IDEA GROUP PUBLISHING
701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com
ITB12732

This paper appears in the book, *Emerging Trends and Challenges in Information Technology Management, Volume 1 and Volume 2* edited by Mehdi Khosrow-Pour © 2006, Idea Group Inc.

The Next Wave in IT Infrastructure Risk Management: A Causal Modeling Approach with Bayesian Belief Networks

Daniel J. Hinz, J.W. Goethe University, E-Finance Lab, Mertonstrasse 17, 60325 Frankfurt/Main, Germany, Phone: +49 69 7162-5354, Fax -5355, dhinz@wiwi.uni-frankfurt.de

Heiko Gewald, J.W. Goethe University, E-Finance Lab, Mertonstrasse 17, 60325 Frankfurt/Main, Germany, Phone: +49 69 4272-6016, gewald@wiwi.uni-frankfurt.de

ABSTRACT

The management of risks associated with information technology (IT) infrastructure becomes increasingly important, as companies may face severe negative outcomes in case of failures. This paper proposes a new approach to manage IT infrastructure risks even in highly dynamic environments. Currently, IT infrastructure and its risks are managed based on historical loss data, which allows very precise forecasts for potential risks in stable environments. However, this is not adequate for the increasing number of firms facing dynamic environments like outsourcing or merger scenarios. Therefore, the next wave in IT infrastructure risk management has to employ more qualitative methodologies. Based on an ongoing case study with two leading IT consultancies and a European service enterprise, this paper demonstrates, how causal modeling with Bayesian Belief Networks enables the prediction and, most important, the proactive management of IT infrastructure risks.

INTRODUCTION

During the last decade, IT infrastructure has become a top priority on the agenda of information systems (IS) top executives as well as IS researchers. In 1990, this topic appeared for the first time in a survey of the Society for Information Management (SIM) and the MIS Research Center (MISRC) to identify the most critical issues in IS management [22] and has not disappeared until today [8, 19].

This high awareness results from the two facts, that first, IT infrastructure is the foundation for any IT enabled business activity [8, 26], and second, losses in case of failures are potentially severe (e.g., power outages, natural disasters, or terrorism [9, 12]).

Current management practice for IT infrastructure, which emerged mainly from the American IT security research history, is typically based on historical loss data, which allow very precise forecasts for potential risks in stable environments. For the increasing number of firms facing dynamic environments, e.g. outsourcing or merger situations, historical data is not an adequate estimator for future events [2, 6]. Currently qualitative, questionnaire based approaches are proposed for these scenarios, but more integrated methodologies have to be employed in order to climb the next s-curve of proactive risk management. Based on an ongoing case study with two leading IT consultancies and a European service enterprise, this paper demonstrates, how causal modeling based on Bayesian statistics enables the prediction and, most important, the proactive management of IT infrastructure risks.

STATUS QUO OF IT INFRASTRUCTURE RISK

Classical IT Infrastructure Risk Management

Losses from operational risks like IT risks arise from two categories of events. High frequency, low impact events are characterized by a huge amount of loss events, each with a relatively low loss. On the contrary, low frequency, high impact events appear very seldom, but may cause severe damage.

A single measurement approach that addresses both event types equally is difficult to implement. Therefore, operational risk management as well as IT risk management typically employs both quantitative and qualitative measurement approaches [25].

Quantitative measurement is appropriate for high frequency, low impact events, where enough data is available. In an enterprise, a main data source for IT incidents is the help desk trouble ticket system. Most errors are reported to these systems, as users fortunately fulfill this reporting obligation automatically by requesting support from the help desk. This procedure enables very detailed ex-post analyses of the collected data, at least if it has been collected thoroughly which may not be the case [23]. An important part is the calculation of service levels, which allow the measurement and control of service quality [23].

Qualitative measurement is applied, where the probability or occurrence of the event is low and therefore historical data is hardly available. One prominent example is outsourcing, where literature has identified an impressive number of outsourcing risks, the most frequent stated ones are cost overruns, quality of service degradation, and loss of innovativeness [15]. All these risks have been identified by expert interviews and by the analysis of existing outsourcing deals, but concrete quantification remains a challenge. Theory offers approaches like Extreme Value Theory (EVT), which can be used to analyze the tail behavior of a distribution [14, 20], however, it does not generate real incident scenarios.

