MODEL-BASED PREDICTION OF IT SERVICE AVAILABILITY – A LITERATURE REVIEW

Sascha Bosse[(a)], Matthias Splieth[(b)], Klaus Turowski[(c)]

Very Large Business Applications Lab, Faculty of Computer Science
Otto von Guericke University Magdeburg
P.O. Box 4120, 39016 Magdeburg

[(a)]sascha.bosse@ovgu.de, [(b)]matthias.splieth@ovgu.de, [(c)]klaus.turowski@ovgu.de

ABSTRACT
In today's world, more and more systems are supported by IT services. High availability of these IT services is often crucial to customers and – in order to avoid penalty costs for service-level violation – also for providers. Decisions affecting IT service availability are mainly made in the service design stage and may be costly to correct afterwards. Model-based prediction models can support decision making in this stage. However, the suitability of the existing approaches is still discussed in literature. In this paper, a quantitative evaluation of the state of the art in IT service availability prediction is performed based on a structured literature review. Results indicate that no standard approaches do exist to this moment in time. One reason for this fact is the limited comparability of these models due to the lack of quantitative and comparable evaluations.

Keywords: IT Service Management, Availability Management, Reliability Analysis

1. INTRODUCTION

The continuing service-orientation of IT has led to an increased use of IT services in both the private and the enterprise sector. This trend is not at least catalyzed by the concept of cloud computing, which allows consumers to obtain computing resources as elastic IT services (Mell and Grance 2011). In order to compensate the lack of control on external IT service providers, service level agreements (SLAs) are concluded between providers and consumers. In addition to legal aspects, these agreements specify quantitative guarantees for non-functional quality metrics. Among these metrics, the availability of a service is the most crucial one to IT service consumers (Franke, Johnson and König 2013).

Availability is defined as "the ability of a service [...] to perform its agreed function when required" (Hunnebeck 2011). An IT Service's availability (A) is usually measured as the ratio of the time an IT service was able to perform its function (uptime) and the time it was required (time), cf. equation 1 (Shooman 2002). Related to availability is the attribute reliability which is often quantized by the failure rate of the service.

$$A = uptime \,/\, time \qquad (1)$$

Since violations of the guaranteed availability levels of SLAs may cause penalty costs and loss of reputation (Emeakaroha et al. 2012), IT service providers have to manage IT service availability with respect to the costs of fault-tolerance mechanisms in order to achieve high availability. The challenge hereby is that new technologies and changing business requirements lead to a continuous demand for changes in the underlying data centers (Eckert et al. 2007), making assumptions for future IT service availability even harder (Miroslaw Malek 2008).

In order to cope with this challenge, good-practices in IT service management such as the IT Infrastructure Library (ITIL) recommend the application of prediction methods in the service design phase (Hunnebeck 2011). Since crucial decisions are made in that stage which may be expensive to correct in the subsequent service transition and operation phase, prediction models are intended to serve as a decision support tool for availability management. However, the industrial application of these approaches is rather low (Hunnebeck 2011). One reason for this fact is the lack of reliable information about the suitability and the accuracy of the existing approaches (Chan, Lyu and M. Malek 2006; Gokhale and Trivedi 2006). A systematic and quantitative overview of these prediction approaches may provide this information, but is missing in literature to this date.

Therefore, in this paper, the current state of the art of IT service availability prediction in the design phase is presented, based on a structured literature review, e.g. defined in (Webster and Watson 2002). On this basis, fields for future research can be identified in order to increase the applicability of prediction approaches in industrial contexts.

In the remainder of this paper, first the setup of the literature review and the material collection is described in section 2. After that, identified literature reviews related to this work are presented in section 3, before the results of the descriptive analysis are discussed in section 4. Section 5 provides insights into the defined analysis categories and their values while section 6 presents the results of the material evaluation. Section 7 concludes this contribution by summarizing the paper and discussing future work in the field of IT service availability prediction.

2. LITERATURE REVIEW PROCESS AND MATERIAL COLLECTION

The structured literature review process that was performed for this paper is derived from (Seuring and Müller 2008), which consists of four steps:

1. Material collection,
2. Descriptive analysis,
3. Category selection and
4. Material evaluation.

The material collection aims to identify the relevant contributions in the research field. In the descriptive analysis, the formal aspects of the identified papers are analyzed in order to characterize the result set. In order to evaluate the material in the last step, content-related categories are identified and selected in step three.

