Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A.A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

FORECASTING CYBER MAINTENANCE COSTS WITH IMPROVED SCAN ANALYTICS USING SIMULATION

Theodore T. Allen and Enhao Liu

Integrated Systems Engineering
The Ohio State University
1971 Neil Avenue – 210 Baker Systems
Columbus, Ohio 43210-1271 USA

ABSTRACT

This article proposes a discrete event simulation model of an organization that maintains computer hosts and incurs several millions of dollars in maintenance and incident response costs. The common maintenance policy is referred to as “out-of-sight is out-of-mind” (OSOM) because the majority of hosts are absent from scans and ignored. Hosts are “dark” (absent) because they are not accessible (turned off or with restricted permissions). The proposed model is used to compare OSOM with alternatives including improved analytics that make dark host vulnerabilities visible. Findings clarify the apparent benefits of OSOM unless indirect costs for intrusions or improved policies are applied. Also, benefits from using Windows operating systems and improved policies are clarified including millions in expected savings (vs. Linux).

1 INTRODUCTION

Cyber-security-related costs are important on multiple levels from national and international politics to electric grids connecting thousands of organizations to expenditures within individual organizations. Discrete event models have explored political effects (Naugle et al. 2016). Models at the power grid level include those described by Nguyen et al. (2015). Also, attack simulation models include Shinet al. (2015) and Case (2016).

In our own research, we have explored Markov decision process models of organizational expenses focusing on the evolutions of single hosts (Afful-Dadzie and Allen 2014; 2016). Computer hosts may be ordinary personal computers, laptops, servers, printers, or even exercise equipment. Here, we focus only on devices connected to the Internet that could be compromised and are scanned and maintained. These devices are used for student, research, and administrative tasks. These devices have so-called “vulnerabilities” which are weaknesses that attackers can exploit. For example, a host might use a weak password, software with an out-of-date encryption, or software without sufficient checks on the size of inputs or outputs. These vulnerabilities are rated by the U.S. National Institute of Standards (NIST) and the common vulnerability scoring system.

Here, we propose to extend the data and assumptions for maintenance policy development to discrete event simulations. This is similar to patch management in electric utilities addressed by Gauci et al. (2017) except that we consider a larger number of past incidents and a broader assortment of policies and host types. Benefits of discrete event simulation include relatively intuitive ways to include the inception and destruction of hosts and finite patching and incident response resources. We argue that host “end of life” issues are important to consider because, anecdotally, we are aware of hosts that were believed to be retired being used and causing incidents.

In our experience, a common policy is to require that staff attempts to patch or mitigate high or critical level vulnerabilities within one month of the time when the vulnerability is observed in the monthly scans. The policy ignores the medium- or low-level vulnerabilities which tend to accumulate. Also, typically 70%

978-1-5386-6572-5/18/$31.00 ©2018 IEEE

of the almost 50,000 distinct hosts that we studied were missing from the scans in any given month. This can occur because the host is turned off during the scan or permissions are lacking. Some methods to impute the vulnerabilities missing in the scan data are described by Afful-Dadzie and Allen (2014; 2016). Recently, we have methods that can predict with high accuracy (0.05% errors) the vulnerabilities on hosts which are not present ("dark") in the monthly scans.

Here, we consider the implications of 21 months of observed transitions from month to month of approximately 50,000 hosts. The resulting transition probability estimates are shown in Table 1. The probabilities reflect the combined effects of at least four factors. First, users of the hosts are constantly adding software and the software they already added is aging. Second, hackers are constantly searching for vulnerabilities, observing the acknowledgement of vulnerabilities that are publically reported, and obtaining exploits (which are also often freely published). Third, vendors are constantly attempting to automatically patch their software remotely. Fourth, staff is attempting to patch vulnerabilities according to organization policy with lists of vulnerabilities obtained from scans and the results of their own searches for available patches, testing patches obtained for not destroying functionality, and applying patches found and tested (if any).

