Process-Aware Attack-Graphs for Risk Quantification and Mitigation in Industrial Infrastructures

Gal Engelberg [1,2]

[1] *University of Haifa, Department of Information Systems, Haifa, Israel*
[2] *Accenture Labs, Tel Aviv, Israel*

Abstract

As connectivity constantly increases, business processes are vulnerable to external cyber-attacks, which may hamper their continuity. As the frequency and derived impacts of these attacks increase, there is a need to prioritize and mitigate risks, considering their impact on business processes, in order of importance to the business. Addressing this need arises several challenges, starting with how to quantify the cyber-security risk over the infrastructure abstract level, map it to the business abstract level, and then propagate it across all process dependencies, and ending with how to prioritize issues to be addressed first. We identified that a holistic approach to answer these challenges in a process-aware manner is still missing. Therefore, the research aims to develop the following framework. First, we will form a process-aware attack-graph that stands for the potential behavior of an attacker within an industrial infrastructure and its impact over the business processes. Second, we will develop a risk inferencing method to quantify the risk over the infrastructure level, map it to the business level and propagate it across different process dependencies. Finally, we will develop a method to identify the risk root causes and recommend for risk mitigation steps. The framework will be evaluated based on real-life event logs and simulated settings of a smart manufacturing factory. The resulted artifacts will be evaluated by a panel of subject matter experts from the areas of cyber-security and business process management.

Keywords

Process-Mining, Attack-Graph, Risk management

1. Introduction

Today, enterprises in general and industrial infrastructures in particular, are increasingly connected to external networks. As such, business processes that were once isolated, are now vulnerable to external cyber-attacks. As the frequency and derived impact of these attacks increase, there is a need to prioritize and mitigate risks in order of importance to the business [1]. However, this need arises several challenges. First, how to quantify the cyber-security risk over the different assets within the industrial infrastructure. Second, how to map this risk to the business abstract level and then, how to propagate it across all the dependencies associated with process elements. Finally, how to recommend and prioritize which issues should be addressed to mitigate the risk over the different process elements.

Cyber-security risk quantification and mitigation is a well-researched area, usually based on a representation of the potential adversary behavior within a system [2], [3]. Recently, an emerging effort is being invested to assess the impact of a cyber-attack over a higher abstract concept, such as a mission [4] or a business process [5]. This is usually done by encoding the dependencies among the different abstract levels into the attack graph in a manner that enables propagating the impact from the infrastructure level to the higher abstract level. We identified that the existing literature is missing of methods that support all the above challenges holistically, and in a complete process-aware manner. We define process-awareness in this context, as the ability to represent the potential adversary behavior

Proceedings of the Doctoral Consortium Papers Presented at the 34th International Conference on Advanced Information Systems Engineering (CAiSE 2022), June 06–10, 2022, Leuven, Belgium
EMAIL: gal.engelberg@accenture.com (G. Engelberg)
ORCID: **0000-0001-9021-9740** (G. Engelberg)

within a system, augmented by its impact and dependency on the different business process perspectives, including: control flow, data flow, resources, and time. In this work we aim to address the identified gaps by developing a process-aware framework for risk quantification and mitigation in industrial infrastructures.

The remainder of the paper is structured as follows: section 2 provides an analysis of the related work, then research questions are posed in section 3. Section 4 depicts the methodology and the proposed approach. Finally, current state of research and conclusions about expected contribution are drawn.

2. Related Work

Cyber-security risk quantification and mitigation is generally based on an attack graph, which simulates an adversary potential behavior within a system. [2], [3] presented MulVAL, an approach to build an attack graph based on a logical reasoning method that uses collected evidence, and inference rules gathered by cyber-security experts to infer what logical steps an adversary should perform to reach a pre-defined target. Then the resulted graph could be used to quantify the cyber-security risk of a system. For instance [6] proposed a methodology to prioritize security-controls implementation to mitigate the system's risk. They used graph centrality measures to rank target nodes' importance in the attack graph, then quantified the cyber-security risk as a function of target nodes' importance, and the difficulty rank of an adversary to reach these nodes. Then they simulated which security controls would reduce the risk effectively.

