has mastered that concept. Leaves may be defined as any concept without a child. Concepts at the leaf level may be rolled up to determine proficiency in higher level concepts. For example, the assessments for concepts noted as circles, parabolas, ellipses, and hyperbolas may be used to assess the parent concept "Examples of Conic Sections and their properties." This means that each leaf in the hierarchy or item on a list that is assessed contributes to authoring workload as it requires the development of an assessment (e.g., knowledge, skill test, or real-time assessment coupled with an external environment). Currently, this task is fully performed by the human (i.e., Level 1 automation). Even if the system provided ontologies for high-level concepts to form hierarchies, the manner in which those hierarchies may be ultimately structured will likely differ between course designers.

3.2.2. Content curation and creation

Content curation and, subsequently, content sequencing are two tasks that humans are primarily responsible for in GIFT, with little oversight from system. Content creation / curation involves gathering all of the multimedia that will be displayed to the learner, whether created from scratch or taken from another source.

GIFT makes the assumption that content curation is an integrated part of the authoring process. In other words, the efforts of finding, retrieving, and organizing content must be performed within the system. We assume this must be done with all content. Authored content need not be created from scratch, but does have to be managed. Some content (e.g., presentation material, surveys, quizzes, multimedia, or simulation scenarios) can be reused. The need for content is primarily driven by the total number of concepts, which include terminal learning objectives (TLOs) and enabling learning objectives (ELOs). In GIFT, concepts can be organized in a hierarchical or linear structure.

3.2.3. Content sequencing and course-flow

Content sequencing involves building the timeline of events that the learner will encounter (Figure 1); in GIFT this is known as the *course flow*. The course flow is made up of a series of events known as *course objects.*

The current authoring experience is based about a visual course building interface within the Course Creator. From within this interface, all other core aspects of course authoring are accessible to the user. The course flow timeline interface was redesigned based on a flow chart (or discrete event process) metaphor with simple drag-and-drop functionality. The visual structure of the course more accurately suggests the sequencing functions that are available to course authors. The design intent was to evoke a *mental model* of similar, more familiar interfaces in order to make tutor authoring more learnable for new users. The system provides a set of all possible course objects, displayed in the toolbox on the left-hand side of the interface (similar to automation level 2). Authors can drag and drop objects onto the timeline in any position. Objects already on the timeline can be re-ordered or deleted as needed.

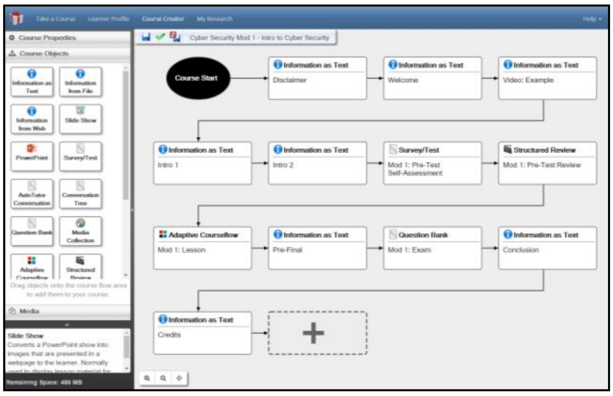


Figure 1: Human authors are responsible for sequencing course objects, which contain multimedia, assessments, digital game scenarios, and other instructional events.

The current authoring experience is based about a visual course building interface within the Course Creator. From within this interface, all other core aspects of course authoring are accessible to the user. The course flow timeline interface was redesigned based on a flow chart (or discrete event process) metaphor with simple drag-and-drop functionality. The visual structure of the course more accurately suggests the sequencing functions that are available to course authors. The design intent was to evoke a *mental model* of similar, more familiar interfaces in order to make tutor authoring more learnable for new users. The system provides a set of all possible course objects, displayed in the toolbox on the left-hand side of the interface (similar to automation level 2). Authors can drag and drop objects onto the timeline in any position. Objects already on the timeline can be re-ordered or deleted as needed.

3.2.4. Author discrete-time assessments

Creating content for assessment is another task that is, in part, delegated to the author; however, parts of this process are in planning to increase in level of automation. Discrete-time assessments include surveys and question banks, as well as the logic that supports adaptive branching through a course. In the case of the former, that is similar to a media-content creation task (Figure 2). Human authors must generate a set of questions (and responses) for a concept(s) within the course. The author must also designate the correct response(s) and manner in which the question will be scored. If the question is contained within a question bank, the author may also assign a difficulty level to the question.

Within a question bank, the human author must also specific the logic for the number and type of questions that should be randomly displayed to the learner at runtime. That is the point at which we are developing more automation to handle the logic set by the author. In the past would have been responsible for calculating and setting the question selection and scoring criteria for the survey, but usability analysis determined that this created a burden on the author to remember the type and count of each question type (e.g., easy, medium, hard) as well as the possible points available for a particular survey administration. Now, the GAT is moving toward an automated process by which these sums are calculated on-the-fly and dynamically update when changes are made to the survey content. This seemingly superficial change significantly reduces the possibility for display and/or scoring errors during runtime.

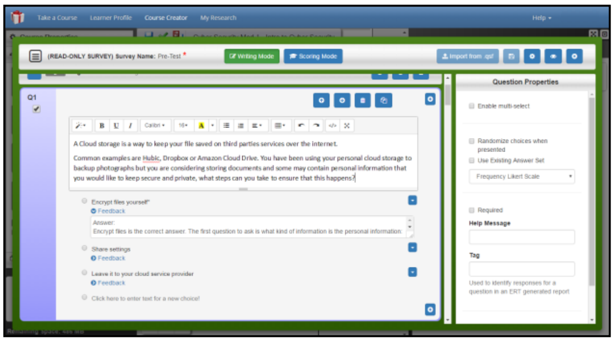


Figure 2: Human authors create questions, the system selects questions to display and calculates scores

One way in which GIFT provides adaptive tutoring is via the *adaptive courseflow course object*. In that object, each concept must be tied to content presented to the learner as part of Merrill's component display theory (CDT; Merrill, 1983) implemented within the GIFT authoring schema. For a set of concepts, this CDT content includes information about *rules* (facts, principles), *examples* (models of successful behavior), recall (an assessment also known as a check on learning or a knowledge test), and *practice* (opportunities to apply knowledge and develop skill). Recall and practice are relevant to this current paper.

For the recall quadrant, GIFT assesses domain knowledge. For the practice quadrant, both knowledge and skill may be assessed as part of an interactive experience (e.g., simulation, serious game; see sections 3.2.5 and 3.3.2, respectively). The human author is responsible for setting thresholds at both the recall and practice phases at which the learner will either move onto the next activity, or return to either the rules or examples phases for remediation in the form of additional content. That content comes from the same pool of content described earlier (in 3.2.2), however the content must be tagged with appropriate metadata in order for the tutor to dynamically select content for presentation at runtime. Currently, much of this selection logic is transparent to authors, and was coded based on empirical research and review of the relevant literature. It would be similarly beneficial if agents within the tutoring system were able to make intelligent suggestions regarding the metadata tagging of media content (i.e., automation level 5). Recommendations for such agents can be found in other work (Brawner, 2015).

3.2.5. Author real-time assessments

Real-time assessment in GIFT refers to the monitoring of learners while engaged in an interactive experience, which could be anything from dynamic slide show content, to an immersive game or simulation environment. Those experiences may occur within the practice phase of an adaptive course flow course object (see 3.2.4) or as a standalone activity. The process of creating real-time assessments is inherently complex and contains a number of moving parts. The technical details of how GIFT communicates with external applications is outside the scope of this discussion; however, the relevant aspect of that function is that authors are responsible for interpreting messages and creating assessment logic from inputs provided by the interactive application. Real-time assessment of an interactive application includes four steps to be defined by the author (Figure 3): 1) scenario properties, 2) tasks and concepts 3) instructional strategies and 4) state transitions. An example of this for a virtual excavator simulator is shown in the following figure.

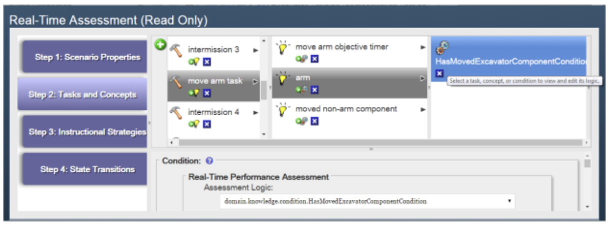


Figure 3: The four step process to create real time assessments for external learning environments is currently a highly manual process.

There, the GAT provides support around automation level 3 (refer back to Table 1 for a refresher), in which options available to the human author in the latter steps of this process are constrained based on human inputs in the earlier part of the authoring process. However, this may be a point at which *balanced work* is disturbed by the sheer complexity of this process, and the potential for errors during configuration. To that end, the GAT development team is currently working toward a more streamlined workflow for creating real-time assessment of practice environments and other external applications. By providing more of a guided experience through this process, and reaching level 4/5 of automation, the memory burden on the human author can be reduced. In other words, instead of the human author working from a blank slate, an intelligent set of options can be presented from which the author can evaluate. Thus, authors free up cognitive resources to work on other tasks, such as designing feedback or remedial content to improve the quality of the instruction for the learner.

3.2.6. Author tutor interactions

Tutor interactions refer to the tutor's responses to the learner as the learner interacts with the instructional content. For a survey or quiz, tutor interactions take the form of feedback based on the learner's responses. The tutor may also take action to provide remediation to the learner, in the case of an adaptive courseflow sequence (see 3.2.4). Real-time assessment tutor interactions are much more open ended. Tutors can alter the scenario in an interactive simulation in response to learner behavior, provide encouragement or coaching via a synthetic avatar, or suggest additional instruction on previous concepts. While tutors are expected to interact with the learner without direct human intervention, human authors are currently responsible for manually specifying most of the feedback messages, scenario adaptations, etc. during the authoring process.

3.3. ITS authoring tasks for computers

Automating different functions in GIFT associated with authoring has been explored through different approaches with varying degrees of success including, automated content development from text sources and the development of wizards to guide inexperienced authors. This section discusses current functions in GIFT that seem appropriate for allocation to automation at higher levels. Early portions of the project maintained many of these tasks as human tasks – the exact tasks, formats, structures, and other details were still being defined. The modern case study of GIFT involves many tasks which were originally human-performed, but are now automated.

3.3.1. System-level functions

While not a major step of authoring, *per se*, we identify system level functions as those tasks that are necessary to the operation of ITS platform, but not necessarily relevant to the goal of creating a tutor. Those include, for example, file management on a server, maintaining users and files within a database, and generating configuration scripts to power the tutor. We will take a closer look at that last example.

At the system-level of GIFT, a set of extensible markup language (XML) files provide the configuration for a particular GIFT course, including the use of learning content, assessments, and external training applications (Sottilare et al., 2013). The current version of the GAT is a cloud-hosted, web-based productivity application (known as GIFT Cloud) that uses flow charts and forms to help users to structure and configure tutors. In terms of function allocation, the construction of the XML code is being automated entirely by the GAT, with virtually no oversight by the author to verify the syntax of the code is error-free. In earlier, downloadable versions of GIFT, however, humans were responsible for writing and maintaining the underlying XML code.

In most cases, save for administrators and power users, system level functions should be automated by the system, and only informing the human when there is an error (automation level 7). This helps to provide balanced work, by reducing the burden on the human author, and allocating more human resources to tasks relevant to tutoring. In writing, this may seem obvious, however many ITSs are borne from development projects, where authoring interfaces (if available at all) are driven by engineering needs, not user-centered design. ITS development, by its nature, outpaces the development of authoring interfaces and tools. Constant monitoring throughout the ITS development project is therefore necessary to ensure that humans are not accidently assigned system-level tasks.

3.3.2. Creating practice and training content

Much of Army training differs from traditional ITS content (e.g., problem-based mathematics and physics tutors) in that it often requires conceptual knowledge (why you are doing something) in addition to procedural knowledge (what to do). At the U.S. Army Research Laboratory (ARL), we are seeking new methods to reduce the skill and time required to author scenario-based simulations and serious games to allow GIFT to automatically author variants of existing training scenarios which are relevant to the authors defined learning objectives. That is similar to content creation (see 3.2.2), however the workload scales exponentially in comparison when learners require an abundance of unique opportunities for interactive practice and evaluation.

The method to address that challenge is called automated scenario generation (ASG; Zook et al., 2012) or evolutionary scenario generation (ESG; Luo, Yin, Cai, Zhong, & Lees, 2017). This method focuses on how to use information from a "parent" scenario to generate hundreds or thousands of "child" scenarios and then rank order the child scenarios according to their relevance to a set of author-defined learning objectives (see 3.2.1).

The automated scenario generation method described would allow a GIFT-based tutor to customize (e.g., change difficulty level of the scenario) in real-time based on the learner's state (e.g., performance or emotion) or traits (e.g., personality) to optimize their learning, retention, and transfer of skills from training to the operational or work environment. This method would allow ITS developers who want to integrate GIFT with training simulation or serious games (e.g., Virtual Battle Space) to expand existing training capabilities to facilitate adaptive instruction with minimal additional burden on the scenario author.

ASG and ESG are good examples of ideal collaboration between humans and automated authoring tools. At automation level 5/6 the system is providing intelligent suggestions to the human regarding potential scenarios, and the human is able to review those suggestions and make manual changes, as needed.

3.3.3. Reporting and Analytics

There are two authoring-related tasks that occur after a tutor has been created: reporting and analytics. The tasks are described as authoring-related because either may result in changes being made to an existing tutor, or providing a foundation for a new tutor. For purposes of the current discussion, *reporting* refers to the collection and presentation of data resulting from learners interacting with a course (e.g., instructor dashboard). *Analytics*, here, refers to the aggregation of learner and system data, perhaps in a historical context, over a longer-term (e.g., training effectiveness, return on investment).

Those tasks are combined here because, within GIFT, they are currently accomplished from within the same interfaces in the GAT. At present, these processes are highly manual (automation levels 1 / 2). An event reporting tool is available within the *My Research* section of GIFT. This tool allows humans to build customized reports by selecting from a set of available data features to compile into a spreadsheet file (Figure 4). Currently, the human is responsible for processing the data in whatever manner they choose. This task has remained mostly manual because GIFT is highly domain independent, meaning that tutors may generate potentially many different types of data.

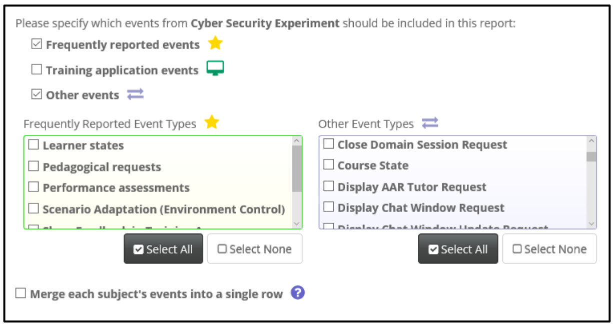


Figure 4: GIFT Cloud's event reporting tool allows authors to select from either suggested or all possible options.

We are currently investigating best practices in automating this process in the form of dynamically configurable instructor dashboards, and enhanced visualization tools for analytics. The design intent is to reduce the amount of time needed to generate insights from data, as well as reduce the skill level required in order for humans to leverage insights.

4. DISCUSSION

Using GIFT as a case study, we have identified major functions related to adaptive tutor authoring. Each function is evaluated through the lens of function allocation between humans and the system, and the degree to which automated functions within the system support human authors. A primary aim of the present paper was to present evidence and make recommendations for the ongoing design of ITS authoring tools, in general.

4.1. Summary of GIFT authoring automation

The current case study was presented in the context of where the GAT is now with respect to automation. Target levels of automation for those tasks, near or on the horizon, were also described. Those are summarized in Figure 5.

With the exception of system-level functions, it is suggested that most ITS authoring tasks be automated no higher than level 5. This is the point at which the human still approves all decisions and/or suggestions made by the system. Stated differently, it is recommended that a human remains "in the loop" regarding all content and experiences that a learner may encounter when interacting with the adaptive tutor. There are several reasons for this: 1.) Intelligent tutoring, while supported by decades of research, is still relatively novel to some involved in the process of creating and managing instruction. Therefore, keeping the human in the loop allows them to build knowledge of an otherwise novel system, while also developing trust in the efficiency and effectiveness of adaptive tutors. 2.) The concept of what an ITS is (and could be) is still evolving, in both features and intelligence. Automating too much of the process may inhibit authors from making technical and creative contributions to the community at-large. 3.) Existing educational models rely upon instructors who are responsible for the student experiences – learning, or failure to learn, is the responsibility of the instructor. The current instructional model is more "sage on the stage" than "guide on the side." In the future, especially in alternative models of education, the level of automation recommendation may change to reflect higher levels of automation, however more empirical evidence is required to allow systems to autonomously act on system suggestions without human supervisory approval.

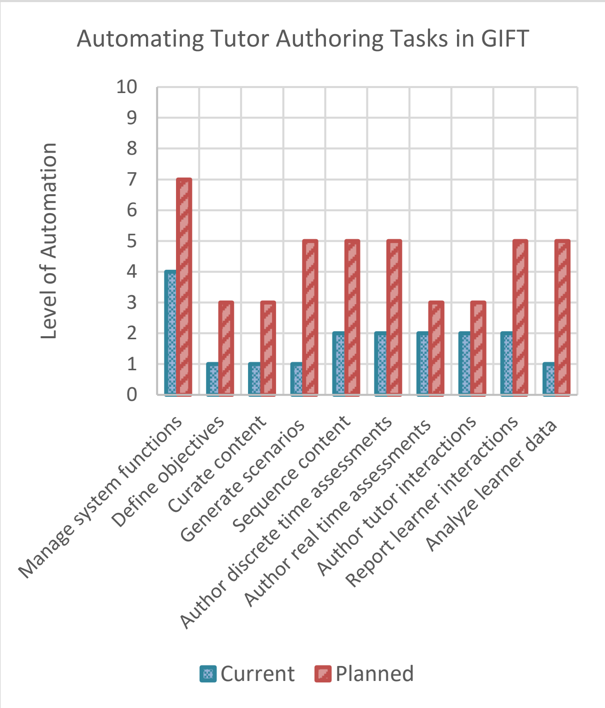


Figure 5: Current vs. planned tutor authoring function level of automation in GIFT

4.2. Looking toward the future

The authors would like to end on a discussion of the "art of the possible" – what is the maximum amount of automation possible at each level and what are the probable technologies involved at the high level? This paper presented a view of how a human author affects the system operation, while previous work outlines the operation of system-level functions (Brawner, Sinatra, & Sottilare, in press), or existing system-level function already developed (Brawner, Heylmun, & Hoffman, 2017).

First, it is worthwhile to mention that a discussion of learner and instructional modeling has been removed from this work. Other than the configuration of a small number of settings, for which defaults exist, there is nothing that the human user needs to do in order to capture emotional state data from a Microsoft Kinect, or to select the type of instruction used in remediation from a failed quiz. Learner-specific instructional strategies can now be applied directly from provided content without *any* interaction from the human user – removing this functionality from the discussion of instructor roles.

The GIFT system, and the base of instructors / instructional designers that use it, is not necessarily ready for this approach, but it is technically possible to automate the author out of these functions. As an example, if we assume that the author can provide or point to seed content (such as a textbook, deck of PowerPoint slides, etc.), the identification of instructional objectives can be automated via natural language processing technologies. Using the same technologies, the content can be curated into which portions of it address which objectives. Further use of text processing technologies such as Latent Dirichlet Allocation (LDA) and Latent Semantic Allocation (LSA) can personalize learning paths based on observed evidence or relation to other topics (Lahti, 2010). In total, a fully automatic authoring process (level 5+) may be possible, as discussed elsewhere (Olney, Brawner, Pavlik, & Koedinger, 2015). Further, an example of "full" (level 5) automation may consist of the techniques and associated technologies described in Table 2.

Table 2: Summary of higher automation authoring tasks and associated technologies

Authoring technique	Enabling technology
An analysis of a subset of Wikipedia pages, including the linked-in references	Learning objectives identified LSA / LDA technologies used
An analysis of the learner profile and mapping to the learning objectives	Simple difference calculations
A path mapping of learning objectives through content	A* ("best first") search algorithm
Generation of discrete time assessments	Generation of assessment questions Generation of distractors / incorrect LDA, Ontology search technologies such as syntactic tree kernels
Generation of simulation events	Using seed scenarios, scenario generation techniques, simulated students for scenario assessments Multiple AI/search technologies, no clear choices

ISBN 978-88-97999-90-4; Bruzzone, Cayirci and Sottilare Eds.

Authoring technique	Enabling technology
Generation of realtime assessments	Deviation from expert performance criterion Simple metrics such as standard deviations
Analysis of learner data	Typical reporting – grades, trends, etc. Typical algorithms, clustering of student groups, etc.
Self-Improving systems	Learning appropriate instructional interventions over time Markov decision processes, policy adjustments

5. CLOSING THOUGHTS

This paper presented a view of the three questions stated earlier: How should ITS authoring tasks be delegated between humans and software? What level of automation is appropriate for tasks that are delegated to software? How might tutor authoring be automated in the future? The answer to all questions is that the system (automation) and user have a collaborative relationship – the system shouldn't ever present unknown materials to a learner, but suggestions, automation, and assistance is valuable during the authoring process. While there is potential to fully automate ITS creation, human tools to create ITSs will be in use and practice for the foreseeable future. Level 5 automation, where the human has approval ability, but not time-sensitive veto ability, appears to be the "sweet spot" for the various activities of content creation. Simultaneously, the management of pedagogy, learner profiles, reporting, and other tasks can be left to the machine alone.

ACKNOWLEDGMENTS

The research described herein has been sponsored by the U.S. Army Research Laboratory. The statements and opinions expressed in this paper do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

REFERENCES

Aleven, V., & Sewall, J. (2010). Hands-on introduction to creating intelligent tutoring systems without programming using the cognitive tutor authoring tools (CTAT). *Proceedings of the 9th International Conference of the Learning Sciences-Volume 2*, 511-512.

Bill & Melinda Gates Foundation. (2015). Teachers know best: What educators want from digital instructional tools 2.0. Retrieved from http://collegeready.gatesfoundation.org/2015/12/what-educators-want-from-digital-tools-2-0/

Brawner, K. (2015). Authoring agent based tutors *Design Recommendations for ITS* (Vol. 3). Orlando, FL: U.S. Army Research Laboratory.

Brawner, K., Heylmun, Z., & Hoffman, M. (2017). The GIFT 2017 Architecture Report. *5th Annual GIFT Users Symposium*.

Brawner, K., Sinatra, A., & Sottilare, R. (in press). Motivation and Research In Architectural Intelligent Tutoring. *Simulation and Process Modelling Special Issue of "New trends of simulation and process modeling in multiple domains"*.

Fitts, P. M. (1951). Human engineering for an effective air-navigation and traffic-control system.

Grote, G., Weik, S., Wäfler, T., & Zölch, M. (1995). Complementary allocation of functions in automated work systems. *Advances in Human Factors/Ergonomics, 20*, 989-994.

Hollnagel, E., & Bye, A. (2000). Principles for modelling function allocation. *International Journal of Human-Computer Studies*, *52*(2), 253-265.

Lahti, L. (2010). Personalized learning paths based on Wikipedia article statistics. *2nd International Conference on Computer Supported Education (CSEDU 2009)*.

Luo, L., Yin, H., Cai, W., Zhong, J., & Lees, M. (2017). Design and evaluation of a data-driven scenario generation framework for game-based training. *IEEE Transactions on Computational Intelligence and AI in Games*.

Merrill, M. D. (1983). Component display theory. *Instructional-design theories and models: An overview of their current status, 1*, 282-333.

Mitrovic, A., Martin, B., Suraweera, P., Zakharov, K., Milik, N., Holland, J., & McGuigan, N. (2009). ASPIRE: an authoring system and deployment environment for constraint-based tutors. *International Journal of Artificial Intelligence in Education*, *19*, 155-188.

Murray, T. (1996). Having it all, maybe: Design tradeoffs in ITS authoring tools. *Intelligent Tutoring Systems*, 93-101.

Murray, T. (2004). Design tradeoffs in usability and power for advanced educational software authoring tools. *Educational Technology*, *44*(5), 10-16.

Murray, T. (2014). Theory-based authoring tool design: considering the complexity of tasks and mental models. In R. A. Sottilare, A. C. Graesser, X. Hu & K. Brawner (Eds.), *Design Recommendations for Intelligent Tutoring Systems* (Vol. 3 Authoring Tools and Expert Modeling Techniques, pp. 9-29). Orlando, FL: U.S. Army Research Laboratory

Norman, D. A. (1986). Cognitive engineering. In D. A. Norman & S. W. Draper (Eds.), *User centered system design: New perspectives on human-computer interaction* (pp. 31-61). Hillsdale, NJ: Lawrence Erlbaum Associates.

Olney, A. M., Brawner, K., Pavlik, P., & Koedinger, K. R. (2015). Emerging trends in automated authoring. In R. A. Sottilare, A. Graesser, X. Hu & K. Brawner

(Eds.), *Design recommendations for adaptive intelligent tutoring systems: Learner modeling* (Vol. 3, pp. 227-242).
Olsen, J. K., Belenky, D. M., Aleven, V., & Rummel, N. (2013). Intelligent tutoring systems for collaborative learning: Enhancements to authoring tools. *Artificial Intelligence in Education*, 900-903.
Ososky, S. (2013). *Influence of task-role mental models on human interpretation of robot motion behavior.* (Doctoral Dissertation). University of Central Florida, Orlando, FL.
Ososky, S. (2016). Designing the user experience of the GIFT Cloud authoring tools. In R. Sottilare & S. Ososky (Eds.), *Proceedings of the 4th Annual GIFT Users Symposium* (Vol. 4, pp. 145-156). Orlando, FL: U.S. Army Research Laboratory.
Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 39*(2), 230-253.
Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, *30*(3), 286-297.
Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin, 100*(3), 349.
Sottilare, R., Brawner, K. W., Goldberg, B. S., & Holden, H. A. (2012). *The Generalized Intelligent Framework for Tutoring (GIFT). Concept paper released as part of GIFT software documentation.* Orlando, FL: U.S. Army Research Laboratory – Human Research & Engineering Directorate (ARL-HRED) Retrieved from https://gifttutoring.org/attachments/152/GIFTDescription_0.pdf.
Sottilare, R., Graesser, A. C., Hu, X., & Holden, H. (2013). *Design Recommendations for Intelligent Tutoring Systems: Volume 1 Learner Modeling*. Orlando: U.S. Army Research Laboratory.
Suraweera, P., Mitrovic, A., & Martin, B. (2010). Widening the knowledge acquisition bottleneck for constraint-based tutors. *International Journal of Artificial Intelligence in Education (IJAIED), 20*, 137-173. doi: 10.3233/JAI-2010-0005
VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems and other tutoring systems. *Educational Psychologist*, *46*(4), 197-221.
Zook, A., Lee-Urban, S., Riedl, M. O., Holden, H. K., Sottilare, R. A., & Brawner, K. W. (2012). Automated scenario generation: toward tailored and optimized military training in virtual environments. *Proceedings of the international conference on the foundations of digital games*, 164-171.

AUTHORS' BIOGRAPHIES

Scott Ososky, Ph.D. is a Postdoctoral Research Fellow assigned to the U.S. Army Research Laboratory (ARL) within the Human Research and Engineering Directorate (HRED) at ARL Orlando. His current research examines mental models of adaptive tutor authoring, including user experience issues related to productivity tools and interfaces within the adaptive tutor authoring workflow. His prior work regarding mental models of human interaction with intelligent robotic teammates has been published in the proceedings of the Human Factors and Ergonomics Society, HCI International, and SPIE Defense & Security annual meetings. Dr. Ososky received his Ph.D. and M.S. in Modeling & Simulation, as well as a B.S. in Management Information Systems, from the University of Central Florida.

Dr. Robert A. Sottilare leads adaptive training research at the US Army Research Laboratory where the focus of his research is automated authoring, instructional management, and analysis tools and methods for intelligent tutoring systems (ITSs). He is a co-creator of the Generalized Intelligent Framework for Tutoring (GIFT), an open source, AI-based adaptive instructional architecture. He is the lead editor for the *Design Recommendations for Intelligent Tutoring Systems* book series and the founding chair of the GIFT Users Symposia. He is a program committee member and frequent speaker at the Defense & Homeland Security Simulation, Augmented Cognition, and AI in Education conferences. Dr. Sottilare is a member of the AI in Education Society, the Florida AI Research Society, and the American Education Research Association. He is a faculty scholar and adjunct professor at the University of Central Florida where he teaches a graduate level course in ITS design. Dr. Sottilare is also a frequent lecturer at the United States Military Academy (USMA) where he teaches a senior level colloquium on adaptive training and ITS design. He has a long history of participation in international scientific fora including NATO and the Technical Cooperation Program. Dr. Sottilare is the recipient of the Army Achievement Medal for Civilian Service (2008), and two lifetime achievement awards in Modeling & Simulation: US Army RDECOM (2012) and National Training & Simulation Association (2015).

Keith Brawner, Ph.D. is a researcher for the Learning in Intelligent Tutoring Environments (LITE) Lab within the U. S. Army Research Laboratory's Human Research & Engineering Directorate (ARL-HRED). He has 11 years of experience within U.S. Army and Navy acquisition, development, and research agencies. He holds a Masters and PhD degree in Computer Engineering with a focus on Intelligent Systems and Machine Learning from the University of Central Florida. His current research is in machine learning, cognitive architectures, learning technologies, and ITS architecture. He manages and advises research in adaptive training and architectural programs towards next-generation training.

ADAPTIVE INSTRUCTIONAL METHODS TO ACCELERATE LEARNING AND ENHANCE LEARNING CAPACITY

Robert A. Sottilare, Ph.D. and Gregory A. Goodwin, Ph.D.
US Army Research Laboratory
robert.a.sottilare.civ@mail.mil; gregory.a.goodwin.civ@mail.mil

ABSTRACT
This paper examines adaptive instructional methods to accelerate learning and improve learning capacity. Adaptive instruction provides tailored, computer-based learning experiences based on the needs and preferences of the learner. Often the goal is to optimize learning, performance, retention, and transfer of skills from instructional environments to work/operational environments. In this case, we shall examine methods to accelerate learning (improve instructional efficiency) and enhance learning capacity (improve instructional effectiveness) during adaptive instruction using Intelligent Tutoring Systems (ITSs). Specifically, we will examine best practices incorporated or emerging in ITSs authored by the Generalized Intelligent Framework for Tutoring (GIFT; Sottilare, Brawner, Goldberg, Holden 2012; Sottilare, Brawner, Sinatra, and Johnston, 2017) GIFT is a prototype, free, open-source architecture for authoring, managing, and evaluating ITSs and adaptive instruction.

Keywords: adaptive instruction, accelerated learning, instructional efficiency, instructional effectiveness, Intelligent Tutoring Systems (ITSs), Generalized Intelligent Framework for Tutoring (GIFT)

1. INTRODUCTION

Intelligent Tutoring Systems (ITSs) provide effective one-to-one instruction predominantly in well-defined domains like mathematics, physics, and software programming. ITSs have been shown to be as effective as expert human tutors (VanLehn, 2011) and therefore should be an instructional tool of choice for self-paced, computer-guided learning. The Generalized Intelligent Framework for Tutoring (GIFT; Sottilare, Brawner, Goldberg & Holden, 2012; Sottilare, Brawner, Sinatra & Johnston, 2017) is a prototype, open-source architecture for authoring, managing, and evaluating Intelligent Tutoring Systems (ITSs) and adaptive instruction where computer-based intelligent agents guide learners based on their learning needs, preferences, and progress toward learning objectives, which are called concepts in GIFT. Thorndike (1906) inferred the importance of adaptive instruction in training and education long before ITSs ever existed: "The principal consequence of individual differences is that every general law of teaching has to be applied with consideration of the particular person … [which] will vary with individual capacities, interests, and previous experience." Adaptive instructional systems use learner attributes to tailor instruction for each individual learner or team, and specifically to drive instructional decisions (e.g., selection of future content and experiences or feedback type and frequency). This paper examines how adaptive instruction might be used as a tool to: 1) improve instructional efficiency by reducing the time needed to learn a fixed set of concepts (accelerating learning) and 2) improve instructional effectiveness by increasing the amount of material that can be learned in a fixed amount of time or improving learning capacity.

2. ENHANCING INSTRUCTIONAL EFFICIENCY

The goal of enhancing instructional efficiency is to reduce the time for learners to reach a desired level of knowledge and/or skill. Training efficiency is a relevant measure when the goal of the training is to insure that all learners attain a standard level of proficiency/knowledge. For example, organizations often have recurring training requirements to insure that individuals in the organization maintain critical knowledge and skills. In such situations, adaptive training has the potential to reduce the time needed to train for some percentage of the population (Figure 1).

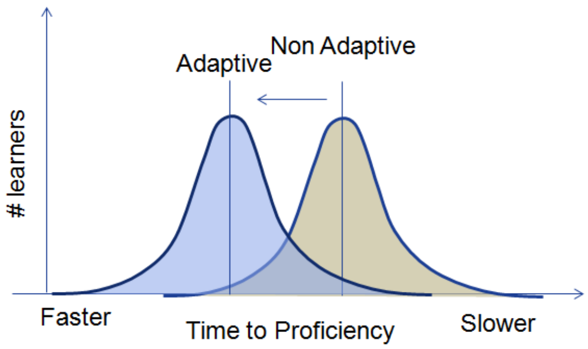


Figure 1: Enhancing Learning Efficiency

A key adaptation provided by ITSs is tailoring of content based on each learner's prior knowledge of the task domain. This reduces the amount of content shown to the learner, but varies with the learner's competence with high competency learners able to skip the most content. While this saves time during instruction, it should be noted that some review of material is required on a periodic basis to maintain proficiency, and skipping content does not accelerate learning. It only saves time spent in training or education that might be used to cover new material.

To improve training efficiency, Goodwin, Kim, and Niehaus (2017) recommend prioritizing design decisions based on a cost-savings comparison. Specifically, the cost of implementing adaptive features should be compared to the savings resulting from improved efficiency. Only features that save more than they cost should be implemented.

These recommendations and accelerated learning as a concept fly in the face of long term, deep learning goals. To overcome this, the tutor must have a highly accurate model of the learner by which to make instructional decisions. Many ITSs today only adapt based on learner performance. This has been a very clear choice. Selecting to adapt on other learner states, traits or preferences imparts some risk in the tutoring process.

3. ENHANCING INSTRUCTIONAL EFFECTIVENESS

The goal of enhancing instructional effectiveness is to increase the learner's capacity to acquire knowledge and/or skill in a fixed time. The assumptions for measuring effectiveness are that the amount of material to be learned is variable and the learning time is fixed (Figure 2). Since what is learned is variable, learners may be below, at, or above expectation in terms of their mastery of the material at the conclusion of the training time. Adaptive instruction has the potential to be more effective because it can address specific learner problems and employ a variety of methods known to be effective for each individual.

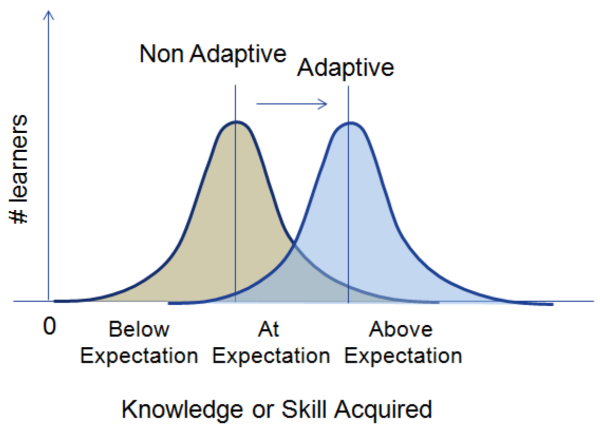


Figure 2: Enhancing Learning Effectiveness or Capacity

To support enhanced learning effectiveness or learning capacity, Goodwin, Kim, and Niehaus (2017), recommend methods to more accurate diagnose learner errors, tailored remediation, and tailored training methods. Each is important, but might include additional sub-goals. Diagnosis of the learner might include accurate classification of all critical learner states. In addition to learner errors, this might include assessment of learner performance trends (short and long term), diagnosis of learner misconceptions indicated by errors. Tailored remediation and training methods might be expanded to include tailored interaction. Not only feedback could be adapted to specific learner performance, but other states, traits, and preferences. Adaptation could also include preference tailoring in which the environment is adapted to the specific learner's cultural background to provide a familiar mental model for learning. This could enhance learner engagement and result in less down time during instruction. Another adaptation to improve effectiveness could include tailoring based on learner interests.

4. DISCUSSION

As the foregoing indicates, training efficiency and training effectiveness are goals that are sometimes at odds with one another. When seeking to minimize time to train, it is sometimes necessary to sacrifice deep learning or overtraining. When seeking to maximize long-term retention or knowledge and proficiency, it will be difficult to simultaneously reduce training time. We discuss these tradeoffs in the design of adaptive instruction in the context of 1) training vs. education, and 2) individual vs. team learning.

4.1. Adaptive Instruction in Training vs. Education

According to Fletcher (2017), training and education serve different purposes. These differences moderate, influence, or limit the effect of adaptive instructional methods and should be considered in the process of instructional design. We will focus on two factors and how these relate to learning efficiency and effectiveness. The first factor is the difference in the objectives for training and education. Training objectives are focused on learning to do a specific task or set of tasks in the operational or work environment. Educational objectives are much broader and focused on preparing to perform in yet unknown environments. If you think of training as a pebble and education as a boulder based on their relative complexity, it is much easier to move a pebble. The sheer differences between the scope of training and education are likely to result in different levels of effect when applying adaptive instructional methods. It is much more likely that adaptive instruction will have an impact on learning efficiency in training given there are smaller, less complex domains/environments encountered during training.

A second consideration in the differences between training and education is the difference in horizon. Training tends to focus on near-term goals while education has a much longer horizon. This temporal difference means the emphasis for training is on "return on investment" while education is more about "cost effectiveness" (Fletcher 2017). Return on investment (ROI) is the ratio of net gain (or benefit) to cost. If you have to invest an hour to acquire a unit of knowledge/skill, it is much more efficient than taking two hours to acquire the same knowledge/skill. Cost effectiveness is about producing optimum results for a fixed expenditure. So training may be more conducive

ISBN 978-88-97999-90-4; Bruzzone, Cayirci and Sottilare Eds.

to adaptive instructional methods that emphasize efficiency (learning as fast as possible) while education may be more compatible with methods that emphasize effectiveness (learning as much as possible).

Finally, ROI or cost effectiveness of adaptive instruction should consider not only the cost to the learner (e.g., time invested in instruction), but also the cost of creating the content and the adaptive tutor (Fletcher and Sottilare, 2014).

4.2. Adaptive Instruction for Individuals vs. Teams

Another area where differences should be considered for the design and application of adaptive instruction is individual and team (also known as collective) instruction. While ITSs adapt instruction based on individual differences (states, traits, and preferences), an ITS that adapted only on the individual differences of the members of the team would likely be less effective (and efficient) than an ITS that also modeled the collective needs of the team. Teams are "two or more people whose tasks are in some way interdependent (i.e. individual efforts are dependent upon the efforts of other members) and who have shared, common goals" (Salas 2015, p.3.; Dyer 1984; Kozlowski & Bell 2003; Salas, Dickenson, Converse, & Tannenbaum 1992).

Considerations for the adaptive instruction of teams should examine the interaction of team members which may be subdivided into teamwork and taskwork. Teamwork is "coordination, cooperation, and communication among individuals to achieve a shared goal" (Salas 2015, p.5.). "Teamwork consists of the interdependent interactions among team members as they work towards completing their objectives" (Salas 2015, p.5.). Taskwork consists of "working on a specific duty of one's job [within a team]" (Salas 2015, p.5.). Team taskwork refers "to those relevant behaviors that directly lead to the successful accomplishment of collective goals" (Salas 2015, p.5.). Teamwork may be considered largely domain-independent while taskwork is specific to a domain.

The point being made here is that teams, their behaviors, and their interactions are much more difficult to assess with respect to learning objectives and teamwork. Therefore adaptive instruction of teams is more complex than adaptive instruction individuals. This makes accelerating learning and improving learning capacity much more difficult than for individuals, and impacts the return on investment and cost effectiveness of adaptive instructional methods.