In order to be considered as relevant, a paper should provide an approach for predicting IT service availability using an architectural model and should furthermore be applicable in the service design phase. Hence, measured data about the service's availability from the past should not be required by the prediction model.

The first step of the material collection was performed by keyword search in major publication databases of the IT domain:

- *IEEE Xplore Digital Library* (http://ieeexplore.ieee.org/),
- *SpringerLink* (http://link.springer.com/),
- *ACM Digital Library* (http://dl.acm.org/),
- *Wiley Online Library* (http://onlinelibrary.wiley.com/),
- *Google Scholar* (http://scholar.google.com/) and
- *AIS Electronic Library (*http://aisel.aisnet.org/).

Only double-blind reviewed contributions written in the English language were considered. For the keyword term the following phrase was used: *"(Reliability OR Availability OR Dependability) AND (IT OR Service OR Computing OR System) AND (Prediction OR Estimation OR Forecast)".*

Using the keyword search, around 6,500 contributions were found. After the results from the databases were filtered by checking title and abstract for relevance, 121 papers remained. Once these papers were downloaded, content filtering could be carried out which resulted in 79 relevant publications and three related literature reviews.

3. RELATED WORK

Based on the keyword search described in the previous section, three papers are identified that aimed to investigate the state of the art in availability prediction by analyzing the literature.

In (Goseva-Popstojanova and Trivedi 2003), the authors give an overview about quantitative methods to assess the reliability of software systems. They distinguish state-based, path-based and additive approaches and present examples from literature for each category. State-based approaches model all possible states of the system while path-based approaches focus on execution paths. Additive models consider the system reliability as a random variable and try to map it to stochastic processes. Furthermore, the authors define requirements for reliability prediction in the software domain and identify the most important problems in the field to that date: the estimation of the model parameters, the validity of the Markov assumption, the incorporation of usage profiles, the ability to map inter-component dependencies and the prediction of multiple quality attributes.

The authors of (Gokhale and Trivedi 2006) continue this research stream, developing a unification framework for state-based models. Therefore, seven continuous- and seven discrete-time Markov chains were analyzed and harmonized.

In (Immonen and Niemelä 2008), the authors provide a survey of reliability and availability prediction methods from the viewpoint of software architecture. They develop a framework for comparing different methods which includes criteria from the context, users, contents and validation of the models. They compare six exemplary approaches from literature using the previously described framework and conclude that none of the investigated methods satisfies all defined criteria.

Although the three identified literature reviews provide interesting insights and enhance the research field, none of them evaluates the state of the art in availability prediction quantitatively. In addition to that, the last literature review to be found was published in 2008. Due to the ongoing discussion of that field, the insights of these papers should be updated with respect to recent publications.

4. DESCRIPTIVE ANALYSIS

In this section, the formal attributes of the publications are assessed. Therefore, for each contribution, the publication year and type are considered.

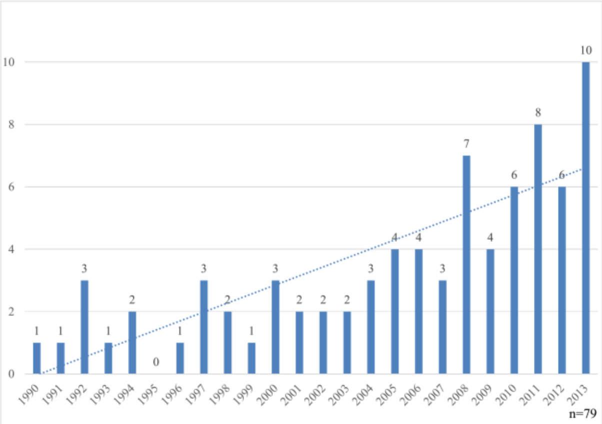

Figure 1: Number of Relevant Publications per Year

Figure 1 presents the number of publications per year, ranging from 1990 to 2013. Although the service-orientation of IT started at the earliest in 1996 with the CORBA standard (OMG 1998), some identified papers had earlier dealt with the development of reliability prediction models for integrated hardware and software systems, e.g. (Laprie et al. 1991). This advancement can be seen as the basis for the following relevant publications, since the developed models can also be applied to IT service availability prediction. Therefore, these papers are included in this literature review.