Even though cyber-security risk quantification is a well-researched topic in the information technology (IT) discipline, this topic is new in the industrial infrastructures landscape (which is based on operational technology - OT). While IT assets deal with information flow, OT assets deal with the operation of physical processes and machinery used to carry them out. Accordingly, IT and OT assets differ by their hardware, software, functionality, network architecture, and communication patterns, and as a result, by their potential adversarial tactics and countermeasures (from offensive and defensive point of view respectively).

[7] collected techniques to detect safety and security risks within industrial infrastructures, [8]–[10] proposed further risk quantification methods. However, that area lacks methods for assessing the impact of cyber-security risks over the higher-level processes, and their organizational business impact. Recently the research topic of cyber mission impact assessment (CMIA) [4] raised to address this need, by projecting the cyber impact measured in the infrastructure level to the abstract level of a mission (e.g. a business activity, a business domain/functionality, etc.). [11] provided a mathematical model for bias and context-free mission impact assessment based on both missions and resources dependency models. [12] used a similar approach to select mitigation actions based on operational and financial impact assessment.

[5] presented an approach to assess cyber-attack impact on business processes by generating an interconnected graph of the dependencies between vulnerabilities on hosts, relations between services to hosts, and tasks to services. The authors encoded the above dependencies with a Datalog[13] as facts and rules and used MulVAL to generate the graph. Then they presented a method to calculate the impact score by propagating the impact generated by the vulnerabilities to the impacted hosts, services, and tasks. [14]–[16] used a similar approach to assess mission impact of cyber-attacks on energy delivery systems, cyber-physical systems, and enterprise applications, respectively. We identified the following limitations in this approach. First, by using a dependency graph or a BPMN model as an input to MulVAL, the business process level facts and rules are limited to short-term control flow dependencies. This approach lacks an ability to identify other dependencies between process elements, which might affect the resulting impact. These include long-term control flow dependencies (as eventually followed by), data flow dependencies, and time constraints. Second, this approach formalizes only the impact of the attacker's behavior concerning the infrastructure layer, and propagates it to the business process layer. However, this impact could be bidirectional, namely, the business process behavior could affect the attacker's capability to compromise infrastructure assets. For example, a vulnerable file could be transferred between multiple endpoints according to some business logic, while some of the endpoints hold existing pre-conditions that enable the attack, and some do not. Third, while this approach provides

a method to quantify the impact over the business abstract level, it does not provide methods to analyze its root cause and recommend risk mitigation actions accordingly. Finally, the impact quantification over the business abstract level keeps the infrastructure perspectives and lacks inferring of the process perspectives. For example, considering an infrastructure impact perspective of an asset's availability, which, according to the proposed approach, triggers an impact over business process availability. However, this approach lacks inferring the expected impact over the process KPIs. For example, an impact over an asset's availability triggers an impact over a KPI of production quantity.

Recent research by [17] presented a method to propagate the risk within a network of business processes and IT services, by transforming the operational risk over the underling IT services to a financial risk over their related business processes. However, the cyber risk quantification is not cyber security exclusive and lacks aspects such as confidentiality and safety. Furthermore, the risk quantification is based on historical events and lacks consideration of events that have not been observed or have only recently emerged, which is a very common phenomenon in the cyber security landscape.

Based on the above, we conclude that there is a need to develop a holistic and a process-aware approach to address the above challenges within the industrial infrastructure domain. This approach will be able to form a process-aware representation of the potential adversary behavior within an industrial infrastructure, quantify the cyber-security risk based on a potential adversary behavior, map it to the process abstract level, propagate it across its dependencies, and prioritize risk mitigation actions according to their business importance.

3. Research Questions

To address the above challenges, we pose the following high-level research question: How should organizations measure cyber-security risk and prioritize risk mitigation steps to protect their industrial infrastructures, and assure business process continuity? To make the question more concrete, we derive the following research questions.