Prioritizing efficiency vs effectiveness in team training may be driven more by the criticality of the team's function than anything else. For example training for a medical team performing a complex surgery will likely prioritize training effectiveness because there is such a low tolerance for error. On the other hand training for a team of food preparers in a fast-food restaurant might prioritize training efficiency since the individuals will be working together on a daily basis and the cost of error is minimal.

5. APPLICATION OF ADAPTIVE INSTRUCTIONAL METHODS IN GIFT

Next, we examine authoring capabilities in GIFT for adapting instruction to accelerate learning and enhance learning effectiveness. GIFT allows the author to adapt instruction based on several factors in two primary areas: learner attributes and content attributes.

5.1. Adapting Instruction Using Learner Attributes

Data sources (e.g., people) emit raw data that is captured by sensors and then processed by a classifier to yield a learner state unless the state is self-reported. Each learner state can be assessed with data from a sensor, a training application or a survey depending on the validated methods available. Each state can be assessed as either a two (high and low) state or a three state (high, moderate, and low) attribute.

A five step process allows GIFT users to create new learner state interpreters as follows:

- Step 1: What learner state interpreter would you like to create? This step includes a dropdown menu that lists the six previously mentioned state interpreters plus off-task behaviors, skill, and understanding.
- Step 2: What data sources will be used to evaluate and predict the learner's state? This step includes a dropdown menu that lists eleven sensors that have been integrated with GIFT and are able to accept and interpret data from each of those sensors. These sensors include: Affectiva Q sensor for electrodermal activity, Microsoft Kinect for motion capture and facial marker mapping, Emotive Epoc wireless headset for brainwave detection, and Zephyr BioHarness for breathing and heart rate detection.
- Step 3: Which translator should be used to manage incoming data? GIFT provides a default translator, but users may build their own to filter or interpret incoming data.
- Step 4: Which classifier can consume the incoming translated data in order to calculate both short and long term learner states? Choices will be limited to a classifier based on the learner state selected in Step 1.
- Step 5: Which predictor can consume the incoming translated data in order to predict future learner states?
- Choices will be limited to a classifier based on the learner state selected in Step 1.

Once the learner state interpreter is configured, it should be validated for accuracy of predictions. The importance of highly accurate state classifiers cannot be understated. Even small errors can multiply if the tutor assumes an incorrect state and begins remediation based on that false assumption. Currently, GIFT adapts instruction based on assessed learner states as follows:

- engagement
- arousal
- motivation
- prior knowledge
- anxiety
- engaged concentration

5.1.1. Engagement and Learning

Engagement is "the degree of attention, curiosity, interest, optimism, and passion that students show when they are learning or being taught, which extends to the level of motivation they have to learn and progress in their education" (Hidden curriculum, 2014, August 26). The value of engagement is predicated on the tenet that learning is enhanced when learners are curious, interested, and/or inspired by the topic, content or instructor. In contrast, learning tends to decrease when students are disengaged for whatever reason (e.g., boredom, disinterest, or lack of purpose or goal). Accurate modeling of the learner and their goals can go a long way toward adapting instruction in a way that results in more efficient learning (accelerated learning) or effective learning.

5.1.2. Arousal and Learning

Arousal is a "physiological and psychological state of being awoken or of sense organs stimulated to a point of perception" (Wikipedia, 2017). Yerkes-Dodson (1908) state that too much or too little arousal can negatively influence task performance, and Sharot & Phelps (2004) note the tight coupling between memory and arousal which affects learning capacity. By understanding the learner's arousal from boredom to interest, a tutor (human or computer-based) might change either the environment (e.g., challenge level of a problem or scenario) or otherwise interact with the learner to optimize their arousal and thereby their learning (Figure 3).

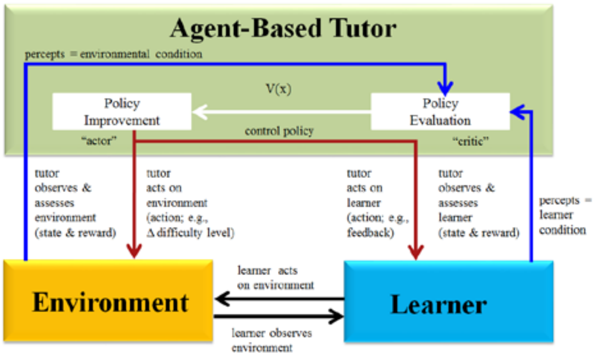


Figure 3: Optimizing the Arousal of the Learner

5.1.3. Motivation and Learning

Motivation can be defined as the purpose or reason driving the plans and actions of an individual or a team (Elliot & Covington, 2001), but it may be thought of simply as an alignment of actions with goals. The more closely aligned actions/activities are with individual or team goals, the more engaged the learner(s) will be in the activity, and the greater the opportunity for learning (knowledge and skill acquisition).

Goals are often driven by values which are shaped by many sources (e.g., family, religion, society, needs, and organizations), but may also be prioritized as in Maslow's (1971) hierarchy of needs. The tie between motivation and goals has a direct impact on learning. Motivation positively influences cognitive processes by increasing the learner's attention time on task, influencing their perseverance in the learning process, and sharpening their focus on achieving their goals (Pintrich & Schunk, 2002). By considering the goals, interests, and values of a learner, a GIFT-based ITS might select content and activities which tap into existing motivational drives and enhance learning.

5.1.4. Prior Knowledge and Learning

Prior knowledge includes the knowledge, skills, beliefs, and attitudes derived from previous experiences, and learners come to new instructional experiences with prior knowledge that influences their attention, interpretation, and organization of new data, information, and knowledge. The ability of the tutor to model and use prior knowledge to inform instructional decisions is directly related to learning efficiency and effectiveness. Instructional strategies that focus too heavily on prior knowledge can lead to boredom while focusing too lightly on prior knowledge may not provide enough of an anchor to tie in new knowledge resulting in learner anxiety. Prior knowledge may be used in GIFT-based tutors as a trigger to skip content that may have been learned previously. Errors and classified learner misconceptions trigger the tutor to review material that may not have been deeply learned.

Prior knowledge may be used differently to achieve training efficiency vs training effectiveness. Assessing the learner's prior knowledge can allow the adaptive training system to skip content which might improve efficiency. When focusing on effectiveness however, the system might give learners with more prior knowledge more advanced training to bring them to a higher level of proficiency.

5.1.5. Anxiety and Learning

Anxiety is "a feeling of worry, nervousness, or unease, typically about an imminent event or something with an uncertain outcome" (English Oxford Living Dictionaries, 2017). Since "learning" is about acquiring knowledge and/or skill through new experiences with uncertain outcomes, it is little wonder that anxiety and learning are generally incompatible. To confirm learning, instruction often includes some type of assessment of the knowledge or skill defined in the learning objectives. This assessment or test can also be a source of anxiety. Computer-based instructional environments can provide a setting for learner anxiety to grow due to lack of trust in the technology or due to the difficulty of the domain content or the use of its interface (O'neil, Spielberger & Hansen, 1969).

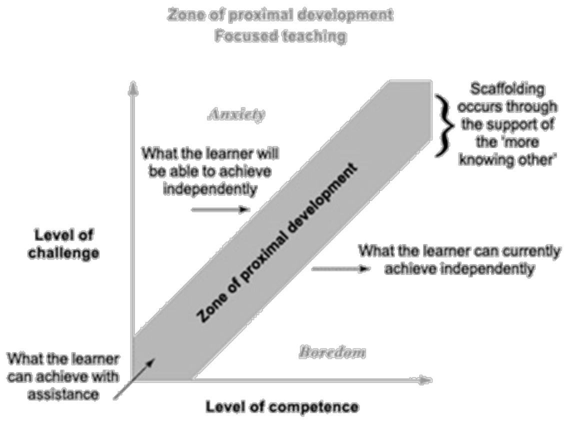


Figure 4: Aligning Content Difficulty with Learner Competence to Reduce Anxiety and Boredom

Per Vygotsky's zone of proximal development (VPD; 1978; Figure 4), an anxious learner who appears to be overwhelmed by the difficulty of the content being presented during instruction is compatible with two instructional strategies. The first strategy is to reduce the difficulty of the content presented to the learner so it is compatible with the learner's capabilities. The second strategy is to have the tutor provide scaffolding or support allow the learner to progress in learning the content at the current difficulty level. The effectiveness of instructional strategies or aids may be quantified by examining task performance with and without the aid at various levels of expertise (e.g., very low, low, moderate, high, very high). The effect of the aid may vary with the level of learner expertise. GIFT allows the author to select these strategies through selected triggering events or automatically through built in logic.

5.1.6. Engaged Concentration and Learning

Baker, D'Mello, Rodrigo, & Graesser (2010) define engaged concentration as a cognitive–affective state that may be of short duration, but more persistent than boredom. Engaged concentration is a state of engagement with a task where the learner is fully immersed in the experience and their "concentration is intense, their attention is focused, and their involvement is complete". In comparison to boredom, frustration, confusion, delight, and surprise, engaged concentration was common (average of 60% of learner time during instruction) and appeared often in computer-based learning environments.

According to Baker et al (2010), engagement concentration is of positive valence and neutral arousal. In addition to immersion, focus, and concentration on the system, Baker et al (2010) also noted additional behaviors associated with engaged concentration: leaning towards the computer, mouthing solutions, and pointing to parts of screen. Engaged concentration has been found to be positively correlated with learning (Craig, Graesser, Sullins, & Gholson, 2004; Graesser, D'Mello, Chipman, King, & McDaniel, 2007). A natural tutoring strategy for a learner in the state of engaged concentration might be to "do nothing" since the learner is already in an ideal state for learning.

5.2. Adapting Instruction Using Content Attributes

The pedagogical configuration in GIFT allows users to adapt instruction based on assessed learner states within the engine for managing adaptive pedagogy (eMAP) and content metadata attributes as follows:

- interactive multimedia instruction (IMI) level
- user control
- difficulty level
- content type
- example type

5.2.1. IMI and Adaptive Instruction

IMI (Galbreath, 1992) includes four levels to describe the interactivity of content where 1 is low interaction (e.g., reading material) and 4 is highly interactive (e.g., a fully immersive virtual simulation). Frear & Hirschbuhl (1999) indicate that the selection of the interactivity level of content has a significant effect on both achievement and problem solving skills. Lee and Boling (1999) advocate guidelines for screen design during IMI to both enhance motivation (expansive guidelines) and reduce poor practices which might negatively impact motivation (restrictive guidelines). Expansive guidelines include the use of fonts to capture the learner's attention to make it easier to navigate content, and the use of standard images to represent the learner's concepts and expectations (e.g., pause, rewind, and fast forward for video controls). Restrictive guidelines include adhering to cultural conventions when selecting images, and considering the learner's prior knowledge when selecting images. While the GIFT authoring tools do not specifically enforce these conventions, future versions of the authoring tools may include agent-based policies/rules or wizards to reinforce good IMI practices which are independent of learner attributes.

5.2.2. User Control and Adaptive Instruction

For our purposes, user control for adaptive instruction may be defined as being synonymous with adaptability in system design where the decisions and actions by the learner mold the look, feel, and function of the learning system. We adopted Oppermann & Rasher's (1997) provisions for user control for adaptive learning systems:

- offer the learner a means to initiate/halt adaptation of the system during every phase of learning
- allow the learner to accept, modify or reject every or any part of proposed adaptation
- enable the learner to specify adaptation parameters
- inform user about the proposed changes due to adaptation before actual changes take place

- giving the learner access and sole control over his/her behavior records and their evaluation (open learner model)

In GIFT, user control is defined at three levels (high, moderate, and low) where high user control would be modeled per Oppermann & Rasher's (1997) provisions. While GIFT does not yet provide a high level of learner control, triggers have been integrated to implement a moderated level of learner control (specifically, an open learner model). GIFT allows the author to select and link levels of user control to a variety of learner and content attributes with the goal of influencing learning and transfer. Hassan, Ali, & Hamdan (2015) evaluated several user control strategies for instruction with animation content, and found random user control strategies had a larger effect on achievement than other user control strategies (e.g., linear, program, free, and no user control). Mayer & Chandler (2001) discovered that learners who were allowed to exercise control over the pace of content presentation performed better in terms of their transfer of skills, but not retention.

As with prior knowledge, the implementation of user control might vary for efficiency vs. effectiveness. If the goal is to improve effectiveness, then users might be able to increase the amount of content available to maximize their knowledge of a domain. If training efficiency is the goal, then learners might be able to determine the training needed to reach the required proficiency level with the least effort.

5.2.3. Difficulty Level and Adaptive Instruction

Difficulty level is also defined at three levels (high, moderate, and low). The author can elect to tag questions or other content to allow a GIFT-based tutor to select content based on learner performance state. This meta-data tagging supports adaptation to match learner competence and content difficulty (see Vygotsky's Zone of Proximal Development; Figure 4).

5.2.4. Content Type and Adaptive Instruction

Content type ranges from animations and graphics to text to video to visual content and may be somewhat redundant with IMI level adaptations, but allows GIFT authors to target and link specific types of media (e.g. video, audio, text, animations) with learner attributes.

5.2.5. Example Type and Adaptive Instruction

Finally, GIFT provides authors with two example types: case studies and worked examples. Case studies present criteria for solving problems and making decisions, and then the learner is given one or more example cases to exercise their decision making. Worked examples allow authors to present problems in a fully worked form and gradually reduce the sequence of the problem, process, or scenario so more information is provided by the learner over time.

5.3. Using Meta-data in GIFT Tutors

As content is added to a GIFT course, it is labeled with one or more of the metadata attributes described previously in Section 5.2. This allows rules in the pedagogical configuration where eMAP is the default to determine what type of content to select for the learner based on their assessed state. Developing new rules is a simple three step process as follows:

- Step 1: In which quadrant will the metadata be used? Since GIFT's theoretical instructional basis is Merrill's (1983) Component Display Theory (CDT), each learner state is assessed in the context of four instructional quadrants: rules, examples, recall, or practice.
- Step 2: Which state must the learner be in to use the metadata? This is a long list of learner attributes that include grit, learner ability, learning style, goal orientation, engagement, and several emotional states. One is selected from a dropdown list along with a state classification (high, medium, low, or unknown).
- Step 3: Which metadata attributes should be used? In this step the author selects from a dropdown list of metadata attributes (content descriptors) as noted above.

This allows the author to link content and adapt content based on changing learner states. Again, a critical element in this process is the accurate classification of learner states. Anything that interferes with data to support classification or affects the accuracy of the classification affects the effectiveness and efficiency of the tutor, and this in turn limits opportunities to improve learning capacity or accelerate learning.

6. CHALLENGES AND NEXT STEPS

A major challenge is to balance acceleration vs effectiveness. For example, if we accelerate learning how do optimize deep learning which usually requires high numbers of cycles and/or long periods of time for learning to set? How might we optimize multiple outcomes like rapid learning, high retention and high rate of skill transfer? Finally, how do we develop authoring tools that allow designers and developers to easily choose the appropriate design features to achieve these goals?

A next step will be to use the experimental testbed within GIFT to analyze learner attributes, adaptive instructional methods, and content to develop methods to balance instructional outcomes as shown in Figure 5 (Hanks, Pollack, and Cohen 1993).

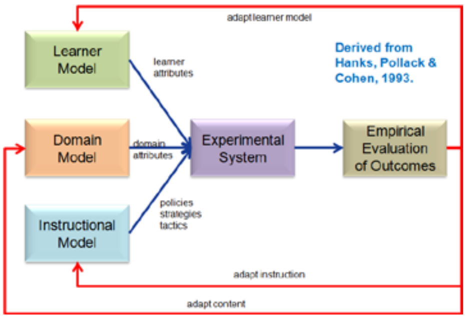


Figure 5: GIFT Evaluation Testbed

Another challenge to accelerating learning is optimizing complex decisions made by the tutor. The ability to make these instructional decisions rapidly is of some importance, but a more impactful capability will be highly effective decisions made by the ITS. This will reduce the amount of time used by the learner interacting with the tutor in activities that are not relevant or influential to learning outcomes. The basic research challenge is to optimize complex instructional decisions which involve multi-dimensional conditions of both the learner (e.g., states/traits) and the environment (e.g., entities, events, options) to select actions that influence learning and the desired outcome of "reducing time to proficiency". Meeting this challenge will likely involve solving other problems including:

- modeling complexity in individuals and teams as systems
- understanding the variability of human traits and behaviors and their relationship to learning

The modeling of the complexity of teams as systems will require investigation into teamwork as an antecedent to team learning and performance. Sottilare et al (2017) developed a model of team learning and performance based on a large scale meta-analysis of the team and tutoring literature. This provides a few initial steps in being able to manage the instruction of teams efficiently. As part of understanding human variability, the potential exists to gain some efficiency and effect through augmentation of learners. While this augmentation could take many forms, it could be as simple as understanding the relationship between learning capacity and the physical well-being of the learner. Research that shows exercise as a method for regulation of emotions (Salmon 2001; Karoly et al 2005), the association of fitness with enhanced fluid intelligence (Hillman, Erickson & Kramer 2008), and connection between exercise and executive attention (Kubesch et al 2009) might be applied in future ITSs to improve learner capacity and reduce "lost" time during instruction.

ACKNOWLEDGMENTS

The research described herein has been sponsored by the U.S. Army Research Laboratory. The statements and opinions expressed in this article do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

REFERENCES

Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, *68*(4), 223-241.

Craig, S., Graesser, A., Sullins, J., & Gholson, B. (2004). Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. *Journal of educational media*, *29*(3), 241-250.

Dyer, J. L., 1984. Team research and team training: A state-of-the-art review. *Human factors review*, *26*, 285-323.

Elliot, A. J., & Covington, M. V. (2001). Approach and avoidance motivation. *Educational Psychology Review*, 13(2), 73-92.

English Oxford Living Dictionaries (2017). Anxiety. Retrieved from: http://en.oxforddictionaries.com/definition/anxiety July 2017.

Fletcher, J. & Sottilare, R., 2014. Cost Analysis for Training & Educational Systems. *Design Recommendations for Intelligent Tutoring Systems: Volume 2 - Instructional Management*. Orlando, Florida: Army Research Laboratory. ISBN: 978-0-9893923-2-7.

Frear, V., & Hirschbuhl, J. J. (1999). Does interactive multimedia promote achievement and higher level thinking skills for today's science students?. *British Journal of Educational Technology*, 30(4), 323-329.

Galbreath, J. (1992). The Educational Buzzword of the 1990's: Multimedia, or Is It Hypermedia, or Interactive Multimedia, or...?. *Educational Technology*, *32*(4), 15-19.

Goodwin, G., Kim, J., and Niehaus, J., 2017. Modeling Training Efficiency and Return on Investment for Adaptive Training. In Proceedings of the 5th Annual GIFT Users Symposium. Orlando, FL: US Army Research Laboratory.

Graesser, A., D'Mello, S., Chipman, P., King, B., & McDaniel, B. (2007, July). Exploring relationships between affect and learning with AutoTutor. In *Proceedings of the International Conference of Artificial Intelligence in Education (AIED).*

Hanks, S., Pollack, M.E. & Cohen, P.R., 1993. Benchmarks, Test Beds, Controlled Experimentation, and the Design of Agent Architectures. AI Magazine Volume 14 Number 4.

Hassan, A., Ali, A. Z. M., & Hamdan, M. N. (2015, October). Instructional animation, segmentation and user control strategies. In *Science in Information Technology (ICSITech)*, 2015 International Conference on (pp. 85-88). IEEE.

Hidden curriculum (2014, August 26). In S. Abbott (Ed.), The glossary of education reform. Retrieved from http://edglossary.org/hidden-curriculum, 29 June 2017
Hillman, C.H., Erickson, K. I., & Kramer, F. (2008). Be smart, exercise your heart: Exercise effects on brain and cognition. Nature Reviews. *Neuroscience*, 9, 58–65.
Karoly, P., Ruehlman, L. S., Okun, M. A., Lutz, R. S., Newton, C., & Fairholme, C., 2005. Perceived self-regulation of exercise goals and interfering goals among regular and irregular exercisers: A life space analysis. *Psychology of Sport and Exercise*, 6(4), 427-442.
Kozlowski, S. W., & Bell, B. S. 2003. Work groups and teams in organizations. In W.C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology (Vol. 12): Industrial and Organizational Psychology* (pp. 333-375). New York: Wiley.
Kubesch, S., Walk, L., Spitzer, M., Kammer, T., Lainburg, A., Heim, R., & Hille, K. (2009). A 30-Minute Physical Education Program Improves Students' Executive Attention. *Mind, Brain, and Education*, 3(4), 235-242.
Lee, S. H., & Boling, E. (1999). Screen design guidelines for motivation in interactive multimedia instruction: A survey and framework for designers. Educational technology, 39(3), 19-26.
Maslow, A. H. (1971). The farther reaches of human nature.
Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages?. *Journal of educational psychology*, *93*(2), 390.
Merrill, M. D., 1983. Component display theory. *Instructional-design theories and models: An overview of their current status*, 1, 282-333.
O'neil, H. F., Spielberger, C. D., & Hansen, D. N. (1969). Effects of state anxiety and task difficulty on computer-assisted learning. Journal of Educational Psychology, 60(5), 343-350.
Oppermann, R., & Rasher, R. (1997). Adaptability and adaptivity in learning systems. Knowledge transfer, 2, 173-179.
Pintrich, P. R., & Schunk, D. H. (2002). Motivation in education: Theory. Research, and Applications, Second Edition, Merrill Prentice Hall, Columbus, Ohio.
Salmon, P., 2001. Effects of physical exercise on anxiety, depression, and sensitivity to stress: a unifying theory. *Clinical psychology review*, 21(1), 33-61.
Salas, E., 2015. *Team training essentials: A research-based guide*. London, England: Routledge.
Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I., 1992. Toward an understanding of team performance and training. In R.J. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (p.3-29). Norwood, N.J.: Ablex.
Sharot, T., & Phelps, E. A. (2004). How arousal modulates memory: Disentangling the effects of attention and retention. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(3), 294-306.
Sottilare, R.A., Brawner, K.W., Goldberg, B.S. & Holden, H.K., 2012. The Generalized Intelligent Framework for Tutoring (GIFT). Concept paper released as part of GIFT software documentation. Orlando, FL: *U.S. Army Research Laboratory*. Retrieved from: https://gifttutoring.org/attachments/152/GIFTDescription_0.pdf
Sottilare, R., Brawner, K., Sinatra, A. & Johnston, J., 2017. An Updated Concept for a Generalized Intelligent Framework for Tutoring (GIFT). Orlando, FL: *US Army Research Laboratory*.
Sottilare, R.A., Burke, C.S., Salas, E., Sinatra, A.M., Johnston, J.H. & Gilbert, S.B., 2017. Towards a Design Process for Adaptive Instruction of Teams: A Meta-Analysis. *International Journal of Artificial Intelligence in Education.*
Vygotsky, L. (1978). Interaction between learning and development. *Readings on the development of children*, *23*(3), 34-41.
Wikipedia. (2017). Arousal. Retrieved from https://en.wikipedia.org/wiki/Arousal, 29 June 2017.
Yerkes, R. M. & Dodson, J. D. (1908) The relation of strength of stimulus to rapidity of habit formation, *Journal of Comparative Neurology and Psychology*, *18*, 459–282 doi:10.1002/cne.920180503.

AUTHOR BIOGRAPHIES

Dr. Robert A. Sottilare leads adaptive training research at the US Army Research Laboratory where the focus of his research is automated authoring, instructional management, and analysis tools and methods for intelligent tutoring systems (ITSs). He is a co-creator of the Generalized Intelligent Framework for Tutoring (GIFT), an open source, AI-based adaptive instructional architecture. He is the lead editor for the *Design Recommendations for Intelligent Tutoring Systems* book series and the founding chair of the GIFT Users Symposia. He is a program committee member and frequent speaker at the Defense & Homeland Security Simulation, Augmented Cognition, and AI in Education conferences. Dr. Sottilare is a member of the AI in Education Society, the Florida AI Research Society, and the American Education Research Association. He is a faculty scholar and adjunct professor at the University of Central Florida where he teaches a graduate level course in ITS design. Dr. Sottilare is also a frequent lecturer at the United States Military Academy (USMA) where he teaches a senior level colloquium on adaptive training and ITS design. He has a long history of participation in international scientific fora including NATO and the Technical Cooperation Program. Dr. Sottilare is the recipient of the Army Achievement Medal for Civilian Service (2008), and two lifetime achievement awards in Modeling & Simulation: US Army RDECOM (2012) and National Training & Simulation Association (2015).

Dr. Gregory Goodwin is a senior research scientist and acting Branch Chief at the Army Research Laboratory-Human Research and Engineering Directorate, Simulation and Training Technology Center (STTC) in Orlando, Florida. His research focuses on methods and tools to maximize the effectiveness of training technologies. After completing his Ph.D. at the State University of New York at Binghamton in 1994, Dr. Goodwin spent three years in a post-doctoral fellowship at the Columbia University College of Physicians and Surgeons followed by a year as a research associate at Duke University Medical Center before joining the faculty at Skidmore College. In 2005, Dr. Goodwin left academia and began working at the Army Research Institute (ARI) field unit at Fort Benning Georgia and six years later, he came to the ARI field unit in Orlando, FL where he has been examining ways to leverage technologies to reduce the cost and improve the effectiveness of training.

TEST CASE BASE FOR DYNAMIC VERIFICATION OF DEFENSE SIMULATION MODEL BASED ON SYSTEM ENTITY STRUCTURE AND MODEL BASE

Jaewoong Choi[(a)], Jiyong Yang[(b)], Inho Jung[(c)], and Changbeom Choi[(d)]

[(a),(b),(c)]School of Information and Communication Engineering, Handong Global University,
Handong-ro 558, Pohand-si, Kyungbuk, Rep.of Korea
[(d)]School of Global Entrepreneurship and ICT, Handong Global University,
Handong-ro 558, Pohand-si, Kyungbuk, Rep.of Korea

[(a)]21631005@handong.edu, [(b)]21631004@handong.edu, [(c)]21731002@handong.edu
[(d)]cbchoi@handong.edu

ABSTRACT
The development process of the defense simulation consists of developing simulators and models for purpose of simulation. To increase the efficiency of the existing defense simulation development process, it is possible to apply SES / MB to defense simulation. When SES / MB is applied to the development process of defense simulation, defense simulation can be developed by importing and synthesizing models corresponding to nodes of SES from MB. In this paper, we propose Test case Base for dynamic verification of defense simulation model based on SES / MB. The proposed TB can be used for dynamic verification of defense simulation based on SES / MB. As a case study, we verify dynamic verification of simulation models modeled in the AddSIM environment being used for modeling in the National Defense Science Institute of Korea.

Keywords: Dynamic Verification, Defense Simulation, SES / MB, Test case Base

1. INTRODUCTION

Defense simulation, also known as a war game, is a simulation for testing and improving the war theory without actual war. Defense simulation is widely used for training, analysis, and acquisition. In addition to the defense simulation that is used for various purposes, the general simulation development process consists of two processes: determining the purpose of the simulation and developing a model that fits the purpose.
The two steps of the simulation development process, establishing the simulation purpose and modeling process, have significant meaning in the M&S. Establishment of the simulation purpose involves setting the target system of the simulation and determining the simulation purpose for the target system. First, since modeling does not represent the entire system but abstracts the system according to the purpose of simulation, it is critical to establish the purpose of simulation. Second, since the resolution of the models that are configured according to the purpose of the simulation is determined, and the model has changed accordingly, it is important to establish the purpose accurately. Third, because the target system should be modeled exactly to express the characteristics used in M&S, modeling process is important.
In conventional defense simulation, developers had to re-develop and verify simulators and models that depend on the target system and simulation purpose. Although there are efforts to improve reusability such as using reusable parts of the developed model in the process of development, it does not have a significant influence on the increase of the reusability of the model.
The System Entity Structure (SES) formalism and the Model Base (MB) concept can be applied to improve the reusability of the model. SES formalism is a formalism that expresses all the alternatives of one system using tree structure, and it includes three kinds of information such as system configuration, classification, and connection relation (Kim, 1990). The MB is a database for managing models constituting the system represented by the SES formalism. By using SES / MB appropriately, a defense simulation, which is developing the model by designing a model that is suitable for the simulation purpose, pruning the system represented by the SES formalism, and combining the sub-models taken from MB, can be constructed. Details of SES / MB are described in Chapter 2.
As described briefly above, applying the SES / MB to the defense simulation makes it possible to build a model management system that improves the reusability of the model. In this paper, we propose Test case Base for dynamic verification of defense simulation model. TB manages the test cases saved in the MB.
This paper is composed as follows. Section 2 explains SES / MB, which is a background for understanding this paper. Section 3 describes the TB. Section 4 describes the case study of dynamic verification of defense simulation using the proposed TB and SES / MB and concludes in Section 5.

2. SYSTEM ENTITY STRUCTURE FORMALISM AND MODEL BASE

The SES formalism is that expresses the target system in the form of a total solution including all the alternatives of the target system using the tree structure. The SES formalism includes three pieces of information: configuration relation, classification relation, and connection relation of the target system. Configuration relation indicates that the system is composed of any model, classification relation indicates that the system is a model may have any Alternatively, the connection relation is expressed with respect to the connection relation of the model.

The SES formalism has three nodes to represent this information. First, entity node refers to models that make up the system and, aspect indicates whether the node is drawn is one of the model consists of any sub-model line was solid in the tree. A specialization node is a node that indicates which alternatives a model can have, drawn in solid double-dashed lines in the tree. The multiple aspect node, which is a type of aspect node, draws three solid lines in the tree, representing the subordinate models constituting a single model, all of which are the same kind of special aspect nodes.

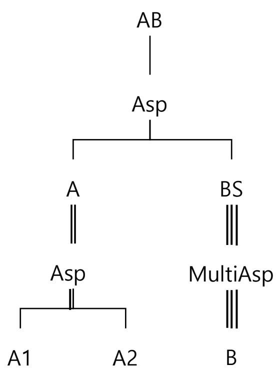
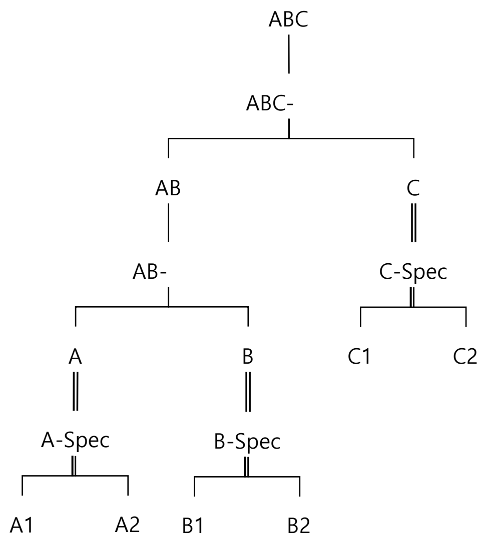


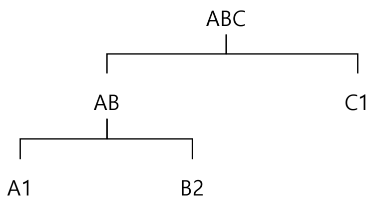
Figure 1 SES example

Above Figure 1 is a simple example to illustrate SES formalism. In the figure above, AB, A, BS, A1, and B correspond to entity nodes, and Asp and Spec correspond to aspect nodes and specialization nodes, respectively. MultiAsp is a multiple aspect node. In Figure 1, the system AB is represented by AES and BS. A has an alternative of A1 or A2. The BS consists of several B models.

The target system expressed through the SES formalism represents all the alternatives that the target system can have. The pruning process is required to re-express the model in a suitable model structure after selecting the appropriate nodes according to the simulation purpose. The result is a Pruned Entity Structure (PES).



(a) SES



(b) PES

Figure 2 SES pruning example

For example, if an arbitrary user conducts pruning to implement a system expressed as SES in Figure 2 (a), the PES in Figure 2 (b) is the result that selects A, B, and C as A1, B2, and C1, respectively. The PES can be generated variously according to the pruning algorithm or according to the user's selection in the pruning process.

MB is a database that manages the verified models through unit testing, and user can implement the PES as a real simulation model by taking each model constituting the PES from the MB and synthesizing them (Park, 1997; Zeigler, 1991).

3. TEST CASE BASE FOR DYNAMIC VERIFICATION OF DEFENSE SIMULATION MODEL BASED ON SYSTEM ENTITY STRUCTURE AND MODEL BASE

3.1. Test case Base

The TB proposed in this paper is a database that manages test cases for the dynamic verification of each model, which is corresponding to each node constituting the SES of defense simulation. TB manages test cases of models existing in MB, and test cases can be written in various formats. The user can use the test cases existing in TB in some ways, and can add/delete/edit test cases. However, to increase the reliability of TB, only the user who has privileges must be able to add/delete/edit test cases.

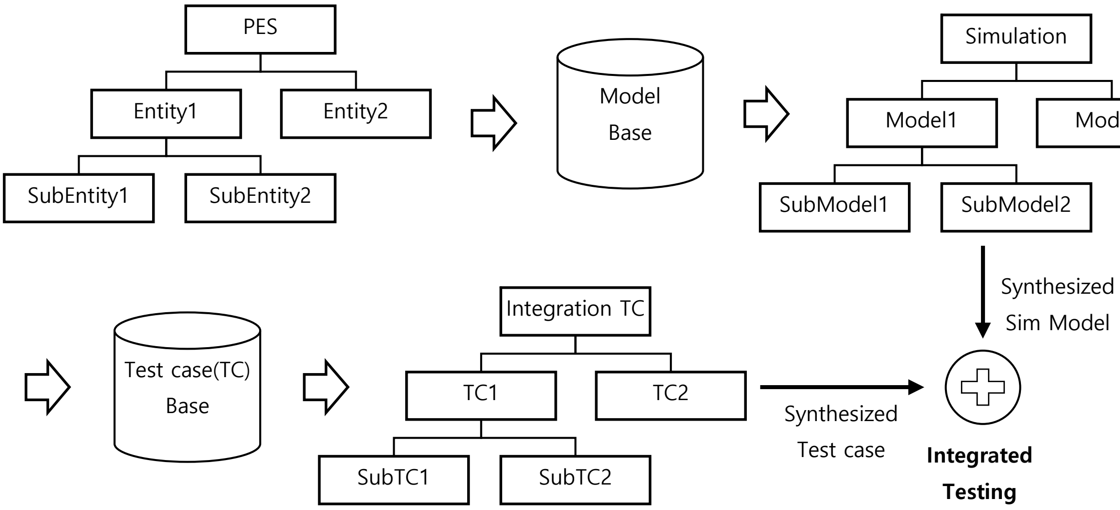


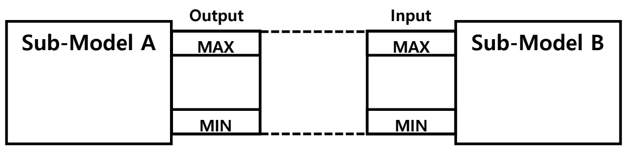
Figure 3 Blueprint for dynamic verification method

3.2. Test case Base

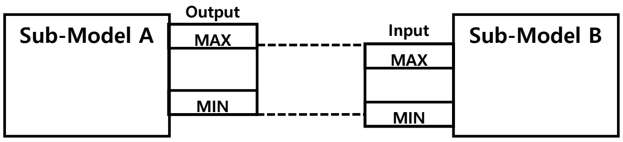
The dynamic verification of the defense simulation model consists of two parts: unit testing of the sub-models that make up the defense simulation model and integration testing of the synthesized model. Unit testing is necessary to ensure the reliability of the sub-models. Even if the defense simulation model is composed of verified sub-models, integration testing of the synthesized simulation model is required due to various errors that may occur during the synthesis process of the simulation model. (Gao, 2003)
As mentioned above, the target of the dynamic verification method using TB is the defense simulation model developed by using the defense simulation model management system using SES / MB. Therefore, it means that the sub-models constituting the developed defense simulation model are the models taken from the MB, and the MB is the database managing the verified models through the unit test, so unit testing for the sub-models do not need to be performed separately. Therefore, dynamic verification of defense simulation model using TB is integration testing of simulation model composed of verified sub-models.
Figure 3 shows the dynamic verification process using the TB of the simulation model developed using the defense simulation described in the SES formalism. The process of development and dynamic verification of simulation is as follows. The user takes the models corresponding to the entities of the PES generated through appropriate pruning process from the model base and synthesizes them to create a simulation model suitable for the simulation purpose. Then, we take the test cases corresponding to each model from TB and synthesize them or add user-defined test cases to create integration test cases. And we apply the integration test case to the synthesized simulation model to perform dynamic verification of the synthesized simulation model.

3.3. Use of Testcase Base

The defense simulation represented using SES is hard to describe the range of values of input and output of sub-models due to the limitation of representation of SES. Therefore, it was difficult to judge whether synthesis is possible between certain sub-models. For example, if two models are model A that outputs a value between 1 and 5 and model B that operates with input value between 6 and 10, the model A and B cannot be combined without a separate process of converting values. However, since the model A and the model B represented using the SES do not describe information about the input and the output range, the user cannot know that the model A and the model B can not be combined with each other. However, by using TB, we can reverse-trace the range of the output value of the model A and the range of the input value of the model B through the test case of the model A and the model B existing in the TB. Thus, TB can be used as a way of diagnosing the minimum possible combination between any sub-models.
In the process of importing and combining the test cases corresponding to each model from the TB, two cases occur depending on the attributes of the sub-models to be integrated. Figure 4 shows two cases that occur according to the input/output range of sub-models A and B when we try to make an integration test case for the integrated model AB by synthesizing the test cases of sub-models A and B.



(a) IO range matching case



(b) IO range not matching case

Figure 4 Two case of integrating sub-models

Figure 4 (a) shows the case where the output of sub-model A matches the range of input of sub-model B. Suppose that one of the test cases for sub-model A, A_Testcase$_n$, exists and its input and output are input$_n$ and output$_n$, respectively. At this time, A_Testcase$_n$ can be expressed as (A_input$_n$ / A_output$_n$). The output of A_Testcase$_n$, output$_n$, is the input of B, and the resulting output is AB_output$_n$. At this time, AB_Testcase$_n$, which is an integration testcase for the integrated model AB, can be defined as (A_input$_n$ / AB_output$_n$). Figure 5 is simplifying of above description.

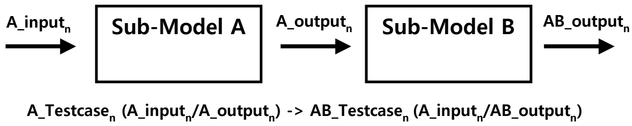


Figure 5 Integration test case generation in the case of IO matching case

Thus, in case of IO matching case, various integration test cases can be created by using existing test cases. Also, if the user determines that the generated testcase AB_Testcase$_n$ has a meaningful result, it can be registered in the TB and can be recommended as a testcase for the future integrated model AB.

Figure 4 (b) shows the case where the output of sub-model A and the input of sub-model B do not agree with each other. In this case, integration test cannot be performed using existing testcase stored in TB, so user must add user-defined integration test case directly.

4. CASE STUDY: DYNAMIC VERIFICATION O F SIMULATION MODEL MODELED IN ADD SIM ENVIRONMENT

As the case study of proposed dynamic verification method, dynamic verification of defense simulation model developed in AddSIM2.0 environment which is used at Korea Defense Science Institute proceeds.

The target model of this case study consists of 8 components in total and consists of aircraft component, detection radar component, launcher component, and four missile components. Simulation scenario is as follows: when the enemy aircraft is within the range of the allied detection radar, the allied detection radar detects it and transmits the coordinates to the launcher, and the allied launcher fires the missile to shoot down the enemy aircraft. In this case study, an integrated model combining the launcher model and the four missile models in the simulation model was verified.

The launcher model receives the location information and speed of the enemy aircraft as inputs and transmits the information received as input to the available missile model. And the missile model calculates and shoots the enemy aircraft based on the information of the enemy aircraft. The input and output formats of the launcher model and the missile model used in the case study are as follows.

- Launcher_input: (X, Y, Z, Vd)
  Launcher_output: (Mn, X, Y, Z, Vd)
- Missile input: (Mn, X, Y, Z, Vd)
  Missile output: ifSuccess

The output format of the launcher and the input format of the missile, Mn, are the id of the missile to which the enemy airplane is being hit. In this scenario, a total of four missile models are assigned, so Mn is a value between 1 and 4. Also, ifSuccess, the output format of the missile, compares the coordinate value of the enemy aircraft with the target coordinate value of the missile and applies the probability to indicate whether the enemy aircraft is shot down as True / False.