In the late 1990s, IT systems grew more and more complex, leading to a need of novel prediction models, which can be determined by the increasing number of publications during these years. With the breakthrough of service-oriented computing in the middle of the 2000s, even more approaches have been published in recent years. In figure 1, the linear regression model of the paper number per year (represented by the dotted line) shows that the number of publications increased year by year.

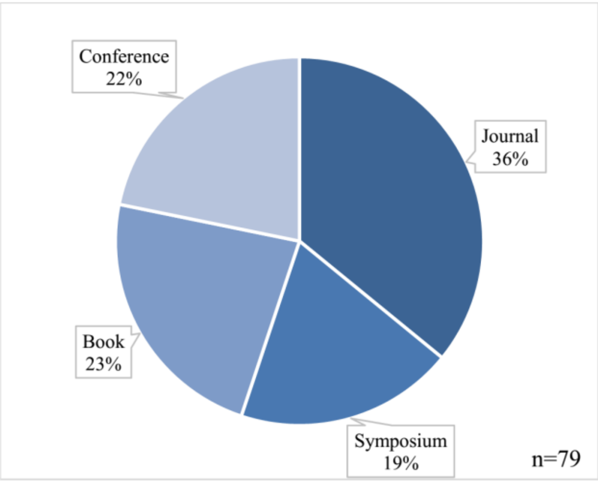

Figure 2: Publication Types of Relevant Contributions

On the one hand, this illustrates the importance of IT service availability prediction. On the other hand, it can be concluded that the evolution of prediction approaches is not finished yet and will receive further attention in the scientific community.

The identified papers were published in four types (cf. figure 2): journal articles (28), book chapters (18) as well as conference (17) and symposium (15) papers. The high number of journal articles indicates that completed research projects in this field do exist, promising validated approaches for IT service availability prediction. However, even more papers were published in symposium or conference proceedings, which shows the lively discussion of the topic. Book chapters often provide a more detailed theoretical foundation of the developed approaches than other publication types due to more available space. The most frequent publication outlets are the *Reliability and Maintainability Symposium* and the *IEEE International Conference on Systems, Man, and Cybernetics* (both four publications), followed by the journals *IEEE Transactions on Reliability*, *IEEE Transactions on Software Engineering*, and the *Journal of Software and Systems Modeling* (each three publications).

Considering both publication year and type, it can be concluded that the field of IT service availability prediction is still heavily discussed in many papers and different contexts. This indicates that current research still lacks of approaches that are widely approved by scientist and practitioners.

5. CATEGORY SELECTION

Since this paper aims to investigate approaches for predicting the availability of IT services, the categories for material evaluation are derived from the identified approaches. These are the followed approach, the used method, the computation technique and the evaluation type and result.

Approaches for predicting IT service availability quantitatively base on measured data (black-box approaches) or on architectural information (white-box approaches), cf. figure 3. Since the investigated approaches should be applicable in the design stage where measured data is not available, only white-box approaches were considered here. These approaches model the components and their interactions that are required for service provisioning. They can be classified further into combinatorial, state-space-based and hierarchical methods (Trivedi et al. 2008). Path-based approaches using information about execution paths are not considered here as a special class because their solution methods are in general either combinatorial or state-space-based.

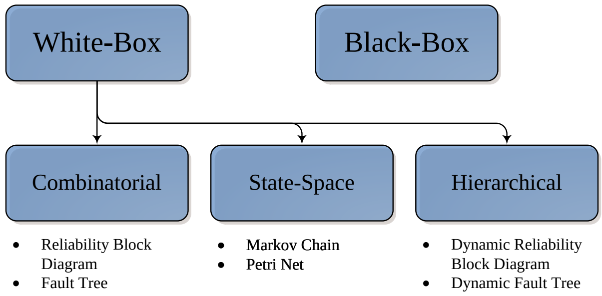

Figure 3: Classification Frame for Availability Prediction Approaches with Example Methods

In combinatorial approaches, an IT service is modeled as a combination of independent components, for instance, in serial, parallel or k-out-of-n systems (Milanovic and Milic 2011). Examples for combinatorial methods are reliability block diagrams (RBD) and fault trees (FT). Due to the assumption of independent components, these models allow simple modeling and fast results' computation. However, complex relations between components such as maintenance and standby configurations cannot be modeled, which is why the accuracy of these approaches for IT service availability prediction is limited (Trivedi et al. 2008).