RQ1: How to form a process-aware representation of the potential adversary behavior within an industrial infrastructure? To answer this question, we first need to identify the relevant types of dependencies among process elements, considering control-flow, data-flow, resources, and time perspectives. Second, we need to consider how these dependencies could be discovered in an event log. Third, we need to investigate how these dependencies affect the capability of an attacker to compromise infrastructure assets and business process goals. Finally, we intend to express these dependencies as facts and rules to be used as part of an attack-graph inference engine, and need to formalize them as such.

RQ2: How to infer what is the cyber-security risk over the business abstract level? To answer this question, we first need to determine how to quantify the cyber-security risk over the different assets within the industrial infrastructure. Second, we need to establish a mapping between the cyber-security risk and the business abstract level. Finally, we need to develop a method for propagating the risk across all process dependencies.

RQ3: How to identify the root causes of a risk, prioritize the issues, and suggest a relevant plan for remediation actions? This question will be addressed by considering the following risk mitigation strategy dimensions. First, we will consider different business objectives for the risk mitigation task (reducing the risk for a single activity, a single process, a production line, a factory, and more). Second, we will examine different configurations of risk aspects for the risk mitigation task (reducing the risk for a single risk aspect (e.g., availability), multiple risk aspects, or considering an overall risk measure). Third, we will assess different objective functions for risk reduction (e.g., minimal cost). Finally, we need to address different possible remediation strategies.

4. Methodology

Towards answering the research questions, we take a design science approach, with a main envisioned artifact: a process-aware framework for a cyber-security risk quantification and mitigation in industrial infrastructures, as depicted in Figure 1. This framework gets as an input an event log of the business process, instances of cyber-security facts of an industrial infrastructure, and a knowledge base that holds inference rule templates. First, we intend to develop a method to create a process-aware attack-graph, answering RQ1. Based on the process-aware attack-graph we propose to develop a risk inferencing method to answer RQ2, and methods for a root cause analysis and risk mitigation recommendation to answer RQ3.

The rest of the section is structured as follows. First, we will describe the proposed approach for addressing each of the research questions. Then, we will describe the proposed evaluation method.

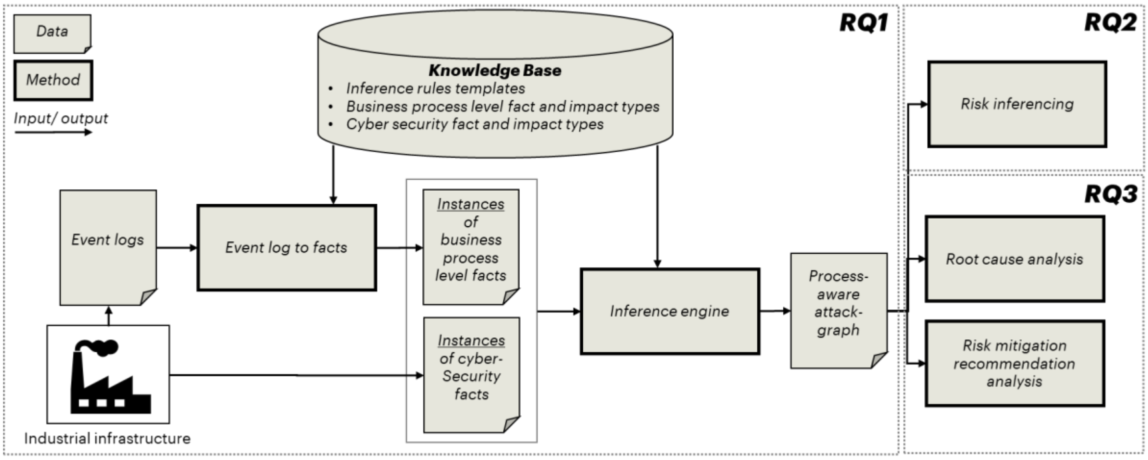
Data
Method
Input/ output
RQ1
RQ2
RQ3
Knowledge Base
Inference rules templates
Business process level fact and impact types
Cyber security fact and impact types
Event logs
Event log to facts
Instances of business process level facts
Instances of cyber-Security facts
Industrial infrastructure
Inference engine
Process-aware attack-graph
Risk inferencing
Root cause analysis
Risk mitigation recommendation analysis