The target model of the case study belongs to the IO matching case described above because the output of the launcher and the input range of the missile are matched. The integration testcase is created as many times as the number of test cases for the launcher model as described in the previous chapter. For each generated integration testcase, an experiment that entering the input of the launcher according to the generated integration test case, recording the output of the missile using the probe code, and compare the recorded output of missile with the output of generated integration test case was performed n times.

As the value of n increases, the dynamic verification result of the target model for each integration test case converged to a similar value of the collision probability value specified by itself in the missile in the error rate of about 3%. Therefore, the integration testcase generated based on the existing testcase stored in TB is valid, and it means dynamic verification can be performed using TB.

5. CONCLUSION

In this paper, we propose test case base for dynamic verification of defense simulation model based on system entity structure and model base. The proposed TB can increase the reusability of the test cases of models by efficiently managing the test cases of the models developed for defense simulation. Also, the user can create various values by reusing test cases managed in TB. Specifically, we can perform dynamic verification by generating an integration test case for a specific case among synthesized simulation models by combining test cases of existing models managed through TB. Also, TB can be the basis for backtracking the attributes of the model that cannot be described due to the limitation of SES, and it is possible to determine whether synthesis can be performed between specific models using the properties of the backward traced model.

In this paper, only the verification of the case of implementing the defense simulation by synthesizing the model stored in the MB is performed. However, the test case stored in the TB can be usefully used even when the stored model is modified. For example, if the user has a model A stored in the MB and user want to verify it by defining a model A, which is similar to A, user can get a hint from A's existing testcase stored in TB. This will be discussed in future research.

ACKNOWLEDGMENTS
This work was supported by BSM-based verification and user code verification module(UD160075BD) funded by the Agency for Defense Development.

REFERENCES

T. G. KIM, C. LEE, E. R. CHRISTENSEN and B. P. ZEIGLER, "System Entity Structuring and Model Base Management," IEEE Transactions on Systems, Man and Cybernetics, vol. 20, no. 3, pp. 1013-1024, Oct., 1990.

Park, Hyu Chan, Wan Bok Lee, and Tag Gon Kim. "RASES: A database supported framework for structured model base management." Simulation Practice and Theory 5.4 (1997): 289-313.

Zeigler, Bernard P., Cheng-Jye Luh, and Tag-Gon Kim. "Model base management for multifacetted systems." ACM Transactions on Modeling and Computer Simulation (TOMACS) 1.3 (1991): 195-218.

Gao, Jerry, H-SJ Tsao, and Ye Wu. Testing and quality assurance for component-based software. Artech House, 2003.

REQUIREMENT SPECIFICATION METHOD BASED ON SES AND DEVS FORMALISM IN DEFENSE SYSTEM SIMULATION

Jiyong Yang[a], Jaewoong Choi[b], Inho Jung[c], Changbeom Choi[d]

[a],[b],[c] Dept. of Information and Communication, Handong Global University
Handong-ro 558, Pohang-si, Kyungbuk, Rep.of Korea
[d] School of Global Entrepreneurship and ICT, Handong Global University
Handong-ro 558, Pohang-si, Kyungbuk, Rep.of Korea

[a]21731002@handong.edu, [b]21631005@handong.edu, [c]21631004@handong.edu
[d]cbchoi@handong.edu

ABSTRACT

A precise analysis of the requirements in the development of the software and verification of the software is essential. Moreover, for classified solutions, the importance of requirements analysis is even greater. Defense simulation is a representative example of a classified solution. It is essential to set up appropriate requirements, analyze them, and perform verification in simulation configurations such as simulation resolution, simulation target setting, damage evaluation model, and engagement scenarios. However, the requirements of the existing defense system simulation are not specified in the requirement specification method based on the theoretical background, so it is hard to analyze through the formal approach. Therefore, this study establishes the basic structure of the simulation model used in the defense simulation configuration through the SES formalism, and introduces the specification method based on the DEVS formalism about the interaction between the models.

Keywords: requirement specification, system entity structure, discrete event system specification, base model for combat system

1. INTRODUCTION

The requirements analysis phase during the software development is a major step in establishing the overall design and the functions to be implemented. Therefore, it is necessary to accurately analyze and specify the requirements in the process to verify that the later developed products meet the requirements. Specification techniques are needed to perform a precise analysis and verification of the requirements. To make a formal specification of the requirements, the structure of the system should be represented by the formalism, and a mathematical model that can express the interaction of the components constituting the system is needed. The defense system is a complex system of sub-systems with various types of functions and structures depending on the experience of personnel, organization, equipment, materials, doctrine, military and commander. Therefore, in the case of defense system simulation, a methodology is required as a combination of elements that perform independent functions rather than a general simulation methodology. The defense system can be classified as a complex system.

To construct a defense system simulation, a particular role model is created as a combination of independent functional sub-models, and simulations such as engagement, detection, and reconnaissance are performed through a combination of models. Therefore, a complex system constructs a simulation based on a combination of previously defined models. Complex system simulation of models consisting of a mix of independent sub-models forms a hierarchical structure. In the case of such a complex system simulation, reusability becomes important because it is made up of a combination of sub-models, and a structural expression of the system for ensuring reusability is needed.

In this study, simulation of combat system participating directly in the engagement of the defense system is targeted. The basic model structure that is the basis of engagement is established through SES. SES is a formalism that defines a hierarchical structure based on a formalism and organizes the model based on reuse. The Pruned Entity Structure is derived through pruning, which is a process of obtaining one behavioral model through System Entity Structure which collectively represents the alternatives of one system. The basic model consists of a combination of a physical layer, a logical layer, and an information layer, which correspond to the Pruned Entity Structure, which is a behavioral model derived from the basic structure of System Entity Structure. In this study, the rules about the structural characteristics are formulated regarding inheritance and combination. The simulation is constructed based on the basic model and protocol thus established.

Therefore, this study establishes the basic model structure of simulation for the battle system which is directly applied to engage in the defense system through

System Entity Structure, and introduces Discrete Event System Specification formalism based on the part of the interaction between models. This paper introduces the protocols required for the simulation configuration, the specification of the XML-based model structure and the model generation based on the specifications. Finally, This paper introduces the case study of the simulation which is composed through the basic model and simulation configuration method finally introduced.

2. RELATED WORKS & BACKGROUNDS

This chapter introduces related research and background knowledge related to this study. The related work deals with DEVS-based combat modeling for engagement-level simulation and introduces SES formalism and DEVS formalism which are basic knowledge of this study.

2.1. Related Works

A study related to this study is called "DEVS-based combat modeling for engagement-level simulation" by Dr. Seo, Kyung-min. This research is based on the DEVS formalism for the basic model of the combat model participating in the engaging class simulation. The basic model of the combat is classified according to the platform and the weapon system, and the discrete event model and the object model. It is meaningful that DEVS is used to form a composite of a single model with a combination of sub-models by expressing the structure of the engagement model according to the level. The overall structure of the engagement system model is shown in Figure 1.

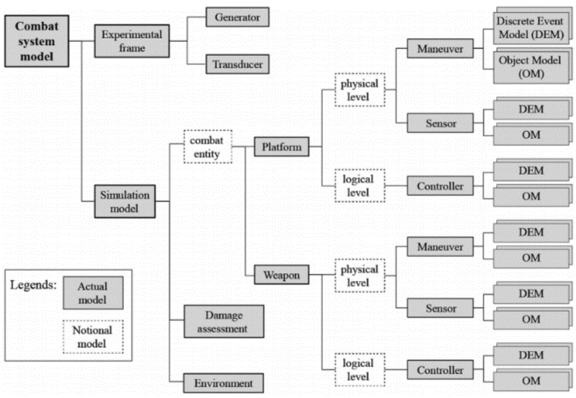


Figure 1: Overall model structure of combat system model

Figure 1 represents the decomposition tree for the overall model structure which systematically organizes a family of models.

2.2. Backgrounds

This section introduces DEVS formalism and SES formalism, which is related to this study.

2.2.1. DEVS

As it is well-known, the classical DEVS formalism can specify a system in two aspects: one for the behavior of a basic component, and the other for the overall structure of a system. An atomic DEVS formalism describes the behavior of a unit component not further decomposable, which consists of three sets and four functions.

$$AM = \langle X, Y, S, \delta ext, \delta int, \lambda, ta \rangle$$

, where

X : input event set,

Y : output event set,

S : sequential state set, and total state set

$Q = \{(s,e) \mid s \in S, 0 \leq e \leq ta(s) \}$,

$\delta ext : Q \to Q$:external transition function, for $\delta ext(s,e,x) = (s', e'), e < ta(s), e' = 0$

$\delta int : Q \times X \to Q$: internal transition function, for $\delta int(s, e) = (s', e'), e = ta(s), e' = 0$

$\lambda : Q \to Y$: output function, for$(s, e) \in Q, e = ta(s)$

ta : S → : time advance function, is the non-negative real number set.

There are two types of transitions of a model: 1) external transitions entailed by external events; and 2) internal transitions in the case of no event occurrence until current state sojourn time has elapsed. In the latter case, just before the internal transition, an output event is produced at the state. In analogy to the continuous systems, external transitions would correspond to the input is driven state
transition and internal ones the input-free state transition.

The coupled DEVS formalism specifies the structure of discrete event systems composed of components communicating with each other through event couplings,

$$CM = \langle X, Y, M, EIC, EOC, IC, SELECT \rangle$$

, where

X : input event set,

Y : output event set,

M : component model set, either atomic models or coupled models,

$EIC \subseteq CM.X \times \cup_i Mi.Yi$: external input coupling relation,

$EOC \subseteq \cup_j Mj.Yj \times CM.Y$: external output coupling relation,

$IC \subseteq \cup_j Mj.Yj \times \cup_i Mi.Xi$: internal coupling relation

SELECT : 2M — ø → M : select function

Notice that the coupled DEVS formalism above has the closure property, i.e., a coupled model may contain another coupled models as well as atomic models as its components. It captures the structure of a system, the components hierarchy and the interfaces between components. The SELECT function relates to the simultaneous scheduling problem of simulation that arranges the priorities of components when more than one component is to be scheduled at the same time.

2.2.2. SES

System Entity Structure is a representation scheme which contains the decomposition, coupling, and taxonomy information for a system(Zeigler, 1984). Formally, SES is a labeled tree with attached variable types that satisfy five axioms – alternating mode, uniformity, strict hierarchy, valid brothers, and attached variables. Detail description of the axiom is available in (Zeigler, 1984).

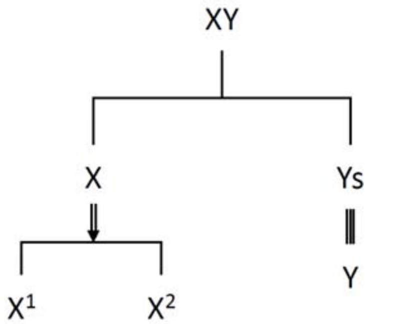


Figure 2: System Entity Structure example

There are three types of nodes in SES – entity, aspect, and specialization – which represent three types of knowledge about systems. The entity node, having several aspects and/or specializations, corresponds to a model component that represents a real world object. The aspect node (a single vertical line in the labeled tree of Figure 2) represents one decomposition, out of many possible, of an entity.

Thus, the children of an aspect node are entities, distinct components of the decomposition. The specialization node(a double vertical arrows in the labeled tree of Figure 2) represents a way in which a general entity can be specialized into special entities.

A multiple entity is an entity that represents a collection of homogeneous. We call such components a multiple decomposition of the multiple entity. The aspect of such a multiple entity is called multiple aspect(triple vertical lines in the labeled tree of Figure 2). Note that instead of presenting all Ys's components, only one B is placed in the labeled tree.

Pruning extracts a sub-structure of the SES by selecting one aspect and /or one specialization for each entity in the SES. The pruning operation also expands multiple entities as well as assigning values of attributes attached to entities in the SES.

3. REQURIREMENT SPECIFICATION METHOD IN DEFENSE SIMULATION

To specify the requirements, it is necessary to understand the domain. In the case of defense simulation, the simulation is made up of a combat system and a non-combat system. Different models are used depending on the resolution of the target simulation. The model of simulation differs in the way of expressing even the same object according to abstraction. Therefore, the process of constructing the simulation environment every time is a process of creating new models and combining them. Therefore, in this chapter, a basic model is set up for the engagement models participating directly in combat in the defense simulation system, and simulation configuration and requirement specification technique based on this model are introduced.

3.1. Base Combat Model for Requirement Specification in Defense Simulation

The basic model of the combat system, which is operated directly in the engagement of the defense system, can be composed of the information layer, the physical layer, and the logical layer. The information layer deals with the information related to the weapon system and manages information such as the characteristics of the weapon system, acquisition information, spatial information (position and attitude), movement path, attribute value, and scenario list. When generating the initial model, the values are initialized to a predetermined value and then act as a condition or restriction on the selection decision of the behavior of the model in the logical layer as necessary.

The physical layer is a part that manages the physical behavior or characteristics of the weapon system and consists of sub-models such as sensor detection, communication, movement, shooting, detection, and survival response. When creating a single engagement model, a physical layer is constructed using only the necessary sub-models. For example, in the case of a ground weapon system, there is no part of a flight, so there is no need to add a sub-model related to flight when constructing a real model.

The Logical Layer is a part that determines what action to take according to the characteristics of the weapon system based on the data from the information layer and is compatible with the sub-models constituting the Physical Layer. In other words, if the physical layer is composed of sensor sensing, communication, and movement models, the Logical Layer is composed of sensor sensing control, communication control, and movement control models.

The base model, consisting of three basic layers, is expressed on an SES basis. SES was introduced as a formalism that defined hierarchical models and organized them for reuse on a model basis. SES, which represents the alternatives that a system can have, is a tree structure expressed by three nodes: entity, aspect, and specialization. The basic model expressed by SES is shown in Figure 3.

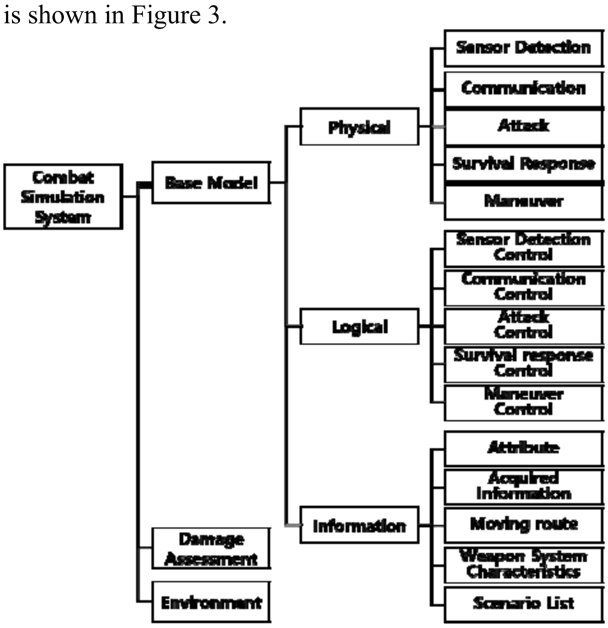


Figure 3: Representative structure based on the SES formalism

The base models that make up the engagement simulation have a multiple aspect relationships. And we can confirm that there are a Physical layer, a Logical layer, and an Information layer as sub-models that constitute Base Model. The pruned entity structure is derived from pruning process through basic SES, and the simulation is constructed through the combination of these.

In the process of constructing an engagement simulation, there are parts that cannot be expressed because they simulate the reality, and these parts are omitted as abstract values or considered as constant values. In other words, it is necessary to specify a rule for parts such as constraints because it is not enough to express the basic model using only SES. This part is considered in the information layer.

In the case of constructing the simulation through the modeling process based on the Base Model, it is advantageous that the reuse, requirement verification, and testing are easy in the multi-layer simulation environment because the role of the Base Model is sure to be performed even in the different environment.

3.2. Simulation Composition

The configuration of the simulation automatically generates and assembles the model according to a predefined protocol. This process allows specification of the requirements. Except for the performance part of the software in the process of specifying the requirements of the simulation, the process of modeling takes up most of the requirements specification process. That is, if the initial modeling process is described according to a predefined protocol, then the simulation is configured to satisfy the requirements.

In this study, model structure is established through XML. In case of atomic model based on DEVS formalism, model name, port information, state information, and attribute value are read through XML file. Coupled model reads coupling information of EIC(External Input Coupling), EOC(External Output Coupling), IC(Internal Coupling), etc., and reads information about sub-component.

Each PES from SES consists of hierarchically combined atomic models and coupled models. Since PES basically consists of physical, logical, and information layers, it has a tree structure with two depths. Thus, it forms a combined structure of complex coupled models.

The specification of the requirements is specified as a tag that enters the attribute in describing the DEVS model in XML. Because modeling is done within predefined protocols, modeling-related requirements are automatically satisfied if the initial modeling process describes XML with the protocol. For other additional requirements, we will proceed to validate the XML in the test scenario tag and verify that it is satisfied, which we will study later.

4. CASE STUDY

We developed a prototype program that automatically generates the model based on SES structure information in XML format. Information about the model was assigned to the attribute in the model tag in XML. The models are divided into an atomic model and coupled model based on DEVS. The program is based on C ++ and is implemented based on the DEVSim ++ module. An example of a basic model XML structure is shown in Figure 4.

```
<ModelStructure>
  <Model name="Outmost" type="Coupled" lib_path="./lib/outmost.so">
    <Ports>
      <Port type="input" name="in1"/>
      <Port type="output" name="out"/>
    </Ports>
    <Componets>
      <Compnent name="ABC" file_path="./model/abc.xml"/>
    </Componets>
    <Coupling>
      <Coupling from="Outmost" o_port="out1" to="ABC" i_port="in1" />
      <Coupling from="ABC" o_port="out1" to="Outmost" i_port="in1" />
    </Coupling>
  </Model>
</ModelStructure>
```

Figure 4: XML Base Model Structure example

Because it is based on DEVS, the type of model is composed of coupled and atomic, and consists of basic information such as <Ports> tag, <Components> tag, and <Coupling> tag which is necessary for a model. Coupled model consists of <Coupling> tag simply for EIC, EOC, IC. The state information of the atomic model, the attribute value, etc. are specified in the <Info> tag.

5. CONCLUSION

In this paper, we introduce the basic model used in the defense simulation environment based on the SES formalism and the DEVS formalism. Models created according to predefined protocols can satisfy the requirements in the modeling process if the XML file is described well in structure. Future work will be based on this study, unit testing according to the actual scenarios of the test scenarios, and testing to perform interactions between the models automatically.

ACKNOWLEDGMENTS

This work was supported by BSM-based verification and user code verification module(UD160075BD) funded by the Agency for Defense Development.

REFERENCES

Seo, Kyung-Min, et al. "DEVS-based combat modeling for engagement-level simulation." Simulation 90.7 (2014): 759-781.

IEEE, IEEE Standard for Modeling and Simulation(M\&S) High Level Architecture(HLA) - Framework and Rules, IEEE Std 1516-2000, 2000.

IEEE, IEEE Standard for Modeling and Simulation(M\&S) High Level Architecture(HLA) - Federation Development and Execution Process(FEDEP), IEEE Std 1516.3-2000, 2003.

T. G. KIM, C. LEE, E. R. CHRISTENSEN and B. P. ZEIGLER, "System Entity Structuring and Model Base Management," IEEE Transactions on Systems, Man and Cybernetics, vol. 20, no. 3, pp. 1013-1024, Oct., 1990.

Kim, Tag Gon, and B. P. Zeigler. "The DEVS-SCHEME simulation and modelling environment." Knowledge based simulation: Methodology and application (1990).

Tag Gon Kim and B. P. Zeigler, "ESP-Scheme: A Realization of System Entity Structure in a LISP Environment,"in Advances in AI and Simulation, Simulation Series, Tampa, USA, pp. 135 - 140, Mar., 1989.

Zeigler, Bernard P., Herbert Praehofer, and Tag Gon Kim. Theory of modeling and simulation: integrating discrete event and continuous complex dynamic systems. Academic press, 2000.

THE SOFTWARE FOR THE SECURITY MANAGEMENT IN THE SOFT TARGETS

Lucia Duricova[(a)], Martin Hromada[(b)], Jan Mrazek[(c)]

[(a),(b),(c)] Tomas Bata University in Zlin, Faculty of applied informatics

[(a)]duricova@fai.utb.cz, [(b)]hromada@fai.utb.cz, [(c)]jmrazek@fai.utb.cz

ABSTRACT

The soft targets and crowded places are defined as the objects or the places without the integration of the special security measures, which are oriented in opposed to the threats of the violent criminal attacks. These objects have the similar characteristics, which can be identified according to the features of the soft targets. One of the main features is that these objects have a lot of the uncontrolled visitors per day. The purpose of these objects is to provide the function or the services to the whole range of the visitors without monitoring access to these objects. The current system doesn't require the special security options for the soft targets. The recent security state of the objects is caused by different security and safety requirements to the objects. For example: need to buy a ticket to the football match, installed cameras for the monitoring stealing. This paper proposes the software which can evaluate the objects, calculate the current state of the object and defined preventive actions, which have to do with the object. The proposed software can also have the dynamic part, which can use the planned data and simulate the state of the objects.

Keywords: workstation design, work measurement, ergonomics, decision support system

1. INTRODUCTION

The soft targets or crowded places are the objects (buildings, areas, complex, free spaces and others), in which are concentrated a lot of people in the same place. These objects have not integrated the special security measures, which can secure the people before the violent criminal attacks. The violent criminal attacks can be defined as criminal attacks, which are oriented to human life and can cause the fatal consequences on population. The attack for these objects can cause the fatal consequences for the visitors but not only for them. The losses can be caused in the closed area of these objects too. The measures can increase the ability to secure the object or can increase the ability to react to the attack faster.

According to the **Marquise (2015)** are defined indicators of the attacks to the soft targets as:

- Airport terminals – vulnerable to insider threats, multiple locations to place explosives, open accessibility (outside security), large number of people in a small area, numerous targets and adjacent structures, size and remote location of some facilities.
- Shopping Malls – unrestricted public access, large number of access point, unrestricted access to adjacent buildings, access to suppliers, vendors and maintains workers, limited employee background checks, and limited security force.
- Stadiums/ Arenas – large number of people entering with varying levels of inspection, limited control of vehicles entering area, limited or no inspection of items carried in by event participants, vendors, contractors, limited facility security between events, large number of people at scheduled and publicly announced events.
- Convention Centres – open access, large urban locations, limited background checks on employees vendors, little or no screening of patrons, little standoff for parking areas, infrequent use of intrusion detection systems, deliveries are unmonitored.
- Hotels – unrestricted public access, unrestricted access to areas adjacent to buildings, limited employee background checks, limited security force, unprotected HVAC systems, and building designs are not security- oriented, multiple locations to place explosives or hazardous agents.

In the document Basics of Soft Targets Protection – Guidelines **Kalvach (2016)** defined the soft targets as:

- Schools, educational facilities, colleges, dining rooms, libraries.
- Church memorials and places dedicated to worship.
- Shopping centres and complexes.
- Cinemas, theatres, concerts halls, entertainment centres.
- Meetings, processions and demonstrations.
- Bars, clubs, discos, restaurants and hotels.
- Parks and squares, tourist sights and attractions, museums and galleries.
- Sports halls and stadiums.
- Major transport hubs, train and bus stations, airport terminal.

- Hospitals, polyclinics and other medical facilities
- Cultural and social events.
- Community centres.

According to these specifications, we can define a lot of kinds of criteria for the security analysis. However, if we want to develop the system, which can evaluate the degree of security of the object, then we need to know the threats. If we don't know what threat is, we can't know how the level of security is. We need to identify and understand the threats and according to them, we can define criteria. And according to the criteria, we can calculate the security level.
This paper is organized as follows. The second part describes the current global situation in security of private citizens. The next part is focused on the methodology of the proposed assessment of the soft targets. The concrete description of the specific and general analysis is in Section 4. Section 5 describes the proposed dynamical software support. And the last section of the paper explains the current results of the research and suggests the next development of the research.

2. THE GLOBAL SITUATION

According to the Global Terrorist Index, the private citizens and property is the group facing the highest number of deaths from terrorist attacks, although deaths from attacks on civilians and other non-governmental targets declined in 2015. We can see this progress in Figure 1.

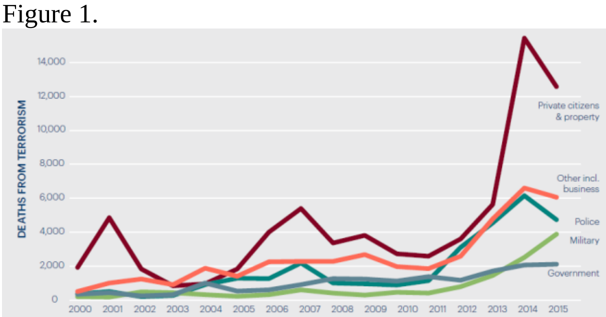


Figure 1: Sample Figure Caption

The majority of the attacks were against private citizens and property, accounting for 43 per cent of all deaths in 2015 according to the Global Terrorist Index 2016. According to these facts, we can constate that the risk of attack on human life is high. However, not only terrorism can endanger us and our life. The same characteristics of terrorism are very complicated defined. We can say, that Czech republic didn't have terrorist attack yet. However, in the last 3 years, we registered 2 violent attacks, which are described in Table 1. On the other hand, Virginia Fusion Center in document **Educational Facilities Threat Assessment (2008**) mentioned, that the schools and educational facilities were targeted in 285 terrorist attacks in 2008, representing 3% of the targets attacked. But according to **Crockett (2005)** other types of attacks were not classified as terrorist but as violent and criminal attacks. We can suppose, that number of the violent and terrorist attacks is higher than the only number of the terrorist attacks. As a result of these facts, we need to develop the system, which can analyze the object features according to the closely specified threats.

Table 1: The numbers of the attacks in Czech Republic

Type of object	Town	Year	Number of death / number of injured	Method
School	Žďár nad Sazavou	2014	1 / 0	Attack by knife
Restaurant	Uherský Brod	2015	8 / 0	Shooting

According to the global situation, but according to the situation in Czech republic too, we need to develop the methodology or system, which can have the significant impact on the security situation in these objects. This methodology and system are based on the definition of the potential attackers and threats. The potential attacker can be defined as the attacker, who want to attack to human life. The system will be developed for the assessment the current situation in the object in relation to violent criminal or terrorist attack. The assessment has some differences in evaluated parameters according to the chosen threats.

3. THE METHODOLOGY OF THE ASSESSMENT

The next part of the paper describes the basics of proposed methodology on which software is based. In the proposed phase of the research, we build on the authors **Chunlin (2012)** and **Raspotnig (2013).** The first part of the paper is aimed to the primary process, which is applied to the decision making process.

3.1. The Assessment of the State of the Security

The whole process of the assessment is based on the simple process, which is described in Figure 2. The process of the assessment is divided into 5 steps.

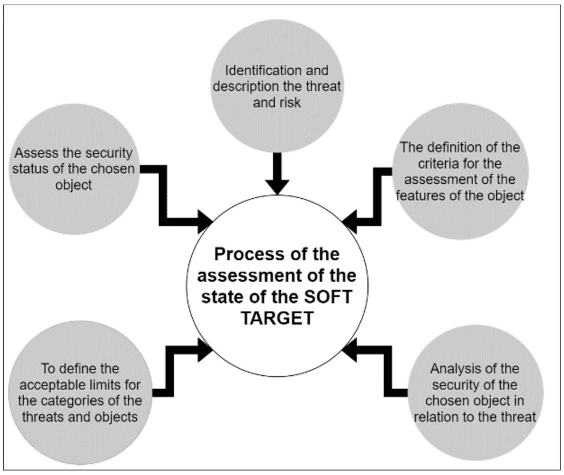


Figure 2: The Process of the Soft Target Assessment

The assessment of the state of the soft targets should be defined according to the assessment of the security in relation to the threat. The first step of the process in Figure 2 is the identification and the description of the threat and the risks. The methodology can examine the state of the objects according to the identified threat and risk. The examination of the state of the object is realized according to the second step of the process and it is the definition of the criteria for the assessment. If the system knows how the features of the target are then the system can evaluate current state. The aim of the whole process is step number 3. The third step is analysis. The whole process of the analysis is described in the next section of the paper. The next step needs to know acceptable limits for the decision making process. The system can make the decision according to the limits for each of these analyzed data. The last step is to define the final assessment. If the current security state is ok, then the system can propose the recommended actions. On the other hand, if the current state is not ok, then the system defined immediate corrective actions and permanent corrective actions.

The identification and description of the threats and risks can be seen in Figure 3. Each of these threats can be described according to the scenarios. The next step is based on the evaluation and the calculation the risk of these defined scenarios.

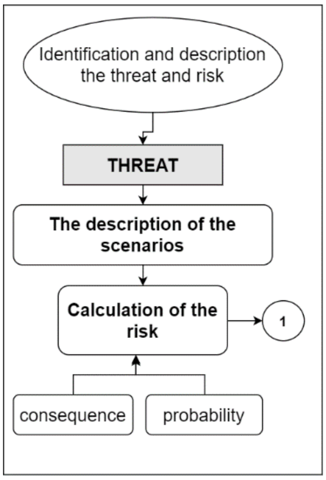


Figure 3: The Description of the Threat

The process continues into the step of the definition the criteria, which can be seen in Figure 4. The criteria examine the significant features of the object. The significant features are chosen according to the previous step – according to the chosen the threat.

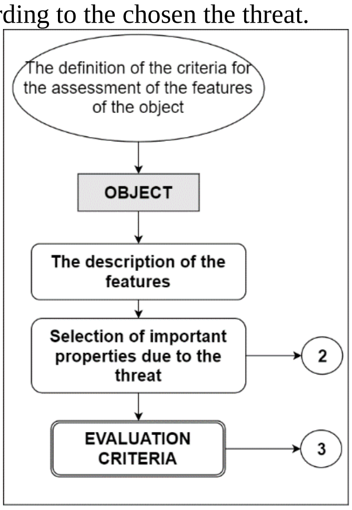


Figure 4: The Definition of the Criteria

The next step of the process is the analysis. The analysis is based on the chosen threat (previous step) and defined criteria, which examine the features of the building. The basics of this step are described in Figure 5.

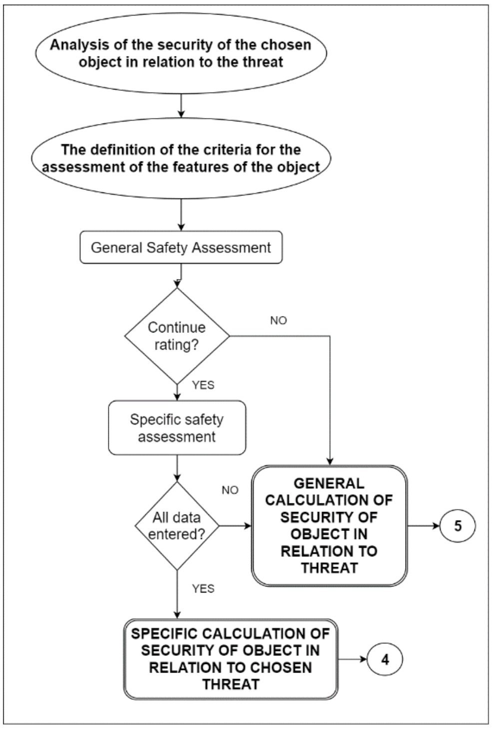


Figure 5: The Analysis of the Object Security

The system analyses the features of the object according to the threat, but the system doesn't know if the result is ok or not. The definition of the result is based on the next step, which is described in Figure 6.

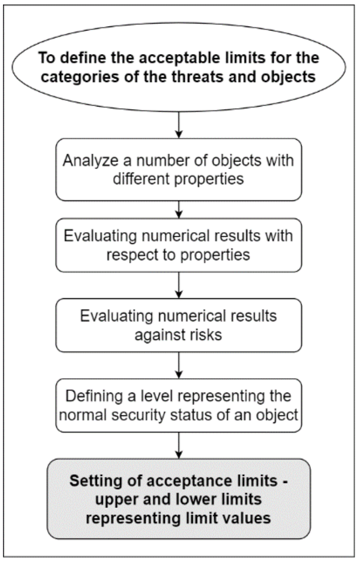


Figure 6: The Definition of the Acceptable Limits

The system can assess the result of the analysis according to the limits. We called these limits acceptable limits. If the system compares evaluated value with the required value or acceptable value, then the system can make the final decision. The process of the final decision is described in Figure 7.

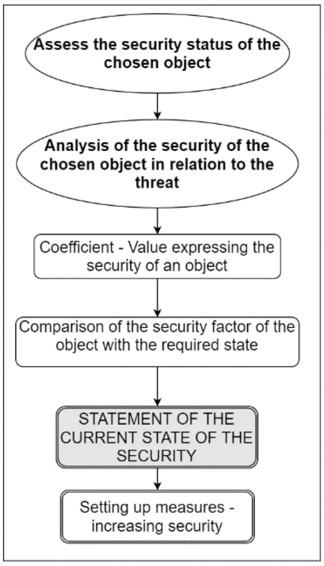


Figure 7: The Final Assessment of the Security

The last step of the proposed methodology is to define the final result about the security situation in the evaluated object according to the chosen threat. The important step of the process is to define the acceptable limits for each of these threats. The system can make a decision according to these limits, however, we need to examine a lot of objects if we want to know these limits for each of these threats.

4. THE GENERAL AND SPECIFIC ANALYSIS

The concrete assessment is based on the general and the specific analysis. The general analysis is faster than specific but not as exact as the specific analysis. The process of the general analysis can be seen in Figure 8. In Figure 8 can be seen also the process of the assessment. This assessment is based on the calculation of the risk (chosen threat), the selection of the important properties and features according to the threat, and the next is the definition of the evaluation criteria. The last step is to set the weights of criteria according to the threat. The aim of this process is the calculation of the general security coefficient.

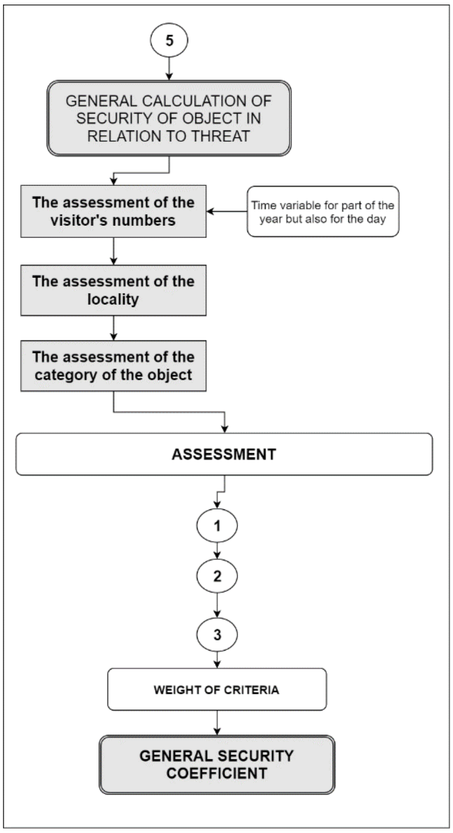


Figure 8: The General Assessment

The process of the calculation of the general security coefficient is based on three types of general coefficients, which must be defined in each type of the analyzed object. In Figure 9 can be seen the process of the calculation the visitor's number factor. This process is based on the general information about locality and at the attractivity of the city and place, where the object is situated. However, we can use the better methods for more exact monitoring the number of visitors in the object. In the software is proposed to use the monitoring of sim cards (mobile telephone sim cards). This monitoring can provide for the software very exact and actual data about the number of visitors in the object. The data from this method of monitoring are not to the public access, and that is the problem, why we don’t use this type of data in the proposed meteorological standard. However, in the dynamical software tool must be used this type of data.

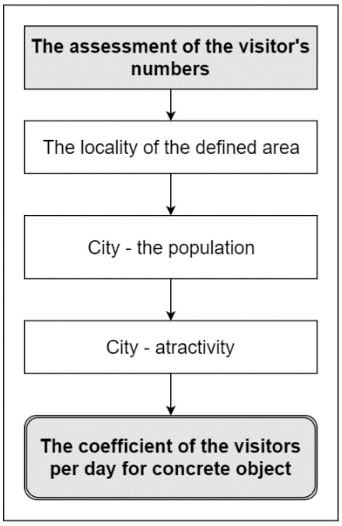


Figure 9: The Assessment of the Visitors

In Figure 10 can be seen the process of the assessment of the locality. The first step is to select the locality and identifies the interested objects. The next step is to identify the coefficients of these interesting objects. This coefficient expresses the impact on the security on the other objects in the same locality. For example, the localization of the synagogue, localization of the ministries, or other objects, which are attractive for a potential attacker can have the significant impact on the security of the object.

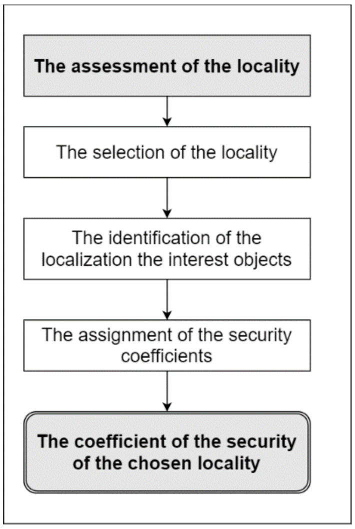


Figure 10: The Assessment of the Locality

In Figure 11 can be seen the process of the assessment of the category of the object. This process is the important part of the general security assessment. According to the other similar objects, we can predict the results of the specific concrete assessments.

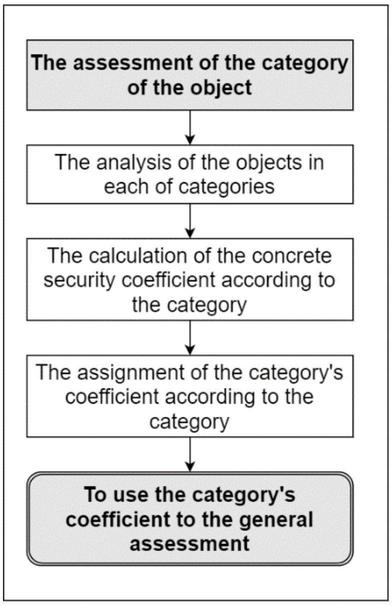


Figure 11: The Assessment of the Category

The next process of the identification of the chosen threats is described in Figure 12. This process describes the assessment of the threats and evaluates the impact on the evaluation criteria. The system can evaluate the state of the object according to the threat. The system has to know, which the features are important for the decision making and evaluation process.

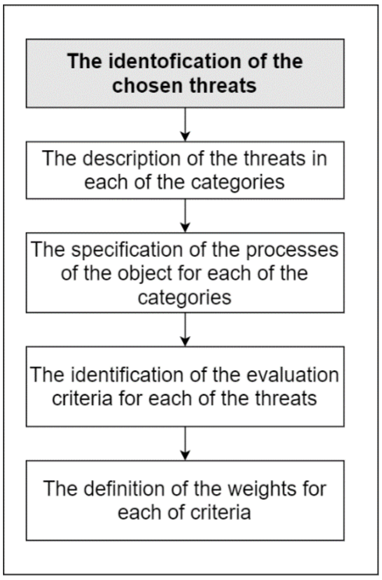


Figure 12: The Identification of the Chosen Threats

The main exact assessment is described in Figure 13. This assessment is based on the exterior, interior and processes analysis. In each of these categories are chosen the criteria according to the chosen threats.

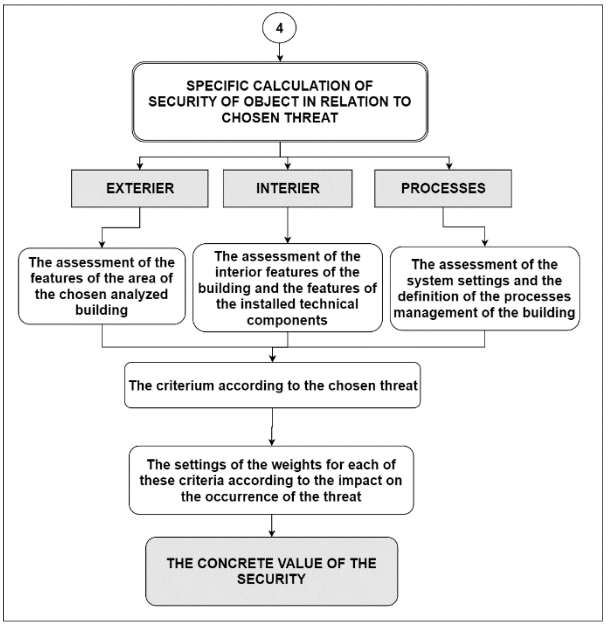


Figure 13: The Specific Assessment

The whole process of the assessment can be seen in Figure 14. We need to observe, that the proposed software solution is oriented only to the buildings and centers of the buildings. Although, we suppose the assessment of the events will be developed in the future.