State-space-based approaches, on the other hand, are able to model inter-dependencies between the components by mapping all possible states of the system and their transition probabilities into a single model (Callou et al. 2012). Markov chains or processes are often applied methods in this class of approaches since the assumption of the Markov property simplifies the evaluation of these models, for example, in discrete- or continuous-time Markov chains (DTMC, CTMC).

However, restricting the state transition time to exponential distributions conflicts with reality where especially recovery times are non-exponentially distributed (Chellappan and Vijayalakshmi 2009). Therefore, other techniques such as semi-Markov processes (Huang, Lin and Kong 2011) or phase-type distributions (Distefano, Longo and Scarpa 2010) are proposed to overcome this assumption.

In comparison to combinatorial approaches, state-space-based approaches not only increase the modeling power drastically, but also the model's complexity. The problem of exponential complexity growth is known as state-space-explosion, leading to challenges in constructing, storing and solving complex models (Trivedi et al. 2008). Some approaches address this problem by encoding the state-space in a Petri net, e.g. in generalized stochastic Petri nets (GSPN) (Kanoun and Ortalo-Borrel 2000) or stochastic reward nets (SRN) (Muppala, Ciado and Trivedi 1994).

Another option to reduce model complexity and increase usability is the use of hierarchical approaches where service and component availability are modeled separately. Often combinatorial models are used for modeling service availability due to their simplicity, while state-space approaches are applied to model components and their dynamic behavior. Examples for methods in this class are dynamic reliability block diagrams (DRBD) or dynamic fault trees (DFT) (Distefano 2009).

A further important aspect of a prediction model is the computation technique, which can be a mathematical or a simulation technique. In the former, closed formula are derived from the model, which are normally solved numerically for complex problems. Since combinatorial approaches are very simple, these models are solved mathematically. However, in state-space approaches, the state-space can grow so complex that numerical methods become error-prone (Sachdeva, D. Kumar and P. Kumar 2008). As an alternative, state-space-based models can be simulated (Xia et al. 2011). On the one hand, this increases the scalability of the approach (Sachdeva, D. Kumar and P. Kumar 2008), on the other hand, only approximate results can be obtained.

Simulation methods also enable the dynamic analysis of availability, providing not only mean values but also information about the availability's variance. This leads to better assumptions about the probability of SLA violations in comparison to steady-state approaches (Franke 2012). Some approaches also provide both computation techniques for maximum flexibility.

Since a developed prediction model is a design artifact, it should be evaluated according to the design science research methodology in order to determine its suitability (Hevner et al. 2004). Therefore, the evaluation type and the prediction accuracy are identified as review categories. In this analysis, the evaluations are classified according to the following scheme:

- Example evaluation would be conducted, if a hypothetic scenario was modeled and availability was computed to show the basic applicability of the developed model. In this evaluation type, results can only be compared to those of other approaches.
- Experimental evaluation would be conducted, if a realistic scenario was modeled and its parameter were varied for experiments which shows the ability of the model to compare different configurations. The results can be evaluated against lab experiments or simulations.
- Case-study evaluation would be conducted, if a real scenario was modeled and analyzed. This proofs the applicability of the model to real-world problems. Evaluation can take place by comparing predicted values to observed ones.
- No evaluation would be conducted, if none of the former evaluation types was carried out.

If a comparison was performed, the accuracy of the approach is computed according to equation 2 where A_m

stands for the measured and A_p for the predicted availability.

$$Acc = |A_m - A_p| / A_m \quad (2)$$

In table 1, the defined categories and their common values found in the results are summarized.