Here also, we consider only two types of hosts. These are Linux and Windows hosts for which the user has administrator privilege to install new software and the host is not controlled by administrators. (Controlled hosts are generally much safer.) Here, we refer to the common maintenance policy in which dark hosts are ignored as "out-of-site is out-of-mind" (OSOM). A major objective of this article is to clarify issues with the OSOM policy and the possible benefits of more sophisticated policies.

Table 1: Estimated transition data from a major university (a) Linux hosts, (b) changed transitions reflecting improved informatics, (c) Windows hosts, and (d) changes from improved informatics.

(a)

	Low-Med.	Low-Med.-Dark	High-Crit.	High-Crit.-Dark	Comp.	Comp.-Dark
Low-Med.	0.2820	0.6580	0.0177	0.0413	0.0005	0.0005
Low-Med.-Dark	0.2820	0.6580	0.0177	0.0413	0.0005	0.0005
High-Crit.	0.1290	0.3010	0.1560	0.3640	0.0250	0.0250
High-Crit.-Dark	0.0000	0.0000	0.2250	0.7000	0.0250	0.0500
Comp.	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Comp.-Dark	0.0000	0.0000	0.0000	0.0000	0.8000	0.2000

(b)

High-Crit.-Dark	0.1290	0.3010	0.1560	0.3640	0.0250	0.0250

(c)

	Low-Med.	Low-Med.-Dark	High-Crit.	High-Crit.-Dark	Comp..	Comp..-Dark
Low-Med.	0.2760	0.6440	0.0239	0.0559	0.0001	0.0001
Low-Med.-Dark	0.2760	0.6440	0.0239	0.0559	0.0001	0.0001
High-Crit.	0.1444	0.3369	0.1554	0.3627	0.0003	0.0003
High-Crit.-Dark	0.0000	0.0000	0.2988	0.7000	0.0006	0.0006
Comp.	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Comp.-Dark	0.0000	0.0000	0.0000	0.0000	0.8000	0.2000

(d)

High-Crit.-Dark	0.1444	0.3369	0.1554	0.3627	0.0003	0.0003

The remainder of this paper is organized as follows. In Section 2, the structure of the proposed model is described. Section 3 documents a computational experiment involving six alternative systems. In Section 4, the implications for decision-makers and opportunities for future research are given.

2 THE PROPOSED MODEL

2.1 Unit Size and Time Period

Our discrete event simulation model necessarily specifies the number of servers and entities typically within the system (Allen 2011; Law and Kelton 2000). We observed that a large university is generally organized as multiple, largely independent departments, each with typically 100 hosts. Each organization has an administrator principally responsible for repairing vulnerabilities and facilitating responses to known incidents. Therefore, the model includes somewhat more than 100 hosts (on average) over a period of more than 100 years to approximately capture maintenance and response costs for a university. As noted in Afful-Dadzie and Allen (2016), we assume that patching vulnerabilities costs are on average $150 and responding to known incidents costs on average $2,000. Therefore, impacts of vulnerabilities are counted but only in relation to direct costs for legally addressing known incidents.

2.2 States

Following Afful-Dadzie and Allen (2016), we categorize hosts by the highest risk vulnerability, e.g., a host with any critical vulnerability is categorized as critical. In the common policy, low- and medium- risk hosts are generally ignored. Hosts can also be compromised, e.g., the host has malware that is attempting to contact the hacker or hacker team but is intercepted by the intrusion prevention system. Because some hosts are "dark" in the scan and some intrusions are unknown, we consider states in addition to the trashed or recycled host state. States include visible and dark combinations of low-medium, high-critical, and compromised. Low and medium and high and critical are paired because they are often treated as equivalent in organizational policies.

Note that knowing about the vulnerabilities or the intrusions may not help the perceived goals of the organization. Yet, observability is clearly a desirable property of "resilient" systems (Allen et al. 2016). A major objective of this article is to clarify the possible benefits of improved observability.

2.3 SIMIO Model

The model is implemented in SIMIO software. The "NewHosts" in the upper left of Figure 1 below is the source with hosts going to the low-medium vulnerability node where there is no processing. This lack of processing (research, testing, and applying patches if they exist) is a common cost-saving measure in which lowly rated cyber vulnerabilities are ignored. Until recently, because of inspection difficulties all non-network cyber vulnerabilities were largely ignored also by many universities and other organizations. Therefore, they are ignored here also. All paths are fixed "time paths" which correspond to one month.