Figure 1: a process-aware framework for a cyber-security risk quantification and mitigation in industrial infrastructures

4.1. RQ1: A Process-Aware Attack-Graph

As described in the related work section, attack-graphs are usually based on MulVAL as a tool that models the interaction between software vulnerabilities, the system, and network configuration. It uses Datalog [13] to model the interaction of various system components to the form of interaction rules (denoted with ellipses in Figure 2). Each interaction rule has preconditions and a derivation. A derivation stands for a new information that is discovered when all preconditions of a rule are met. A precondition of a rule could be either a primitive evidence (denoted with boxes in Figure 2, can be referred to as a fact) or a derivation (denoted with diamonds in Figure 2, can be referred to as an impact) that was discovered before, satisfying a prior interaction rule in the attack-graph. The reasoning process is completed when a predefined impact is discovered, and this impact is denoted as an attack goal.

Figure 2. shows a simple example attack-graph containing 10 nodes. If host1 has a malware and his antivirus is not updated (nodes 1,2), the attacker can run a code on host1 (node 3). Since host1 and activity1 hold a resource-activity dependency (node 4), the attacker can cause a direct denial of service over activity1 (node 5). Finally, since activity1 and activity2 holds a precedence dependency (node 6), an indirect denial of service will affect activity2 (node 7).

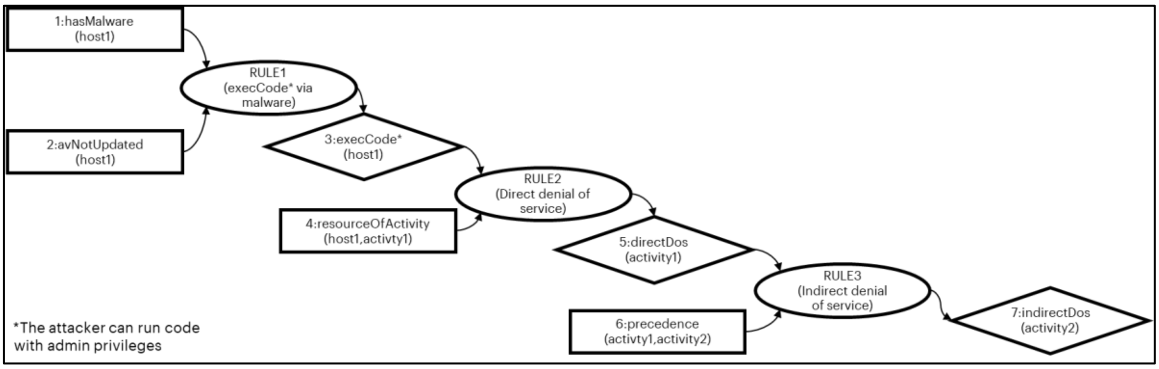
1:hasMalware (host1)
RULE1 (execCode* via malware)
2:avNotUpdated (host1)
3:execCode* (host1)
RULE2 (Direct denial of service)
4:resourceOfActivity (host1,activty1)
5:directDos (activity1)
RULE3 (Indirect denial of service)
6:precedence (activty1,activity2)
7:indirectDos (activity2)
*The attacker can run code with admin privileges

Figure 2: an attack-graph example

We intend to construct a process-aware attack-graph based on the MulVAL approach, as follows. First, we aim to develop a model (denoted as a knowledge base) which will hold information regarding cyber-security and business process fact and impact types, and their corresponding inference rules. We expect to formulate three types of inference rules:

- Cyber-attack rules (e.g., RULE1), based on public knowledge as ATT&CK® for ICS[2].
- Dependencies between the infrastructure and the business level (e.g., RULE2).
- Different process dependencies (e.g., RULE3), considering the perspectives of control-flow, data-flow, resources, and time constraints.