Table 1: Categories and Common Values for Material Evaluation

Category	Common Values
Approach	State-Space, Combinatorial, Hierarchical
Method	RBD, FT, CTMC, DTMC, GSPN, SRN
Computation	Mathematical, Simulation
Evaluation Type	None, Example, Experiment, Case-Study
Accuracy	Quantitative

6. MATERIAL EVALUATION

After the categories have been identified, the quantitative analysis of the 79 selected papers can be carried out. Figure 4 presents the classification of the found approaches.

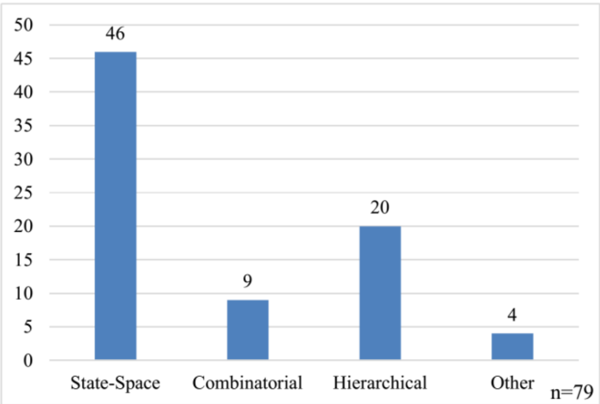

Figure 4: Number of Contributions per Approach

Nearly all approaches could be classified into the three categories state-space, combinatorial and hierarchical. Only four publications dealt with other approaches, e.g. hybrids of state-space- and path-based techniques (Mansour and Dillon 2011). The state-space approaches are most popular because of their modeling power. Nevertheless, some combinatorial approaches were also developed to provide better usability than state-space approaches. Hierarchical approaches are the second most frequent approaches, however, most of the publications presenting hierarchical methods have been published recently. This indicates that hierarchical approaches become more popular since they combine modeling power and usability. Nonetheless, the correlation factor between year and approach is no more than 0.071, which means no trend could be identified over the years. Hence it can be said that no approach is predominant in scientific literature.

By analyzing the applied methods, it could be identified that the combinatorial approaches use reliability block diagrams (RBD) and fault trees (FT) alike. On the other hand, the service availability in hierarchical approaches is mainly modeled with RBDs. This can be explained by the fact that in FTs, failure events are modeled and the component level is not described explicitly. On the contrary, the RBD method model defective components directly, hence these models can be easily connected to the formulated component level state-space models. These models are built in Markov chains and Petri nets in literature without any preference for either of the methods.

In models that are purely state-space-based this is not the case. As figure 5 shows Markov chains and especially continuous-time Markov chains (CTMC) are the dominant method in this approach. Discrete-time Markov chains (DTMC) are used in nearly a quarter, Petri net-based methods in nearly a fifth of the publications. Other used methods are mainly based on Bayesian networks (Roshandel, Medvidovic and Golubchik 2007) or other probabilistic methods.

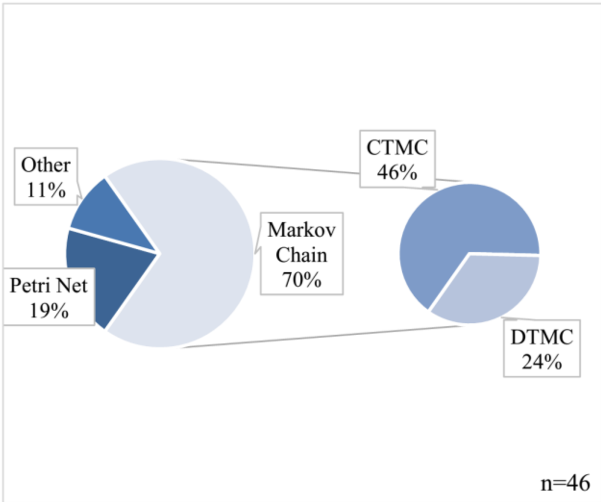

Figure 5: State-Space Approaches classified by the Used Method

The popularity of Markov chain methods is principally founded in the formal specification and the standardized solving algorithms. However, the definition of the state-space is not intuitive and hence error-prone. Petri net models are more intuitive due to the implicit mapping of the state-space (Kanoun and Ortalo-Borrel 2000) and allow the modeling of concurrency. On the other hand, some high-level Petri nets can only be solved via simulation.