The weights are proportional to the probabilities in Table 1. The nodes with no processing correspond to states 1, 2, and 4. The servers are states 3, 5, and 6. Even though the dark compromised state does not require work from the internal staff; a server is used to record cost-related information from that state. The retirement node is on the right in Figure 1 in which hosts are recycled or sent to landfills. Overall, hosts are created on the left and flow to destruction on the right. They move from safe states at the top to vulnerability and compromise at the bottom.

Of course, in the real world, the computers reside in offices or cafes and experience minimal movement (with the exception of laptops and cell phones). Therefore, the usual logic of moving hosts is applied as indicated in Figure 2. Hosts do move at inception and at the end of their "lives" when they enter landfills.

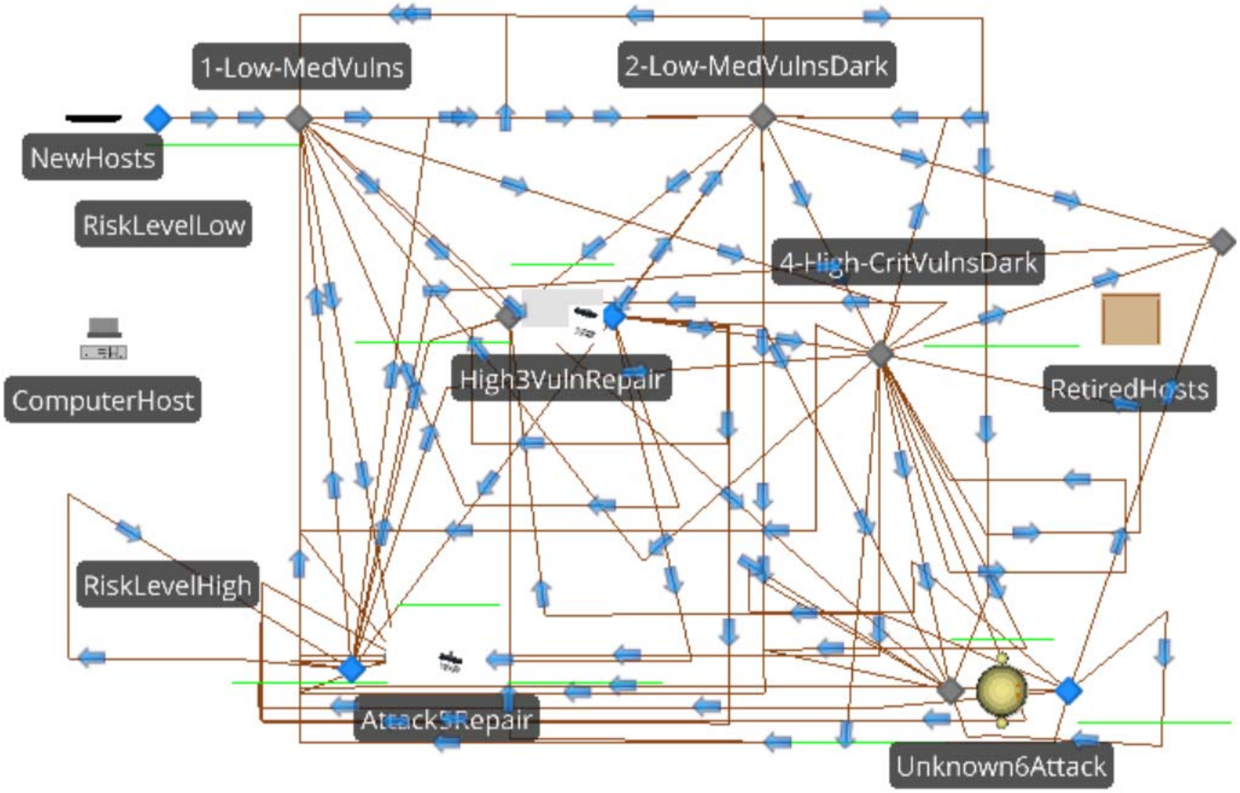

Figure 1: The SIMIO model for organizational or departmental cost forecasting.