Limitations

As both methods have their respective merits and disadvantages, none of them should be employed solely. Both methods should be used in conjunction [7]. It is even more important that the classical approaches do not allow for the simulation of major future events. As described above, they are well suited for day-to-day management based on both the quantitative and qualitative approaches. For future change scenarios like an upcoming outsourcing deal, assessments can only be qualitative, which makes the calculation of a business case incorporating the risk component extremely difficult.

In order to overcome this deficiency, a "perfect solution" [21] to combine quantitative and qualitative analysis is the employment of causal modeling methods based in Bayesian statistics, as this method provides convincing advantages for the measurement of operational risk.

BAYES – THE ENABLER FOR THE NEXT WAVE

In this section we investigate the principles and merits of Bayesian statistics and introduce Bayesian Belief Networks as a way to formally incorporate expert's know-how (qualitative method) into classical risk measurement activities (quantitative method).

Bayesian Statistics

Followers of "traditional" statistical thinking insist that the true but unknown parameters of a distribution can be estimated based on the data available [1]. "Bayesians" accept some degree of subjectivity in estimating the parameters of a distribution, as from their perspective there is no essential distinction between observables and parameters of a statistical model, since they are all deemed to be random [10]. Therefore Bayesian statistics allow for a combination of the empirically gathered data and experts belief in the model determining parameters [2].

The acceptance of subjectivity in modeling the distribution gives Bayesian statistics a substantial advantage in operational risk measurement. Major data problems can be addressed without losing statistical rigor:

- **Data quality**. Filtering rules have direct implications on the parameters of the distribution, they act like subjectivity.
- **Lack of data**. Whenever data is not available in the desired quantities, structured workshops with experts on the questionable topic can be used to gain insights into (subjective) parameter expectations (prior information).
- **Historicity of data**. As data loses its value for the calculation of the loss distribution once fundamental changes have been applied to the organization, experts can evaluate those impacts on the parameters thus anticipating the impact of the change.

Bayesian interference describes the process of fitting a probability model to a set of data and summarizing the result by a probability distribution on the parameters of the model and on unobserved quantities such as predictions for new observations [13].

Bayesian statistics regard the process of statistical estimation as one of continuously revising the subjective beliefs about the state of an issue as more data become available [1]. The very basis of Bayesian methods is the Bayes' rule:

$$p(\text{Parameter} \mid \text{Data}) \propto p(\text{Data} \mid \text{Parameter}) \cdot p(\text{Parameter})$$

The posterior distribution of a parameter is proportional to the likelihood times the prior distribution of the parameter. The prior distribution of the parameter represents the subjective prior beliefs, whereas the likelihood is the probability of the observations given the unobserved parameters. The resulting posterior distribution is the conditional probability of the parameter given the observations.

Bayesian Belief Networks

Bayesian Belief Networks (BBN) have been studied for management purposes for some time now and have been successfully applied within several disciplines, amongst them decision sciences and the artificial intelligence community (e.g. [17, 24]). Their use for operational risk management has been advocated [1, 21, 28].

BBNs are appropriate for decision making in systems where uncertainty is present and it is represented by n-dimensional discrete random variables $X_1, \ldots, X_n$, where the discrete values of the random variables are dependent. The dependencies between the random variables are represented by conditional probabilities or conditional densities, respectively (chain rule):

$$p(x_1,\ldots,x_n \mid C) = p(x_n \mid x_1, x_2,\ldots,x_{n-1},C) \cdot p(x_{n-1} \mid x_1,x_2,\ldots,x_{n-2},C) \cdot \ldots \cdot p(x_2 \mid x_1,C) \cdot p(x_1 \mid C)$$

Those conditional probabilities form a directed acyclic graph – the basic structure of a BBN. The nodes of the graph represent the random variables and the edges represent the relationships (conditional probabilities) between those random variables. The head of an edge points to the random variable whose conditional probability is described. The source of an edge represents a random variable that conditions the probability of the target node.

When observations exist, they represent information about the random variables and determine the likelihood. Thus in a Bayesian Belief Network a priori information as well as observations and deterministic know-how can be considered.

THE CAUSAL MODEL FOR DESKTOP INFRASTRUCTURE RISKS

After having introduced the concept of Bayesian statistics and Bayesian Belief Networks, this chapter presents an actual BBN for the risks of desktop infrastructure (desktop service providing).