This also explains the relatively high number of simulation computation techniques for Petri net approaches as presented in table 2. While a third of the Petri net models can be solved by simulation, for Markov chains and all approaches in total the mathematical computation technique is more dominating.

Table 2: Computation Technique Frequency

	Markov Chains	Petri Nets	Total
Mathematical	84.0 %	66.7 %	81.0 %
Simulation	16.0 %	22.2 %	12.7 %
Dual	0.0 %	11.1 %	6.3 %

Analyzing the evaluations carried out in the identified publications it stands out that only 8.9 % of the papers spare the evaluation, cf. figure 6. In the other contributions, no evaluation type is dominating. Simple numerical examples are used in 27.8 %, experimental setups in 25.3 % of the cases. In nearly 38 % of the papers, a real-world case-study was performed for evaluation. However, just a very small number of publications (7.6 %) provide a quantitative comparison with measured values.

In (Boudali, Sözer and Stoelinga 2009), the authors use fault-injection in their case-study to produce comparable availability results. Their developed CTMC model achieves an accuracy of 0.8 %. (Janevski and Goseva-Popstojanova 2013) evaluated their hybrid prediction model combining state-space- and path-based methods against real data with an accuracy of 1.91 %. Other comparisons were carried out in experimental environments using, for example, testbeds (Reussner, Schmidt and Poernomo 2003). The accuracy achieved in these evaluations varies from 0.36 % to 2.1 %.

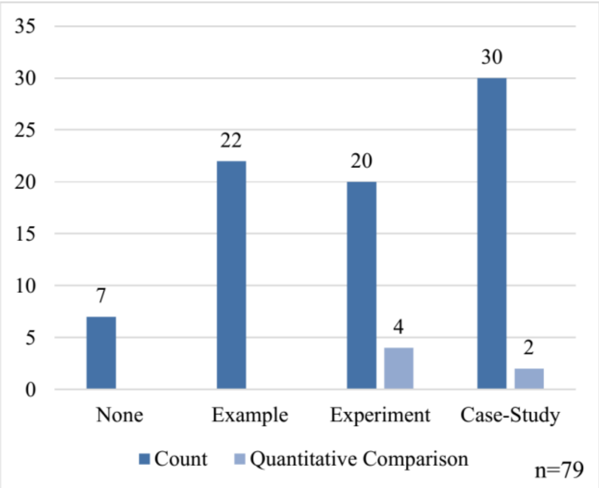

Figure 6: Number of Publications by Evaluation Type

7. CONCLUSIONS

In this paper, a structured literature review was carried out in order to analyze the state of the art in model-based IT service availability prediction quantitatively. Therefore, 79 relevant publications were identified by keyword search. After the descriptive analysis pointed out that the scientific discussion of this research field is still ongoing, the categories for the content analysis of the identified papers were defined. On that basis the material evaluation could be performed. It revealed that approaches based on the state-space of the availability problem are very popular in literature, although hierarchical approaches combining combinatorial and usable service level with state-space-based component level availability models also come into focus.

The detailed analysis of the state-space approaches revealed that continuous- and discrete-time Markov chains are mostly used for modeling the state-space. Petri net-based approaches are developed in nearly a fifth of the identified publications, although the encoding of the state-space decreases the model complexity. Together with simulation methods for solving these models, Petri net approaches provide better scalability than the mathematical evaluation of Markov chains (Sachdeva, D. Kumar and P. Kumar 2008).

While modeling power, applicability and usability of the selected approaches are discussed in the publications, the scalability and even the accuracy of the approaches are evaluated hardly. This is because the evaluation of these models is often not conducted quantitatively in comparison to real-world values. From the 79 identified publications, only six ones provided a quantitative evaluation and compared its result with the real world. Furthermore, only two of these six papers used a real-world case to validate their prediction model. Nevertheless, the evaluation results in these cases could only be produced by manipulating the experiment, e.g. with fault injection, reducing the confidence in the evaluation results.