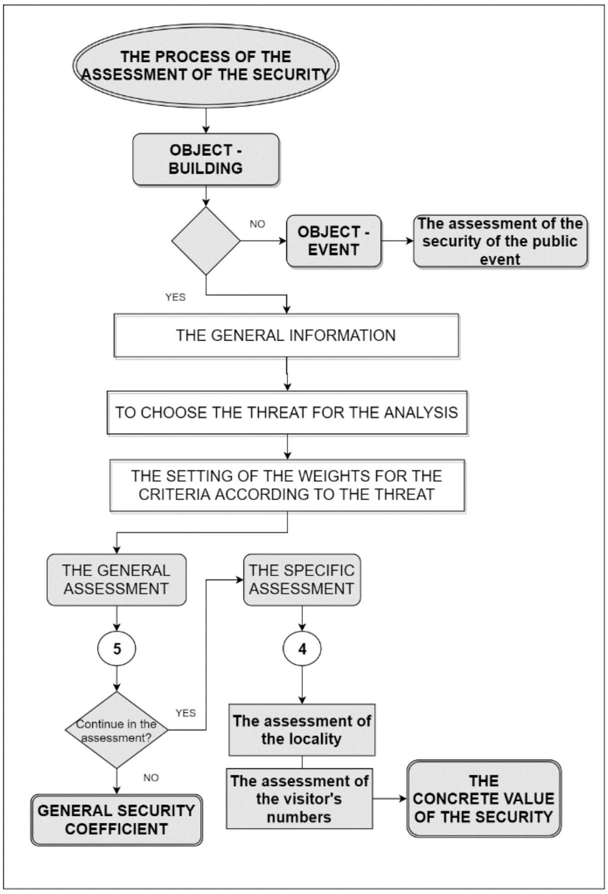


Figure 14: The Final Assessment of the Security

The whole methodological standard was developed and verified in this part of the research. However, in this phase of the research, we need to verify and confirm the proposed methodological standards on the different kinds of objects and different kinds of threats too. The first results influenced proposed software, but we suppose that the system can be also changed in the next verified phase of the research.

5. THE DYNAMICAL ASSESSMENT OF THE STATE OF THE SECURITY

In this research, we cooperate with Ministry of interior and Integrated Rescue System of Czech republic too. These institutes sense need to develop this proposed software as a dynamic software tool, which can help our society with the prediction of the risks and also with the proposal the effective immediate and permanent corrective actions. The process of the function of the dynamical assessment can be seen in Figure 15.

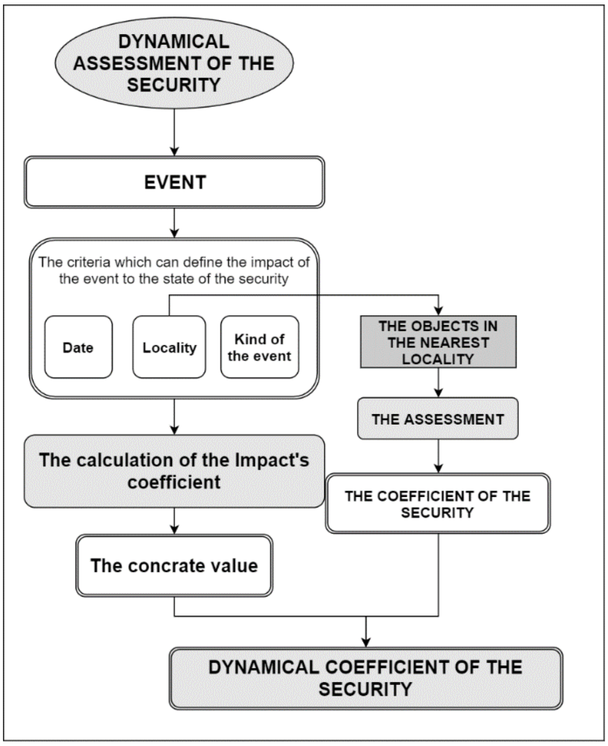


Figure 15: The Proposed Dynamical Assessment

The system needs to know the information about the planned event. This information can yield the data about the impact of the event on the other objects security in the same area. Between this information belong date, locality of the event and closed area of the event, and kind of the event. Kind of the event is closely related to the risk value of the threats. The kind of the events can have the impact on the probability of the security incidents.

6. RESULTS OF THE CURRENT RESEARCH

The current state of the research is to repair weaknesses of the programmed software, which can automatically analyze the soft targets according to the proposed methodological standard. Although, this current software has not programmed analyzes according to the chosen threats. This software is done for the verifying the proposed methodological standard and to analyze the objects according to this mathematical operations. Each of these proposed criteria evaluates the state of the features of the analyzed object. The evaluation process is aimed at the correct methods for the transformation the features to numerical values.

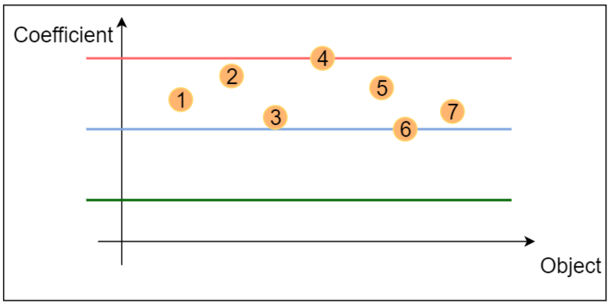


Figure 16: The Proposed Dynamical Assessment

In Figure 16, we can see the results of the static security analyses. This type of the analysis is the general. That means, this analyzes is not realized according to the chosen threat, but only according to the chosen criteria. In Figure 16, we can see three types of the lines. The green line marks the optimal security coefficient, the blue line marks the normal security coefficient and the red line marks limited security coefficient. The objects belong to the same category of the objects. We suppose, that we need to know these free types of the limits (lines) for each of defined threats. As a result, we can consider, that dynamical software can provide us more specific data. The system of the dynamical software can examine the progress in the time and the decisions can be more exact as in the static part.

7. CONCLUISION

This research is aimed to the assessment the security of soft targets. The current research was aimed to the building the system and methodological standard to the automatic software analyzes. The first phase of the verification optimized the criteria and mathematical processes for the relevant results. The next part of the research will be aimed at the verification whole range of categories of the objects. However, we need to define the threats and according to the threats, we need to select the relevant criteria. One of the last challenges for the research is to develop the dynamical part of the software and to develop the system, which can automatically propose the corrective actions to the soft targets.

ACKNOWLEDGMENTS

This project is realized as the research with a doctoral student, and it is the primary input for next research, which we will develop in next term. It was realized with the support of the university. This work was supported by Internal Grant Agency of Tomas Bata University under the project No. IGA/FAI/2017/017 and by the research project VI20172019073 "Identification and methods of protection of Czech soft targets against violent acts with elaboration of a warning system", supported by the Ministry of the Interior of the Czech Republic in the years 2017-2019 and by the research project VI20172019054 "An analytical software module for the real-time resilience evaluation from point of the converged security ", supported by the Ministry of the Interior of the Czech Republic in the years 2017-2019.

REFERENCES

Marquise R.A., 2015. Preventing Soft Target Terrorist Attacks.

Kalvach Z., 2016. Basics of Soft Targets Protection – Guidelines.

Global Terrorism Index 2016, 2016, Institute for Economics and Peace, pp.26.

Europion Union Terrorism Situation and Trend Report 2017, 2017, European Union Agency for Law Enforcement Cooperation 2017, pp. 5-17.

Educational Facilities Threat Assessment, 2008, Virginia Fusion Center.

Crockett S., et al. Protecting Soft Targets: A Case Study and Evaluation of High School Security, 2005.

Alexander F., 2012. Safety and Security Target Levels: Opportunities and Challenges for Risk Management and Risk Communication, In: International Journal of Disaster Risk Reduction 2: pp. 67-76.

Chunlin L. et all., 2012. The Security Risk Assessment Methodology, In: Procedia Engineering 43, pp. 600-609.

Raspotnig Ch. and Opdahl A., 2013. Comparing Risk Identification Techniques for Safety and Security Requirements, In: Journal of Systems and Software 86 (4): 1124-1151.

AUTHORS BIOGRAPHY

The authors work as research scientists on several projects under security research of ministry of interior (Czech Republic). Simultaneously, authors also works on the projects for commercial sphere in category of the soft targets.

KNOWLEDGE ACQUISITION VERSUS SKILL DEVELOPMENT: PEDAGOGICAL CONSIDERATIONS FROM AN INTELLIGENT TUTORING PERSPECTIVE

Benjamin Goldberg[(a)], Benjamin Bell[(b)], Debbie Brown[(c)]

[(a)]U.S. Army Research Laboratory—Human Research and Engineering Directorate
[(b,c)]Eduworks Corporation

[(a)]Benjamin.s.goldberg.civ@mail.mil, [(b)]Benjamin.bell@eduworks.com, [(c)]Debbie.brown@eduworks.com

ABSTRACT
Intelligent Tutoring Systems (ITS) are traditionally developed within cognitive domain spaces that associate with problem solving and procedural application of knowledge. As such, much of the research and literature on pedagogical management in ITS environments is confined to managing cognitive impasses and misconceptions as they relate to declarative and procedural information within the context of a scenario/problem. A gap in the literature that needs to be addressed is how best to apply ITS methods in a skill development domain (i.e., a domain incorporating psychomotor task elements that require consistent and precise execution to meet task objectives). In this paper we present pedagogical guidelines as they relate to ITS coaching in a psychomotor task environment, and how those guidelines are informed through theoretical underpinnings of skill acquisition and techniques of practice. These guidelines will inform the design of skills-oriented tutoring systems and are useful constructs in defining requirements for skill-oriented ITS authoring tools.

Keywords: intelligent tutoring systems, psychomotor, skill acquisition, feedback, pedagogy

1. INTRODUCTION

Recent interest in the application of Intelligent Tutoring Systems (ITS) in psychomotor skill domains has raised fascinating questions centered on pedagogy. With a preponderance of ITS literature situated within cognitive domains that focus on problem solving and knowledge application, much of the findings with respect to instructional management do not apply outside those problem types. For this reason, it is important to re-conceptualize the role of ITS pedagogy in skill domains that involve muscle movement and hand-eye coordination.

In this paper, we tease apart the components as they relate to knowledge acquisition versus skill development, and how the identified distinctions should drive pedagogical configurations as they adhere to instructional design and learning theory. Following, guidelines will be presented for feedback management and coaching in psychomotor skill related domains. These guidelines will be informed initially through theoretical underpinnings of kinesiology, sports psychology, and cognitive psychology. The outcome will be a set of competing pedagogical approaches that vary the level of support an ITS will provide, with distinctions in learner individual differences driving which approach to enact. These approaches will ultimately require empirical evaluations to gauge their utility in applied settings.

2. KNOWLEDGE ACQUSITION VERSUS SKILL DEVELOPMENT

For the purpose of this discussion, we distinguish between *knowledge acquisition* and *skill development*.

2.1. Knowledge Acquisition Basics

In its most basic theoretical form, knowledge acquisition associates with the process of perceiving, processing and storing new information in memory. Moreover, this involves the ability for knowledge retrieval when the situation warrants its application (Baddeley 2004).

Knowledge is typically categorized as declarative (memory in the form of concepts, facts, or episodes) or procedural (ability to perform tasks through proceduralized associations), with an encoding in memory that proceduralizes knowledge following multiple applicable uses of that information (Anderson 1982). To further break down knowledge acquisition, Nunes and Karpicke (2015) present the following guidelines based on how knowledge is organized and how learning materials should be created: (1) process material semantically, (2) process and retrieve information frequently, (3) learning and retrieval conditions should be similar, (4) connect new information to prior knowledge whenever possible, and (5) create cognitive procedures as procedural knowledge is better retained in memory and more easily accessible. These principles associate with how information is effectively stored in memory for easy retrieval when the time warrants its application. The development of human expertise is attributed to knowledge stored within schemata where simple ideas are combined or chunked into more complex ones, not through the processing and arrangement of elements unorganized within long-term memory (Van Merriënboer and Sweller 2005).

Based on these foundations, ITSs are traditionally developed through the application of models that target the domain knowledge space based on the declarative and procedural properties required to solve defined problem sets. There are numerous modeling approaches applied over the years (e.g., Bayesian, neural nets, etc.; see Pavlik, Brawner, Olney and Mitrovic 2012 for a thorough review), each designed for the purpose of tracking learner progression and to identify any impasses or misconceptions in their procedural understanding of that topic space. In addition, these systems are designed with pedagogical underpinnings driving problems/scenario selection and guidance/feedback variations based on individual differences during a practice event (Kulik and Fletcher 2016, Woolf 2009). What is missing in the ITS community is an extension of these methods to account for physical properties of a domain space that associate with skill development, both from the modeling perspective, which is required to inform contextualized assessments to drive feedback and remediation, along with pedagogical considerations that adhere to learning psychomotor skills that go beyond a cognitive understanding of what to do.

2.2. Skill Acquisition Basics

As a defining characteristic, we associate physical skill development as an interplay between cognitive understanding (declarative and procedural) and psychomotor application, whereas the utility of skill requires gross control of fine motor movements when performing a task. This association is critical when conceptualizing the role of ITS in military contexts, as majority of the tasks performed across the spectrum of warfare incorporate physical interactions, with varying degrees of complexity, frequency and duration (Department of the Army 2011).

There are common tenets expressed in the literature associated with learning a new physical skill. The first and foremost is that experience and practice trumps all, with a definitive association between time spent in practice and improvement in proficiency (Côté, Baker and Abernethy 2007). However, simply practicing a skill over multiple repetitions seldom leads to expert performance. Without formalized structure around the type of practice and the feedback received, performance eventually plateaus below what is considered optimal/expert (Ericsson 2008).

Developing a new skill follows three primary phases of acquisition: (1) cognitive novice phase where an individual tries to understand the cognitive and physical requirements of the activity to generate actions while avoiding errors; (2) the associative intermediate phase where focused attention on task performance is no longer required and noticeable errors become increasingly rare; and (3) the autonomous expert phase where the execution of a skill becomes automated with minimal cognitive and physical effort (Goldberg 2016, Fitts 1967). From the coaching perspective, the most critical stage for directive instruction is in the initial cognitive phase where mental models are being established that link motor control to objective outcomes. How can an individual modify or reinforce behavior if there is no way to effectively link actions to performance? During this critical phase of learning, behavioral tendencies are established and schemas are formed in memory, making feedback to instill proper habits critical. In the traditional sense, a coach/instructor with knowledge in the domain will closely observe a learner, identify errors in his or her behavior as determined by a performance outcome, and provide feedback to correct errors and reinforce proper technique. This process is repeated until evidence is acquired that supports stable development of skill execution. The goal is for an ITS to mimic this interplay in an automated fashion.

2.2.1. Deconstructing a Skill into Fundamental Components

Utilizing technology to facilitate these described inference procedures is challenging. It requires perceptual oriented data corresponding to behaviors that an expert human would assess, and the application of models to determine how the captured data relates to a representation of desired behavior. This proposed process warrants a deconstructed task analysis to break a skill set down into a hierarchical structure of varying fundamental concepts and procedural applications.

This representation is critical as it becomes the foundation by which an ITS is developed. In other words, identifying the fundamental components of a task, and the skill sets required to successfully perform the task, informs the design of instructional materials and practice opportunities that target specific subsets of skills to drive an individual's acquisition curve. This top-down approach also establishes criteria for measuring performance and designating thresholds for gauging success.

In terms of relating what's already been discussed to a real-world example, take the domain of rifle marksmanship. When someone is attempting to learn how to shoot a rifle for the first time, the initial approach to instruction is focused on a set of fundamentals. These fundamentals provide a foundation of required skills to successfully perform as an elite marksman. In this instance you can decompose marksmanship into four physical fundamental skills: (1) body stability, (2) breathing, (3) breath control, and (4) site alignment. Each of these break-down further into a set of sub-skills that ascend in complexity as you progress through practice opportunities (e.g., body stability in prone, body stability while kneeling, body stability while standing, body stability while on the move, etc.).

The desired end state is the development of muscle memory to automatically perform a task without dedicating cognitive function to make it happen. When you establish automated execution of fundamental behaviors, then an individual can progress to more complex scenarios requiring advanced application of a skill (e.g., hitting a moving target). This is followed by practice opportunities to combine the application of disparate skill sets (firing a rifle and communicating tactical decisions) for more enriched scenarios to better instill autonomous execution.

This analogy can associate with almost all psychomotor domains of instruction, regardless if its association with job-related activities or athletics. Each domain can be deconstructed into a set of fundamental components that are performed when a situation warrants their execution. The goal of an automated ITS is to establish models of fundamental behaviors to make the assessment space manageable and to establish training and practice opportunities that target the development of specific fundamental skill sets.

2.2.2. The Link between Practice and Skill

While there is a large base of research focused on understanding how conditions of practice, including variability, distribution, and segmentation influence performance and retention of skill (see Lee, Chamberlin and Hodges 2001 for a thorough review), each practice event provides critical data points from which initial ITS development should be based.

Anders Ericsson's theory of deliberate practice highlights the following attributes of an effective practice event: (1) the event is designed to improve performance; (2) the individual has the ability to repeat the application over multiple trials; (3) the task requires high mental engagement; and (4) feedback is continuously made available that is designed to serve in a coaching capacity (Ericsson, Krampe, and Tesch-Römer 1993; Ericsson 1996). While these attributes provide generalizable guidelines when constructing specific scenarios and interactions, they do not provide measureable constructs that can be used to guide actual scenario development. Combining these guidelines with a task analysis provides the pedagogical building blocks to configure specific interactions to be managed by the ITS. The goal is to balance challenge with guidance, as outlined in the Zone of Proximal Development (Vygotsky 1987), so as to promote skill development that retains over time and transfers to novel situations.

In addition, there are three further distinctions that must be addressed to support self-regulated deliberate practice events: (1) the content and material applied in support of the practice events, (2) the sequence of interactions leading up to and following a practice event, and (3) the coaching approach applied within that practice event. As the concept of psychomotor ITSs are framed within a self-regulated learning model (Department of the Army 2011), the initialization of a practice event is preceded by the delivery of upfront instructional materials that prepares an individual learner for the set of tasks they will be asked to perform. In the following sections, we highlight the role of ITS technologies to facilitate the sequence of interaction that focuses on self-regulated skill development. We present the considerations in place that are applied to configure the pre-practice materials (i.e., multimedia content and other forms of instruction), the practice event itself with real-time assessment and feedback functions, followed by an after-action review component that serves as a form of remediation.

3. ITS APPLICATION FOR SKILL DEVELOPMENT

Each factor presented above is critical when determining the implications of using ITSs to replace human counterparts to train psychomotor skills. They should act as guiding principles when establishing practice opportunities within an ITS framework, whereas each attribute serves as a validation check to ensure the training experience is grounded in human performance heuristics. To base the discussion, we present design considerations as they adhere to the U.S. Army Research Laboratory's Generalized Intelligent Framework for Tutoring (GIFT; Sottilare, Brawner, Goldberg and Holden 2013).

3.1. Generalized Intelligent Framework for Tutoring

GIFT is an open-source architecture project with an evolving set of standardized software modules for authoring and configuring adaptive training materials across an array of supported training applications. Each application represents externally developed systems designed for educational and training purposes, each integrated through GIFT's gateway module for assessment and pedagogy practices. GIFT provides the ability to apply common modeling techniques against application generated data to inform real-time assessments in any supporting system, including game-based applications (Goldberg, Brawner, Holden and Sottilare 2012, Shute, Ventura, Small and Goldberg 2013) and now psychomotor supported simulation environments (Bell, Brown and Goldberg 2017, Goldberg, Amburn, Ragusa and Chen, 2017).

With integration technologies in place, GIFT also provides the mechanisms for configuring a learner's experience within and across disparate training applications within a single lesson experience (Goldberg, Davis, Riley and Boyce 2017). In this instance, GIFT provides a platform to enable the interchange of multiple training environments to instruct a single set of concepts. This highlights the open robustness of the architecture in that it can be repurposed for any conceivable domain, assuming the content and assessments are in place to support lesson creation. This also shows how GIFT promotes interoperability between platforms, as the lesson could incorporate multiple training environments to manage the attainment of multiple complimentary skill sets.

A driving requirement within the GIFT program is providing a set of tools to Army training developers that enable the creation of customized ITS applications outside of laboratory settings. For this purpose, a focus of the GIFT program is establishing pedagogical frameworks and authoring workflows that guide the development experience and ensures sound learning science principles are being leveraged. This entails applying instructional management principles and theories into the interactions supported when authoring lessons in the GIFT authoring tool. A GIFT lesson is configured at two levels: (1) the macro level where a structure of GIFT course objects are established that dictates the overall sequence a learner will experience, and (2) the micro level where each course object is configured to designate the interactions within, the assessments applied, and the pedagogical logic designated for managing assessment outcomes. Each of these associations will be discussed, with a focus on extending GIFT's pedagogical logic to support theory derived from skill acquisition and psychomotor learning.

3.2. ITS Pedagogy at the Macro Level

As an organizing function, instructional design models of psychomotor skill progression should be incorporated to guide the interactions and practice events managed by the ITS. A model is required to structure and guide configured interactions where specific pedagogical considerations can be authored within. The goal is to apply a theoretically based schema that dictates what type of lesson/training material should be applied at what time in support of self-regulated skill development.

An example of this nature developed for cognitive associated problem spaces is GIFT's Engine for Management of Adaptive Pedagogy (Goldberg et al. 2014, Goldberg, Hoffman and Tarr 2015). The EMAP serves as the first domain-independent closed-loop model where an instructional design approach outlined in David Merrill's component display theory was modeled as an organizational schema to reference learning interactions (see Figure 1). In its most simplistic form, the EMAP guides an author in establishing content to cover the declarative and procedural rules of a domain with examples provided, recall oriented questions to act as checks on learning to confirm comprehension, and practice opportunities supported by external applications with assessment and guidance configured. This structure is encapsulated in GIFT's Adaptive Courseflow course object.

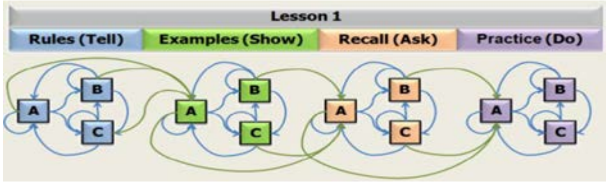


Figure 1. GIFT's EMAP highlighting the sequence and flow of instruction and remediation.

In Brown, Bell and Goldberg (2017), the authors review five fundamental theories related to psychomotor learning, with search results identifying a synthesized theoretical model to feed ITS development efforts (see Table 1). The project, the Psychomotor Skills Training Agent-based Authoring Tool (PSTAAT), is leveraging this synthesized model to build a new course object designed for psychomotor skill domains. The PSTAAT applies workflows to guide an ITS developer in configuring materials within the model components of observation, imitation and practice, with adaptation practices enabled throughout.. Based on architectural foundations established during EMAP development, the PSTAAT serves as an abstraction of that model in that it applies the theoretical framework described in Table 1 as the building blocks for system configuration. To simplify the cognitive load with authoring a psychomotor ITS, the PSTAAT incorporates an intelligent agent that provides guidance and examples during the design and configuration process with built-in tools to establish modeling techniques for assessment purposes.

Table 1. Synthesized skills domain model, derived from Brown, Bell, & Goldberg (2017)

Level	Definition	Example
Observing	Active mental attending of a physical event.	The learner watches a more experienced person. Other mental activity, such as reading may be a part of the observation process.
Imitating	Attempted copying of a physical behavior.	The first steps in learning a skill. The learner is observed and given direction and feedback on performance. Movement is not automatic or smooth.
Practicing	Trying a specific physical activity over and over.	The skill is repeated over and over. The entire sequence is performed repeatedly. Movement is moving towards becoming automatic and smooth.
Adapting	Fine tuning. Making minor adjustments in the physical activity in order to perfect it.	The skill is perfected. A mentor or a coach is often needed to provide an outside perspective on how to improve or adjust as needed for the situation.

ISBN 978-88-97999-90-4; Bruzzone, Cayirci and Sottilare Eds.

3.3. ITS at the Micro Level

At runtime, an ITS operates on three primary models: (1) a model of the task being performed, (2) a model of the learner being instructed, and (3) a model of pedagogy and instructional practice as it relates to the domain and learner being instructed. The lynch-pin of these systems is the ability for an ITS to gauge performance as it relates to the assessments being captured during interaction. The pedagogical reasoning in an ITS is dependent on the information it has as it relates to the context of the scenario being performed, and the environment from which the interaction is taking place (Goldberg 2017).

As mentioned above, each domain can be deconstructed into a set of fundamental components that are performed when a situation warrants their execution. The goal of an automated ITS is to establish models of fundamental behaviors to make the assessment space manageable. While the assessment space of a domain is defined around a set of concepts and objectives, it is inherently dictated by the data one can collect.

3.3.1. Understanding the Assessment Space

There are two fundamental applications of domain modeling in ITS environments. These approaches establish how learner interaction will be assessed, and how a system responds pedagogically is dependent on how the assessment is implemented. In the area of ITS for skill development, the two modeling applications we focus on are (1) expert models and (2) buggy-libraries.

The goal of these models is to classify real-time performance; with a granular enough representation to dictate what objective and/or concept requires feedback and/or remediation.

3.3.2. Expert Models vs. Buggy-Libraries

Expert models are statistical representations of ideal behavior within a designated scenario or problem. In cognitive problem-spaces, common models are based on model-tracing approaches that can track a learner's actions as they relate to procedural steps, and identify impasses and misconceptions when a deviation from a desired path is recognized. Modeling techniques continue to evolve over time, with new methods capturing more diagnostic information to better inform feedback practices, but these methods do not translate to physical task spaces where assessment is not bound by well-defined procedural steps.

For psychomotor skill domains, current practices of expert modeling leverage task performance data feeds collected within a simulated or augmented environment (Goldberg 2016). The model is dependent on the data produced during the interaction and the characteristics of the task being performed. In initial efforts, the assessment space was based on descriptive models of behavioral signals over specified windows of time surrounding a designated event within the practice scenario (e.g., modeling trigger data for two seconds leading up to a shot). Models are established on expert data sets, with thresholds established based on observed model properties.

Current efforts to support these models in GIFT require configurations of sensor inputs to inform assessment state representations. The sensor inputs are mapped to fundamental concepts that data feed supports. To aid in this authoring process, PSTAAT provides a simpler abstraction of the sensor configurations, so that one could adjust the performance threshold expectations at different phases of instruction. This supports transition between imitating/practicing as well as could enforce difficulty level if the activity warrants it as an individual progresses in their skill acquisition curve.

As the initial expert models generated can be used to dictate what an individual is doing differently, these models do not have the ability to accurately determine what is truly causing differences in performance, which limits the system's ability to provide detailed feedback.

3.4. Building a Closed Loop System

With assessment models in place, the goal is to develop a completely closed-loop system that supports self-regulated skill development based on individual performance and their progression within a skill development curve. The assessment models are designated to specific application oriented exercises that specify what action a learner is to take. These actions are represented through an ontological breakdown of the task features and their derived concepts. The models designed should map directly to each concept representation, so that during the execution of a task, the system is able to monitor and grade performance as it relates to the standards built within the models' scoring logic. This domain representation is established and ontologically translated to the remaining ITS modules used to drive coaching and personalization. With performance scores available, these states are communicated to a learner model to create a full picture of the learner state (i.e., includes information as it pertains to performance, individual differences, and affective response if available). A performance state update, classified as above-expectation, at-expecation, below-expectation, and unknown at a concept by concept basis, are the events communicated to the learner module. The system is observing transitions in performance so as to drive pedagogical interventions when shifts in learner state trigger predefined actions.

With a closed-loop capability in place, and a theoretical foundation to configure lesson interaction, the next function in developing a generalized GIFT course object for training psychomotor skills is establishing pedagogical principles that drive personalization practices. The goal is to create training experiences that adhere deliberate practice. In the remainder of the paper, we will present guidelines, along with a use case of application, that synthesize elements of deliberate practice with the synthesized model of psychomotor skill development to create a highly adaptive training experience that personalizes feedback and remediation based on real-time performance and individual characteristics.

3.5. ITS and Post-Exercise Remediation

Following the completion of a problem, scenario, or exercise, a well-designed ITS contains the assessment and performance data required to drive post training events that target deficiencies in ability. In essence, a system can prescribe follow-on exercises that directly instruct specific concepts or misconceptions detected during a trainee's interaction.

Current components in GIFT support customized remediation paths based on logic configured within the EMAP. Development efforts are extending this rule-based approach by implementing stochastic modeling processes that adhere to an additional model of interaction based on Chi's (2009) learning activity framework that categorizes activities as being constructive, active, or passive (CAP). This framework is providing the theoretical foundation to apply Markov Decision Processes (MDPs) for the purpose of targeting individualized learning materials following an assessment event based on individual differences and the needs of a given trainee (Rowe et al. 2015). In an effort to use this existing infrastructure, an additional goal of the PSTAAT tool is to incorporate the CAP components in a physical skill domain, which will require reconceptualization on the type of activities to configure based on task characteristics and the environment from which the training event takes place.

4. PEDAGOGICAL GUIDELINES FOR SKILL DEVELOPMENT

There are two levels of instructional considerations when developing an ITS for a psychomotor skill domain. These include the application of pedagogical design theories that adhere to ways in which individuals learn a new skill, and the use of sports psychology and coaching practices that associate with improvement-oriented feedback and motivation management. To guide the discussion within this section, we will present a use case to ground the defined guidelines with contextualized examples.

4.1. The Adaptive Marksmanship Trainer (AMT)

As a means to provide a grounding function for the guidelines presented below, we will use the AMT use case to give specific examples of pedagogical strategy implementations. The AMT is a well-suited exemplar, as it has all the components in place to provide individualized assessment on the functional elements of firing a rifle (Goldberg, Amburn, Ragusa and Chen 2017, Department of the Army 2016). This capability is enabled through sensor technologies embedded within a simulated carbine rifle (e.g., barrel movement, trigger pressure, and breathing waveform), and associated models of expert behavior as deemed through a cross-fold validation descriptive modeling technique. With a closed-loop function, the AMT can collect shooter performance and behavior data in real-time, process that data to compute behavioral metrics, and apply those metrics within established models to determine if particular behaviors were being erroneously executed.

Table 2. Pedagogical guidelines embedded within synthesized skills domain model to inform personalization practices.

	Skill Acquisition Phases		
	Cognitive	Associative	Autonomous
Observing	Learner observes expert task performance with provided narrative on fundamental application of behavior.	Learner observes expert task performance and describes behavior in own words.	Learner observes own behavior and provides subjective interpretation of technique and application.
Imitating	ITS provides focused practice opportunities designed to elicit imitation of specific micro-behaviors (e.g., exercise focused solely on trigger control, without any other component of marksmanship addressed)	Imitation phase for associative learner is focused on process incorporating all fundamental behaviors required to complete task. Focus is on task set-up, rather than execution.	Imitation phase for autonomous learner is represented in practice, where trend analysis is used to determine consistent imitation of proper technique while practicing.
Practicing/ Adapting	The practice/adapt loop is the primary component of an ITS pedagogical function. Coaching is process-oriented, with focus on repetition and exhibited consistency through data-driven assessment.	Coaching is error-based, as procedural process is in associative phase. Adaptation is specific to behavioral models, with buggy-library as preferred mechanism.	Coaching is error-based, but provided in After-Action Review format, with focus on behavior trends as learner observes their own performance outcomes.

Current logic in GIFT can monitor behavior on a shot-by-shot basis, where performance is derived on the completion of a 5-shot group (i.e., the trainee executes five consecutive shots with the goal of striking the same location on a target with each round). Following the 5-shot group GIFT computes performance measures and classifies behavioral assessments (i.e., breathing, trigger control, and body stability). If performance is below a designated threshold, the behavioral models are applied to decide on which element of the task to remediate. The current AMT will provide feedback on a selected fundamental, with the same material provided to each trainee, regardless of experience and current skill level. An overarching goal is to extend the current AMT to support a more robust self-regulated experience. This includes applying the PSTAAT task model to drive training interactions, and to develop further pedagogical guidelines that personalize and adapt a training event as skill progression is observed.

4.2. Coaching Considerations for Psychomotor Skills

A major distinction this work addresses is the role pedagogy plays in skill development across the three phases of acquisition. Just as in any instructional setting, the level of support and challenge of the task should be adjusted to the needs of a given individual. In this instance, learning a psychomotor skill involves a number of elements, including: (1) knowing at a declarative and procedural level what one should do, (2) knowing at the application level how to apply physical actions to meet a task objective, (3) making the link between physical action and cognitive understanding, enabling an adaptation loop on mental schemas as one continues to perform a task, and (4) managing emotional and affective states related to performance outcomes and task characteristics/dependencies (Colvin 2008, Gladwell 2008).

From here, rather than apply common scaffolding techniques applied in ITSs, we employ a psychomotor pedagogical model grounded in sports psychology. This theoretical basis introduces new techniques related to the development of expertise, with elements of scheduling, coaching, and repetitive practice opportunities (Ericsson 2008). While these high-level descriptors are defined in common human performance oriented literature, we focus on a further layer of decomposition as it relates to coaching in an ITS. For this purpose, we focus on two forms of feedback support, process-oriented and error-oriented. Following the breakdown of these support functions, we describe how all the macro- and micro-adaptive functions are combined, along with rules for skill development progression, to create an advanced psychomotor training experience using ITS technologies.

4.2.1. Process-Oriented Coaching

The foundation of a process-oriented coach is focused on instilling a proceduralized set of fundamental actions that should be conducted when performing a task. From a pedagogical perspective, this approach to instruction is most appropriate for novices learning a new set of skills. It is in this instance where performers are most likely to commit an abundance of errors as it relates to the process of executing a task. While an ITS should be developed with models in place to detect errors in skill performance, during initial skill acquisition, the primary focus of instruction should be centered on process and technique, rather than reactive to specific errors being observed. Rather than the system directing the learner's attention to a specific violation, the ITS should log the error information but apply pedagogy that reinforces specific fundamental principles that a learner needs to master.

In the example of the AMT system, there are multiple models in place to determine if a learner is properly executing fundamental behaviors while firing a rifle. While these models include diagnostics to provide focused coaching on specific errors, this approach to pedagogy should be postponed until that learner is deemed to be in the associative phase of acquisition. For novice learners, the models might determine that the individual has poor body stability and is quick on the trigger. While the ITS can provide feedback directly linked to those errors, the approach applied by expert instructors aims to instill process over behavior. The errors in fundamental are noted, but the coaching approach focuses at the process level of doing everything correctly in unison rather than focusing on a specific element of the task.

4.2.2. Error Response-Oriented Coaching

Error-response oriented coaching is a form of feedback support that is complementary to process-oriented coaching. This kind of feedback is activated by deviations detected during skill performance, and it is intended to draw a user's attention specifically to that error. This approach to coaching is most effective in the associative phase of skill development (see Table 2), as mental schemas of proper execution are already established, yet errors in technique are consistent. In these instances, it is important to focus attention on specific elements of behavior. Rather than harp on procedure, feedback should be applied to link a specific function of their behavior with the observed performance outcomes. The goal is to create micro corrections in their application so as to build a connection between physical action and cognitive understanding (Kim, Dancy, Goldberg and Sottilare 2017). This varies from process-oriented coaching as it requires specific feedback strings or content (e.g., videos, slides, etc.) that is linked to each fundamental component modeled within that task. It is in this portion of the coaching paradigm within ITSs where the CAP approach described above can be instantiated within.

4.3. Bringing it All Together

With a synthesized task model of psychomotor learning in place, along with mechanisms to track progress within training as it relates to the phases of skill acquisition, and some high-level pedagogical guidelines based on process and error, the building blocks are in place to create a customized training experience at the individual level.

This is captured in Table 2, with specific references within that associate with the AMT use case. Balancing observation, imitation, and practice with varying degrees of complexity as one progresses provides the foundation for any psychomotor ITS.

Moving beyond the AMT use case, there are some assumptions related with the implementation of the approaches described. The primary assumption is that the task environment produces data sources at a granular enough level to drive assessment methods on fundamental components captured within a task analysis.

5. CONCLUSIONS

Rebranding ITS applications to support physical skill development in psychomotor domains requires new pedagogical principles to drive instructional design and coaching methods. In this paper, we present initial high level guidelines related to activity types and feedback strategies as they relate to a synthesized model of skill development and the activities that occur within. While research is required to provide empirical evidence of these guidelines, they provide a good starting point to establish a theoretical basis around. An advantage of applying these methods within GIFT is the nature in which GIFT can serve as a testbed to configure controlled experimental conditions to determine pedagogical effect on performance and skill retention.

REFERENCES

Anderson, J. R., 1982. Acquisition of a cognitive skill. Psychological Review, 89, 369-406.

Baddeley, A. D.,1995. The psychology of memory. In: A. D. Baddeley, B. A. Wilson, F. N. Watts eds. Handbook of Memory Disorders. New York: Wiley and Sons, 3-25.

Bell, B., Brown, D. and Goldberg, B., 2017. Focused authoring for building GIFT tutors in specialized domains: a case study of psychomotor skills training. Proceedings of 5th Annual GIFT Users Symposium, Orlando, FL.

Brown, D., Bell, B. and Goldberg, B., 2017. Authoring Adaptive Tutors for Simulations in Psychomotor Skills Domains. Proceedings of MODSIM World 2017. Virginia Beach (Virginia, USA).

Colvin, G., 2008. Talent is overrated: what really separated world-class performers from everybody else. London, England: Penguin Group.

Côté, J., Baker, J. and Abernethy, B. 2007. Practice and play in the development of sport expertise. In: R. Eklund and G. Tenenbaum, eds. Handbook of Sport Psychology. Hoboken, NJ: Wiley, 184-202.

Department of the Army, 2016. Rifle and carbine training circular (TC 3-22.9). Washington, D.C.: Army Publishing Directorate.

Department of the Army, 2011. The U.S. Army Learning Concept for 2015. Fort Monroe, VA.

Desmarais, M. C. and Baker, R. S., 2012. A review of recent advances in learner and skill modeling in intelligent learning environments. User Modeling and User-Adapted Interaction, 22(1-2), 9-38.

Ericsson, A., 1996. The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games. New York, NY: Psychology Press.

Ericsson, K. A., 2008. Deliberate practice and acquisition of expert performance: a general overview. Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine, 15(11), 988-994.

Ericsson, K. A., Krampe, R. T., and Tesch-Römer, C. ,1993. The role of deliberate practice in the acquisition of expert performance. Psychological Review, 100 (3), 363-406.

Fitts, P. M. and Posner, M. I., 1967. Human Performance. Belmont, CA: Brooks/Cole Publishing.

Gladwell, M., 2008. Outliers: The story of success: Hachette, UK.

Goldberg, B., Brawner, K. W., Sottilare, R., Tarr, R., Billings, D. R. and Malone, N., 2012. Use of Evidence-based Strategies to Enhance the Extensibility of Adaptive Tutoring Technologies. Proceedings of Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), Orlando (Florida, USA).

Goldberg, B., Brawner, K. W., Holden, H. and Sottilare, R., 2012. Adaptive Game-Based Tutoring: Mechanisms for Real-Time Feedback and Adaptation. Proceedings of DHSS 2012, Vienna (Austria).

Goldberg, B., Hoffman, M. and Tarr, R., 2015. Authoring instructional management logic in GIFT using the engine for management of adaptive pedagogy (EMAP). In: R. Sottilare, A. Graesser, X. Hu and K. Brawner, eds. Design Recommendations for Intelligent Tutoring Systems: Authoring Tools: Volume 3. Aberdeen Proving Grounds: U.S. Army Research Laboratory.

Goldberg, B., 2016. Intelligent Tutoring Gets Physical: Coaching the Physical Learner by Modeling the Physical World. Proceedings of 10th International Conference on Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience-Vol. 9744. July 19-21, Toronto (Ontario, Canada).

Goldberg, B., Davis, F., Riley, J. M. and Boyce, M. W., 2017. Adaptive Training across Simulations in Support of a Crawl-Walk-Run Model of Interaction. Proceedings of the 11th International Conference on Augmented Cognition. July 12-14, Vancouver (British Columbia, Canada).