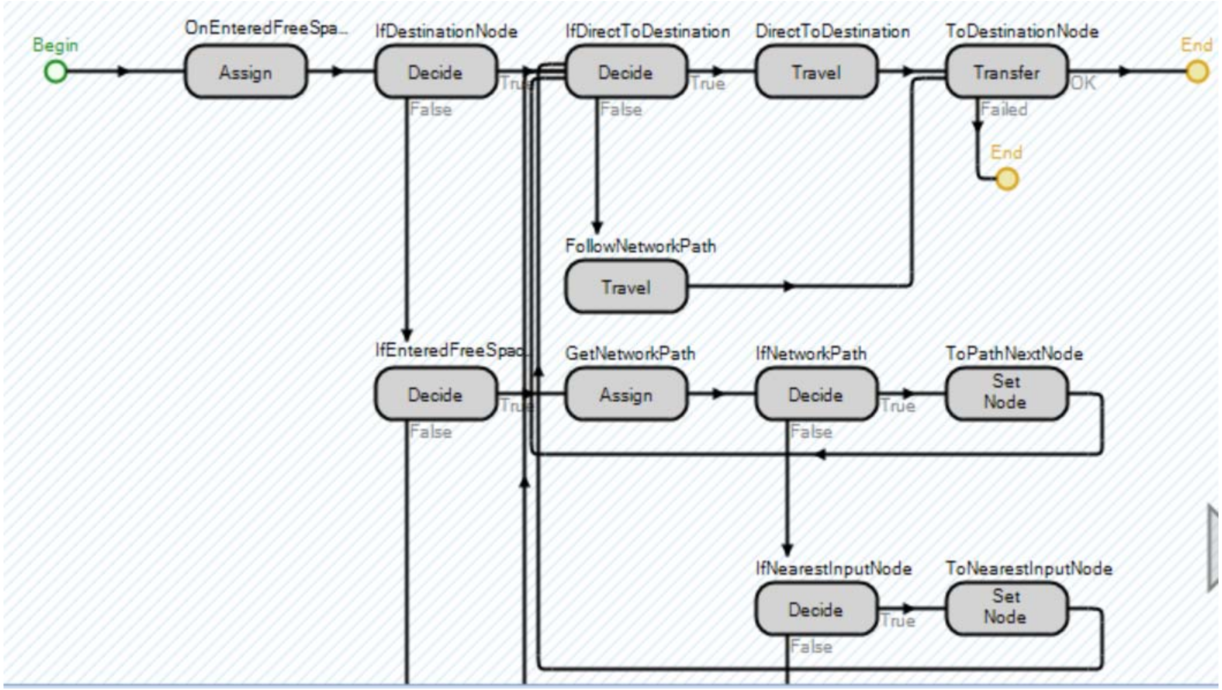

Figure 2: Part of the common built-in SIMIO logic for the computer hosts.

The primary differences between the model in Figure 1 and the Markov Decision Process model in Afful-Dadzie and Allen (2016) are the inclusion of the birth and death of hosts here and the relatively less thorough exploration of optimal policies here. A major strength of Markov Decision Processes is the ability to generate optimal control policies. Yet, the quality of these "optimal" policies is limited by the associated assumptions. Also, unknown attacks are considered here in the simulation model and not previously.

3 RESULTS

3.1 Raw Outputs

The raw SIMIO outputs are shown in Table 2. In the results, 100 replications are used to keep 95% confidence interval half widths to less than 1% of estimated quantities. The results include "H3VRStation1" to clarify that they account only for visits at the visible repair station and not for the dark or unknown vulnerabilities. These hypothetical costs are added in the output analysis derivations so that they do not derived directly from the simulations. Linux scenarios derived from Table 1(a) and (b) probabilities and Windows derived from Table 1(c) and (d) probabilities.

The results in Table 2 relate to the numbers of hosts visiting each node. Visiting a repair or incident node directly results in a cost incurred as a staff member needs to attempt to patch related vulnerabilities or respond to relevant incidents. Therefore, the scenario costs are $150 × (Avg. #Repairs) + $2,000 × (Avg. #Incidents).