This lack of comparable evaluation leads to the situation that the accuracy, scalability and suitability of most approaches cannot be assessed exactly. One reason for this might be the lack of accessible data for IT service availability for a constant system configuration over a long period of time. Therefore, it is hard to decide which approach is the most feasible one for a specific scenario. Further research could provide a benchmark where configurations were observed during a long period of time so that the measured availability data can be used for the evaluation of prediction models.

Besides the discussion about approaches there are two major barriers that make the application of design phase prediction models difficult. First, the model creation process requires deep knowledge of the data center. If human resources are involved in this process, it may be costly and error-prone. Some papers have started to address this problem with automatic model generation procedures using monitored infrastructure data, for example in (Milanovic and Milic 2011).

The second problem is the parameterization of the prediction models in order to quantify component availability. In case of hardware components, manufacturers' data is often over-optimistic (Pinheiro, Weber and Barroso 2007) and field-test data must be generated. In the case of software components, this problem is even more crucial due to the lack of reliable failure data, especially for novel software components that are developed in the design phase or require frequent software updates.

Consequently, the low industrial penetration of availability prediction models may be caused by these two problems in addition to the fact that the existing approaches are barely comparable due to lacking evaluation. In order to ease SLA management by providing assumptions about availability in the design phase, these drawbacks need to be addressed in future research.

REFERENCES

Boudali, H., Sözer, H. and Stoelinga, M., 2009. Architectural Availability Analysis of Software Decomposition for Local Recovery. *Proceedings of the 3rd IEEE International Conference on Secure Software Integration and Reliability Improvement*,

Callou, G., Maciel, P., Tutsch, D., Araújo, J., Ferreira, J. and Souza, R., 2012. A Petri Net-Based Approach to the Quantification of Data Center Dependability. *Petri Nets - Manufacturing and Computer Science*, 313–336, InTech.

Chan, P.P.W., Lyu, M.R. and Malek, M., 2006. Making Services Fault Tolerant. *Proceedings of the 3rd International Service Availability Symposium*, 43–61, Lecture Notes in Computer Science, Berlin, Heidelberg: Springer.

Chellappan, C. and Vijayalakshmi, G., 2009. Dependability modeling and analysis of hybrid redundancy systems. *International Journal of Quality & Reliability Management*, 26, 76–96.

Distefano, S., 2009. Reliability and Dependability Modeling and Analysisof Dynamic Aspects in Complex Systems. *8th IEEE International Conference on Dependable, Autonomic and Secure Computing*, 43–48, Chengdu, China: IEEE Computer Society.

Distefano, S., Longo, F. and Scarpa, M., 2010. Availability Assessment of HA Standby Redundant Clusters. *29th IEEE International Symposium on Reliable Distributed Systems*, 265–274, New Delhi, India: IEEE Computer Society.

Eckert, J., Repp, N., Schulte, S., Berbner, R. and Steinmetz, R., 2007. An Approach for Capacity Planning of Web Service Workflows. *Proceedings of the 13th Americas Conference on Information Systems (AMCIS)*,

Emeakaroha, V.C., Netto, M.A.S., Calheiros, R.N., Brandic, I., Buyya, R. and Rose, C.A.F.D., 2012. Towards autonomic detection of SLA violations in Cloud infrastructures. *Future Generation Computer Systems*, 28, 1017–1029.

Franke, U., 2012. Optimal IT Service Availability: Shorter Outages, or Fewer? *IEEE Transactions on Network and Service Management*, 9, 22–33.

Franke, U., Johnson, P. and König, J., 2013. An architecture framework for enterprise IT service availability analysis. *Software and Systems Modeling*. Available at: http://dx.doi.org/10.1007/s10270-012-0307-3.

Gokhale, S.S. and Trivedi, K.S., 2006. Analytical Models for Architecture-Based Software Reliability Prediction: A Unification Framework. *IEEE Transactions on Reliability*, 55, 578–590.

Goseva-Popstojanova, K. and Trivedi, K.S., 2003. Architecture-Based ApApproach to Software Reliability Prediction. *Computers and Mathematics with Applications*, 46, 1023–1036.

Hevner, A.R., March, S.T., Park, J. and Ram, S., 2004. Design science in information systems research. *MIS Quarterly*, 28, 75–105.