Kim, J. W., Dancy, C., Goldberg, B. and Sottilare, R., 2017. A Cognitive Modeling Approach-Does Tactical Breathing in a Psychomotor Task Influence Skill Development during Adaptive Instruction? Proceedings of International Conference on Augmented Cognition, Vancouver (British Columbia, Canada).

Kulik, J. A. and Fletcher, J., 2016. Effectiveness of intelligent tutoring systems: A meta-analytic review. Review of Educational Research, 86(1), 42-78.

Nunes, L. D. and Karpicke, J. D., 2015. Retrieval-based learning: Research at the interface between cognitive science and education. In R. A. Scott & S. M. Kosslyn (Eds.), Emerging trends in the social and behavioral sciences. West Lafayette, Indiana: John Wiley & Sons, Inc.

Pavlik, P., Brawner, K., Olney, A. and Mitrovic, A., 2012. A review of student models used in intelligent tutoring systems. In: R. Sottilare, A. Graesser, X. Hu and H. Holden, eds. Design Recommendations for Intelligent Tutoring Systems - Volume 1: Learner Modeling. Aberdeen Proving Grounds, MD: U.S. Army Research Laboratory, 39-67.

Rowe, J., Mott, B., Lester, J., Pokorny, B., Peng, W. and Goldberg, B., 2015. Toward a modular reinforcement learning framework for tutorial planning in GIFT. Proceedings of Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium (GIFTSym3). Orlando (Florida, USA).

Shute, V., Ventura, M., Small, M. and Goldberg, B., 2013. Modeling student competencies in video games using stealth assessment. In: R. Sottilare, A. Graesser, X. Hu and H. Holden, eds. Design Recommendations for Intelligent Tutoring Systems, Volume 1: Learner Modeling. Aberdeen Proving Grounds, MD: Army Research Lab, 141-152.

Sottilare, R., Brawner, K. W., Goldberg, B. and Holden, H., 2013. The generalized intelligent framework for tutoring (GIFT). In: C. Best, G. Galanis, J. Kerry and R. Sottilare, eds. Fundamental Issues in Defense Training and Simulation. Burlington, VT: Ashgate Publishing Company, 223-234

Van Merriënboer, J. J. G. and Sweller, J., 2005. Cognitive load theory and complex learning: Recent developments and future directions. Educational Psychology Review, 17(2), 147-177.

Vygotsky, L., 1987. Zone of proximal development. Mind in society: The development of higher psychological processes, 52-91.

Woolf, B. P., 2009. Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning. Burlington, MA: Morgan Kaufmann.

AUTHORS BIOGRAPHY

Benjamin Goldberg is a member of the Learning in Intelligent Tutoring Environments (LITE) Lab at the U.S. Army Research Laboratory's (ARL) Human Research and Engineering Directorate (HRED) in Orlando, FL. He has been conducting research in intelligent tutoring for the past eight years with a focus on adaptive learning in simulation-based environments and how to leverage Artificial Intelligence tools and methods to create personalized learning experiences. Dr. Goldberg holds a Ph.D. from the University of Central Florida in Modeling & Simulation.

Benjamin Bell is the president of Eduworks Government Solutions and an expert in the application of artificial intelligence to decision support, simulation, training, and human-machine interaction. He has practiced in the field of AI for over twenty years, leading funded research and development in applied settings, largely for military applications. He holds a Ph.D. from Northwestern University and is a graduate of the University of Pennsylvania.

Debbie Brown is a software engineer and senior learning technologist at Eduworks Corporation focusing on web application implementations that use AI and machine learning to semi-automate user-centered workflows and enable advanced adaptive eLearning tools and experiences. She has been conducting eLearning R&D projects for 20 years, and operating as a software engineer in government, academic, and workforce settings for 30 years. She holds an MS in Instructional Systems and a BS in Computer Engineering from Mississippi State University.

MODELING AND SIMULATION OF FORCED POPULATION DISPLACEMENT FLOWS USING COLORED PETRI NETS

Jenaro Nosedal-Sanchez[(a)], Ali Asgary[(b)], Adriano O. Solis[(c)], Francesco Longo[(d)]

[(a),(b),(c)] School of Administrative Studies, York University, Toronto, Ontario, Canada
[(d)]DIMEG, University of Calabria, Rende (CS), Italy

[(a)]jnosedal@yorku.ca, [(b)]asgary@yorku.ca, [(c)]asolis@yorku.ca, [(d)]francesco.longo@unical.it

ABSTRACT
In this paper, we present a discrete event model, developed and implemented using colored Petri nets. The proposed model uses empirical data distributions to simulate the forced population displacement flows observed during the Mosul military operations in 2016-2017. The simulation model allows estimation of the magnitude and destination of the flows within identified routes (between districts in Iraq) and the simulation results appear to reproduce the aggregate displacements actually experienced.

Keywords: colored Petri nets, stochastic modeling, discrete event simulation, forced population displacement, Mosul

1. INTRODUCTION

An internally displaced person (IDP) is someone who has moved within the bounds of his or her own country, as a result of threats due to natural or man-made events such as earthquakes, floods, pandemics or war.
In June 2014, the Islamic State of Iraq and the Levant (ISIL), also known as the Islamic State of Iraq and Syria (ISIS), invaded and took control over the city of Mosul (Figure 1) in Northern Iraq. Since then, the city had become a major hub for terrorist activities and subsequently became ISIS' de facto capital in Iraq (Reddy 2017). Mosul's population, among the most diverse in Iraq, includes substantial populations of Sunni Arabs, Assyrian Christians, Kurds, Yazidis, Turkmen, and other minorities.
After the invasion of Mosul by ISIL/ISIS in 2014, millions of Iraqi people have been displaced. In the case of Mosul and adjacent areas, a large fraction of the displaced population have decided to flee from their homes to other areas inside the country (including other areas within Mosul).
Iraqi armed forces started the military operation to retake the city of Mosul on October 16, 2016 (BBC.com 2016; CNN.com 2016). At that time, there were about 1.5 million people living in the city.
The Displacement Tracking Matrix (DTM) emergency tracking (ET) system for Iraq is used by the International Organization for Migration (IOM, formerly the UNHCR), to track "sudden displacement or return movements triggered by [the Mosul crisis]" (http://iraqdtm.iom.int/EmergencyTracking.aspx).

Figure 1: City of Mosul, Northern Iraq *(encircled in red)*

2. DATA SET AND ANALYSIS

Based on the information available from the ET system, it is possible to track the evolution of the displacements in terms of number of people displaced on a specific date and the accumulated number of IDPs.
Figure 2 provides a view of the displacements observed over the time interval from October 27, 2016 to May 04, 2017 (observations are reported every 3-4 days on average, twice a week). A quick glance at the slope of the *SUM* function seems to provide a clear and predictable pattern of accumulated IDPs. However, by checking exclusively the accumulated IDPs, the magnitudes of the variations are diminished. Measuring the variations (IDP_t – IDP_{t+1}) provides a view of the fluctuations between successive observations (number of IDPs from date/observation t to $t+1$), represented as N.
These fluctuations in population displacements over time are triggered by specific circumstances, such as changes in the current situation in terms of security and

other social, political and demographic indicators (Davenport, Moore, and Poe 2003; Shellman and Stewart 2007; Asgary, Solis, Longo, Nosedal, Curinga, and Alessio 2016). In certain situations, the identification of appropriate explanatory variables and development of causal models may be possible (e.g., Collmann, Blake, Bridgeland, Kinne, Yossinger, Dillon, and Zou 2016; Engel and Ibáñez 2007). However, as in the current case, there is an inherent complexity in identifying relevant explanatory variables and having access to their values over sufficiently long time periods.

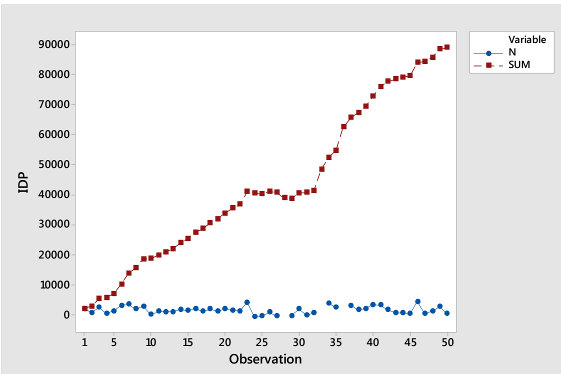


Figure 2: Number of Persons Displaced (per Observation and Cumulative)

In this paper, rather than focusing on the causal approach, the model is developed using a stochastic approach, in order to reproduce the evolution and uncertainty of displacements over time.
Figure 3 shows the distribution of values of *N*, while Figure 4 shows the cumulative distribution function (CDF). Besides the variability observed for *N*, in this case it is also critical to determine the direction of the displacement described by the origin and destination of the individual that decides to flee (in terms of routes taken).

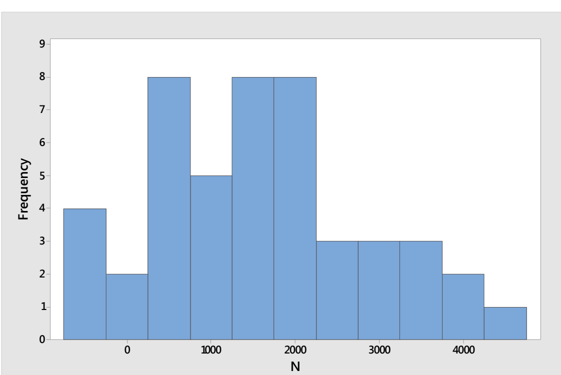


Figure 3: Distribution of Values of *N*

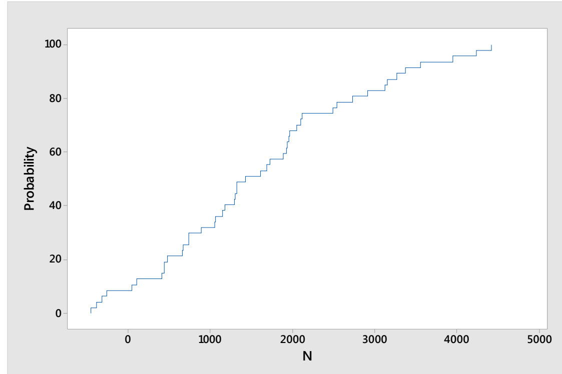


Figure 4: CDF for *N*

From the available data, a set of 52 routes of displacement – both out of and into Mosul – are observed, including the cases of displacement within Mosul itself (denoted as R37). However, the vast majority (a total of 86%) of the displacements are associated with only three of the 52 routes, while 14% fall into the 49 other routes (see Figure 5).

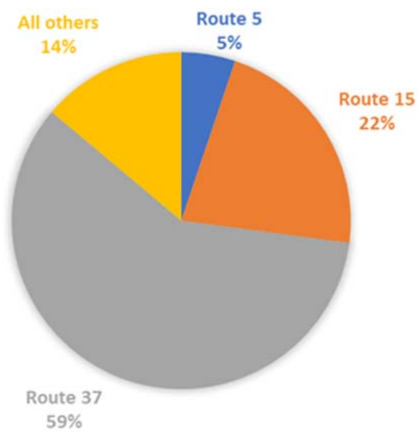


Figure 5: Distribution of IDP Routes

Table 1 summarizes the three principal routes taken by families, while the 49 others are lumped into 'alternative routes' (R_Alt). However, we point out that a more detailed analysis is required of the variation of IDP flows over time.

Table 1: Main Routes Taken by IDPs

Routes	Origin	Destination
R37	Mosul	Mosul
R15	Mosul	Al-Hamdaniya
R5	Hatra	Mosul
R_Alt	Various	Various

Instead of trying to build a regression model for the aggregate flows, in this case, based on observed data the empirical distributions for values of interest are obtained (number of people displaced, percentage of IDPs taking each route).
The empirical distributions of the percentage of IDPs observed for each of the listed routes are presented in

Figure 6. In general, these empirical distributions do not fit with theoretical distributions. These empirical distributions will be implemented as inputs in the simulation model (explained in the next section).

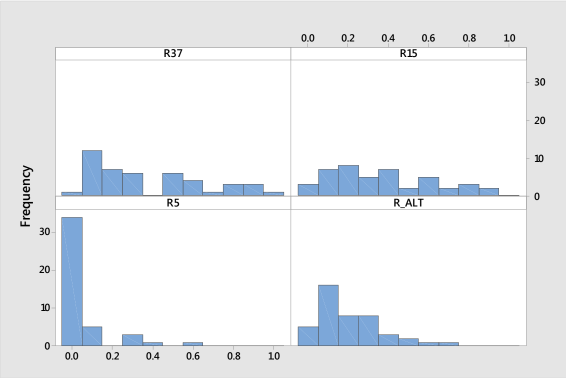


Figure 6: Empirical Distribution of the Percentage of IDPs taking Main Routes

Additionally, temporal fluctuations are illustrated by means of time series for each route (Figure 7). These values exhibit the variation over time without showing any clear trend or tendency.
Finally, to show the probability to observe a specific percentage of displacement for each case the CDF for each of the main routes is presented in Figure 8.

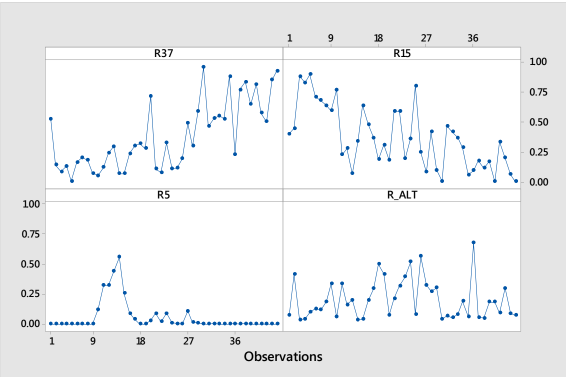


Figure 7: Time Series for Main Routes

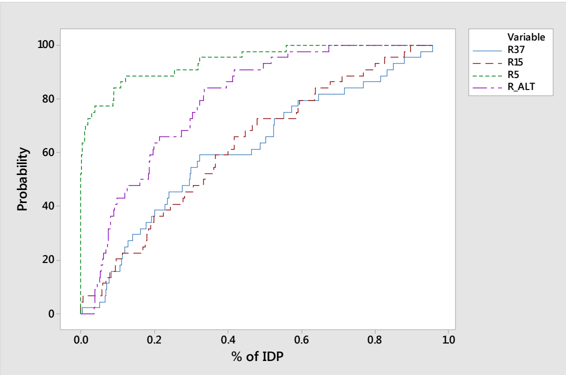


Figure 8: CDFs for Main Routes

3. MODEL DESCRIPTION AND SIMULATION RESULTS

3.1. Model Description

Petri nets (PNs) are an extension of state graphs which provide a compact and effective representation of real systems (Jiménez and Pérez 2004).
The graphical representation by directed bipartite schemas and the mathematical expression by state matrix or state equation support their wide application for modeling, simulating and analyzing the evolution of discrete event systems (Nosedal, Baruwa, and Piera 2013).
This paper presents the application of colored Petri nets (CPNs), which are an enriched type of PN. CPNs allow a higher level of modeling by using colors that enable the representation of entity attributes (Pennisi, Cavalieri, Motta, and Pappalardo 2016).
A full description of the CPN modeling approach and its applications can be found in Jensen, Kristensen, and Wells (2007) and the description of a graphical language for its representation and simulation in Jensen and Kristensen (2015).
In this section, our model is introduced by first listing and providing descriptions of its main elements: colors (Table 2), transitions (Table 3), and places (Table 4). The model (see Figure 9) has been implemented using CPN Tools (Jensen and Kristensen 2015).

Table 2: Colors

Color	Description
K37	Value for % of flow in route 37
K15	Value for % of flow in route 15
K5	Value for % of flow in route 5
Ka	Value for % of flow in the remaining routes
D	Current available % of flow to be assigned
N	Value of the overall flow magnitude simulated
c	Replications sequential number

Table 3: Transitions

Transition	Description
S_R37	Simulation of the flow in route 37
S_R15	Simulation of the flow in route 15
S_R5	Simulation of the flow in route 5
S_RALT	Simulation of the flow in the remaining routes (accumulate)
INTEGRA-TION	Standardization for the values to assure consistent % distribution of the flows and generation of simulation records.

Table 4: Places

Place	Description
ED_R37	Contains the values that describes the empirical frequency distribution for route 37
ED_R15	Contains the values that describes the empirical frequency distribution for route 15
ED_R5	Contains the values that describes the empirical frequency distribution for route 5
ED_RALT	Contains the values that describes the empirical frequency distribution for the 49 remaining routes (aggregate)
S_R37	Contains the simulated value for the flow using route 37
S_R15	Contains the simulated value for the flow using route 15
S_R5	Contains the simulated value for the flow using route 5
S_RALT	Contains the simulated value for the flow using the 49 remaining routes (aggregate)
PERCENT-AGE	Contains the memory for the current available % of flow to be assigned during the simulations
ED1	Contains the values that describes the empirical frequency distribution for the aggregate value of the flow (overall magnitude)
counter	Assigns a sequential number for each replication
S_RESULTS	Keeps the records of the simulated values (sequence, overall magnitude, % for route37, % for route 15, % for route 5, and % in the 49 remaining routes).

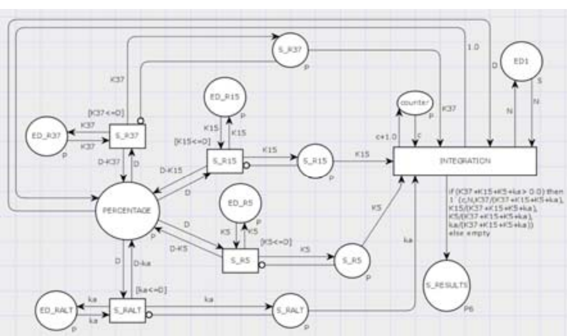


Figure 9: CPN Model for Simulation of Displacement Flows

3.2. Simulation Results

For aggregate IDP flows and flows along each of the four main routes as listed in Table 1, 30 replications of the model were run, with 50 time periods in each replication. For instance, in Figure 10, comparisons are provided concerning the time series for the aggregate magnitude the actual vs. simulated flows in the 11th replication. N_11 corresponds to the values for the entire simulated time frame of 50 periods, SUM_11 corresponds to the cumulative value of simulated flows, while N and SUM correspond to the real values (obtained from ET data set).

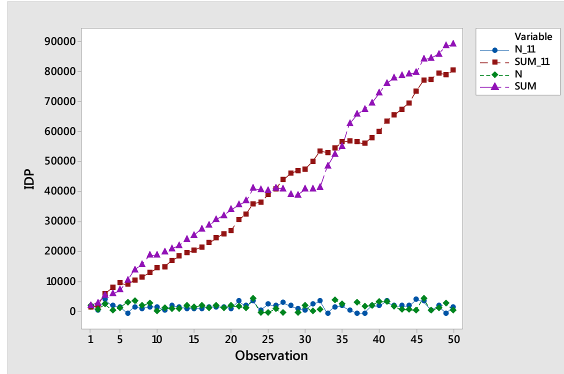


Figure 10: Actual vs. Simulated Flows (Replication 11)

Similarly, in Figure 11, comparison of the time series are shown for values obtained during the 26th replication.

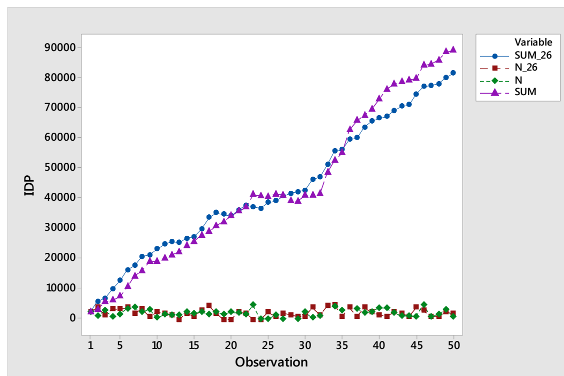


Figure 11: Actual vs. Simulated Flows (Replication 26)

The distributions of the values of *N* for each of the 30 replications (N_1 to N_30) are shown in Figure 12. The actual distribution is included in the top-left corner (indicated by N).

The distribution comparisons for the simulated values for percentage of flow within route R37 (IDP flows within Mosul) are presented in Figure 13.

Despite the fact that routes taken by IDPs follow a very different distribution (as illustrated in Figure 6), the actual and simulated distributions are very similar. Variations among simulations are expected due to the inherent uncertainty of the real data, and it is the intent of this model to reproduce this behaviour.

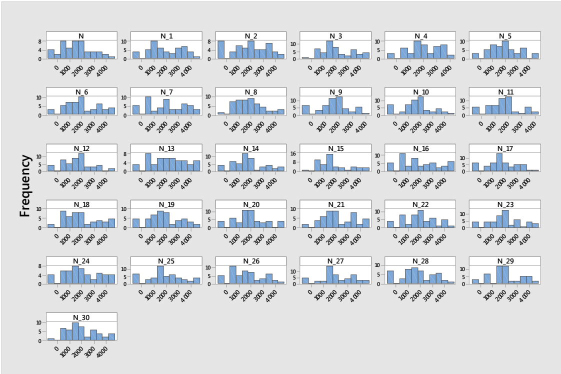


Figure 12: Distributions of Simulated Values of *N* (Replications 1,...,30)

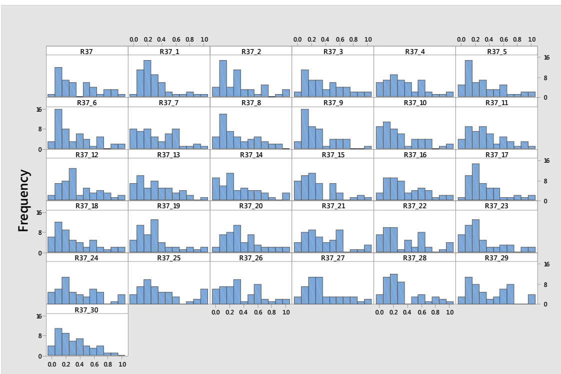


Figure 13: Distributions of Simulated Values of Percentage of IDPs taking Route 37 (Replications 1,...,30)

These simulated values and their stochasticity are illustrated and compared with the actual values by means of cumulative distribution diagrams (Figures 14 to 18).

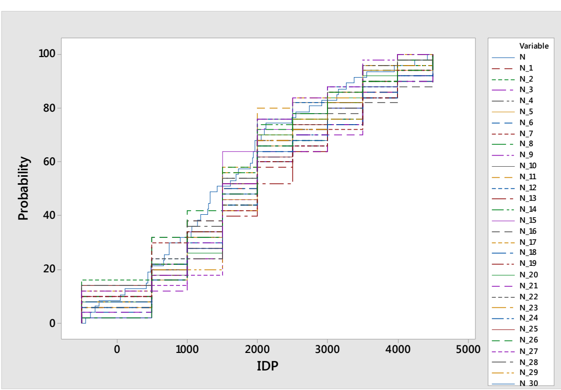


Figure 14: Cumulative Distribution Diagram for *N* – Actual vs. Simulated Values (Replications 1,...,30)

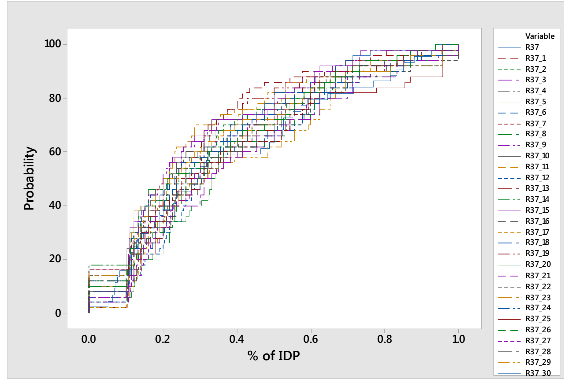


Figure 15: Cumulative Distribution Diagram for Percentage of IDPs taking Route 37 – Actual vs. Simulated Values (Replications 1,...,30)

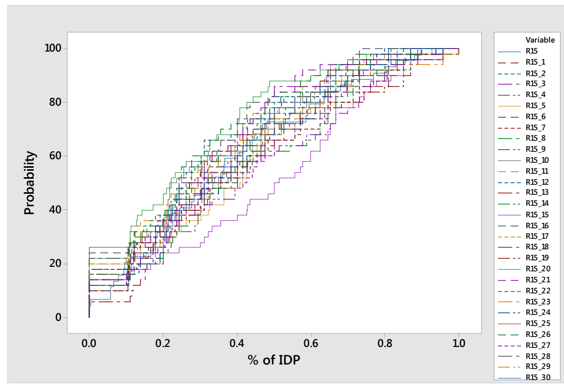


Figure 16: Cumulative Distribution Diagram for Percentage of IDPs taking Route 15 – Actual vs. Simulated Values (Replications 1,...,30)

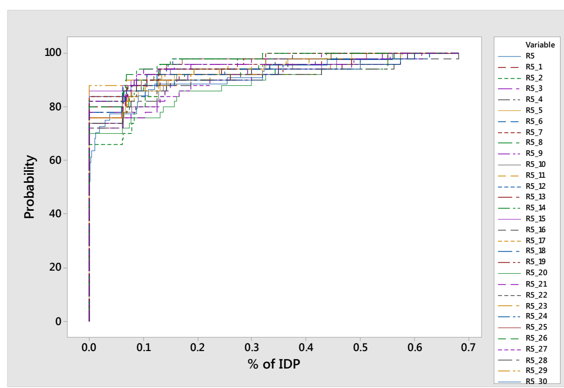


Figure 17: Cumulative Distribution Diagram for Percentage of IDPs taking Route 5 – Actual vs. Simulated Values (Replications 1,...,30)

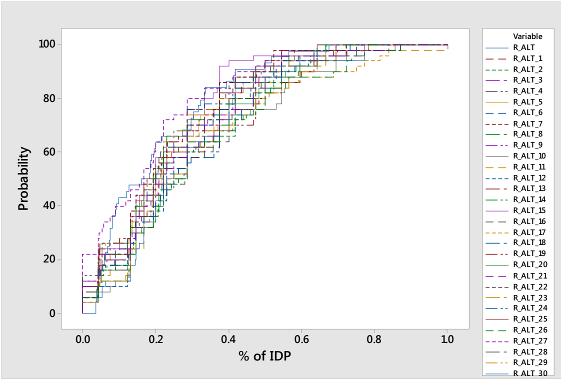


Figure 18: Cumulative Distribution Diagram for Percentage of IDPs taking Alternative Routes – Actual vs. Simulated Values (Replications 1,...,30)

In terms of the reproducibility of the temporal evolution of numbers of IDPs (actual time series were illustrated in Figure 7), Figures 19 and 20 provide examples of time series for the simulated values (in specific replications) for Route 5 and Alternative Routes, (both exhibiting fairly complex actual time series patterns).

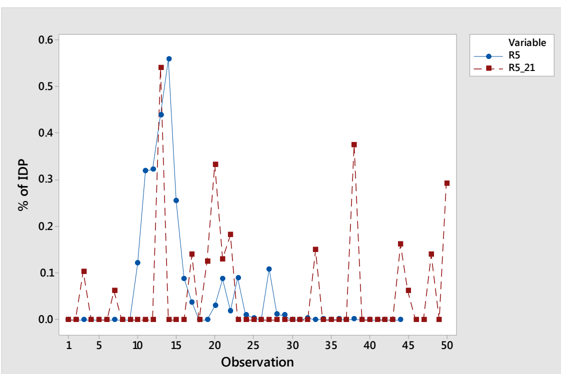


Figure 19: Actual vs. Simulated Values for Route 5 (Replication 21)

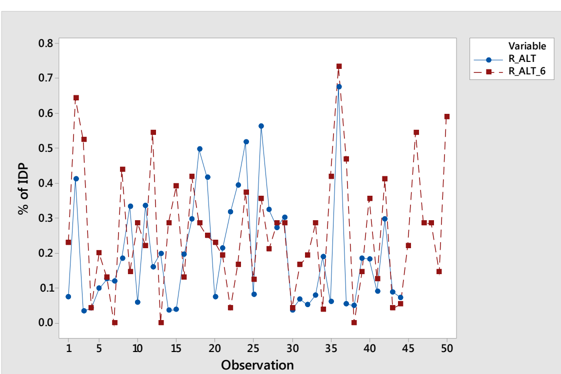


Figure 20: Actual vs. Simulated Values for Alternative Routes (Replication 6)

In both cases, despite the inherent variability and uncertainty of the observed values, our model appears to provide simulated time series that more or less follow the actual patterns as determined from the values reported in the ET system.

4. CONCLUSIONS AND FUTURE WORK

The current model involves stochastic modeling and simulation of forced population displacements based on discrete event modeling with the use of CPN.

The current model represents a single component that may be integrated into a more comprehensive model. The main components and structure for such a comprehensive framework, which integrates other elements and other simulation approaches (e.g., agent-based modeling), is reported and explained by Asgary, Solis, Longo, Nosedal, Curinga, and Alessio (2016).

Based on the initial results obtained, further work will include extending the CPN model to expand the set of possible routes. A further refinement would include updating the input data set (empirical distribution of IDP flows and the observed percentage distribution among routes). Since the military operation at Mosul has been completed as of July 2017 (CNN 2017), the input data may readily be updated.

ACKNOWLEDGMENTS

This research has been partially funded through the Collaborative Research and Training Experience (CREATE) Program in Advanced Disaster, Emergency and Rapid-response Simulation (ADERSIM) at York University, which is funded by the Natural Science and Engineering Research Council of Canada (NSERC). The research is also part of a Partnership Development Grant (Building Bridges across Social and Computational Sciences: Using Big Data to Inform Humanitarian Policy and Interventions) funded by the Social Science and Humanities Research Council of Canada (SSHRC). The Modeling and Simulation Center–Laboratory of Enterprise Solutions (MSC-LES) of the University of Calabria, Italy, is an international collaborator in the ADERSIM program.

REFERENCES

Asgary A., Solis A.O., Longo F., Nosedal J., Curinga M.C., and Alessio L.E., 2016. An agent-based modeling and simulation tool for estimation of forced population displacement flows in Iraq. Proceedings of the 6th International Defense and Homeland Security Simulation Workshop, pp. 75-81, September 26-28, Larnaca, Cyprus.

BBC.com, 2016. Mosul: Iraq and Kurdish troops make gains in battle (article dated 17 October 2016). http://www.bbc.com/news/world-middle-east-37679325 (accessed 16 April 2017).

Collmann J., Blake J., Bridgeland D., Kinne L., Yossinger N.S., Dillon R., and Zou K., 2016. Measuring the potential for mass displacement in menacing contexts. Journal of Refugee Studies, 29 (3), 273-294.

CNN.com, 2016. Battle for Mosul begins with gunfire and car bombs (article dated 17 October 2016). http://www.cnn.com/2016/10/17/middleeast/mosul

-isis-operation-begins-iraq/ (accessed 16 April 2017).
CNN.com, 2017. Mosul: Iraqi-led forces push into key city (article dated 10 July 2017). http://www.cnn.com/2016/10/17/world/gallery/mosul/index.html (accessed 19 July 2017).
Davenport C., Moore W., and Poe S., 2003. Sometimes you just have to leave: domestic threats and forced migration, 1964-1989. International Interactions, 29 (1), 27-55.
Engel S. and Ibáñez A.M., 2007. Displacement due to violence in Colombia: A household-level analysis. Economic Development and Cultural Change, 55 (2), 335-365.
google.ca, 2017. Map of Iraq and surrounding region. https://www.google.ca/maps/@34.5047751,45.6161276,6z (accessed 19 July 2017).
Jiménez E. and Pérez M., 2004. Simulation and optimization of logistic and production systems using discrete and continuous Petri nets. Simulation, 80 (3), 143-152.
Jensen K., Kristensen L.M., and Wells L., 2007. Coloured Petri nets and CPN Tools for modelling and validation of concurrent systems. International Journal on Software Tools for Technology Transfer, 9 (3-4), 213-254.
Jensen K. and Kristensen L.M., 2015. Colored Petri nets: a graphical language for formal modeling and validation of concurrent systems. Communications of the ACM, 58 (6), 61-70.
Nosedal J., Piera M.A., and Baruwa O., 2013. Concurrent and distributed systems analysis using colored Petri nets. Proceedings of XXXIV Jornadas de Automática, pp. 538-544, September 4-6, Terrassa, Spain.
Pennisi M., Cavalieri S., Motta S., and Pappalardo F., 2016. A methodological approach for using high-level Petri nets to model the immune system response. Proceedings of the 15th International Conference on Bioinformatics (InCoB 2016), pp. 91-105, September 21-23, Queenstown, Singapore.
Reddy P.A., 2017. Battle of Mosul: mass displacement of natives and a blatant violation of international humanitarian laws. LSE Human Rights Blog: http://blogs.lse.ac.uk/humanrights/2017/02/22/battle-of-mosul-mass-displacement-of-natives-and-a-blatant-violation-of-international-humanitarian-laws/ (accessed 19 July 2017).
Shellman S.M. and Stewart B.M., 2007. Predicting risk factors associated with forced migration: an early warning model of Haitian flight. Civil Wars, 9 (2), 174-199.

METHODOLOGY OF SHOOTING TRAINING USING MODERN IT TECHNIQUES

Grzegorz Gudzbeler [a], Jarosław Struniawski [b]

[a],[b] Police Academy in Szczytno, ul. Marszałka Józefa Piłsudskiego 111, 12-100 Szczytno, Poland

[a]gudzbeler@gmail.com, [b]j.struniawski@wspol.edu.pl

ABSTRACT
Mastering, improvement, shaping and preservation of skills of safe, efficient and effective use of the firearm requires the use of relevant methodology of conducting the shooting training. However reality of police trainings does not usually allow for intensive training shooting with the use of ammunition. An alternative solution is the use of modern training technologies .Example of this is the "Virtual system of improvement tactics of intervention services responsible for security and shooting training. "Introduction of stimulator to police trainings will enable complete stuff preparation to achieve its tasks, creating potential of knowledge and experience in many areas, far exceeding the capabilities of conventional training.

Keywords: training, weapon, security, Police, simulator

1. INTRODUCTION

Protection of life, health and property of citizens are constitutionally protected values. Their protection creates the conditions for the safe functioning and development of society, the country and its institutions. For the effective protection of these values, in addition to good law, it is necessary to have well-functioning system of administrative bodies equipped with the appropriate legal, material and technical approaches and technology for the prevention and countering the dangers [1].Within this entire system the Police plays special role, bearing the greatest responsibility for the state of public order and security [2].

The police is equipped with a power competence realized in specific forms of actions, ensuring a fast and efficient elimination of risk associated with different manifestations of behaviors violating the legal order of social co-existence. This privilege provides the ability to take emergency action, including the use of the means of direct coercion by the police. The aim is to submit to the recipients issued under the law commands, prohibitions or orders. The choice of the means of direct coercion must comply with the needs of the situation, which are necessary to achieve this goal .In case, where these means have proved insufficient or their use, due to the circumstances of the incident is not possible, a police officer is allowed to use firearms. It is used primarily against persons directly threatening life and health and the perpetrators of physical aggression, often with the use of dangerous objects. As the mean of direct coercion of extraordinary nature, firearm is considered as the mean of extremity.

The effectiveness of interventions and conducting them in accordance with police procedures and rules require from the officers a continuous improvement of their skills. Regulations, techniques and tactics of intervention with the use of force and firearms are the parts of vocational trainings and vocational training courses. Only systematic improvement allows to master the ability to make the right decisions and make them effectively, and most importantly – to safely carry out intervention.

2. SHOOTING TRAINING

Development of police officers trainings includes: central improvement- organized by police schools and the Higher Police School in Szczytno, local development - organized by the Police organizational units or cells of these units and the development of external ones- organized by entities outside the police.

Basic forms and purposes of conducting the training in shooting, including shooting training realized in the Police determines the Instruction on the organization and the police shooting from firearms and shoots (throws) of chemical agents incapacitating agents appended to annex No. 1 to Decision No 713 of the Chief of Police of 30 December 2005 on the shooting training of police officers [3].

The aim of the shooting training of police officers is:

- mastering and improving the skills of safe, efficient and effective use of firearms in different situations and conditions by the police officers;
- the formation and preserving compliant with law policeman's behavior in similar situations to the actual realities of the use of firearms;
- develop a skill of proper assessment of the situation, identification and selection of goal, the decision to use or abandon the use of firearms.

On the other hand, shooting range training includes:

- theoretical instruction (lectures on topics of firearms safety, basic knowledge of weapons and ammunition, basic marksmanship, shot theory, jams of firearms);

- practical instruction (non-shooting training, shooting training and programming shootings);
- exams and tests of knowledge and skills.

All forms of the training are subject to a specific methodology. Methodology of teaching is referred to as "the discipline dealing with the practical teaching methods of selected contents, most of various school subjects. Methodology of teaching comes down to the search for ways (methods) of rational action possible by a thorough analysis of the subject content and experience-based teaching methods and means of diagnosis enabling students to master the content "[4].Such an understanding of the methodology used in the control process, improvement, forming and preservation of skills of safe, effective and efficient use of firearms allows to isolate and create the following types of shooting scenarios:

- from short firearms - preparatory, static, fast, dynamic, situational; it is realized for a variety of targets, for a different distance, the distance (5, 10, 15, 20, 25 m), with different shooting postures (standing, kneeling, lying, out from the curtain), for different purposes;
- with submachine guns - static, fast, dynamic, situational;
- with smoothbore weapon - static, fast, dynamic, situational.

For the purpose of training technical measures are used, including targets which depending on the specificity of shooting can be rotating mechanisms showing the goal in a specified period of time or hanger that gives the possibility to suspend the blade at a certain height and distance convenient for the shooter. The most common targets used in shooting paper targets are designed for a specific type of shooting. They can also be annular discs for learning shooting and precision shooting, as well as shields physique primarily designed for combat shooting. Lecturers shooting may use other purposes, for example. Poppers, figures silhouettes of people or animals [5].Element to ensure the safety of participants in training shooting are sight and hearing protection. No such protection can lead to permanent hearing or vision.

Each form of training requires applying very strict rules of safe use of weapon which, among others, specify in detail the obligations of all participants, that is the shooting instructor and the shooters.

The process of shooting training of police officers covers several stages and the effect itself-the use of ammunition- ought to be proceeded by systematic and correctly organized and realized non-shooting training[6], which is the basic element of realizing shooting training regardless of the degree of advancement of the shooter. It covers elements which will be used during the training as well as remembering and consolidating ones which have already been learnt by the participants.

The objective of the non-shooting training is to educate a police officer as far as skills guaranteeing safe and efficient use of firearms is concerned, as well as elaborating adequate habits and response reactions and acquiring the skills enabling reaching precision of a shot. Consolidating the string of adequate actions which the trainee must execute requires imposing them on him as instructor's command. Due to the fact that the basic weapon of a police officer is short firearms, during the training one must place special emphasis on actions related to reaching this type of gun (taking the gun out of the holster), reloading, assuming shooting postures as well as actions related to correct keeping the finger off the trigger connected to observation of gunpoint devices and combining these activities with breathing.

The key element of shooting training is to learn the correct shooting posture. This is such placing of the body which ensures good statics of a gun, efficiency of directed fire and own safety. Shooting posture is the most variable element of shooting technique, mostly determined by situation in which a given police officer finds themselves and depending on the type of shooting (static, dynamic, situational), used gun (short, long) and distance of shooting. Another element of training is ranging, which consists of placing gun barrel in such a location on vertical and horizontal plane, so that the flight path of bullet passed through the designated target. These actions are realized through the so called syncing sights, that is notch and gunpoint. Also the skill of keeping the finger off the trigger has a significant impact on the quality of shot, since wrong move of an index finger during this activity results in a missed shot. The last stage of shooting technique is not shooting but withstanding it. This consists in maintaining the state of all elements of body and gun in location prior to the shot. The key reason for the necessity to maintain full concentration of attention and muscle tension for 1-2 seconds after the shot is to avoid, already before or during the shot, psychological and physical relaxation. In the context of shooting training it is very important to regulate breathing which corresponds to the accuracy of a shot. Not holding one's breath during control of synced sights and keeping finger off the trigger decreases the results of shooting.

Well-elaborated and correctly realized shooting training facilitates controlling the entire set of elements which form a single shot, leading to acquiring and maintaining correct habits and automating technical actions and, consequently, reaching among others high results in shooting.