Table 2: Raw SIMIO outputs from 100 replications for the numbers of arrivals at the 3 key stations and associated expected or mean costs. The four key "objects" or servers are "Active5Repair" (A5R), "High3VulnRepair" (H3VR), and "Unknown6Attack" (U6A).

Scenario	Object Name	Average #	Half Width	Stdev.	Exp. Cost	Stdev.	Scen. Totals
Linux	A5R	1241.5	8.2	40.8	$2,482,980	81,557	-
Linux	H3VR	6915.9	30.7	152.8	$1,037,387	22,927	-
Linux	U6A	1529.4	11.1	55.4	$3,058,860	110,719	$6,579,227
Linux No D.	A5R	1051.4	6.7	33.4	$2,102,700	66,825	-
Linux No D.	H3VRStation1	5733.1	25.7	127.6	$2,866,565	63,808	-
Linux No D.	U6A	1201.5	8.4	41.7	$2,402,920	83,436	$7,372,185
Windows	A5R	114.8	2.3	11.3	$229,660	22,511	-
Windows	H3VR	8528.2	33.8	168.0	$1,279,229	25,198	-
Windows	U6A	69.0	2.1	10.6	$137,980	21,133	$1,646,869
Windows No D.	A5R	90.5	2.0	10.0	$180,920	20,049	-
Windows No D.	H3VRStation1	5902.8	25.4	126.3	$2,951,415	63,139	-
Windows No D.	U6A	46.5	1.5	7.6	$93,080	15,102	$3,225,415
Linux No Darkness	H3VRStation1	5733.1	25.7	127.6	$2,866,565	63,808	$7,372,185
Windows Hypoth.	H3VRStation1	5902.8	25.4	126.3	$2,951,415	63,139	$864,283

3.2 Comparison of Alternatives

Six systems are compared in Figure 3 in relation to the predicted expected costs. The outputs for the current Linux and Windows systems derive directly from the simulation with inputs in Table 1 and outputs in Table 2. The so-called "improved analytics" policy for each system relates simply to the probabilities or weights coming from Table 1(b) or Table 1(d) for Linux and Windows operating systems respectively. These changes correspond to making state 4 equivalent to state 3 in performance so that additional patching operations would occur. In other words, the hidden vulnerabilities are revealed. This added 1/0.3 times the cost from the server in state 3 (A3VR).

The "Possible Linux" system estimates are based on elicitation from an expert. Questions about what would be expected and what would plausibly be too high or too little were used to elicit estimates that reasonably include the costs of unknown incidents through a marketing-type elicitation process (Allen and Maybin 2004). The possible Linux results are intended to reflect benefits from knowing the vulnerabilities on dark hosts.

The improved policy estimates are based on the likely results that might occur if only critical vulnerabilities (1/5 vulnerabilities or less) were patched on Windows systems. Because of vigorous automatic patching, our analyses from Markov decision processes indicate that patching high vulnerabilities on certain types of Windows systems is not cost effective (Afful-Dadzie and Allen 2016). Yet, there would almost certainly be benefits from patching critical vulnerabilities on dark hosts. Therefore, some of the results in Figure 3 relate to simulation outputs and others are estimates from elicited expert opinions.