Second, we intend to use a Datalog solver (denoted as an inference engine) to extract the attack-graph given inference rules, business process fact instances (as nodes 4,6), and cyber security fact instances (as nodes 1,2). We aim to develop a method to discover instances of business process facts within an event log (denoted as event log to facts). We propose to achieve that using decelerative process discovery techniques [18]–[20]. Instances of cyber-security facts could be collected via multiple systems (e.g., Nozomi[3], Forescout[4], etc.,), thus, we assume that this information is given and will be simulated as part of the industrial environment settings.

4.2. RQ2: A Process-Aware Risk Inferencing

This part takes a process-aware attack-graph and quantifies the risk per each objective process element, considering all perspectives of process dependencies. This could be achieved via the following approaches.

First, by using graph abstraction and traversal algorithms to transform the process aware attack-graph to a probability network (e.g., a Bayesian network [21]). The network will represent a movement of an attacker within the industrial infrastructure, and its resulting impact and dependencies on the business process. Once we have this representation, we could explicitly calculate the risk associated with an attacker causing an impact within the infrastructure, using transition probabilities and their related impact score. Then risk could be calculated for each cyber asset and be propagated accordingly to the business process level. While the above approach might have a high time complexity at runtime, we suggest to use it to label a training set, and then train a graph neural network (GNN) [22], [23] model or any other probabilistic graph model (PGM) [24] to predict the risk (as a regression problem).

4.3. RQ3: A Process-Aware Risk Root-Cause, and Mitigation Analyses

This part takes a process-aware attack-graph and risk prediction results and searches for cyber-security issues whose remediation would reduce the graph size (and the cyber risk accordingly), considering the dimensions described in RQ3. This could be achieved with the following approaches.

First, by using graph theory algorithms to measure importance of nodes (which represent issues within the graph), such as a reachability analysis [25] , a page-rank analysis [26], an eigenvector centrality analysis [27], and more. Second, by developing a causality-graph-based approach for root cause analysis [28], [29]. Third, by performing a greedy-search simulation that iteratively searches for issues whose risk loss is maximal, then eliminates them from the graph, and recalculates the risk until it reaches a steady state. Finally, by training a reinforcement learning [30] model to optimize the simulation performed in the third approach.

4.4. The Proposed Evaluation

The framework will be evaluated in the following settings. The business abstract level will be discovered based on a real-life public event logs of manufacturing processes, as in [31]. The industrial infrastructure abstract level will be simulated in correspondence to the real-life event log. In addition,

[2] https://collaborate.mitre.org/attackics/index.php/Main_Page

[3] https://nozominetworks.com/

[4] https://forescout.com/

it will hold a sufficient coverage of attack types as specified at ATT&CK® for ICS[2]. Then we intend to implement a prototype of the proposed framework and test its applicability in the described settings. The resulted artifacts (as the attack-graph, and methods for risk inference, root-cause, and mitigation analyses) will be evaluated by a panel of subject matter experts from the areas of cyber-security and business process management, based on the following criteria. First, the extent to which the attack-graph is process-aware. Second, the support given by the proposed framework in obtaining a precise inference of risk, in identifying the root causes of risks, and recommending a risk mitigation plan accordingly.

5. Current State of The Research

To this point, we completed a literature review, and concluded an exploration of existing tools in the scope of the identified problem. Furthermore, we conceptualized the proposed framework and created a research agenda towards answering the research questions. Next steps include: a setup of the research environment, and development of the proposed framework according to the described methodology.

6. Conclusions

While the need for a business centric cyber-security risk management arises, the existing literature is lacking a holistic process-aware approach. We aim to address this need by developing a process-aware framework for a cyber-security risk quantification and mitigation in industrial infrastructures. The expected contribution includes the following. A knowledge base of process-aware attack-graph inference rules, and a method to extract business process facts from an event log. Methods for risk inferencing, root-causes, and mitigation analyses.