Shooting training may be enhanced by situations which the police officer may encounter during his or her work with firearms. For this reason, for the purposes of this type of training scenarios of shootings have been elaborated. They assume realization of conditions of a specific variant of static, fast, dynamic or situational shooting. Reaching the assumed minimum in the form of number of points or shooting holes in shield point field in a specified time marks a given variant as done in formal terms.

Static shooting is a shooting at a distance up to 25 meters, to shields TS – 2, TS – 3 and TS – 9 without specifying time of shooting. First shootings are conducted from the distance from 5m and once a police officer passes them, the distance gradually increases up to 25m.

Fast shooting is characterized by a specified and short period of duration. Depending on the stage of fast shootings, the use of methods impacting the swaying of homeostasis of the officer's organism using the firearms is also possible. It is also possible to use stimuli, situations, simulated atmospheric phenomena or behaviours of third persons towards whom actions are taken and which are unpredictable for the trainee. Shooting may be conducted in conditions of limited visibility, with lightening of targets by means of for example a torch, headlamps of cars, with the use of equipment simulating the battle field (smoke grenades, stun grenades) [7].

Dynamic shooting is conducted after physical effort. It is characterized by one or several targets placed in various distances, with the use of various shooting postures and change of shooting positions. During the shooting the police officer moves in line with the principles of tactics. It is advisable that during shooting natural and artificial covers are used. Shootings of this type may be conducted in conditions of limited visibility, with lighting of targets by means of for instance a torch, headlamps of cars. Dynamic shootings in the initial period of training are conducted individually, then in pairs, ending with shift shooting.

Situational shootings are characterized by the necessity to identify the target and take decision regarding shooting. During shooting the police officer moves with a gun ready for shooting, in line with the rules of tactics. It is advisable that during shooting natural and artificial covers are used. Shootings of this type may be conducted after physical effort, in conditions of limited visibility, with lighting of targets by means of for instance a torch, headlamps of cars etc. Situational shooting is conducted individually or in a team.

The most advanced form of shooting training is training with the use of warfare in objects in which for instance persons performing the function of helpers may be placed. During such training elements of tactics and technics of detention of persons, hostage situations or taking over objects in situations of terrorist attacks are realized. This training is realized mostly on tactical ranges through officers who present very high level of intervention skills, as well as who use firearms on a perfect level.

The moment in which safe situation turns into a deadly dangerous is many times difficult to capture and almost unnoticeable and it so happens that it might be tragic as a result. The skill of correct and safe use of firearms is based mostly on habitual behaviours. In psychological terminology, an automated activity , way of behaving or reacting which is shaped as a result of practising, mainly by repetition, is called in this way. In the course of time, these skills become instinctive to the point, where they really stop being controllable on an ongoing basis. However, the reality of police trainings does not allow for intense shooting training with the use of ammunition. This is of course dictated by economic reasons.

3. VIRTUAL SYSTEM TO IMROVE TACTICS OF INTERVENTION ACTIONS OF SERVICES RESPONSIBLE FOR SAFETY AND SHOOTING TRAINING

In the effort to meet these needs, one must search for new, alternative training solutions, also thanks to the use of morn IT technologies. An example of this is "Virtual system to improve tactics of intervention actions of services responsible for safety and shooting training." The proposed system, due to high level of detail and very wide scope of scenarios and variants adjusted to the specificity of actions of the Police, tactics and provisions of law in place in Poland as well as police internal procedures will be able to properly support the process of improvement training of officers. Complementary, mobile training tool will enable exercising realized in real time events, which might occur during realization of tasks with participation of officers in real places of undertaken interventions. The purpose of conducted trainings on this simulator will be efficient increasing and maintaining of high level of training provided for officers in elaboration by them of correct reactions in the process of taking decisions during realization of a specific task. Such solutions allow to reflect in virtual world of computer training system some real scenarios of actions undertaken by one or several (intervention patrol) police officers and other services responsible for safety. Conduct of trainings in the scope of interventions will be enriched with the possibility of improving technique and shooting tactics. Improving the tactics is shaping of the skills in correct selection of the moment of shooting and a place of the strike, whilst the shooting technique refers to fast and accurate shooting to specific target. A significant element of such a solution are not only the novelty technical solutions, but also appropriately selected scenarios and variants of the course of situation stemming from many years of experiences of police officers of all types of services.

The simulation system consists of subsystem of psychophysical parameter measurement, projection subsystems, guns, assessment of trainees and system which integrates the subsystems (simulation management subsystem), enabling management of the simulation process and creation of new scenarios.

For the purposes of the system, film material has been elaborated in line with the elaborated scenarios and variants stemming from the practice of intervention actions of the Police and stemming from the methodology of shooting training. A full list of possible for occurrence in reality crisis situations with which a police officer may meet during conduct of his or her professional duties will be applied, with the option of its expansion by applying a combination of elaborated variants and creating new ones. Under the project, 50 scenarios and variants of actions for training of tactic of intervention actions and shooting training will be elaborated. The variants of actions are the ways of proceeding targeted at obtaining the assumed goal, which are made on the basis of anticipated scenarios of threats. They may be elaborated at each phase of police actions and their number ought to be adequate to the number of assumed threats to safety during a given event. The variant should indicate a police officer, measures anticipated for restoring the breached public order, and pursuant to the alleged description of events also method-tactics and technique of conduct of actions[8].

Virtual training system for tactics of intervention actions and conduct of shootings with firearms is targeted at placing the trainee officer before a decision of choosing adequate measure of restraint and its adequate use. Proper assessment of the undertaken decisions of the officers and selection of adequate evaluation of subsequent exercises, is possible thanks to the use of monitoring tools of biometric parameters.

In the frames of the system, a subsystem of measurement of psychophysical parameters of the trainee persons (POC) was planned as well as automatic/ semi-automatic assessment of exercises- in line with the so far model of assessment of the way of acting of the officers. This subsystem allows for testing in real time the level of focus of attention and stress level. It enables assessment of mental burdens, accompanying the completion of tasks by the trainees, such as taking decision of the use of firearms or use of other measure of restraint towards an aggressive person and individual research of the usefulness for their realization. This ensures the possibility of immediate verification of actions of the officer and modification of the course of scenario so that the efficiency of training was optimized, considering at the same time individual aptitude of the trainees. Gathering information ought to commence already in the phase of preparation and waiting for the training, as well as after its completion.

In order to test biometric and environmental parameters, one must apply the solution in the form of modules placed on the trainee, which then will be wirelessly sent to the position of operator with adequate diagnostic software at disposal. Sensor modules are used for measuring body temperature, blood rate, breathing activities, physical activity, ECG and blood pressure. For this purpose a probe for measurement of blood and pulse was used (installed on the wrist of a persons who was tested, who was located right next to the glove placed on one's hand, in which measuring electrodes will be located) and sensor of saturation located on the ear top. Obtained measurements are saved and enable observing the results of measurements in real time. Sensor may be connected with adequate electric cables to the main ECG module which will be placed on the main strip of the trainee officer. Its tasks will be to monitor ECG of two independent outflows during the conducted test.

The training system, through displaying the registered picture of places of real interventions, will provide the trainee officer with sight stimuli. It is, among others, on the basis of sight observations that decisions are made as to the use of most adequate to situation measure of restraint.

The key elements which form part of the measurement system are oculometer and scene camera. Oculometer is a device using the source of infra-red light and in-built camera for defining the position of eyes in motion and defining the direction of looking. Light emitted by oculometer in infra-red moves outwards from the cornea of an eye of the trainee, causing obtaining in this way information used for its further processing and making available in the form suitable for interpretation. The second element of measurement system is scene camera the task of which is to record the plan of image seen by the trainee on which points and areas of so called heat maps will be depicted. In combination with adequate software, one may view how a trainee observes for instance the course of intervention or in which way he or she synchronizes sights when shooting from a gun to shooting shield displayed on projection screen. Quite often the enclosed software provides the results in graphic form of heat maps, view paths, statistical data, for instance, how long the trainee looked at a dangerous tool held by a criminal displayed on the projection screen.

The applied measurement tools enable conduct of measurement and gathering of measurement material. This might significantly impact evaluation process of education in tactical and technical application of means of direct restraint located in individual equipment of the officer during intervention.

An integral part of training computer system will be the equipment in elements refine measures applied during service means of direct restraint being on individual disposal of an intervening police officer, and in particular, training replica of firearms.

For the purposes of simulator the following were adopted: short gun “Glock 17”, smooth-bore weapon ,,Mossberg” or gas pitcher and electric stun gun, as well as, to increase realistic nature of exercises- police baton and handcuffs. The process of adaptation was conducted with the use of the existing on the market replicas of firearms, the look of which and the method of use is very similar to original guns.

Replicas of these means of restraint were equipped in infrared emitters enabling the mapping of place of strike which will be recorded by a detection camera. These solutions reflect the most realistic way of acting of standard gun and its functions. Therefore, the key aspect of selected solutions is tactical efficiency, mobility of set (gun plus power supply), way of energy supply (gas, combustion, electric, air) with the least harmful exit emission and the cost of use of the set. An important element of acting of the gun is verification in terms of impact of objects within the interaction with operator, that is police officer. Mapping of behaviour of a gun during shooting is possible through applying the so called gunshot recoil, causing after pressing the trigger a move of the lock into the back, extreme location and its return to the front, closing the bullet chamber, making a specific rumble during shooting. The movement of the lock is caused by applying compressed gas of blow-back system specified with GBB abbreviation (Gas Blow Back), the container of which is placed in gun magazine. Gas container may be very simply loaded with gas from the external , larger bottle, through using an adequate disjunctive power connector.

Elaborated and adequately adjusted components are integrated in such a way, so as to enable correct functioning of the gun modified in the frames of the monitoring system shooting efficiency.

Another element of the system is a subsystem of video projection, enabling display on the projection screen of registered earlier scenario and identification of activity conducted by trained persons and elaboration of the system of positioning of point of shooting. In this scope a module of projection was elaborated, equipped in short-throw projectors with high resolution (minimum of 2K) and screens, on which a screen will be displayed, showing real events. The element of subsystem is projection screen ensuring optimum display of elaborated image, registered in real conditions, furthermore it thoroughly maps the size of postures of persons and objects which occur. The module will be integrated with a camera, tracking the screen and enabling receipt of signal (shot) on a specific spot on the screen. The information of the point of signal is relocated to the displayed at that time image, thanks to which the instructor obtains information on whether and how an object on the screen has been shot. The system of video projection enables display of elaborated earlier scenarios of exercises, on the basis of which trainees, equipped in modified gun and measures of direct restraint, take adequate actions.

In the system, elements enabling recreating without the use of live ammunition of shooting from firearms and in particular, generating and indicating a place of shot on the displayed projection screen were placed. For this purpose the detection camera covering with its scope entire area of projection screen will be used, whilst the bullet from firearms will be replaced by light emitter in IR infra-red placed inside the replica of firearms of gun barrel. The infra-red emitter must be placed also in the replica of electric stun gun, replacing shot out probes with emitter light. The system enables registration of the fact of selection of a given means of direct restraint depending on the type of practiced scenario.

During conduct of dynamic shooting training the laser of applied emitter, should be launched only for the time and at the time of giving the shot, mapping thoroughly the moment of shooting. Whilst, during realization of static training focused on increasing competencies of accurate targeting, laser ought to generate constant bundle, enabling registering and mapping the technique of targeting. During mapping by means of emitters the places of striking a target, parameters of the pathway of bullet flight should be considered, which vary depending on distance, as well as the applied ammunition and changeable atmospheric condition.

Apart from the displayed on the projection screen video image of prototype of the created training system must provide should elements. During recording in real places of undertaken interventions, sound material will be registered, enabling recreation during realization of future shooting trainings, sound elements occurring at the time of conduct of real tasks by officers. For accurate recreation of sound applying adequate digital multi-channel sound system is required as well as applying and placing devices emitting the sound.

Managing entire process of simulation is possible thanks to the platform integrating the subsystems. It was equipped in a number of elements, enabling gathering of information on behaviours of trainees. Data obtained from the system are subjected to analysis and reporting to the instructor who conducts the training.

4. CONCLUSIONS

Introduction to the system of police trainings Virtual system to improve the tactics of intervention actions of services responsible for safety and shooting training will allow for full preparation of staff for realization of tasks standing before them, creating the potential of knowledge and experience in many areas, significantly exceeding the possibilities of conventional training. Replacing real exercises with a training on simulator will contribute to decreasing their costs and a simultaneous increase of its scope. Interactive training will eliminate material losses, risk of loss of life and health, caused by an error stemming from lack of adequate high level of skills of the trainee. Furthermore, it will limit the volume of negative factors impacting instructors and trainees which occur, i.e. during shooting training. Additional feature of the system will be the possibility of finding application while elaborating universal, pattern profile of police officer, realizing the tasks at intervention level. It will facilitate the selection of candidates for specific types of police services, increasing through this the level of realization of tasks by entire formation.

ACKNOWLEDGMENTS

This work was partially supported by National Centre for Research and Development of Poland – grant DOB-BIO7/19/01/2015. Construction of the system is conducted by consortium formed by: Police Academy in Szczytno (project leader), Department of Electronics and Information Technology of the Warsaw University of Technology, ETC-PZL Aerospace Industries Sp. z o.o., Dynamic Safety Corporation Sp. z o.o., Alex Stern Sp. z o.o.

REFERENCES

Pieprzny S., 2012. Administracja bezpieczeństwa i porządku publicznego, Rzeszów.
Pieprzny S., 2011. Policja. Organizacja i funkcjonowanie, wyd. 3, Warszawa.
Decyzja nr 713 komendanta głównego Policji z 30 grudnia 2005 r. w sprawie szkolenia strzeleckiego policjantów, Dz. Urz. KGP z 2006 r., nr 3, poz. 9 ze zm.
Okoń W., 2003. Wprowadzenie do dydaktyki ogólnej, Warszawa.
Idzikowski A., 2014. Szkolenie strzeleckie. Bezpieczeństwo – charakterystyka zagadnienia, Legionowo.
Stechnij T., Grzebieluch M., Kukuła A., 2011. Trening techniczny bezstrzałowy jako podstawa szkolenia strzeleckiego policjantów, Katowice.
Decyzja nr 143 komendanta głównego Policji z 2 maja 2016 r., w sprawie programu nauczania na kursie specjalistycznym w zakresie posługiwania się bronią palną podczas pełnienia służby w ubraniu cywilnym, Dz. Urz. KGP z 2016 r., poz. 23.
Struniawski J., 2014. Dowodzenie w trakcie operacji policyjnych w zakresie imprez masowych i zgromadzeń publicznych, Szczytno.

INTEROPERABLE SIMULATION AND SERIOUS GAMES FOR CREATING AN OPEN CYBER RANGE

Agostino G. Bruzzone [(a)], Riccardo Di Matteo [(b)], Marina Massei [(c)], Enrico Russo [(d)], Mirko Cantilli [(e)], Kirill Sinelshchikov [(f)], Giovanni Luca Maglione [(g)]

[(a)] [(c)] DIME University of Genoa,
[(b)] Simulation Team SIM4Future, [(d)] DIBRIS University of Genoa,
[(e)] Liophant Simulation, [(f)] [(g)] Simulation Team

[(a)] agostino@itim.unige.it, [(b)] riccardo.dimatteo@simulationteam.com, [(c)] massei@itim.unige.it,
[(d)] enrico.russo@unige.it, [(e)] mirko.cantilli@liophant.org,
[(f)] kirill.sinelshchikov@simulationteam.com, [(g)] maglione@simulationteam.com

[(a)] [(c)] www.itim.unige.it, [(b)] www.sim4future.com, [(d)] www.dibris.unige.it,
[(e)] www.liophant.org, [(f)] [(g)] www.simulationteam.com

ABSTRACT

The paper proposes an open architecture to support the creation of a synthetic environment devoted to simulate complex scenarios related to the protection of cyber-physical systems. The proposed approach is based on applying the combination of interoperable simulation and serious games to develop a framework where different models, as well as real equipment, could interoperate based on High Level Architecture standard. By this approach, it becomes possible to create a federation reproducing a scenario including multiple physical and cyber layers interacting dynamically and reproducing complex situations. The authors propose an example of specific case study conceptually developed to apply this approach.

Keywords: Cyber Range, Multi Layer Modeling, Interoperable Simulation, Serious Games

1 INTRODUCTION

Today "Cyber" is an hot spot, as well as a buzz word, and there are many discussions about it; therefore despite the general considerations, the experienced along recent years confirm the impact of threats acting on the cyber layer and the escalation of attacks in this "space" that is constantly growing in terms of extension, impact, tactics and strategies (Wilhoit & Hara 2015; Page et al. 2017). Due to these reasons, the development of Cyber ranges as a synthetic environments devoted to address cyber defense is currently evolving quickly as a necessity (Pridmore et al. 2010; Winter 2012; Ferguson et al. 2014). In facts along last years the cyberspace impact on physical system exploded and the cyber attacks confirmed the need to develop new capabilities able to support the development of new systems and policies in this area (Cashell et al.2004; Kunder et al. 2010; Sgouras et al. 2014). A major additional issue is represented by the Education and Training, another set of activities that require a framework where to conduct test and exercises (Pham et al. 2016; Törngren et al. 2017).

In this sense a Cyber Range is expected to provide the opportunity to experience the use of tools and techniques able to improve the stability, security and performances of cyber physical systems.

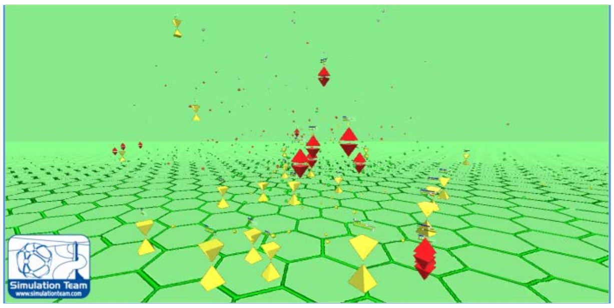


Fig.1 – Cyber Space Representation coupled with real layer within Simulation Team Synthetic Environment

The Cyber Range concept is similar to that one of the shooting and kinetic ranges, in use by military force, that allows to train warfighters in using weapons and conduct operations. In similar way, the cyberwarfare actors and players are expected to use the Cyber Range for training as well as for development and testing of new solutions and tactics in order to ensure consistent operations and readiness for real fighting. From this point of view it is evident that the cyber warfare is evolving from simple attacks to become part of more evolved and comprehensive strategies based on the use of real "cyber weapons" that could compromise critical infrastructures, society policies and endanger human life (Chakhchoukh & Ishii 2015); to face these challenges there is a growing need of Cyber Ranges able to reproduce multiple layers that are closely interconnected as it happen in the real world (Bruzzone et al.2016). Due to theses reasons and considering the complexity of the proposed scenarios, the authors suggest the adoption of an open architecture able to integrate different components including specific models, meta-models, interoperable simulators as well as real equipment. A review of the literature shows that the integration of interoperable simulators, analytical models and real equipment have been also successfully used also in the physical space in sectors like Industry, Logistics, Healthcare. Examples of review articles and applications in different areas also including security issues can be found in (Michael et al., 2014; Zhou et al., 2016; Longo, 2015; Longo 2012).

2 STATE OF ART IN MODELLING CYBER WARFARE & EXISTING CYBER RANGES

It is evident since several years that the cyber warfare is a crucial element to simulate in order to support decision makers to face challenges posed by cyber warfare itself (Stytz & Banks 2012).
Indeed, since the beginning of this decade, it has been proposed this concept as following: "a cyber range is a facility allowing a model of an IT system to run in a simulated environment to perform tests and measurements that are applicable to the real world" (Winter, 2012).
Even statistics turn to be an useful "tool" for analyzing cyber attacks along those years; for instance a pretty interesting study presents a statistical framework for investigating cyber attack data and predicting them in term of attack rate with reasonable accuracy (Zhan, Z. et al. 2013).
US DoD (United States Department of Defense) is deeply involved in cyber range investigation through the National Cyber Range (NCR) operated by the Test Resource Management Center (TRMC) and the project has been seen as a major event in this context (Pridmore et al. 2010). In facts the evolution of this program allowed to analyze the DoD PMs (program managers) necessities and to refine objectives: "Internet-like environment by employing a multitude of virtual machines and physical hardware" have to be created "to find the best approach for testing the cyberspace resiliency of the systems under development" in order to show how incorporating cyber security at early stage of the development life cycle "helps to avoid high cost integration" (Ferguson et al. 2014)
A more recent definition of cyber range in training perspective says "Cyber-Range is a vehicle used to train in offensive and defensive Information Operations and Information Warfare" (Lawless et al. 2014) In this case it is presented a modular approach to the Cyber-Range Framework development, tailored for being adaptable and to meet current and future needs.
Again, statistics play a key role: advanced forecasting technique such as Exponentially Weighted Moving Average (EWMA) are applied to investigate Distributed Denial of Service (DDoS) attacks (Olabelurin et al. 2015) and it is proposed a methodology for reducing the number of alerts and false positive alarms.
Typically, training in this context is very crucial for military personnel: an interesting work on NATO MSG-117 activities (assessing which areas of Modeling and Simulation could contribute to cyber defense) and SISO standards summarizes the current position in this sense (Croom-Johnson, S., 2015).
Obviously the personnel is a key factor in managing cyber security issues and social engineering should be properly addressed in this field (Granger 2001; Evans & Wallner 2005; Goodchild 2012). Indeed, despite a large numbers of certifications, applicants assessment is quite challenging. From this point of view gamefication is used to evaluate skills and technical abilities (Cherinka & Prezzama 2015).

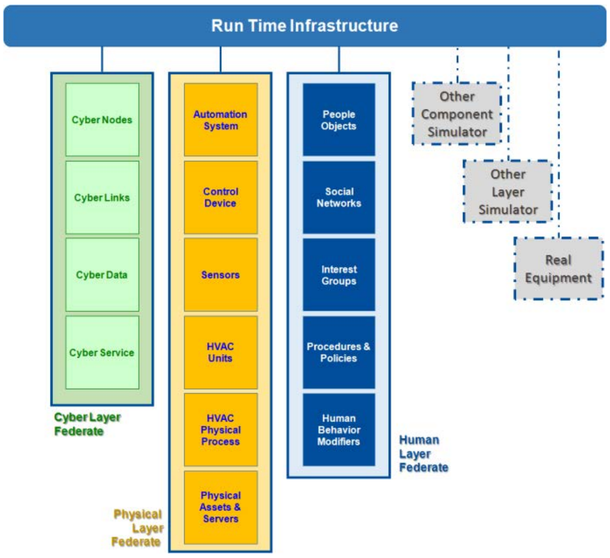


Fig.2 – Example of the Open Proposed Architecture

End users are once more involved in the process, providing an ad hoc experience of cyber threats seems to be an helpful approach in the framework of Internet of Everything (IOE) for securing resilience of devices (Lawless et al. 2015). On the same track it is possible to apply Cyber Range to Concept of Operations (CONOPS) in order to recover from a disaster affecting a large company; this approach had allowed to develop useful and realistic scenarios for disaster testing (Ali & Santos 2015; Kisekka & Baham 2015).
Testing and training are again core elements of scientific research on cyber range: "Cyber Security Training Range, a simulation infrastructure where various scenarios can be recreated and tested, to educate Mission planners, Mission engineers and System administrators on the possible attacks at the different mission phases, shall be available as daily working asset." (Mann & Zatti 2016).
In this sense the use of proper modeling techniques to reproduce heterogeneous networks incorporating multiple entities to create a realistic cyber-physical scenario is fundamental (Bruzzone et al.2013). A large number of trainees performing training activities in specific environment that contains virtual machines, network topology, security-related content etc is quite challenging, a possible solution is the use of automatic generated cyber ranges for education and training based on specifications defined by the instructors (Pham et al 2016; Bruzzone et al. 2016). DDoS are one of the most "effective" and "efficient" threat to deal with. Modeling is used to develop a model of normal user behavior with weighted fuzzy clustering (Zolotukhin et al.2016).
Miming the real systems is another M&S application "mimetic environments, which mimic actual networks including personal computers, network assets, etc., are required for cyber range or malware analysis" (Yasuda,

S. 2016; Yasuda et al. 2017). In facts decision makers face cyber attacks and military commanders (as decision makers themselves) need models able to improve awareness of threats development; from this point of view interoperability is a major issues as suggested by recent works from NATO MSG-117 (Modelling and Simulation to Support Cyber Defence) that address, as anticipated, this aspect with special focus on standardization process (Croom-Johnson & Couretas 2016). It is important to outline that, from this point of view, time-cost reduction is always a goal to be reached while developing cyber range environments through avoidance of errors in its configuration (Damodaran & Tidmarsh 2016). Again, cyber range and test at low cost are heavily involved while analyzing threats of actual automated systems, such as SCADA (Supervisory control and data acquisition), along with related potential issues and criticalities (Hallaq et al. 2016). Lately, research is being even more focused on education issues; for instance, in order to evaluate the success rate of cyber security students, especially while performing lab activities in their master classes, have been identified “metrics, like the number of IDS alerts, network sessions, or top destination IP addresses, …, indicators of success or failure for final grade” (Caliskan et al. 2017). This is expected to be helpful to manage the growing number of cyber security exercises and evaluate students across different institutions and academic years.

3 SIMULACRA ARCHITECTURE

SIMULACRA (Simulation Multi Layer Architecture for Cyber Range) is an open interoperable architecture based on HLA, designed by authors, that is allowing to combining different simulators, models and real systems in order to create a common synthetic environment. The approach is based on MS2G (Modeling, interoperable Simulation and Serious Games) paradigm and emphasize the importance to be able to combine together multi resolution models and to cover specific layers also by metamodels when more high fidelity solutions are not available or usable (Bruzzone et al.2014). In this case, the development process should be strongly integrated with VV&A (Verification, Validation and Accreditation) in order to guarantee a constant control of the simulation confidence bands and fidelity even in relaxed conditions devoted to investigate complex scenarios.

Indeed the use of engagement techniques and immersive solutions common in Serious Games are considered fundamental due to their capability to provide an intuitive environment where the players could quickly understand the situation evolution and consequences of different operations in a scenario where the dimension and complexity is very hard. In facts, today, it could be interesting to evolve in this framework from the traditional man-in-the-loop approach, often based often in using hackers as players in the cyber defense training, to condition where it is applied the concept of man-on-the-loop supervising several assets and driving the operations at high level while low level actions are simulated by agents reproducing humans or artificial intelligences (Magrassi 2013; Bruzzone et al.2015). In this sense the authors have already experienced several applications and developed solutions such as in the case of T-REX or JESSI, an agent driven stochastic simulation able to combine real and cyber components of a multidimensional space that include air, land, sea, underwater, space and cyberspace (Bruzzone et al.2016); this framework includes population, legacy and autonomous assets as well critical infrastructures to develop complex scenarios and it is able to cover social engineering, human behaviors and ICT network simulation (see figure 1). In this case it is proposed an architecture able to incorporate some of the capabilities addressed by T-REX with others provided by other federates as proposed in following figure 2.

The proposed architecture deals to the use of specific objects within the Federation Object Model devoted to share the entities, attributes and interaction among the different simulators as proposed in figure 2. The objects and attributes include among the others:

Cyber Node
ID
Operating Status
ICT Connections
Physical Object Connections
Level of Confidentiality
Level of Availability
Level of Integrity

Cyber Link
ID
Operating Status
Cyber Node Connection
Level of Confidentiality
Level of Availability
Level of Integrity

Physical Object
ID
Location
Operating Status
Cyber Node Connections

People Object
ID
Current Location
Reference Locations
Operating Status
People Object Connections
Interest Group Object Conn.
Cyber Node Connections
Physical Object Connections
Human Behavior Modifiers

Interest Group Object
ID
Reference Locations
Operating Status
Interest Group Relat.
Cyber Node Connections
Physical Object Conn.

Cyber Service / Data
ID
Operating Status
Cyber Node Connections
Level of Confidentiality
Level of Availability
Level of Integrity

The interactions in this case address issues such as:

- Cyber Attack Action against a Target
- Cyber Defensive Action on a Target
- Routing for accessing a Cyber Service
- Sending / Receiving Data from a Cyber Service

Indeed it is important to define also performance indexes to evaluate the pending potential of cyber threats on physical systems; for instance respect an comprised information shared between *A* and *B* as proposed below:

$$OL^{c}_{G,A,B}(t) = \left[1 - \prod_{i=1}^{n} Mn(i,G,A,B,t)\right]\left[1 - \prod_{j=1}^{m} Ml(j,G,A,B,t)\right] \quad (1)$$

$$Mn(i,G,A,B,t) = \begin{cases} i \in G & Ln^{c}_{i}(t) \cdot (1 - P(t,G,A,B)) - 1 \\ i \notin G & 1 \end{cases} \quad (2)$$

$$Ml(j,G,A,B,t) = \begin{cases} j \in G & Ll^{c}_{i}(t) \cdot (1 - P(t,G,A,B)) - 1 \\ j \notin G & 1 \end{cases} \quad (3)$$

$$OLg^{A}_{B}(t) = \max(OL^{c}_{G} \ \forall G \in Path(A,B,t) \quad (4)$$

$OL^{C}_{G}(t)$ Confidentiality Level on the G-th Path a t-time
$OLg^{A}_{B}(t)$ Global Confidentiality between A & B at t time.

n number of cyber nodes
m number of cyber links
G path among two cyber assets
Path(A,B) Set including all alternative path between A and B at t time
P(t,G,A,B) Probability to use G-th Path from *A* to *B* at t time

Ln^{C}_{i}	Level of confidentiality of the i-th node
Ll^{C}_{j}	Level of confidentiality of the j-th link

The authors are currently working in adapting previous simulator for this purposes in order to extend their capabilities for addressing the specific new scenario proposed in the following.

4 GRID ATTACK AND AUTOMATION

One of the critical factors is the presence of nodes of different independent networks in the same geographic locations, for instance an office could have its own cable network with or without connection to the internet, while in the same time there are several mobile and WiFi networks covering the same area. Obviously, a bridge between them could be created; indeed this could happen occasionally or purposely, in the first case, for example the cause could an employee that could decide to use internet on a work place by activating a 3G modem or WiFi hotspot on a phone, connecting internal and external networks. In the second case, depending of available resources and intruder background and skills, there are different types of communication channels that could be established, starting from that ones mentioned before up to data transmission techniques which could allow to communicate even with air gaped devices, for example using PC speakers (Lazic & Aarabi 2006), noise of fan changing its rpm (Guri et al. 2016), transmitting RF (radio frequency) signals using memory data bus (Guri et al. 2015a) and even regulating thermal pattern of a PC (Guri et al. 2015b). Obviously all these techniques could be used not only separately, but also in combination to create mesh network of compromised systems, containing even air gaped devices in the case they are already infected. It is important to mention also that growing amount of IoT (Internet of Things) devices that introduce often insecure and outdated software, creating additional vulnerabilities and vectors of attack (Barcena & Wueest 2015). From the point of view of creating a cyber range, it is convenient to simulate several independent networks, not originally connected, as well as the establishment of occasional temporary or permanent link between them. In facts, for instance a possible targets of hackers is a very critical network: power grid (Adams, 2015); in this sense a very notable case of cyber blackout happened at the end of 2015 in Ukraine (Sullivan 2017). Obviously these facts create a lot of concerns about reliability of perspective smart electric grid; for example as consequence of an attack the power could be simply switched off as well as reprogrammed and invalid switching of electric devices could result in unsafe connections which may lead to fire in the target place (Mo et al. 2012).

There are even some other vectors of attacks which have theoretical foundation, but, fortunately, limited experimental exploitation; for instance it's possible to 'brick' or even explode the battery of an Apple's laptop flashing microcontroller's firmware (Miller 2011). So it is evident that several different kinds of interactions could affect physical and cyber systems and assets.

5 SCENARIO & EXERCISE ON THE SYNTHETIC ENVIRONMENT

It is currently possible to define a testing scenario covering multiple layers including among the others:

- ICT Infrastructures
- Business Processes & Real Operations
- Facilities
- Entity & Units
- Population & Users

To create this reliable and useful scenario, it is proposed to define a specific case to be investigated in relation to an infrastructure providing business services. Obviously, it is useful to finalize the case in order to be realistic and easily extendable to a wide range of applications; in the paper the mission environment suggested is related to a College/University Department facing different kind of internal and external threats as well as malfunctions.

In facts, a University Department represents a good example of a division within a larger organization. While the Department shares a different number of units, services and procedures coordinated centrally, it may decide to adopt or extend its own. For this reason, the Department depends also on internal administrative or technical employees who shall assist its Academic staff. Employees belong to internal units according to their functions, namely, if they work for accounting, research and teaching support or for the IT and technical field support.

As an example of an internal administrative process, we could consider a “purchase request” concerning acquisition of material for a “research project”. The accounting staff of the Department applies internal arrangements on everything related to the tender specifications, the choice of potential suppliers; in this way the staff is supposed to record and verify all the support documentation. When the procurement process is finalized in terms of purchase request and winner selection, it is created and added the final order to the central accounting software, managed by the central University Informatics Centre. At the time of ordered goods arrival at the destination (the Department), it is conducted an acceptance check respect original requirements; in positive case, the administrative staff confirms the invoice receipt and authorize the payment through the accounting software.

The internal IT staff also supports this process by making available different tools that they implement and manage locally. Among these tools, a file server allows the administrative staff to save and share, with the correct authorization rights, all the necessary documentation. Moreover, a calendar system helps in maintaining the deadlines of procurements processes and a ticketing system helps in keeping tracks of their status and of all the related communications.

In our case, as it happen commonly, the Academic staff also access the file server and is enabled to use it to save all the data belonging to research and teaching activity.

Considering the importance of mentioned services and of the saved data, the IT staff should also implement a

reliable backup system and related procedures that could ensure business continuity in the event of problems. In order to keep these IT services active internally, the Department is expected to be equipped with a server room that is powered and equipped by air conditioning through a central power grid and a central HVAC (Heating, ventilation and air conditioning) system supporting all the buildings of the campus were is located the Department. In the proposed case, an external provider is currently responsible for maintenance of these facilities and it mainly operates using the Internet connection to access the control units of such installations. The Department has also a computer lab where students are able to follow lectures, conduct exercises and/or access Internet. A dedicated network segment of the infrastructure of the Department enables desktop computer and students' laptops to access the campus network as well as Internet.
In terms of threats, it is decided to focus on the following aspects:

– Social Engineering: invoice with cryptolocker
– Denial of Service (DoS): Students in the Laboratories (voluntarily/accidentally) block access to the campus network for the whole Department, so it turns impossible to access accounting program; this event could be coordinated in case of voluntary action with some other attack initiative carried out locally
– Hacking of the Service Provider Network resulting in enabling access to infrastructure services

It is identified a set of possible alternative in terms of ICT configurations in order to check vulnerabilities and impacts, for instance between on-premises vs cloud solutions (So 2011).

5.1 Social Engineering Attacks

Social Engineering typically involve human factors; usually psychological manipulation is used to fool users or employees into handing over access privileges, confidential and/or sensitive data (Lord 2017). In facts, the social engineering techniques are often based on simple email or other kind of communication distributed over a wide number of users that devoted to cheat on them; sometime reference to general issues and urgency are used to solicit emotions in the potential victims that often react instinctively or superficially by clicking over malicious links and/or malicious files. In facts social engineering also use the weakness in the processes related to human elements (e.g. password recording or selection), in general it is evident that it is pretty challenging to defend large organizations and enterprises from these threats (Granger 2001; Evans & Wallner 2005; Goodchild 2012).
As anticipated, often social engineering attacks are based on email specifically designed to look like a communication from a contact or a reliable organization (e.g. service messages, bank communication, request for information about some public work); in case the user clicks on these malicious attachments, usually, he install, unconsciously, some malware or ransomware (Abraham et al.2010); in facts, in general these email are pretty generic and include even errors and mistakes in order to discriminate smart and aware employers from superficial ones in order to maximize the penetration capability; however sometime the emails are tailored for specific users, just to look like originated from someone inside their organization or in their contact list. In facts often for most attackers is more easy to rely on social engineering respect to work hard in vulnerabilities of the Operating Systems (O.S.): in facts this approach address temporal weakness or superficial attitude of users and does not require much lower skills and efforts as well as not need to deal with the continuous advances and upgrades of the security systems (Mitnick 2001; Pettey & Goasduff 2010). Based on some statistics it is noted that technical weakness are addressed just by a small percentageof the cyber attacks while the remaining large majority uses social engineering methods; due to these reasons it is evident that the different O.S. does not guarantee too much respect these aspects (Saini et al. 2012). A very common social engineering attack adopts the phishing (statistically around 91% of data breaches) therefore currently ransomware is beginning to turning very popular in this area. In facts, it is not possible to define a single line of defense against social engineering, but education and training are important, in facts there are also solution to address specific typologies, such as ransomware. In general a good practice is to use reliable antivirus software, to make regular backups, to update software and to make sure that email attachments are scanned (especially compressed files and all document formats that support macros); it is also useful to disable the possibility to install unnecessary browser plugins and to teach personnel to pay attention before to click. It is evident that in our case study the Social Engineering represents a fundamental layer to be simulated.

5.2 Denial of Services

DoS attacks accomplish this by flooding the target with traffic, or sending them information that causes an accident. In facts the denial of Service (DoS) is very popular and aims to block a machine, or network, making it inaccessible to its users through saturation.
These attack are usually focusing on denying the service for the legitimate users (employees, members, or account holders). DoS attack victims often are subjects used to manage web services for different organizations, private or public as well as to internal different divisions. Practically the DoS attacks block temporary the services and the server access, so in general they does not cause loss of information nor theft of sensible data, however the service deny could generate extended damages in terms of time delays and costs and could even be a vector to coordinate other cyber attacks (Peng et al. 2007; Gupta & Badve 2016).
Most popular DoS methods deals with flood and crash services. In case of flood attacks the target system receives too much traffic for a server buffer, causing it to saturate, slow down and eventually stop. Examples of

flood attacks include among the others: buffer overflow attacks, ICMP Flood, SYN flood; other DoS attacks simply exploit the technical vulnerabilities that cause a system, a server or target service blocking. A specific and popular alternative type of DoS attack is defined Distributed Denial of Service (DDoS). In facts the DDoS attacks are carried out by multiple systems that conduct synchronized DoS attacks against a single target. DDoS main difference is that uses distributed resources to carry out the attack, instead of being originated by a single location. In general, the identification of attack source is pretty challenging considering that often it is carried out by a *botnet*: compromised systems that have been already corrupted and that serve as operative support for the attackers (Abu Rajab et al. 2006). Today there are several solutions devoted to defend against most traditional forms of DoS attacks, however DDoS are still one of the major threats for many organizations even if usually require strong capabilities by the attackers. Cyber offensive actions are currently very effective and limited capabilities are available for the defenders; in general it is not possible to share resources over the web being attack proof, therefore it is evident that training and experimentation allows to develop technological and procedural solutions as well as cyber defense tactics that could reduce the cyber vulnerability. Some of these methods are simple such as to install and maintain valid antivirus and keep them updated, or to install personal/centralized firewalls and configure it properly to limit outbound and inbound traffic to the desired traffic. In our case study the DoS is expected to be used at different levels: as a standalone attack by students for fun, as well as coordinated attack to block administration processes and controls.

5.3 Hacking

Hacking exploiting the gaps in the service providers network is a consolidated approach by hackers, therefore usually, these shortfalls are quickly solved directly by ISP technicians (Internet Service Providers).
Therefore the web services based on MSPs (managed services provider) introduce specific vulnerabilities considering that these subjects are often responsible for remotely accessing and managing their ICT resources and user systems of their customers; these capability relies usually on direct and privileged accesses to the customer networks; the case of recent Teamviewer hacking is a very good example (Dunn 2017). In facts, when MSPs are based on cloud or hosting solutions they hold a very large amount of data, often sensitive and/or confidential. So by targeting just a single MSP an hacker could obtain several accesses to different networks and organizations. In many cases, often the most popular methodology to conduct the attack is based on phishing emails containing executable attachments. In facts it is common for the hacker to register spoofed domains in order to send emails from them pretending to belong to reliable organizations (e.g. academic organizations, charity, etc.). In case the malicious link is clicked the attachment delivers its payload to access to target network. In general, the stolen MSP credentials usually could provide administrator or domain administrator privileges to the hackers; in addition the attackers trades and shares often accesses and credentials to move through different MSP networks and their users. Major issues to defend MSP networks and services against attackers are based on several main principia such as to avoid that users share single account credential to access the services, two factor authentication, conducting continuous security tests, adopt endpoint security approach. MSP in our case represent a potential source for obtaining credential and accesses to the services for the attackers even if it is not hypothesized to organize a specific hacking of the MSP just for planning an offensive action against the Department.