Study Design

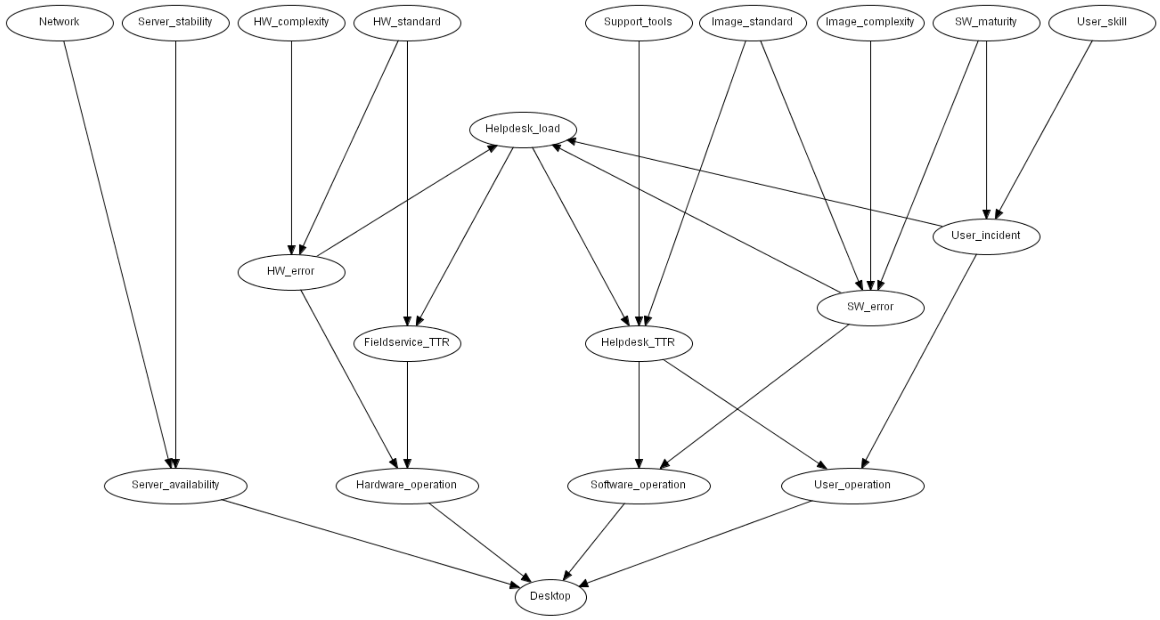
The following causal model was developed together with practice experts on management and partner level of two leading IT consultancies. As it is end-user centric, the model covers many IT infrastructure risks, however it is not all-embracing. Further research might expand the focus of this model to capture more breadth and detail. First, a preliminary model that incorporated aspects of end-user computing risks was designed based on literature research. Even for expert assessments, it is advisable not to start totally from scratch. Although biasing the experts with the initial model, this ensures the provision for existing theory. The model was then refined in further interviews and the interview results communicated back and forth using an iterative approach, until an integrated, harmonized model evolved. This ensured a common understanding and acceptance of the model. The procedure followed the suggestions for building theories from case study research of Eisenhardt [11] and Yin [27]. In this ongoing study, the next step of validation will be to match this network with real enterprise data. By using different subsets of data, e.g. from different business units or time periods, it is possible to both validate and calibrate the model in the enterprise context to allow for better simulation. Already in the current state, the model produces valid results according to the experts' judgments, as it represents the state in which only prior information is available. Adding likelihood information from historical incident data will presumably increase the prognostic power.

Risk Elements and Dependencies

The objective of this network is to calculate the percentage of desktop uptime. In this case, we define uptime as the time when the desktop can be used to perform its designated tasks at that point of time. Figure 1 shows the resulting Bayesian Belief Network, the following explanations are structured according to the lower five nodes.

Desktop. This node represents the measuring objective, i.e. desktop uptime. Therefore, it has only two possible states: "functional" and "defective". The percentage values for each state may be understood either as the share of desktop systems which is up or down on average, or the share of an average desktop system being up or down. Whether a desktop system is functional or not, depends on the following four direct influencing factors in this model.

Figure 1. Bayesian belief network for IT infrastructure risks (desktop focus)



Servers ("Server_availability"). This node describes the population of servers in an enterprise. It is assumed that servers themselves may fail with a certain probability (0.1%), and that network failures render servers defective with a higher probability of 0.5%. In this model, the network node has no direct influence on the desktop itself, which means that all communication in this company happens via servers.