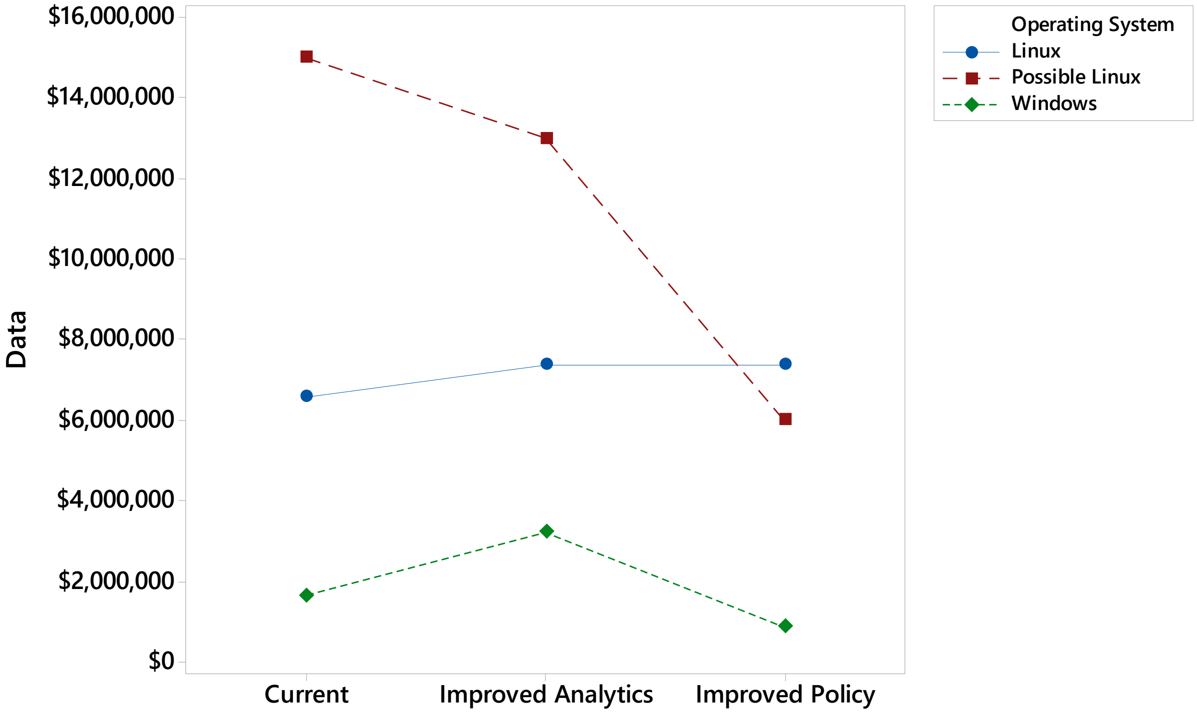

Figure 3: Mean predictions for costs for alternative systems. Half-width intervals are generally less than 1% of the expected costs.

4 CONCLUSIONS AND FUTURE WORK

This article proposes a discrete event simulation model to forecast costs of patching and incident costs. The models are based on hundreds of thousands of recorded transitions. Yet, there are also considerable extrapolations including the effective cost of improved policies or of losses including unknown incidents. With these limitations, the following findings emerge:

1. *Windows hosts require substantially lower maintenance costs in our dataset and simulation predictions than Linux hosts*. This assumes that the host owners had administrator privileges

making these hosts relatively risky to operate. Yet, the vigorous automatic patching carried out by Microsoft likely is associated with lower organizational maintenance costs.

2. *Making dark Windows hosts visible with improved analytics appears not to be cost justified*. This occurs because the cost of dealing with the likely 70% of vulnerabilities ignored by the out-of-site is out-of-mind policy would not be offset by the reduction in known incidents. Yet, if losses to the broader society could be accurately estimated, then the reduced incidents from patching the dark vulnerabilities might be compensated.
3. Making dark Linux hosts visible with improved analytics *is* approximately cost justified and would likely benefit the system with improved resilience and societal benefits.
4. *Making dark hosts of all types visible is likely cost justified if the improved analytics are combined with an improved policy*. For example, for Windows hosts many or all high vulnerabilities might be ignored since auto patching likely addresses many, but the critical vulnerabilities on dark hosts could be predicted and patched to reduce incident costs.

Key limitations of the proposed model relate to features which are unsupported. Multi-fidelity metamodels could provide improved prescriptive ability (e.g., using planning and analysis methods in Allen and Bernstheyn 2005 or Allen et al. 2003). The concepts of partial observability and limited observations can generate useful maintenance recommendations. Also, using automatic control systems based on Bayesian Reinforcement Learning can be applied to direct maintenance and incident response actions that recruit data optimally addressing data limitations.

ACKNOWLEDGEMENTS

We thank LTC Cade Saie and NSF Grant # 1409214 for financial support and Helen Patton, Steven Romig, and Rajiv Ramnath for general support for this and related research.