Huang, J., Lin, C. and Kong, X., 2011. Modeling and Analysis of Dependability Attributes of Service Computing Systems. *2011 IEEE International Conference on Services Computing*, 184–191, Washington D.C., United States: IEEE Computer Society.

Hunnebeck, L., 2011. *ITIL Service Design 2011 Edition*. Norwich, UK: The Stationery Office.

Immonen, A. and Niemelä, E., 2008. Survey of reliability and availability prediction methods from the viewpoint of software architecture. *Software and Systems Modeling*, 7, 49–65.

Janevski, N. and Goseva-Popstojanova, K., 2013. Session Reliability of Web Systems under Heavy-Tailed Workloads: An Approach Based on Design and Analysis of Experiments. *IEEE Transactions of Software Engineering*, 39, 1157–1178.

Kanoun, K. and Ortalo-Borrel, M., 2000. Fault-Tolerant System Dependability—Explicit Modeling of Hardware and Software Component-Interactions. *IEEE Transactions on Reliability*, 49, 363–376.

Laprie, J.-C., Kanoun, K., Béounes, C. and Kaâniche, M., 1991. The KAT (Knowledge-Action-Transformation) Approach to the Modeling and Evaluation of Reliability and Availability Growth. *IEEE Transactions of Software Engineering*, 17, 370–382.

Malek, Miroslaw, 2008. Predictive Algorithms and Technologies for Availability Enhancement. *5th International Service Availability Symposium (ISAS)*, 17–19, Lecture Notes in Computer Science, Tokyo, Japan: Springer Verlag Berlin Heidelberg.

Mansour, H.E. and Dillon, T., 2011. Dependability and Rollback Recovery for Composite Web Services. *IEEE Transactions on Service Computing*, 4, 328–339.

Mell, P. and Grance, T., 2011. *The NIST Definition of Cloud Clomputing*. National Istitute of Standards and Technology.

Milanovic, N. and Milic, B., 2011. Automatic Generation of Service Availability Models. *IEEE Transactions on Service Computing*, 4, 56–69.

Muppala, J.K., Ciado, G. and Trivedi, K.S., 1994. Stochastic Reward Nets for Reliability Prediction. *Communications in Reliability, Maintainability and Serviceability*, 1, 9–20.

OMG, 1998. *CORBAservices: Common Object Services Specification*. Object Management Group (OMG).

Pinheiro, E., Weber, W.-D. and Barroso, L.A., 2007. Failure Trends in a Large Disk Drive Population. *Proceedings of the 5th USENIX Conference on File and Storage Technologies (FAST)*,

Reussner, R.H., Schmidt, H.W. and Poernomo, I.H., 2003. Reliability prediction for component-based software architectures. *The Journal of Systems and Software*, 66, 241–252.

Roshandel, R., Medvidovic, N. and Golubchik, L., 2007. A Bayesian Model for Predicting Reliability of Software Systems at the Architectural Level. *Proceedings of the 3rd International Conference on the Quality of Software-Architectures*, 108–126, Lecture Notes in Computer Science, Springer Heidelberg.

Sachdeva, A., Kumar, D. and Kumar, P., 2008. Reliability analysis of pulping system using Petri nets. *International Journal of Quality & Reliability Management*, 25, 860–877.

Seuring, S. and Müller, M., 2008. From a literature review to a conceptual framework for sustainable supply chain management. *Journal of Cleaner Production*, 16, 1699–1710.

Shooman, M.L., 2002. *Reliability of Computer Systems and Networks – Fault Tolerance, Analysis, and Design*. New York, NY, USA: John Wiley & Sons New York.

Trivedi, K.S., Ciardo, G., Dasarathy, B., Grottke, M., Matias, R., Rindos, A. and Vashaw, B., 2008. Achieving and Assuring High Availability. *5th International Service Availability Symposium (ISAS)*, 20–25, Lecture Notes in Computer Science, Tokyo, Japan: Springer Verlag Berlin Heidelberg.

Webster, J. and Watson, R.T., 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quaterly*, 26, xiii–xxiii.

Xia, Y., Dai, G., Li, J., Sun, T. and Zhu, Q., 2011. A model-driven approach to predicting dependability of WS-CDL based service composition. *Concurrency and Computation: Practise and Experience*, 23, 1127–1145.