Acknowledgements

This PhD thesis proposal is supervised by Professor Pnina Soffer, part of a collaboration with Accenture Labs, Israel.

References

[1] E. Klein, Dan Gal, “Get Ahead of Cyberattacks with Digital Twins | Accenture.” https://www.accenture.com/us-en/blogs/technology-innovation/klein-engelberg-get-ahead-of-cyberattacks-with-digital-twins (accessed Aug. 11, 2021).

[2] X. Ou, W. F. Boyer, and M. A. McQueen, “A scalable approach to attack graph generation,” in *Proceedings of the 13th ACM conference on Computer and communications security*, 2006, pp. 336–345.

[3] “14th USENIX Security Symposium — Technical Paper.” https://www.usenix.org/legacy/event/sec05/tech/full_papers/ou/ou_html/ (accessed Jul. 16, 2021).

[4] S. Musman, “Evaluating the Impact of Cyber Attacks on Missions,” p. 15.

[5] C. Cao, L.-P. Yuan, A. Singhal, P. Liu, X. Sun, and S. Zhu, “Assessing Attack Impact on Business Processes by Interconnecting Attack Graphs and Entity Dependency Graphs,” in *Data and Applications Security and Privacy XXXII*, vol. 10980, F. Kerschbaum and S. Paraboschi, Eds. Cham: Springer International Publishing, 2018, pp. 330–348. doi: 10.1007/978-3-319-95729-6_21.

[6] E. Hadar, D. Kravchenko, and A. Basovskiy, “Cyber Digital Twin Simulator for Automatic Gathering and Prioritization of Security Controls’ Requirements,” in *2020 IEEE 28th International Requirements Engineering Conference (RE)*, Aug. 2020, pp. 250–259. doi: 10.1109/RE48521.2020.00035.

[7] S. Kriaa, L. Pietre-Cambacedes, M. Bouissou, and Y. Halgand, “A survey of approaches combining safety and security for industrial control systems,” *Reliability Engineering & System Safety*, vol. 139, pp. 156–178, Jul. 2015, doi: 10.1016/j.ress.2015.02.008.

[8] A. Teixeira, K. C. Sou, H. Sandberg, and K. H. Johansson, “Secure Control Systems: A Quantitative Risk Management Approach,” *IEEE Control Systems Magazine*, vol. 35, no. 1, pp. 24–45, Feb. 2015, doi: 10.1109/MCS.2014.2364709.

[9] X. Li, C. Zhou, Y.-C. Tian, N. Xiong, and Y. Qin, “Asset-Based Dynamic Impact Assessment of Cyberattacks for Risk Analysis in Industrial Control Systems,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 2, pp. 608–618, Feb. 2018, doi: 10.1109/TII.2017.2740571.

[10] H. Abdo, M. Kaouk, J.-M. Flaus, and F. Masse, “A safety/security risk analysis approach of Industrial Control Systems: A cyber bowtie – combining new version of attack tree with bowtie analysis,” *Computers & Security*, vol. 72, pp. 175–195, Jan. 2018, doi: 10.1016/j.cose.2017.09.004.

[11] A. Motzek and R. Möller, “Context- and bias-free probabilistic mission impact assessment,” *Computers & Security*, vol. 65, pp. 166–186, Mar. 2017, doi: 10.1016/j.cose.2016.11.005.

[12] G. G. Granadillo, A. Motzek, J. Garcia-Alfaro, and H. Debar, “Selection of Mitigation Actions Based on Financial and Operational Impact Assessments,” in *2016 11th International Conference on Availability, Reliability and Security (ARES)*, Salzburg, Austria, Aug. 2016, pp. 137–146. doi: 10.1109/ARES.2016.3.

[13] S. Ceri, G. Gottlob, and L. Tanca, “What you always wanted to know about Datalog (and never dared to ask),” *IEEE Trans. Knowl. Data Eng.*, vol. 1, no. 1, pp. 146–166, Mar. 1989, doi: 10.1109/69.43410.