REFERENCES

Afful-Dadzie, A. and T. T. Allen. 2014. Data-driven Cyber-Vulnerability Maintenance Policies. *Journal of Quality Technology* 46(3):234.

Afful-Dadzie, A. and T. T. Allen. 2016. “Control Charting Methods for Autocorrelated Cyber Vulnerability Data”. *Quality Engineering* 28(3):313-28.

Allen, T. T., 2011. *Introduction to Discrete Event Simulation and Agent-based Modeling: Voting Systems, Health Care, Military, and Manufacturing*. London: Springer Science & Business Media.

Allen, T. and M. Bernshteyn. 2006. “Mitigating Voter Waiting Times”. *Chance* 19(4):25-34.

Allen, T. T. K. M. and Maybin. 2004. “Using Focus Group Data to Set New Product Prices”. *Journal of Product & Brand Management* 13(1):15-24.

Allen, T. T., L. Yu, and J. Schmitz. 2003. “An Experimental Design Criterion for Minimizing Meta-model Prediction Errors Applied to Die Casting Process Design”. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 52(1):103-117.

Allen, T. T., J. Schenk, and D. D. Woods. 2016. “An Initial Comparison of Selected Models of System Resilience”. In *Resilience Engineering Perspectives*, edited by E. Hollnagel and C. Nemeth, Volume 2, 95-116. London: CRC Press.

Case, D. U. 2016. “Analysis of the Cyber Attack on the Ukrainian Power Grid”. Washington, DC: Electricity Information Sharing and Analysis Center (E-ISAC).

Gauci A., S. Michelin, and M. Salles. 2017. “Addressing the Challenge of Cyber Security Maintenance Through Patch Management”. *CIRED-Open Access Proceedings Journal* (1):2599-2601.

Naugle, A., M. Bernard, and I. V. Lochard. 2016. "Simulating Political and Attack Dynamics of the 2007 Estonian Cyber Attacks". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T.M. K. Roeder et al., 3500-3509. Piscataway, New Jersey: IEEE.

Nguyen, C. K. Q., J. E. Dietz, S. Liles, V. Raskin, and J. Springer. 2015. "Cyber Defense Econometric of a Power Grid Distribution Infrastructure". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz et al., 906-911. Piscataway, New Jersey: IEEE.

Huang, D. and T. T. Allen. 2005. "Design and Analysis of Variable Fidelity Experimentation Applied to Engine Valve Heat Treatment Process Design". *Journal of the Royal Statistical Society: Series C* (Applied Statistics) 54(2):443-463.

Law, A. M. and W. D. Kelton. 2000. *Simulation Modeling & Analysis.* 3rd ed. New York: McGraw-Hill.

Shin J., H. Son, and G. Heo. 2015. "Development of a Cyber Security Risk Model using Bayesian Networks". *Reliability Engineering & System Safety* 134:208-217.

AUTHOR BIOGRAPHIES

THEODORE T. ALLEN is an Associate Professor in the Integrated Systems Engineering department at the Ohio State University. He received his B.A. from Princeton, his M.S. from UCLA, and his Ph.D. from the University of Michigan (1997). He is currently the president of the Social Media Analytics section of INFORMS and the simulation area editor of *Computers & Industrial Engineering* (IF: 3.2). He has published over 60 refereed publications and received over 25 grants as PI including from NSF, ARCYBER, and GE Appliances. His research on simulation optimization for voting machine allocation has received national attention and he has contributed to millions of voters avoiding hours of waiting and effective or actual law changes in North Carolina, Ohio, and Michigan. He has also served as associate editor for the *Journal of Manufacturing Systems* and *Quality Approaches in Education* and as a reviewer for *Operations Research*, *Technometrics*, and many other journals (allen.515@osu.edu).

ENHAO LIU is a Ph.D. student in the Integrated Systems Engineering department at the Ohio State University. He received his M.S. from the Ohio State University (2017) and his B.S. from Jinan University in Electrical Engineering and Automation (2015). His interests are related to cyber security, operations research, and reliability engineering (liu.5045@osu.edu).