CONCLUSIONS

The SIMULACRA architecture presented in this paper is inspired by the synthetic environments developed by Simulation Team for different applications and it is very promising to create an interoperable Cyber Range open to be integrated with different models. In facts the authors are currently finalizing the *Academic Department Scenario* in order to conduct dynamic experimentations and to test the interaction of IA with both University and High School Students during the simulation. The following step is to integrate some automation system to verify the interoperability of plant control within the proposed federation of simulators.

REFERENCES

Abraham, S., & Chengalur-Smith, I. (2010) "An overview of social engineering malware: Trends, tactics, and implications", Technology in Society, 32(3), 183-196

Abu Rajab, M., Zarfoss, J., Monrose, F., & Terzis, A. (2006) “A multifaceted approach to understanding the Botnet Phenomenon”, Proc. of the 6th ACM SIGCOMM Conference on Internet Measurement, October, pp. 41-52

Adams Jr, J. A. (2015) “Cyber Blackout: When the Lights Go Out--Nation at Risk”, Friesen Press, Altona, Canada

Ali, J., & Santos, J. R. (2015) “Modeling the Ripple Effects of IT-Based Incidents on Interdependent Economic Systems” Systems Engineering, 18(2), 146-161

Barcena, M., Wueest, C. (2015) “Insecurity in the Internet of Things”, Security Response Symantec Tech Report, Mountain View, CA

Bruzzone A.G., Massei M., Longo F., Cayirci E., di Bella P., Maglione G.L., Di Matteo R. (2016) “Simulation Models for Hybrid Warfare and Population Simulation”, Proc. of NATO Symposium on Ready for the Predictable, Prepared for the Unexpected, M&S for Collective Defence in Hybrid Environments and Conflicts, Bucharest, Romania, Oct.17-21

Bruzzone A.G., Massei M., Longo F., Nicoletti L., Di Matteo R., Maglione G., Agresta M. (2015)“Intelligent Agents & Interoperable Simulation for Strategic Decision Making on Multicoalition Joint Operations”, Proc.of DHSS, Bergeggi, Sept.

Bruzzone A.G., Massei M., Tremori A., Longo F., Nicoletti L., Poggi S., Bartolucci C., Picco E., Poggio G. (2014) "MS2G: simulation as a service for data mining and crowd sourcing in vulnerability reduction", Proc. of WAMS, Istanbul, September

Bruzzone A.G., Merani D., Massei M., Tremori A., Bartolucci C., Ferrando A. (2013) “Modeling Cyber Warfare in Heterogeneous Networks for Protection of Infrastructures and Operations”, Proc.of I3M2013, Athens, Greece, September

Caliskan, E., Tatar, U., Bahsi, H., Ottis, R., Vaarandi, R. (2017) “Capability detection and evaluation metrics for cyber security lab exercises”, Proceedings of the 12th International Conference on Cyber Warfare and Security, ICCWS, pp. 407-414

Cashell, B., Jackson, W. D., Jickling, M., & Webel, B. (2004) “The economic impact of cyber-attacks”, Congressional Research Service Documents, CRS, Washington DC

Chakhchoukh, Y., & Ishii, H. (2015) “Coordinated cyber-attacks on the measurement function in hybrid state estimation”, IEEE Transactions on Power Systems, 30(5), 2487-2497

Cherinka, R., Prezzama, J. (2015) “A model for building a cyber security talent pipeline”, Proc. of 19th WMSCI, Vol.1

Croom-Johnson, S. (2015) “Cyber in simulation - Modelling the invisible threat”, Fall Simulation Interoperability Workshop, SIW, Orlando, August-September

Croom-Johnson, S., Couretas, J.M. (2016) “Cyber tools and standards to improve Situational Awareness”, Simulation Innovation Workshop, SIW, Orlando, September

Damodaran, S.K., Tidmarsh, D. (2016) “Model based verification of cyber range event environments”, Simulation Series, 48 (5)

Dunn J.E. (2017) "Questions linger after ISP blocks Team Viewer over fraud fears", Naked Security, Sophos, March 14

Evans, S., & Wallner, J. (2005) “Risk-based security engineering through the eyes of the adversary”, Proc. of 6th Annual IEEE SMC Information Assurance Workshop, June, pp. 158-165

Ferguson, B., Tall, A., & Olsen, D. (2014) “National cyber range overview” Proc. of IEEE Military Communications Conference (MILCOM), October, pp. 123-128

Goodchild, J. (2012) “Social engineering: The basics”, CSO Online, IDG, August 3

Granger, S. (2001) “Social engineering fundamentals, part I: hacker tactics” Security Focus, December, 18

Gupta, B. B., & Badve, O. P. (2016) “Taxonomy of DoS and DDoS attacks and desirable defense mechanism in a cloud computing environment”, Neural Computing and Applications, 1-28

Guri, M., Kachlon, A., Hasson, O., Kedma, G., Mirsky, Y., & Elovici, Y. (2015a) “GSMem: Data Exfiltration from Air-Gapped Computers over GSM Frequencies”, Proc. of USENIX Security Symposium, pp. 849-864

Guri, M., Monitz, M., Mirski, Y., & Elovici, Y. (2015b) “Bitwhisper: Covert signaling channel between air-gapped computers using thermal manipulations”, Proc. of IEEE 28th Computer Security Foundations Symposium (CSF), July, pp. 276-289

Guri, M., Solewicz, Y., Daidakulov, A., & Elovici, Y. (2016) “Fansmitter: Acoustic Data Exfiltration from (Speakerless) Air-Gapped Computers”, arXiv preprint arXiv:1606.05915

Hallaq, B., Nicholson, A., Smith, R., Maglaras, L., Janicke, H., Jones, K. (2016) “CYRAN: A hybrid cyber range for testing security on ICS/SCADA systems, Security Solutions and Applied Cryptography in Smart Grid Communications, pp. 226-241

Kisekka, V., Baham, C. (2015) “Applying cyber range concepts of operation to disaster recovery testing: A case study” SIGSAND AIS Electronic Library from Proc. of ACIS , Puerto Rico

Kundur, D., Feng, X., Liu, S., Zourntos, T., & Butler-Purry, K. L. (2010) “Towards a framework for cyber attack impact analysis of the electric smart grid”, Proc. of 1st IEEE Int.Conf on Smart Grid Communications (SmartGridComm), October, pp. 244-249

Lawless, B., Flood, J., Keane, A. (2014) “A framework to address challenges encountered when designing a cyber-range”, Proc. of European Conference on Information Warfare and Security, ECCWS, January, pp. 258-263

Lawless, B., Flood, J., Keane, A. (2015) “Analysis of the implementation of an interactive kinetic cyber range component”, Proc. of European Conference on Information Warfare and Security, ECCWS, January, pp. 389-394

Lazic, N., & Aarabi, P. (2006) “Communication over an acoustic channel using data hiding techniques”, IEEE transactions on multimedia, 8(5), 918-924

Lord N. (2017)"Social Engineering Attacks: Common Techniques & How to Prevent an Attack", Digital Guardian, August 29

Longo F., Chiurco A., Musmanno R., Nicoletti L., (2015). Operative and procedural cooperative training in marine ports, Journal of Computational Science, vol. 10, pp. 97-107.

Longo F. 2012. Supply chain security: An integrated framework for container terminal facilities. International Journal of Simulation and Process Modelling, vol. 7, no. 3, pp. 159-167.

Magrassi C. (2013) “Education and Training: Delivering Cost Effective Readiness for Tomor-

row's Operations", Keynote Speech at ITEC2013, Rome, May 22-24
Mann, A., Zatti, S. (2016) "The ESA cyber security training range", Proceedings of the International Astronautical Congress, IAC.
Michael M., Abboudi H., Ker J., Khan M.S., Dasgupta P., Ahmed K. (2014). Performance of technology-driven simulators for medical students - A systematic review. Journal of Surgical Research, vol. 192, no. 2, pp. 531-543.
Miller, C. (2011) "Battery Firmware Hacking", Black Hat, Accuvant Labs Technical Report, USA, June 12
Mitnick, K. (2001) "My first RSA conference" Security Focus, April
Mo, Y., Kim, T. H. J., Brancik, K., Dickinson, D., Lee, H., Perrig, A., & Sinopoli, B. (2012) "Cyber–Physical Security of a Smart Grid Infrastructure", Proc. of the IEEE, vol. 100 (1), 195-209
Olabelurin, A., Kallos, G., Veluru, S., Rajarajan, M. attacks (2015) "Multi-agent based framework for time-correlated alert detection of volume", Lecture Notes in Electrical Eng., 499-507
Page, J., Kaur, M., & Waters, E. (2017) "Directors' liability survey: Cyber attacks and data loss—a growing concern", Journal of Data Protection & Privacy, 1(2), 173-182
Peng, T., Leckie, C., & Ramamohanarao, K. (2007) "Survey of network-based Defense Mechanisms countering the DoS and DDoS Problems", ACM Computing Surveys (CSUR), 39(1), 3
Pettey C., Goasduff L. (2010) "Top Predictions for IT Organizations and Users", Gartner Newsroomm, November 30
Pham, C., Tang, D., Chinen, K. I., & Beuran, R. (2016) "CyRIS: a Cyber Range Instantiation System for facilitating Security Training", Proc. of the 7th ACM Symposium on Information & Communication Technology, Dec., pp. 251-258
Pridmore, L., Lardieri, P., & Hollister, R. (2010) "National Cyber Range (NCR) automated test tools: Implications and application to Network-Centric Support Tools", Proc. of IEEE Autotestcon, September, pp. 1-4
Saini, H., Rao, Y. S., & Panda, T. C. (2012) "Cyber-crimes and their impacts: A review. International", Journal of Engineering Research and Applications, 2(2), 202-9
Sgouras, K. I., Birda, A. D., & Labridis, D. P. (2014). "Cyber Attack Impact on critical Smart Grid infrastructures", Proc. of IEEE Innovative Smart Grid Technologies Conference, Feb., pp. 1-5
So, K. (2011) "Cloud computing security issues and challenges", International Journal of Computer Networks, 3(5), 247-55
Stytz, M. R., & Banks, S. B. (2012) "Information Value Assessment Modeling In Cyber Warfare Simulation", Proc. of Spring Simulation Interoperability Workshop, Orlando, 90-101
Sullivan, J. E., & Kamensky, D. (2017) "How cyber-attacks in Ukraine show the vulnerability of the US power grid", the Electricity Journal, 30(3), 30-35
Törngren, M., Grimheden, M. E., Gustafsson, J., & Birk, W. (2017) "Strategies and considerations in shaping Cyber-Physical Systems education", ACM SIGBED Review, 14(1), 53-60
Winter H. (2012) "System security assessment using a Cyber Range", 7th IET International Conference on System Safety, incorporating the Cyber Security Conference, Edinburgh, Uk, October 15-18
Wilhoit, K., & Hara, S. (2015) "The real World Evaluation of Cyber-attacks against ICS System", Proc. of the 54th IEEE Annual Conference of Society of Instrument and Control Engineers of Japan (SICE), July, pp. 977-979
Yasuda, S. (2016) "4-4 the Design and Application of a Mimetic Network Environment Construction System", Journal of the National Institute of Information & Communications Technology, 63 (2), pp. 93-101
Yasuda, S., Miura, R., Ohta, S., Takano, Y., Miyach, T. Alfons: (2017) "A mimetic network environment construction system Lecture", Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, pp. 59-69
Zhan, Z., Xu, M., & Xu, S. (2013) "Characterizing honeypot-captured cyber attacks: Statistical framework and case study", IEEE Transactions on Information Forensics and Security, 8(11), 1775-1789.
Zhou C., Wang J., Tang G., Moreland J., Fu D., Wu B., (2016). Integration of Advanced Simulation and Visualization for Manufacturing Process Optimization. JOM, vol. 68, no. 5, pp. 1363-1369.
Zolotukhin, M., Kokkonen, T., Hämäläinen, T., Siltanen, J. (2016) "Weighted fuzzy clustering for online detection of application DDoS attacks in encrypted network traffic", Lecture Notes in Computer Science including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, 9870 LNCS, pp. 326-338

An underwater buoyancy-driven glider simulator with Modelling & Simulation as a Service architecture

Manlio Oddone[1*], Agostino Bruzzone[2], Emanuel Coelho[1],
Daniele Cecchi[1], Bartolome Garau[1],

[1]NATO STO Centre for Maritime Research and Experimentation,
Viale S. Bartolomeo 400, 19126 La Spezia, Italy,
[2]DIME, Univ. of Genoa, Genoa, Italy

*E-mail: manlio.oddone@cmre.nato.int.

Abstract

This paper describes the architecture of an underwater glider simulator applying the Modelling & Simulation as a Service (MSaaS) paradigm. The simulator implements a modular and scalable service-oriented architecture where Web services are employed for the underwater glider kinematic and hydrodynamic models, for the Oceanographic models forecast data, for the glider motion control system and for the motion behaviours. The simulator itself is offered as a Web service, using a browser based GUI to present the status and result of the simulation. Two main problems are targeted: 1) the estimation of the underwater trajectory of the glider between two known (measured) surfacing points and 2) the prediction of the glider future surface position given its past history, its motion behaviour and the Oceanographic Model forecast. Considerations about the estimation of the motion of the glider when the corresponding motion control and behaviours algorithms are not disclosed are done, capitalising on the work described in a previous paper and an alternative approach to simulating the unknown behaviours, considering the hardware-in-the-loop, is discussed. The discussion also involves architectural, and technological aspects such inter-process and cross-language communication and hardware-in-the-loop interfacing.

Keywords: Underwater glider, MSaaS, Web services, AUV, Modelling & Simulation, M&S, REST

1 Introduction

This paper describes the architecture of an underwater glider simulator implementing the Modelling & Simulation as a Service (MSaaS) paradigm with Command and Control (C2) and High Level Architecture (HLA) interoperability.
The simulator is provided as a Web service. The Graphical User Interface (GUI) is decoupled from the simulator engine and is provided either in the form of a Web GUI: a Javascript application with mapping display and graphical capabilities, or as an HLA Virtual-Reality graphical console. The glider dynamic and kinematic models, the models of the navigation control system (*front-seat*), the mission behaviours (*back-seat*) and the environment models are implemented as Web services.
The simulator output is provided using open standards for maximum interoperability, with NATO simulators, C2 systems, navigation/mapping/GIS software, Web applications and scientific data analysis/processing tools.
The simulator engine implements a three-phase discrete event processing core with a hybrid continuous/discrete model approach that makes the simulation suitable for constructive as well as virtual simulation applications. It can be configured for a range of applications, including:

- **Decision support tools** for optimisation and risk mitigation during mission planning and execution
- **Post mission analysis and debriefing tool** to analyse, reproduce and explain what happened during a mission in the post mission analysis phase
- **Test bed** for development and tuning of models, path planning and optimisation algorithms
- **Virtual reality applications** for personnel training and mission preparation
- **Hardware in The Loop Virtual Simulation** for mission behaviours safe testing on the bench before real deployment in the field

Despite the apparent complexity, the architecture is an ensemble of simple subsystems that are easy to manage and to secure. The implementation as a Web service makes the tool easily deployable and maintainable. Once the system is deployed on a server (e.g. as a virtual machine or as a container-based virtualisation solution) user may access it from workstations and mobile device with a Web browser. No installation or configuration of the tool is required by the user.
Flexibility and scalability makes the simulator engine suitable for cloud deployment, remotely accessed through a Wide Area Network (WAN) such as the Inter-

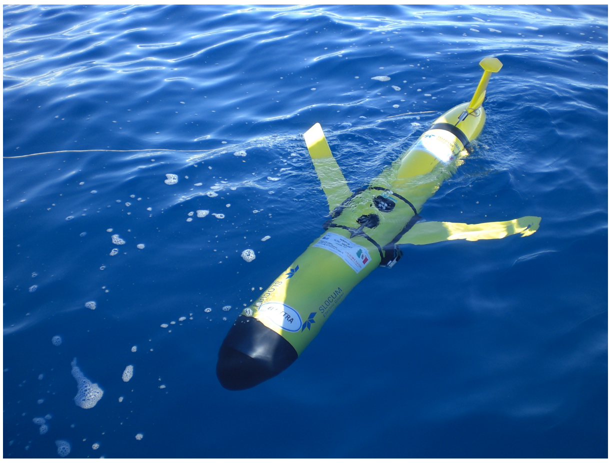


Figure 1: An underwater glider

net, or deployed on a local server and accessed by users on from a Local Area Network (LAN).

1.1 What is an underwater glider

Underwater gliders (see figure 1) are a special class of Autonomous Underwater Vehicle (AUV) that use small changes in their buoyancy to achieve vertical movement. Wings and control surfaces convert the vertical velocity into forward velocity so that the vehicle glides downward when denser than water and glides upward when buoyant. Due to the buoyancy-driven propulsion their motion in the vertical plane follows a saw-tooth profile as depicted in figure 2.

Glider motion is extremely high energy efficient, for this reason, despite the slow speed, they achieve particularly long-range capabilities making them ideally suited for scientific data collection subsurface missions, sampling environmental parameters at regular intervals at a regional or even larger scale.

They can be programmed to autonomously navigate for weeks or months at a time. When they surface they transmit the acquired data to shore, using satellite communication. At the same time they download new missions with substantial cost savings compared to traditional ship-based research methods.

Glider motion is influenced by currents and changes in the water density. Before starting a mission a glider must be carefully ballasted to match the water density of the mission scenario area. An improper calibration of the buoyancy could compromise the glider capability to perform its mission and in the worst cases could even lead to the loss of the glider.

1.2 Why a glider simulator

Underwater gliders operate in ocean environments characterised by a complex spatial variability. Due to their sensitivity to environmental parameter change, monitoring of the glider state is essential to achieve mission success by minimising risks due to adverse environments [19]. By predicting environmental impact on glider navigation, the pilot may decide to modify the mission plan

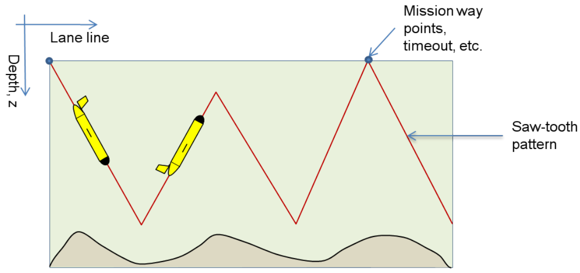


Figure 2: Saw-tooth motion pattern of an underwater glider

in advance to operate with the minimum risk of asset loss [13].

A simulator capable to accurately reproduce a glider behaviour within a scenario with a time-variable environment is considered and in particular two main problems are targeted:

- The estimation of the underwater trajectory of the glider between two known (measured) surfacing points
- The prediction of the glider future surface position given its past history, its motion behaviour and the Oceanographic Model forecast

This work aims to improve a Decision Support Framework[13] implemented at the NATO Science and Technology Organization - Centre for Maritime Research and Experimentation (STO-CMRE) by providing a simulator tool that would allow extending the developed glider model to the 3D case, taking into account the vertical changes of water density, model the glider at the surface, taking into account the effect of waves and winds and finally implement the glider guidance system to compensate for current prediction integration.

The simulator would also serve as a testing Framework for Path Planning algorithms of AUVs in realistic ocean environmental fields [15] and as a Virtual-Reality and environmental simulator with real Hardware in the Loop.

1.3 HLA interoperability

High Level Architecture (HLA)[14] is a general purpose architecture for the interoperability between distributed simulation systems. It is the subject of the NATO standardization agreement (STANAG 4603)[8] for modelling and simulation. In an HLA architecture multiple federates are composed into a federation exchanging object attributes and interactions through a Run-Time Infrastructure (RTI). An RTI is basically a bus connecting the federates with a publish-subscribe model.

Federates joining a federation may subscribe to an object class and then discover objects of such class and receive attribute value updates. Interactions are similar to objects except that an object is persistent (e.g. a tanker) while an interaction is only used once (e.g. an ammunition detonation, or a collision between vehicles).

HLA requires a shared reference data exchange model used as an agreement about the data (objects and interactions along with their attributes and parameters) that will be exchanged within the federation. Such model is called Federation Object Model (FOM). The modular RPR FOM (Real-time Platform Reference FOM) and the more recent NETN FOM (NATO Education Training Network FOM) provide building blocks for creating federation agreements.

HLA interoperability has been achieved by simulators developed at CMRE within the MCWS-MSTPA federation [5].

Interoperability between High Level Architecture (HLA) with Robot Operating System (ROS) [17] based autonomous systems has been achieved in CMRE in the work described in [3] where an HLA-based link between simulation and a SPARUS II AUV has been implemented and the integration into an HLA federation has been tested in a port protection scenario. For such scenario the hardware and software of the AUV has been included in the federation together with a *virtual simulator*. This works aims to extend such architecture by integrating the service-oriented concept.

1.4 Service-Oriented Architecture

The simulator architecture follows MSaaS paradigm that is a means of delivering Modelling & Simulation (M&S) applications, capabilities and associated data on demand by providers to consumers.

It has been designed according with the "MSaaS as a SOA" perspective described in the study on *Service-Oriented Architectures for Modelling and Simulation*[6] provided by the NATO Modelling and Simulation Group (NMSG) Specialist Team MSG-131.

The "MSaaS as a SOA" perspective looks to use Service-Oriented Architecture as the architectural approach for connecting and combining M&S service whose central aspects are:

- Communication following standards
- Loose coupling
- Interoperability and composability of services (to form new and maybe more complex services)

The implementation of the architecture has been done using Web services. The simulation engine itself is implemented as Web service.

Web services are server applications, that can be queried on a network (Internet or private) by client applications in order to provide a service, in the form of an exchange of information, and are not tied to any one operating system or programming language.

1.4.1 Types of Web Services

There are two main implementations of Web services: SOAP based and REST Web services.

Web services based on the Simple Object Access Protocol (SOAP) are commonly known as Web Services (WS or WS-*). WS exchange structured information using an eXented Markup Language (XML) Information Set for its message format, and relies on application layer protocols, most often HyperText Transfer Protocol (HTTP) or Simple Mail Transfer Protocol (SMTP), for message negotiation and transmission.

WS technologies have been successfully used for simplifying interoperability while providing scalability and flexibility in multiple applications, including distributed simulation software[16] however they have been superseded by the more modern and lightweight REpresentational Stat Transfer (REST) or RESTful Web services[23].

REST is not a protocol but an architectural style that allows to manipulating textual representations of Web resources using a uniform and predefined set of stateless operations. A "Web resource" is any entity that can be identified, named, addressed or handled, in any way whatsoever, on the Web, identified by its Unique Resource Identifier (URI).

REST employs the uniform and predefined set of HTTP operations (GET, POST, PUT, DELETE) to allow a Web service to access and manipulate Web resources through their textual representation: requests made to a resource URI will produce a response in XML, HTL or JSON. Use of such well known and widely supported text formats, transmitted over the most widely used internet HTTP protocol is well digested by the most complex network configurations employing firewalls and proxies and allows for the maximum interoperability between systems over the internet.

REST operations are stateless, each request from client to server must contain all of the information necessary to understand the request, and cannot take advantage of any stored context on the server. Session state is therefore kept entirely on the client and has to be passed to the Web service when it's functionality is invoked.

The main advantages of REST over SOAP are:

- REST services are easier to write and use since they implement the architectural style of the Web itself
- a RESTful Web service response may be in XML, HTML, JSON or any other defined format while SOAP only allows XML
- REST doesn't have the overhead of headers and additional layers of SOAP elements on the XML payload and thus requires less bandwidth and is faster to decode

The disadvantages of REST with respect to SOAP are:

- SOAP can use almost any transport protocol while REST uses only HTTP/HTTPs
- REST uses transport level security inheriting the security measures from the transport layer, such as Transport Layer Security (TLS) or Secure Socket Layer (SSL) encryption protocols using HTTPs, while SOAP, in addition, can use its own security mechanism, WS-Security, that offers protection from the creation of the message to its consumption
- SOAP has comprehensive support for ACID (Atomicity, Consistency, Isolation, e Durability) for short-

lived transactions and Compensation Based Transaction Management for long-running transactions and also supports two-phases commit across distributed resources while REST doesn't
- SOAP support WS-ReliableMessaging for a guaranteed level of reliability of message exchanging while REST expects clients to deal with communication failures by retrying
- SOAP has a formal mechanism to describe the service using the Web Service Description Language (WSDL) thah allows to discover the service and generate a usable client proxy automatically

RESTful design is therefore appropriate when the following conditions are met:

- The Web services are completely stateless. A good test is to consider whether the interaction can survive a restart of the server
- The service producer and service consumer have a mutual understanding of the context and content being passed along
- Bandwidth needs to be limited
- A caching infrastructure can be leveraged for performance
- Easy of service delivery or aggregation is a requirement

1.5 Organisation of the paper

In Chapter 2 we will present an overview of the glider model requirements including considerations for Hardware in the Loop simulation, helping to explain the target architecture presented in Chapter 3; in Chapter 4 we will better detail the architecture of a Web service; in Chapter 5 we will see an extension of the architecture from the point of view of security; finally in Chapter 6 we will give an overview to the Graphical User Interface;

2 Glider model requirements

A glider simulator my come to hand in several applications, such as:

- **Decision support tools** where the simulator is used during planning and execution phases of a mission to explore the possible outcomes in a safe simulated environment, in order to optimize the mission goals while minimising the risk of asset loss or mission failure
- **Post mission analysis and debriefing tool** to reproduce and analyse what happened during a mission and to explore what would have happened if different decisions or conditions were applied
- **Test bed** for development and tuning of models, where the simulator is used to reproduce with fidelity the behaviour of a real unit in order to test path planning, motion control and optimisation algorithms under development
- **Virtual reality applications** where the simulator is used to reproduce the behaviour of a real glider introducing at the occurrence particular events, such as hardware/software failures or environmental anomalies, in order to train the personnel for a particular mission or to respond to particular events
- **Hardware in The Loop Virtual Simulation** to safely test missions and behaviours of a real glider hardware and software before deployment in the field

By an analysis of the use-cases provided, it is clear that the simulator has to be able to work either in fast-time mode and in real-time mode. Furthermore it should be able to work either as a *virtual simulator* with Man in The Loop or as a *constructive simulator*. The problems that it should be able to respond to are then ascribable to two main general cases:

- The prediction of glider future surface positions given its past history, its motion behaviour and the Oceanographic Model forecast
- The estimation of the underwater trajectory of the glider between two known (measured) surfacing points

Montecarlo simulation is required to account for the stochastic process in presence of noise or disturbance on the models inputs (e.g. assuming a variability on the environmental inputs due to the estimation errors of the environmental models or to account for the simulated sensors measurements noise).
For the estimation of the motion between two know points, forward and reverse techniques of smoothing could be used, in particular is interesting the application of the Unscented Rauch-Tung-Striebel Smoother[18] in which a separate backward smoothing pass is used for computing suitable corrections to the forward filtering result in order to obtain the smoothing solution.

2.0.1 Glider model

There is not a unique model that fits all the possible use-cases, for example for the accurate motion estimation in the three dimensional ocean environment within a virtual simulator, a 6 degree of freedom non linear hydrodynamic and kinematic model would be suitable to offer the required accuracy, while for a Decision Support tools, a faster, simplified stochastic model as used by [13] may offer better performances.
Depending on the use case the most appropriate model should be used.
Several 3D motion models are available in the literature [2] [11] [10] [22]. These models carefully describe how the glider moves as a result of the forces acting on it.
However to describe accurately the motion of the glider is also necessary to take into account the behaviour of the glider navigation control loop, usually called the *front-seat* and the higher level scripted mission behaviours which, all together, form the mission and are called the *back-seat*.
For the maximum accuracy, it's necessary to include models that, with the maximum degree of fidelity, repro-

duce the front-seat control loop and the back-seat algorithms.
The front-seat acts as a closed loop control that reads the navigation sensors on board the glider and generates control outputs toward the actuators of the rudder, ballast and bladder in order to obtain the desired motion of the glider. Gliders Sensor should also be modelled along with their measurement errors.
On top of the front-seat, an additional control layer, the back-seat is responsible to impose the glider a set of pre-programmed behaviours that together form the glider mission.
For back-seat modelling is also necessary to simulate glider communications and GPS fix readings.
For example communications would only be available after surfacing and GPS fix would only be taken after a variable acquisition time is elapsed during surfacing, with time to fix and communication ranges being affected by the weather conditions.
When the glider surfaces, a different dynamic/kinematic describes the gliders drifting as subject to wind, waves and surface currents.
Therefore a rather complete glider model is represented in figure 3 including the following sub-models:

- Glider Hydrodynamic Model: describe the forces acting on the glider while diving and climbing as a result of its buoyancy.
- Glider Surface Model: describe the motion of the glider while surfacing as a function of wind, waves and currents.
- Glider Kinematics Model: describe the motion of the glider.
- Glider Front Seat Model: describe the navigation control algorithms of the glider
- Glider Back Seat Model: describe high level mission behaviours of the glider: e.g. how it compensate a current.
- Simulated Sensor Outputs: provides the glider sensor outputs. E.g. pressure level, compass reading, Inertial Measurement Unit Output, GPS and the noise associated to the readings.
- Communications: long-range/low-bandwidth satellite communication and short-range/high-bandwidth HF radio communication.

2.1 Hardware in the Loop requirements

Unmanned Vehicles integrate a large number of complex hardware and software subsystem to manage the navigation and the execution of their missions. This level of complexity may be difficult if not impossible to describe with simplified models. Hardware in the Loop simulation comes to hand to test the real vehicle into a virtual environment before deployment in the real environment.
Significant work has been done at CMRE for the integration of HiL within the PARC project[3] for a SPARUS II AUV where the hardware and software of the AUV has been included in an HLA federation together with a virtual simulator. SPARUS II AUV is not a glider but a propeller powered underwater autonomous vehicle, it has a limited time/range autonomy and is less sensible to environmental conditions. Its software is implemented on top of the Robotic Operatin System (ROS)[17]. The approach used in this work was to bridge the ROS bus with the HLA federation by a ROS node which converts the information from ROS to HLA and vice versa.
Our approach to HiL capitalize on the work previously done within the PARC project, extending the concept by introducing a Web service that connecting to the glider hardware, feeds the sensors inputs (e.g. pressure, compass, GPS, IMU...) with simulated data, and uses actuators outputs to calculate the forces acting on the AUV to reproduce its motion in the virtual environment. Changes in water density, sea state, currents, wind and waves intensity and direction are considered by the dynamic and kinematic models to reproduce in the most accurate way the motion of the glider in the simulated environment. How well the HiL simulation reproduce the motion of the glider depends on the accuracy of the environmental and on the motion models of the glider.
HiL simulation could be used to test the correct functionality of the real glider subsystems, including its hardware components; to verify the correct implementation of the behaviours; to understand if a planned mission is feasible or not; and finally to train the personnel to the use of the systems in the most possible realistic conditions.
HiL simulation contributes to speed up development/test/validation time, reduce test/development costs and greatly minimize the risk during real missions execution.
HiL is also of great help when the algorithms of the behaviours and the navigation control loop are not known in the detail (for example for commercial AUV whose software is undisclosed) and have to be reproduced. In such cases, differences between the real and the reproduced behaviours and the control loop could lead to simulation results differ significantly from the real result.
A drawback of HiL simulations is that they are very slow. In most cases, unless the AUV implements a specific system to enable time acceleration, simulations with HiL requires to run in Real-Time. In the case of underwater glider missions, these could last for weeks or even months and the Real-Time execution may be inappropriate. In such cases HiL could still be used to explore a range of particularly significant cases that could be used to tune and then validate the parameters of the equivalent models which could then reproduce the AUV behaviours in fast-time simulations.

3 Simulation Engine architecture

The work done on the definition of the target architecture follows the recommendations of MSG-131[6] on development of MSaaS target architectures. The resulting target architecture is derived from the SD VIntEL NATO Reference Architecture and has the following characteristics:

- Interoperable

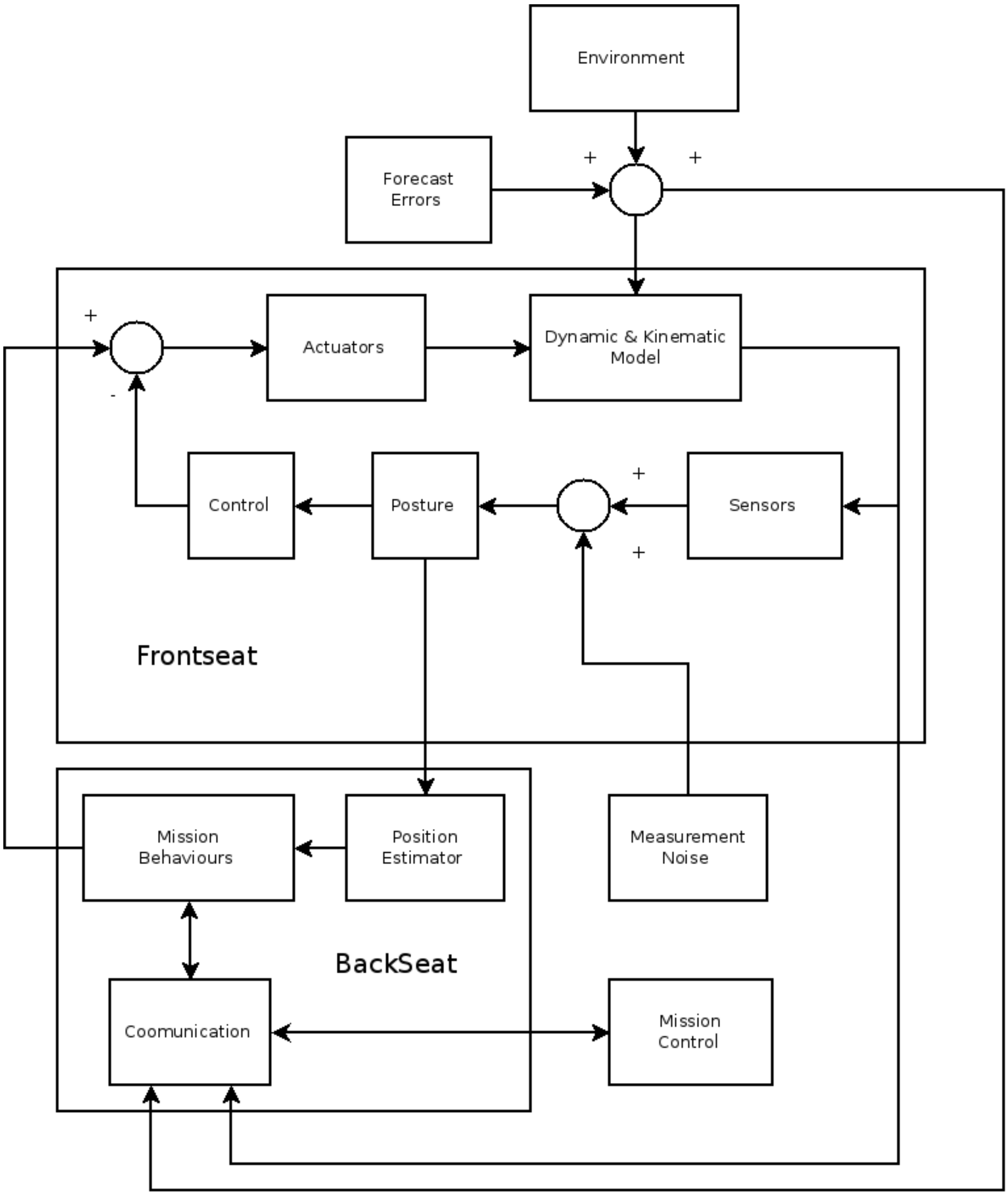


Figure 3: Glider model block diagram

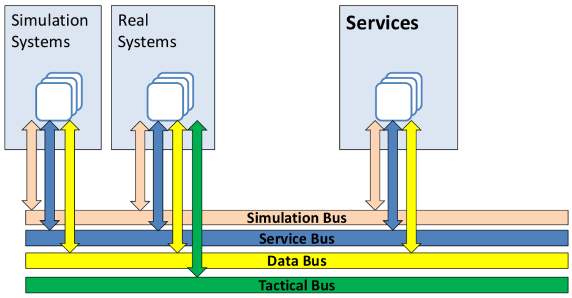


Figure 4: VIntEL Reference Architecture

- Deployable
- Maintainable
- Scalable
- Reusable: services can be used by multiple processes or composed into other services
- Independent units of business functionality: each service provides a business function that is independent of other services.
- Loosely coupled
- Platform-independent
- Based on open standards
- Low-Cost

3.1 SD VIntEL NATO Reference Architecture

MSG-131 final report on MSaaS: New Concepts and Service-Oriented Architectures [6] presents a number of reference architectures to use for MSaaS implementations.
Reference architectures are generic blueprints that may be used for deriving specific target architectures. When possible, building target architectures for specific simulation systems or simulation environments on foundations from established reference architectures will increase not only the efficiency of work in time and budget, but also the quality of the results, and will lead to improved interoperability.
The SD VIntEL reference architecture, depicted in figure 4, defines four buses for Data Exchange:

- Simulation bus, to exchange simulation data using HLA Evolved [14]
- Service bus, to connect with services using an Enterprise Service Bus.
- Data bus, for high volume/bandwidth data (such as camera streams or sensor data)
- Tactical bus, for C2 Interoperability.

It also classifies services into three different types:

- Domain Services: services that fulfil a specific task within the simulation
- Infrastructure Services: services used to control the infrastructure
- Adaptor Services: services that connect features outside the infrastructure

SD VIntEL has been used as a high level reference for the target architecture implementation, however significant deviations have been taken from the reference design.

3.2 CMRE Distributed Simulation Framework - Target Architecture

The target architecture, design fully embeds the recommendations contained in the MSG-131 final report on MSaaS [6] on System Design (SD) and on Simulation Environment Data (DA) and in particular:

- SD-1: Design and Document for Interoperability
- SD-2: Design and Document for Modularity and Composability
- SD-3: Favour Open Standards
- SD-4: Design for Securability
- DA-1: Enforce "Single Source of Truth" Principle

As depicted in figure 5 it has been designed in order to create a distributed simulation framework extending the work previously done in CMRE on interoperable M&S applications [3] [5].
The main feature of this architecture is that the simulator engine is structured as a Web service and is completely decoupled by the GUI that is optional and could be provided as a Web GUI complemented, in case of virtual simulation applications, by a Virtual-Reality Console.
The result is a flexible, modular, scalable and interoperable architecture that could be easily reused and expanded to quickly respond to upcoming requirements.
The architecture of the simulation engine is composed by the following functional blocks:

- Communication buses
- Time Manager and Real Time Clock
- Event Generators
- Discrete Event Processor
- Simulation Engine Core
- Locator
- Storers
- Rest Control Interface
- Services Facade

3.2.1 Communication buses

The target architecture uses four communication buses as the VIntEL architecture:

- Service bus, based on HTTP to connect with RESTful Web services. Ideally the simulation engine, should be connected only to the Service bus, delegating to specific Web services the role of interfacing with the other buses in order to encapsulate specific bus functionality into well defined, reusable Web services. However, mainly for a performances matter this is not always feasible.
- Simulation bus, to exchange simulation data using HLA Evolved [14]
- Tactical Data Link (TDL), using Over The Horizon - Gold (OTH-Gold)[26])messages over TCP for C2 Interoperability

ISBN 978-88-97999-90-4; Bruzzone, Cayirci and Sottilare Eds.