Hardware ("Hardware_operation"). Each hardware error of the desktop computer itself is captured in this node, the state of which can be either "functional" or "defective". The current model assumes that when the state of this node is "defective" the state of its child "Desktop" is also "defective", i.e. the conditional probability is 1. This absolute dependency is set at several nodes of this example in order to reduce complexity. This node is depended on two parent nodes. The first one, "HW_error" (hardware error), represents the desktop hardware failing with a certain probability dependent on the system complexity ("HW_complexity", states "desktop" and "laptop") and the adherence to established hardware standards in the firm [16] ("HW_standard"). Complementary to the failure rate there is the restoration rate, i.e. the average time that field service members need to restore the functionality of the system ("Fieldservice_TTR). This average time is mainly affected by the current level of help desk utilization ("Helpdesk_load").

Software ("Software_operation"). In the same way, errors in the installed software may occur [4], which render a system defective, e.g. cause the system to hang or simply does not perform the operations it is intended to do. Similarly, to the node "Hardware_operation", this node is influence by the error rate of problems ("SW_error") and the corresponding restoration rate of the help desk ("Helpdesk_TTR").

In this model, the three main factors that influence software errors are the adherence to software standards [16], the complexity of a software image (or software profile), and the maturity of the installed software [5] (nodes "Image_standard", "Image_complexity", and "SW_ma-turity").

User ("User_operation"). An important part of a desktop system is the user. If the user does not know how to operate the computer, how to use the installed software, or how to fix small issues by oneself, the desktop system is also non-functional ("defective") even when everything else is working. Similar to the nodes "Hardware_operation" and "Software_operation" the state of this node is mainly dependent on the failure and restoration rate of the user. In other words, what is the likelihood of user incidents, depending mainly on his or her IT skills (node "User_skill"), and how quickly the help desk team can solve these incidents (node "Helpdesk_TTR").

4.3. Indicative feedback and application

With this input, the model is initialized with a priori information based on expert judgment and therefore it returns the prior distribution. Each node could be amended with observations and statistical data to calibrate the network further. The node "Desktop" now calculates the percentage of desktops, which are functional. The model computes a value of 97.57%, meaning a PC is not functional for on average approximately 5.3 working days per year. Consistently, one would expect a value somewhere between 95% and 99% under typical service level agreements. The current state of the model can indicatively be validated by comparing this value to statistics of the help desk and user support ticketing systems.

Practical demand for this approach is high, as it especially helps to better understand, simulate, and communicate the effects of quality measures like improvement of helpdesk quality, enforced standardization, or increased software maturity. Indicative feedback from financial institutions also showed, that especially in the context of Basel II, causal modeling is seen as a promising approach [3], although a working model has not yet been developed.

One simulation scenario may demonstrate the prognostic power of this model.

Scenario: Reducing user support service levels? In times of increased cost cutting initiatives, an overstaffed shared service like user support may be regarded as a promising saving potential. But actions in this case have to be prudent. Currently approximately 89% of all incidents can be solved immediately. If for example this number drops to nearly 85% due to lay-offs and therefore increased help desk load, desktop uptime drops by 1.17 percentage points to 96.40%, meaning that a single desktop is now on average 2.6 days per year longer defective.

CONCLUSION AND FURTHER RESEARCH

Within this paper a new approach to manage the risks of IT infrastructure has been developed. It allows for the proactive management of these highly important risks, thus enabling companies to better prepare for change scenarios like outsourcing or mergers, while at the same time improving classical decision-making. By employing Bayesian statistics, a causal model for desktop risks has been developed in an ongoing study with two major IT consultancies, which allows the prediction of desktop downtimes and demonstrates the predictive power of the chosen methodology.

However, there are some limitations. Although Bayesian statistics perfectly combine qualitative as well as quantitative risk management approaches, it does not completely mitigate the drawbacks of each technique. If severely wrong expert judgments cannot be successfully falsified through existing historical data, the predictive power may be reduced. In the same way, incorrect data may skew model results as well. In addition, the development and the maintenance of a sophisticated causal model require some expertise and resources, which might be scarce.

Further research should focus on two main areas. First, the introduced model for desktop risks should be extended to include other elements of IT infrastructure. Although the current model already includes many aspects due to its user-centric characteristic, especially the network and server nodes have to be refined. Second, more detail could be added to the current model. In the expert interviews, focus was put only on the most important influence factors. Other factors, which may be worth investigating, are the effects of incentives and training on help desk and user support efficiency, hardware quality, and user-friendliness of software. Another interesting issue would be to include the role of problem solving in unofficial networks, as probably many incidents never reach the help desk but are solved elsewhere, e.g., among coworkers [18].

REFERENCES

Alexander, C., Bayesian Methods for Measuring Operational Risks, *IBCI Technical Risk Management Reports, ISMA Centre, Reading University*, 2000.