[14] M. A. Haque, S. Shetty, C. A. Kamhoua, and K. Gold, “Modeling Mission Impact of Cyber Attacks on Energy Delivery Systems,” in *Security and Privacy in Communication Networks*, Cham, 2020, pp. 41–61. doi: 10.1007/978-3-030-63095-9_3.

[15] M. A. Haque, S. Shetty, C. A. Kamhoua, and K. Gold, “Integrating Mission-Centric Impact Assessment to Operational Resiliency in Cyber-Physical Systems,” in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, Dec. 2020, pp. 1–7. doi: 10.1109/GLOBECOM42002.2020.9322321.

[16] M. Bilur, A. Gari, and R. K. Shyamasundar, “Threat Assessment of Enterprise Applications via Graphical Modelling,” in *Network and System Security*, Cham, 2019, pp. 146–166. doi: 10.1007/978-3-030-36938-5_9.

[17] “‘Quantifying Risk Propagation Within a Network of Business Processes an’ by Oscar González-Rojas, Nicolás Castro et al.” https://aisel-aisnet-org.ezproxy.haifa.ac.il/bise/vol63/iss2/5/ (accessed Jul. 17, 2021).

[18] F. M. Maggi, “Declarative Process Mining with the Declare Component of ProM,” p. 5.

[19] C. Sturm, S. Schonig, and C. D. Ciccio, “Distributed Multi-Perspective Declare Discovery,” p. 6.

[20] F. M. Maggi, M. Dumas, L. García-Bañuelos, and M. Montali, “Discovering Data-Aware Declarative Process Models from Event Logs,” in *Business Process Management*, Berlin, Heidelberg, 2013, pp. 81–96. doi: 10.1007/978-3-642-40176-3_8.

[21] T. A. Stephenson, Ed., *An Introduction to Bayesian Network Theory and Usage*. IDIAP, 2000.

[22] L. Lu *et al.*, “Ranking attack graphs with graph neural networks,” in *International Conference on Information Security Practice and Experience*, 2009, pp. 345–359.

[23] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, “The Graph Neural Network Model,” *IEEE Transactions on Neural Networks*, vol. 20, no. 1, pp. 61–80, Jan. 2009, doi: 10.1109/TNN.2008.2005605.

[24] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. MIT Press, 2009.

[25] P. Doreian, “On the connectivity of social networks,” *The Journal of Mathematical Sociology*, vol. 3, no. 2, pp. 245–258, Jan. 1974, doi: 10.1080/0022250X.1974.9989837.

[26] L. Page, S. Brin, R. Motwani, and T. Winograd, “The PageRank Citation Ranking: Bringing Order to the Web.,” Nov. 11, 1999. http://p8090.ilpubs.stanford.edu/422/ (accessed Aug. 04, 2021).

[27] P. Bonacich, “Some unique properties of eigenvector centrality,” *Social Networks*, vol. 29, no. 4, pp. 555–564, Oct. 2007, doi: 10.1016/j.socnet.2007.04.002.

[28] H. Wang *et al.*, “Groot: An Event-graph-based Approach for Root Cause Analysis in Industrial Settings,” *arXiv:2108.00344 [cs]*, Sep. 2021, Accessed: Mar. 04, 2022. [Online]. Available: http://arxiv.org/abs/2108.00344

[29] J. Qiu, Q. Du, K. Yin, S.-L. Zhang, and C. Qian, “A Causality Mining and Knowledge Graph Based Method of Root Cause Diagnosis for Performance Anomaly in Cloud Applications,” *Applied Sciences*, vol. 10, p. 2166, Mar. 2020, doi: 10.3390/app10062166.

[30] L. P. Kaelbling, M. L. Littman, and A. W. Moore, “Reinforcement Learning: A Survey,” *Journal of Artificial Intelligence Research*, vol. 4, pp. 237–285, May 1996, doi: 10.1613/jair.301.

[31] “Production Analysis with Process Mining Technology.” Jan. 28, 2014. doi: 10.4121/uuid:68726926-5ac5-4fab-b873-ee76ea412399.