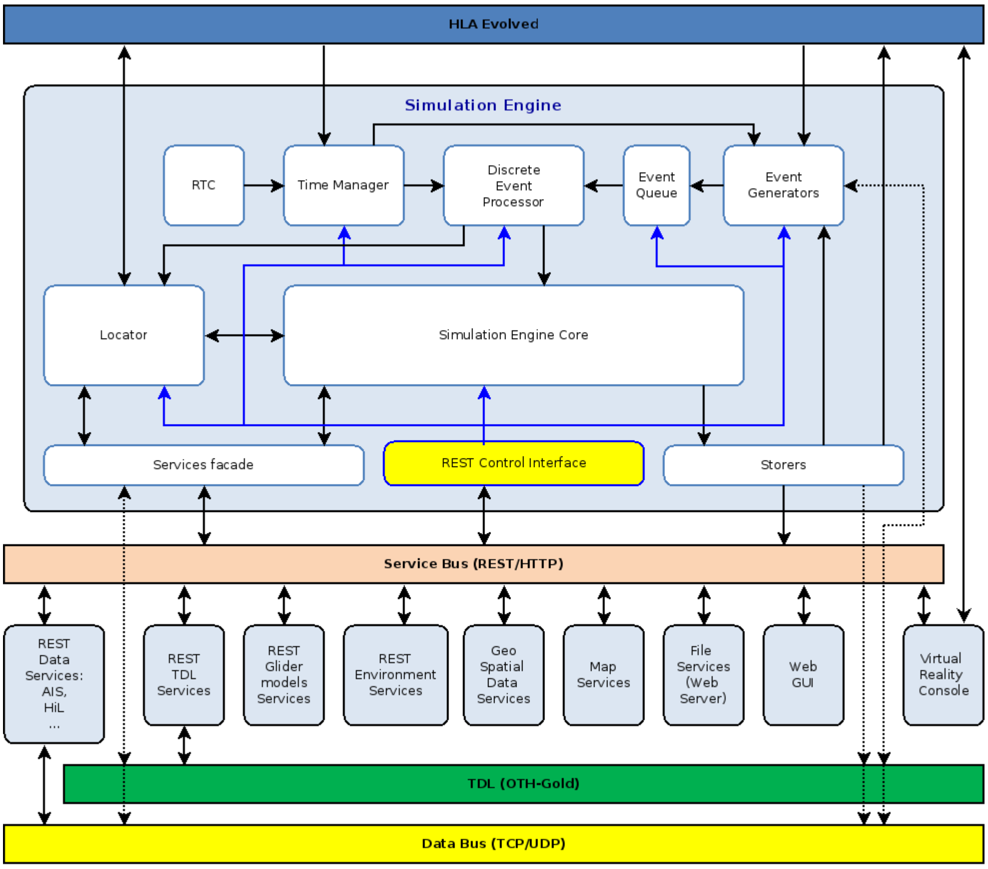


Figure 5: CMRE Distributed Simulation Framework - Target Architecture

- Data bus, TCP/UDP for high volume/bandwidth data (such as streaming sensors data)

Service bus: considering that HTTPs transport provides an adequate level of security for the type of application the Service bus has been implemented using RESTful Web services over HTTP/HTTPs rather than implementing a much more complex and cumbersome Enterprise Service Bus (ESB) solution as in the VIntEL architecture. An alternative to RESTful Web services is to use the more structured and secure SOAP based Web services but, REST has been preferred because of its lightweight and better performances.
Simulation bus: it is using the standard IEEE 1516-2010 known also as HLA Evolved[14]. Ideally HLA connectivity should be provided as a Web service, however it was estimated that, due to the intense interactions between the federate and the RTI (due to time management message exchange, object attributes updating and interactions firing), the HLA/REST conversion would have resulted in a sensible degradation of performances. For this reason the simulation engine is directly connected to HLA rather than through an HLA Web service.
Tactical Data Link bus: it is based on the OTH-Gold[26] protocol, using TCP/IP as transport. OTH-Gold provides a standardized method for transmitting selected data between C2 Systems. It is the primary message format for Tactical Data Processor (TDP) exchange. It is designed to be easily man readable for the non specialized user. Other, more powerful, but also more complex Tactical Data Link systems exist, such as the Link 11/16/22, but are not considered by the current architecture. They could however be included, when requested, by adding appropriate interface modules. Interface to the TDL is done by a specific Web service. The simulation engine connects to such service to subscribe to Tactical Data feeds available on the TDL bus. The TDL Web service then forwards data feeds on the TDL bus to the simulation engine by using a specific Application Programming Interface (API) provided by the simulation engine REST Control Interface.
Data bus: relies on a TCP/IP transport to interface with Real Systems. For example the Automatic Identification System (AIS) that provide streaming data on commercial shipping traffic.
As for the TDL bus, specific Web services would normally interface with the Data bus, however, in case of special requirements a direct connection to these buses from the simulation engine is always possible (dashed lines).

3.2.2 Time Manager and Real-Time Clock

The Time Manager manages simulation time in multiple ways:

- Real-Time: events are executed at the same speed of a Real-Time Clock (RTC),

- Accelerated-Time (2x, 3x...) is as the Real-Time case except that the Real-Time Clock output is accelerated by a factor.
- Fast-Time: events are processed as soon as possible, the Time Manager doesn't respect the interval between two events, but just execute the next event available in the queue as soon as the previous one is completed

Time Manager supports HLA time regulation to provide synchronization of federates progress.
HLA Federates may be designated as regulating federate. Regulating federates regulate the progress in time of federates that are designated as constrained. In general, a federate may be "regulating", "*constrained*", "*regulating and constrained*", or "*neither regulating nor constrained*". By default, federates are neither regulating nor constrained.
The RTI recognizes every federate as adapting one of these four approaches to time management. A federation may be comprised of federates with any combination of time management models. That is, a federation may have several federates that are regulating, several federates that are constrained, or several federates that are regulating and constrained.
A federate that becomes "*time regulating*" may associate some of its activities (e.g., updating instance attribute values and sending interactions) with points on the federation time axis. Such events are said to have a "*time-stamp*". A federate that is interested in discovering events in a federation-wide, time-stamp order is said to be "*time constrained*". The time management services coordinate event exchange among time-regulating and time-constrained federates.
Real-Time application with Man in the Loop, e.g. a mission control console to set the gliders missions, turn the simulator into a virtual simulator.
Time Manager is responsible of triggering the Discrete Event Processor.

3.2.3 Event Generators

The Event Generators is an aggregation of asynchronous (living on different threads) event generators which specialize the `EventGenerator` interface as visible in figure 6.
The available event generators are:

- `HLAEventGenerator`: generates events when objects attributes updates or interactions are fired by a federate on the HLA bus
- `TDLEventGenerator`: generates event upon receiving data on the Tactical Data Link
- `RealTimeEventGenerator`: generates events upon receiving data on the Data Bus
- `HiLEventGenerator`: generates events upon receiving data from the AUV hardware (either though the Data Bus or the HLA)
- `LocalEventGenerator`: generates events from the Simulation Engine Core
- `ScriptedEventGenerator`: generates a predefined set of events from a script file, an example of this could be a predefined list of waypoints to follow at given times, or other scripted information such as behaviours changes, alarms, HW failures...).

Multiple event generators can be active at the same time. Changing the configuration of the event generators allows to match the behaviour of the simulator to the intended use, e.g. as constructive simulator providing information to a decision support suite or as a virtual simulator with Man in the Loop and/or Hardware in the Loop for personnel training and/or HW/SW testing and validation purpose.
Event generators generate events asynchronously. Generated events are pushed into a synchronized (thread safe), time ordered queue so that they can be consumed by the Discrete Event Processor.

3.2.4 Discrete Event Processor

The Discrete Event Processor is designed to process events that multiple, asynchronous event generators push in a queue. Events in the queue are ordered by their time-stamp, not by arrival order, which is not equivalent since events generation time can in general be different from the time they are supposed to be executed (time-stamp).
The Discrete Event Processor is triggered by the Time Manager. At time ticks, provided by the Time Manager it processes timed events using a three-phase approach[25]: in the first phase it jumps to the next chronological event in the queue occurring at the time provided by the Time Manager; in the second phase it executes all events that unconditionally occur at that time (B-events); finally, in the third phase, it executes all the events that conditionally occur at that time (C-events). For maximum performances events can be processed concurrently applying a *Parallel Discrete Event Simulation*[24] approach .
Timed events can be of multiple types such as:

- model-update events, high frequency events that cause an update of the position, models and internal status
- federation-update events, coarser events that cause the Simulation-Engine to fire object attributes updates to the federation
- external events, coming from outside, such as the change of behaviours or mission, the order to release emergency ballasts, the turning on/off of sensors, the occurring of radio transmission, collisions with external assets or bottom structures, commanded hardware/software failures, events, not related to gliders that could be of interest by the simulation community such as ammunition detonation/hit/miss, radio transmission disruption...
- internal events such as collisions with the bottom, pressure alarms, low battery alarms, surfacing/diving of the glider
- timed events such as alarms or specific actions to be executed at a predefined time

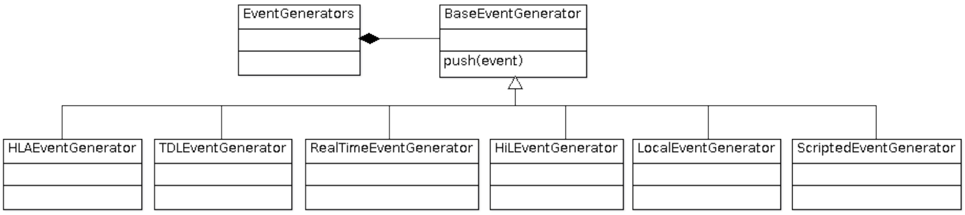


Figure 6: EventGenerators is an aggregation of EventGenerator objects

Events can be inserted with any orders, however they are always executed in order of time-stamp
Events are dispatched to the underlying Simulation Engine Core which updates the simulation status using a combined Discrete/Continuous models approach: the behaviour of the model is simulated by computing the values of the state variables at small time steps while the values of attributes and global variables is calculated at coarser event times.

3.2.5 Simulation Engine Core

The Simulation Engine Core calculates and stores the status of the simulator. It is event driven and uses the underlying models and the event information to update the status of the simulated unit e.g. last position, battery level, sensors/actuators status, HW/SW failures, alarms, damage level, collisions...
UML structure of the Simulation Engine Core and relationship with the Event Generators and the Discrete Event Processor is visible in figure 7.
The Simulation Engine Core is triggered by events received from the Discrete Event Processor and it may in turn generate events (e.g. diving/surfacing, collisions with the Sea bottom, depth/pressure alarms, HW/SW failures....) and dispatch them back to the Event Queue or to the buses through the Storers interface.
Events are dispatched to HLA in form of *HLA interactions*. Updates of the internal status are published to HLA through the Storers interface as *HLA objects attributes updates*.
Events and attribute updates may also cause information dispatching on the TDL or the Data Bus such as platform position updates to C2 or other real systems (e.g. internal status information, sensor performance predicted data calculated by underlying models...)
It has to be noted that The Simulation Engine Core doesn't calculate the location of the platform since this is the main responsibility of the Locator model.

3.2.6 Locator

The Locator model is responsible of calculating the six degrees of freedom location of the simulated platform along with it's bearing and speed. It is implemented as an interface. Multiple implementation of such interface are available, depending on the wished use/application of the simulator. However only one locator instance, per simulator, may be active at the time. The implementations are:

- HLA locator: interface with HLA to provide the location, speed and bearing calculated from an external dynamic/kinematic simulator engine.
- TDL locator: receive surfacing points or last known coordinates of the glider from the Tactical Data Link and uses its internal models to compute the coordinates of the glider (dead-reckoning) in between two known points or to predict the most probable surfacing point given the last know diving point. It works either in Real-Time and in Fast-Time.
- Real-Time locator: it works as the TDL locator except that receives the platform coordinates from the glider when surfacing through a TCP/IP dedicated connection.
- Hardware in the Loop (HiL) locator: interfaces directly with the glider sitting on the bench, to test the glider motion control algorithms (front-seat) and mission programmed beahaviors (Back Seat) complementing and extending the work previously done at CMRE within the context of the Persistent Autonomous Reconfigurable Capability (PARC) project with the Sparus II AUV reported in [3] and as in such implementation it may as well include an HLA interface as a bridge between the Robotic Operating System (ROS) of the AUV and the HLA federation: while the glider believe to be navigating in the ocean, its actuators control values are sent to this locator that uses them to estimate the glider position in the 3D environment, returning to the glider the simulated sensor readings (e.g. compass value, inertial navigation units values and water pressure). In this type of simulation the human glider pilot, controlling remotely the glider from the glider mission control room is as well included in the loop and the simulator can be used either to train the glider pilot as in a real mission but with the glider safely operating in the workshop bench.
- Glider Simulator locator: making use of the dynamic and kinematic models of the glider, reproduce the glider motion in the complex ocean environment. It also employs models to simulate the glider motion control algorithms (front-seat) and the

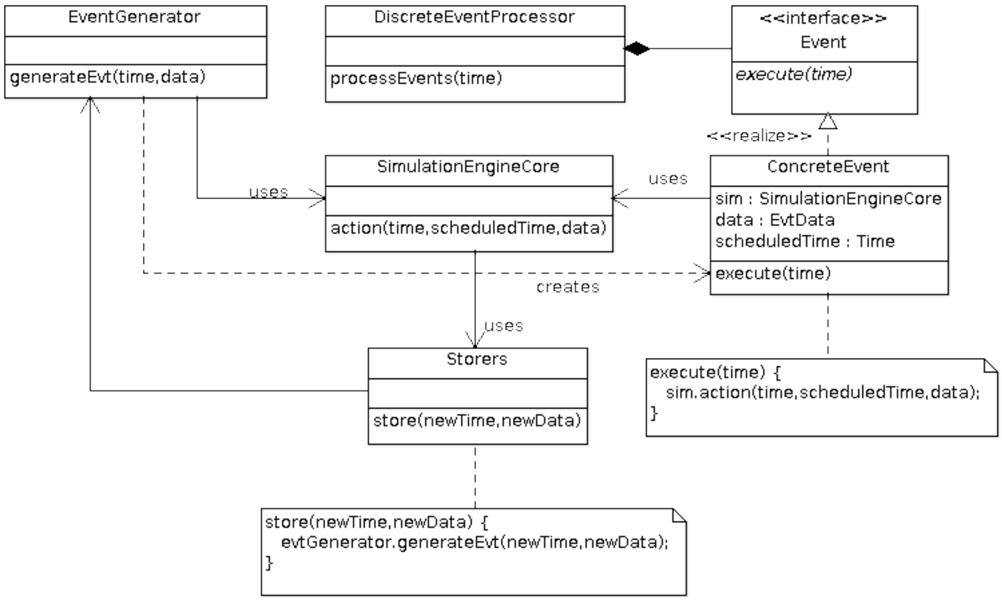


Figure 7: The Simulation Engine Core is event driven and implements a "Command" design pattern

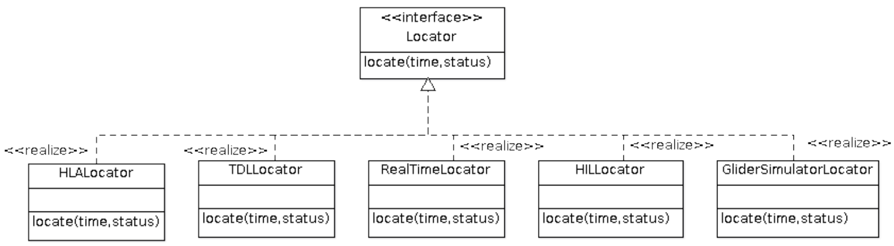


Figure 8: Locator UML diagram

glider mission programmed behaviours (back-seat).

The interchangeable Locator model is a flexible mechanism to adapt the locator to the application requirements. By changing the Locator model the Simulator may behave either as a Constructive Simulator providing information to a decision support suite or as a Virtual Simulator with Man in the Loop and/or Hardware in the Loop for personnel training and/or HW/SW testing and validation purpose.

3.2.7 Storers

The Storers module is in charge of the output of the Simulator and follows recommendation SD-3: Favor Open Standards of the MSG-131 final report on MSaaS [6].
Output data is provided in a variety of formats, however open standards are preferred to favor interoperability. For example the glider position is provided in the KML format to be displayed by any mapping/chartographic system compatible with such format (e.g. GIS software or Google Earth), in HLA format to interface with the majority of NATO Simulators and in the OTH-Gold Tactical Data Link format for compatibility with the majority of NATO Command and Control systems.
Other open data formats provided include the Open Geospatial Consurtium (OGC) data formats, NetCDF, GeoTIFF and the Matlab. HTML dynamically generated Web pages could be also provided.
Storers is a composition of classes which specialize the `Storer` interface as visible in figure 9. A storer is a module that dispatch information received from the Simulation Engine Core onto an output channel. Different storer models specialize the behaviour of the `Storer` interface.
Storers are called in cascading using a flavor of the Chain of Responsibility design pattern, so that only the storer responsible for a certain type of information will be able to forward such information toward the appropriate channel. A channel may be one of the buses, the internal queue or even a file such as a log file, or a specific format file, such as an HTML page to be displayed by a Web browser or a Matlab file to undergo further processing in an external tool.

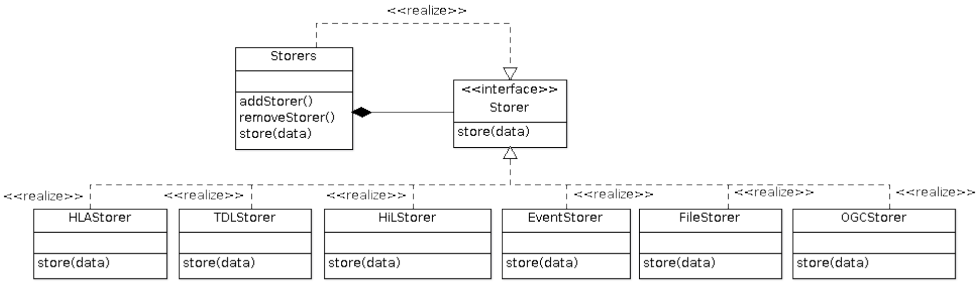


Figure 9: Storers is a composition of Storer objects

The available storers are:

- `HLAStorer`: send objects attributes updates (such as the glider location) or interactions to the HLA federation as a consequence of an internal status change or a generation of an internal event (e.g. a collision with the bottom, a HW fault, an alarm such in case the maximum depth is exceeded...)
- `TDLStorer`: send own track position to the Tactical Data Link
- `HiLStorer`: provide simulated sensors output to the HiL (either though the Data Bus or the HLA)
- `EventStorer`: generates events received from the Simulation Engine Core and sends them to the Event Generator to be inserted in the queue
- `FileStorer`: produces various types of files (e.g. HTML, KML, Matlab, JSON, CSV, XML, NetCDF, GeoTIFF...) and pushes the to a local Web server, available for download from a client application such a Web browser.
- `OGCStorer`: its an extension of the FileStorer that pushes geo-spatial files (e.g. NetCDF GeoTIFF...) to the Geospatial Data Service that in turn will provide them to a client application in the form of as OGC Web services such as Web Feature Services (WFS) or Web Map Services (WMS).

3.2.8 REST Control Interface

This is the REST front-end interface of the simulation engine. It makes the simulation engine available on the Services bus just as any other REST service. API available in the REST Control Interface include functionality to for:

Simulation instances management

- Create a new simulation
- List created simulations
- Destroy a created simulation
- Join a simulation

Simulation execution management

- Start simulation
- Pause simulation
- Stop simulation

Simulation status and parameters editing management

- Get simulation status
- Edit simulation parameters (e.g. real-time/fast-time, montecarlo runs, file/services outputs...)
- Edit models and environment parameters

Feed data from external sources

- Push in data e.g. coming from subscribed services (TDL, AIS...)

REST Control Interface runs in an Independent Process from the simulation engine. In this way it may create or destroy instances of a simulation engine as required.

3.2.9 Services Facade

The services facade module provides a unified interface to the set of underlying services defining a higher-level interface that makes the subsystem easier to use. This module hides the complexities of the larger system and provides a simpler interface to its clients: the Locator and Simulation Engine Core. It includes a single wrapper class that contains the set of member methods required by the clients. These members access the underlying services on behalf of the clients, hiding the implementation details.

3.3 Web Services

The glider models, environmental services and data services required by the simulator engine are provided in the form of Web services. For optimal performances and since the transport layer security provided by HTTP/HTTPs is considered to be satisfactory, RESTful Web services are used.
Basically the types of services that can be found are:

- Glider models such as the glider dynamic and kinematics, front-seat, back-seat, sensor models...
- Environmental services (bathymetry, climatology and weather forecasts)
- Data services to connect with external systems (e.g. Hardware in the Loop connection services; Automatic Identification System (AIS), that returns navigation information of maritime commercial traffic...)

- Geospatial Data services
- Map services

Some of these services have been implemented from scratch, some others are Open Standard services such as the Geospatial Data services or the Map services, based on Open Source Projects (e.g. GeoServer, Thredds or Mapnik).

4 Web services architecture

Web services architecture is represented in figure 10: a REST front-end servlet (an application, running on the server side, capable to respond to HTTP requests) written in Java and running into a Servlet container such as the Apache Tomcat represent the service front-end; it is connected through a Remote Method Invocation (RMI) or a Remote Procedure Call (RPC) to one ore more Micro-Services.

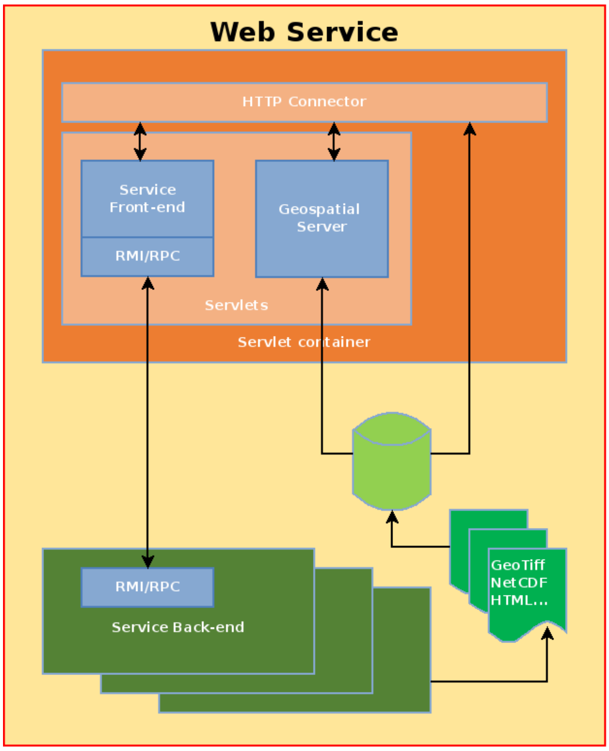


Figure 10: RESTful service architecture

Micro-Services may return directly their result to the Front-end, that would return it to its client or in alternative, in case of GeoSpatial data, they could publish it onto a Geospatial Data Server (such as Geoserver of the Open Source Geospatial Foundation or the Unidata Thredds), returning to the client the Unified Resource Locator (URL) of the published data. In such way the data becomes available to the Client using the OGC standards such as WMS or WFS.
REST is a good compromise to create lightweight, loosely coupled, language and operating system independent, distributed architectures. Behind the REST front-end a number of Micro-Services, written in different languages, represent the back-end of the service.

For example a mix of Matlab and Python processes would provide the glider mathematical models.
Interconnection between the Java front-end and the back-end of the service could be implemented using a number of different technologies:

- for C, C#, C++ (on POSIX systems), D, Delphi, Erlang, Go, Haxe, Haskell, JavaScript, node.js, Ocaml, Perl, PHP, Python, Ruby, Smalltalk back-ends, the Open Source Apache Thrift[21] framework for cross-language remote Procedure Calls (RPC), developed at Facebook to expedite development and implementation of efficient, scalable, cross languages services, has been highlighted as the optimal choice.
- when the back-end is Java, the most appropriate solution is the Java Remote Method Invocation (RMI). Apache Thrift represents however a suitable alternative.
- Matlab back-ends could be integrated with the Java front-end using the Matlab Compiler System Development Kit (SDK) which provide a Java library to interface directly with the Matlab Runtime engine.

In addition to the simulator engine a number of supporting Web services is included in the target architecture.

4.1 Environment Services

Environmental data services provide the simulation scenario with the following information:

- Bathymetry
- Climatology: historical seasonal/monthly average sea water temperature and salinity
- Weather forecast: sea state, waves height/direction, wind, rain, water salinity and temperature

Environmental services can be used either to: reproduce the conditions found during a real mission using recorded historical data; to plan a new mission, using either forecast data or typical seasonal data (climatology); and also to explore what would happen by applying or superimposing certain conditions on the scenario area (e.g. forcing a sea state or wind direction, applying a current field or introducing an anomaly in the salinity of the water at a certain depth that could severely impact on the navigation capabilities of the glider).

4.2 Glider models

Include a number of services that implement glider models or models used by glider models described in Chapter 2. The simulator aim also to become a test-bed and tuning tool for dynamic and kinematic gliders models.

4.3 File Services

Files produced during the simulation execution, such as HTML, KML, Matlab, JSON, CSV, XML, NetCDF, GeoTIFF are pushed to a local Web server and become available for download from a client application such a Web browser and can be further processed by external tools.

ISBN 978-88-97999-90-4; Bruzzone, Cayirci and Sottilare Eds.

4.4 Geo Spatial Data Services

Geo-spatial files (e.g. NetCDF GeoTIFF...) are provided by the Geospatial Data Service in the form of OGC Web services such WFS or WMS.

4.5 Map Services

Map Services are used to provide map tiles to a mapping application. Such services are used by the Web GUI to display the map with the geo-referenced positions of the Glider. The availability of a map-service prevent the need to connect to an online map data service provider making the simulation engine independent from the availability of an internet connection.

5 Notes about securing the system

Exposing Web services to Wide Area Networks such as the Internet requires that a number of precautions are taken against malicious attacks. NATO accreditation must be obtained before the service is made available on the Internet. The main aspects of the security accreditation are based on the NATO Communication and Information Agency (NCIA) requirement to follow the Open Web Application Security Project (OWASP) Application Security Verification Standard including:

- Authentication
- Session Management
- Access Control
- Input Validation
- Cryptography
- Error Handling and Logging
- Data Protection
- Communication Security
- HTTP Security
- Business Logic
- Files and resources
- Miscellaneous requirements

The original architecture shown in figure 10 was subsequently expanded to that shown in figure 11. It may be seen that the original architecture now forms only a sub section of the revised architecture. In particular the revised architecture includes firewalls, Web server, Service Provider, Identity Provider, Directory Server and a users database.
A user that need to access a service connects to the Service Provider hosted on the Web server in the Demilitarized Zone (DMZ), the area comprised between the two firewalls in which no data or software is stored except the Web server own log files. If the user has not been already authenticated it is redirected to the Identity Provider. The Identity Provider performs the user Authentication checking the supplied credentials against those stored in the Users Database through the Directory Server (usually a Lightweight Directory Access Protocol (LDAP) service). Once authenticated the user is redirected to the proper Web service. This process is further filtered according with the Authorizations that each user owns that are contained in the Users Database.
Confidentiality and Integrity of the data are granted on the Client side by the HTTPs transport protocol. Firewalls filter the input/output traffic. The Web server in the DMZ acts as a secure front-line for the incoming requests redirecting clients traffic to the back-standing Web services. Communication between the Web server and the Web services could be done using the lighter and faster HTTP protocol (with no encryption) however communication on the client side is always protected by HTTPs TLS/SSL encryption.
Such architecture allows enforcing the security requirements required to undergo the accreditation process that includes the following phases (extracted from Communication and Information Services (CIS) Security Technical and Implementation Directive for the Security of NATO's Internet Presence):

- Security Level Review
- Design and Architecture Review
- Vulnerability Scans
- Health Checks
- Code Walk-through
- Application Penetration Testing
- Code Review
- Create and Review UML Models:
- Create and Review Threat Models:
- Test Configuration Management

6 Graphical User Interface

The GUI is decoupled from the simulation engine and is optional: for applications where the simulation engine is used as a computation engine the GUI may not be used at all and the hosting application would just interface to the simulation engine directly.
For end user applications where a Graphical User Interface is required a Web application is the optimal solution. Web applications do not need to be installed or configured. The user just need to know the URL address of the application to be able to use it.
The main components of the Web GUI are:

- Simulation Management Controls: to create, list, join or destroy a simulation.
- Simulation Time Management: to start, pause, stop the simulation execution
- Simulation status: display the simulation status in various forms: tables, plots, graphs...
- Simulation parameters editing: to edit simulation parameters
- Simulation models editing: to edit the models parameters
- Map View: to display a Map with the simulator graphical output data overlaid

Web GUI components may be combined into different views. Authorizations provided by an external Identity Provider define which controls and features are available to authenticated users.

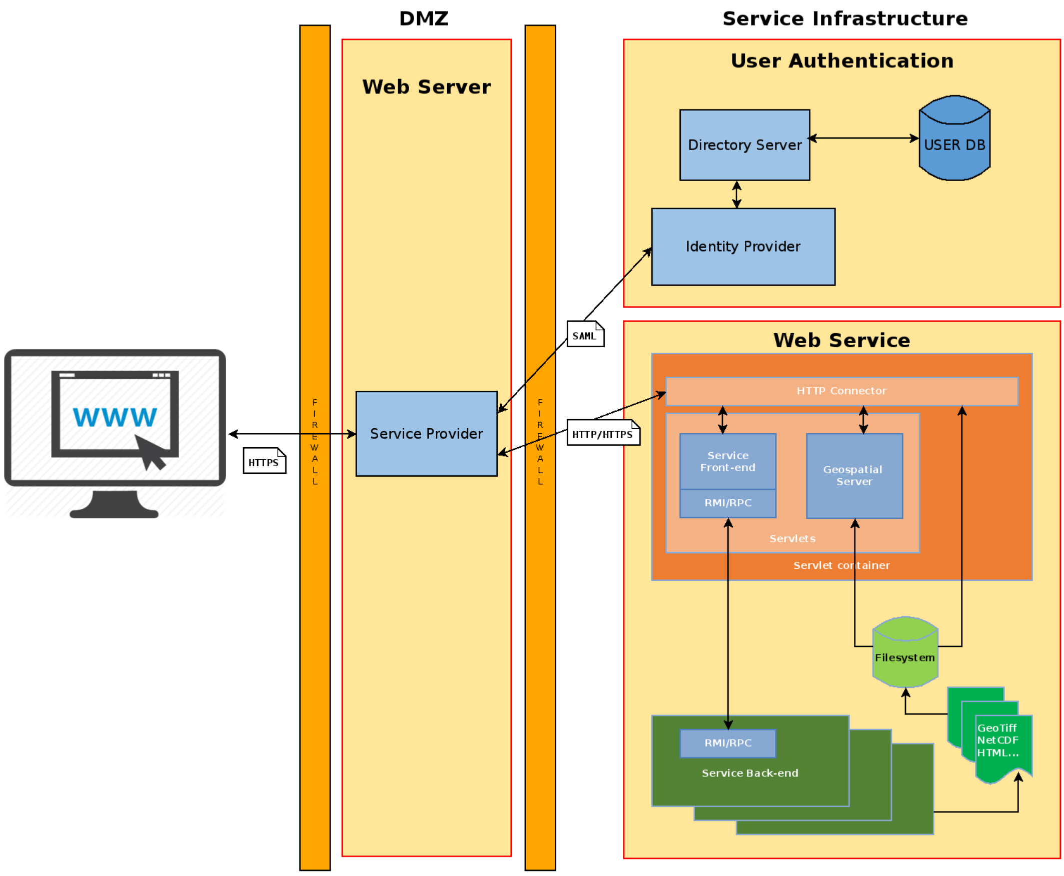


Figure 11: Securing Web Services

6.1 Virtual-Reality console

For Virtual-Simulation use cases, a standalone application with HLA interface represents the most natural solution but moves a little bit out of the MSaaS concept since it would not be provided anymore as a service in the more strict sense of the term but would require the installation of the application package on the client terminal. On the other side, implementing the Virtual-Reality console as a Web application, in order to provide it as a service to the end user, would introduce aspects related to graphics and communication performances from within a browser that are not in the scope of this paper.
The Virtual-Reality Console stand alone application connects as a federate to the HLA federation and subscribe to the objects attribute updates and interactions of interest.
In case of multiple simulator federates, an additional federate would be required to manage the interactions between the federates, such as collisions between the simulation entities and/or static features e.g. an underwater structure.
To avoid interoperability issues between federates due to the use of uncorrelated environmental data, it is necessary to enforce the MSG-131[6] recommendation DA-1 Enforce Single Source of Truth Principle through the whole federation that requires that each environmental data source is unique and all application-specific data items or formats are derived from such source data item.

6.2 The Map View

The Map View allows to display geo-referenced output of the simulator onto a map. In figure 12 it is possible to see the Map View with the AUV icon in Mil STD 2525D[9] AUV icon and a sea velocity field WMS layer overlay.
Information that can be displayed as map overlay include:

- Platforms positions(e.g. the simulated glider icon or other entities received from HLA/TDL/AIS) along with their routes and tracks using the KML format
- environment data layers such currents, winds, waves, water temperature/salinity as WMS data layers
- Models/simulation produced data layers (e.g. sensor ranges, risk areas....)
- Navigation charts
- Elevation and bathymetric contours
- Weather forecast data

Additional layers can be displayed as WMS layers, vector layers (e.g. KML or GeoJSON) or as tiled layers.

References

[1] D. Pochazka, J. Hodiky, “Modelling and Simulation as a Service and Concept Development and Experimentation”, 2017 International Conference on Military Technologies (ICMT)

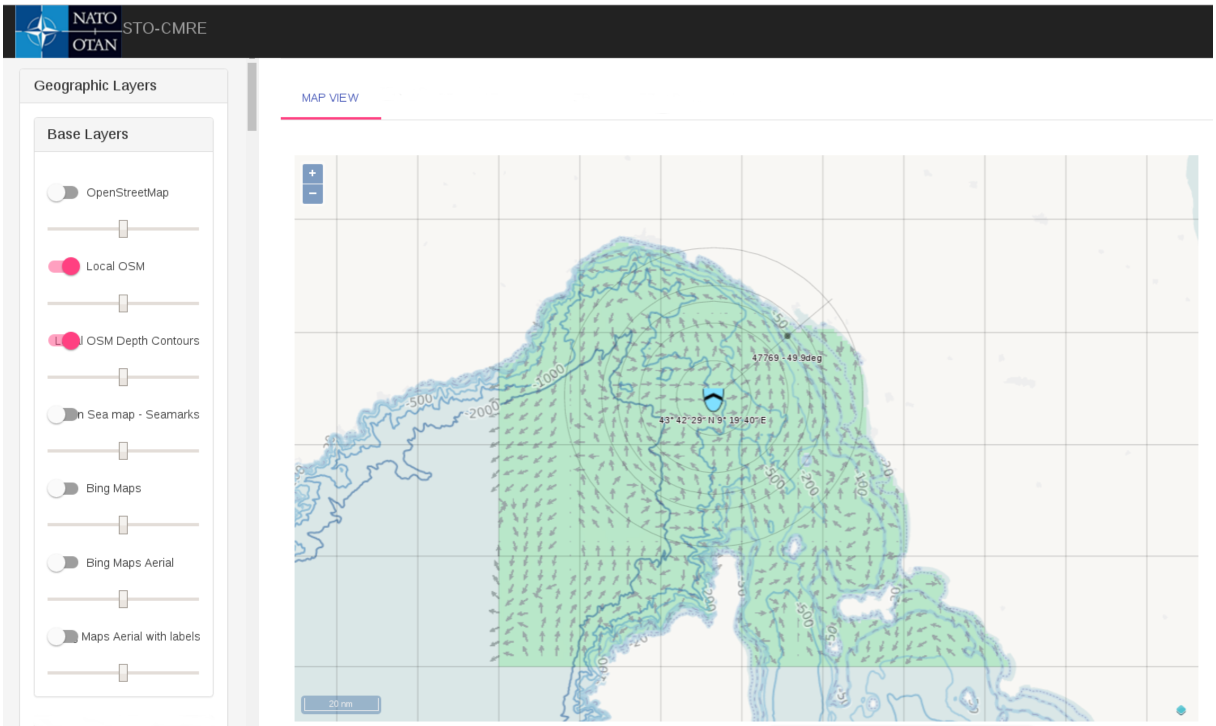


Figure 12: The Graphical User Interface MapView with a sea water velocity field and a Mil STD 2525D AUV icon overlayed

[2] U. Latifa, T. W. O. Putri, B R Trilaksono, E. M. I. Hidayat, "Modelling, Identification and Simulation of Autonomous UnderwATER Glider in Longitudinal Plane for Control Purpose", 2017 2nd International Conference on Control and Robotics Engineering

[3] A. Carrera, A. Tremori, P. Caamao, R. Been, D. C. Pereira, A. G. Bruzzone, "HLA interoperability for ROS-based autonomous systems", 2016, International Workshop on Modelling and Simulation for Autonomous Systems (pp. 128-138). Springer International Publishing.

[4] "Operational Concept Document (OCD) for the Allied Framework for M&S as a Service: DRAFT v0.28", 09 November 2016

[5] A. Bruzzone, A. Tremori, M. Oddone, D. Crespo Perreira, M. Scapparone, S. Vasoli, M. Corso, I. Piccini, "HLA based interoperable simulation: MCWS-MSTPA federation", La Spezia: CMRE, 2015/01. (CMRE Memorandum Report (MR series) CMRE-MR-2015-002

[6] Final Report of Specialist Team MSG-131, "Modelling and Simulation as a Service: New Concepts and Service-Oriented Architectures", 2015, STO Technical Report

[7] NATO STO: Final Report of NATO MSG-086 "Simulation Interoperability", STO Technical Report, STO-TR-MSG-086-Part-I, January 2015.

[8] "Modelling And Simulation Architecture Standards For Technical Interoperability: High Level Architecture (HLA)", 2015, North Atlantic Treaty Organization

[9] "MIL-STD-2525D", 10 June 2014, Department of Defense interface Standard, Joint Military Symbology

[10] E. Cayirci, "Modelling and Simulation as a Cloud Service: a Survey", Proceedings of the 2013 Winter Simulation Conference

[11] K. Isa, M. R. Arshad, "Buoyancy-driven underwater glider modelling and analysis of motion control", Indian Journal of Marine Science, Vol 41(6), December 2012

[12] C. Wang, A. Anvar, "Modelling and Simulation of Motion of an Underwater Robot Glider for Shallow-water Ocean Applications", World Academy of Science, Engineering and Technology, International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering, Vol:6, No 12, 2012

[13] R. Grasso, D. Cecchi, M. Cococcioni, C. Trees, M. Rixen, A. Alvarez, C. Strode, "Model based decision support for underwater glider operation monitoring", OCEAN 2010 MTS/IEEEE, Seattle

[14] "IEEE Standard for Modelling and Simulation (M&S) High level Architecture (HLA) Framework

and Rules”, IEEE std 1516-2010 (Revision of IEEE Std 1516-2000), pp. 1-38, 2010.

[15] B. Garau, M. Bonet, A. Alvarez, S. Ruiz, A. Pascual, “Path planning for authonomous underwater vehicles in realistic oceanic current fields: application to gliders in the western Mediterranean sea”, 2009, Journal of Maritime Research Vol. VI N. II

[16] K. Al-Zoubi, G Wainer, “Using REST Web-Services Architectures for Distributed Simulation”, 2009, ACM/IEEE/SCS 23rd Workshop on Principles of Advanced and Distributed Simulations

[17] M. Quigley, K. Conley, B.P. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, n. Andrew, “ROS: an open-source Robot Operating System”, 2009, ICRA Workshop on Open Source Software

[18] S. Sarkka, “Unscented Rauch–Tung–Striebel Smoother”, 2008, IEEE Transactions on Automatic Control (Volume: 53, Issue: 3, April 2008)

[19] O. Schofield, et al, “Slocum gliders: robust and ready” Journal of Field Robotics, 24(6), 473-485 (2007), Wiley InterScience.

[20] “NATO Consultation, Command and Control Board (NC3B): NATO Architecture Framework”, Version 3, AC/322-D(2007)0048-AS1, 2007

[21] M. Slee, A. Agarwal, M. Kwiatkowski, “Thrift: Scalable Cross-Language Services Implementation”, 2007, Facebook, 156 University Ave, Palo Alto, CA

[22] Bhatta, Pradeep, “Nonlinear stability and control of gliding vehicles”, 2004, Diss. Princeton University

[23] R.T. Fielding, “Architectural Styles and the Design of Network-based Software Architectures”, 2000, UNIVERSITY OF CALIFORNIA, IRVINE, Dissertation

[24] A. Cassel, M. Pidd, ”Distributed discrete event simulation using the three-phase approach and Java”, 2001, Simulation Practice and Theory 8 (2001) 491-507

[25] M. Pidd, “Computer Simulation in Management Science”, 1998, ISBN : 978-0-470-09230-9

[26] “Operational Specification for Over-The-Horizon Targeting Gold”, Revision D, OS-OTG (Rev D), 1 September 2000, published under the direction of CNO (N62) by Navy Center For Tactical System Interoperability

[27] T. Buckman, “NATO NETWORK ENABLED CAPABILITY FEASIBILITY STUDY EXECUTIVE SUMMARY”, Version 2.0, 2005, NATO Consultation, Command and Control Agency

Authors' Index