Alexander, C., Rules and Models, *Risk*, (January), 2002, 18 - 20.

Alexander, C., Managing operational risks with Bayesian networks in: Alexander, Carol (Ed.); *Operational Risk - Regulation, Analysis and Management*, Pearson Education, Edinburgh, 2003, 285-295.

Banker, R.D., Datar, S.M., Kemerer, C.F., and Zweig, D., Software errors and Software Maintenance Management, *Information Technology and Management*, (3), 2002, 25-41.

Barlow, R.E. and Singpurwalla, N.D., Assessing the Reliability of Computer Software and Computer Networks: An Opportunity for Partnership with Computer Science, *The American Statistician*, (39:2), 1985, 88-94.

Benamati, J.H. and Lederer, A.L., Rapid Change: Nine Information Technology Management Challenges, *Information Systems and Operational Research*, (38:4), 2000, 336-358.

BIS, International Convergence of Capital Measures and Capital Standards, *Bank for International Settlements, Basel Committee on Banking Supervision*, 2004.

Brancheau, J.C., Janz, B.D., and Wetherbe, J.C., Key Issues in Information Systems Management: 1994-95 SIM Delphi Results, *Management Information Systems Quarterly*, (20:2), 1996, 225-236.

Cerullo, V. and Cerullo, M.J., Business Continuity Planning: A Comprehensive Approach, *Information Systems Management*, (21:3), 2004, 70-78.

Cruz, M.G., *Modelling, measuring and hedging operational risk*, John Wiley & Sons, Chichester, 2002.

Eisenhardt, K.M., Building Theories from Case Study Research, *Academy of Management Review*, (14:4), 1989, 532 - 550.

Federal Reserve Bank of New York, Summary of "Lessons Learned" and Implications for Business Continuity, *Discussion note, available online at www.newyorkfed.org*, 2002.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B., *Bayesian Data Analysis*, Chapman & Hall, Boca Raton, 1997.

Gencay, R., Selcuk, F., and Ulugülyagci, A., High volatility, thick tails and extreme value theory in value at risk estimation, *Insurance: Mathematics and Economics*, (33), 2003, 337 - 356.

Gewald, H. and Hinz, D., A Framework for Classifying the Operational Risks of Outsourcing, *8th Pacific Asia Conference on Information Systems*, Shanghai, PR China, 2004.

Hanseth, O. and Braa, K., Hunting for the Treasure at the End of the Rainbow: Standardizing Corporate Infrastructure, *Computer Supported Cooperative Work*, (10), 2001, 261-292.

Heckerman, D., Marndani, A., and Wellman, M.P., Real World Applications of Bayesian Networks, *Communications of the ACM*, (38:5), 1995, 24-26.

Heckman, R. and Guskey, A., Sources of Customer Satisfaction and Dissatisfaction with Information Technology Help Desks, *Journal of Market Focused Management*, (3), 1998, 59-89.

Luftmann, J. and McLean, E.R., Key Issues for IT Executives, *Management Information Systems Quarterly Executive*, (3:2), 2004, 89-104.

Medova, E., Extreme Value Theory - Extreme values and the measurement of operational risk, *Operational Risk*, (July), 2000, 11 - 15.

Neil, M. and Tranham, E., Using Bayesian Networks to Predict Op Risk, *Operational Risk*, (August), 2002, 8-9.

Niederman, F., Brancheau, J.C., and Wetherbe, J.C., Information Systems Management Issues for the 1990s, *Management Information Systems Quarterly*, (15:4), 1991, 475-502.

Niessink, F. and Van Vliet, H., Software maintenance from a services perspective, *Journal of Software Maintenance: Research and Practice*, (12), 2000, 103-120.

Pearl, J., *Probablistic Reasoning in Intelligent Systems: Networks of Plausible Interference*, Morgan Kaufmann Publishers, San Mateo (California), 1988.

Rainer, R.K., Jr., Snyder, C.A., and Carr, H.H., Risk Analysis for Information Technology, *Journal of Management Information Systems*, (8:1), 1991, 129-147.

Weill, P. and Vitale, M., What IT Infrastructure Capabilities are Needed to Implement E-Business Models?, *Management Information Systems Quarterly*, (1:1), 2002, 17-34.

Yin, R.K., *Case Study Research: Design and Methods*, Sage Publications, Inc., London, New Delhi, 2003.

Yuji, Y., Application of Bayesian Inference to Operational Risk Management, *University of Tsukuba Working Paper*, 2003.

Related Content