Identifying Adversarial Cyber-Activity in Operational Technology Environments Using Bayesian Networks

Lee T. Maccarone, Dennis M. Buede, Scott T. Bowman, Pawel Ambrozewicz, Charles D. Burdick, J. Connor Grady, and Shaw X. Wen

***Abstract*—Operational technology (OT) systems face increasing cybersecurity risks from adversarial behavior. This paper describes the development of a Bayesian network risk model to enhance the comprehension of observable cyber-events caused by malicious activity in OT environments. The core of the Bayesian network is a process model that characterizes the stages of adversary behavior. The remainder of the model leverages the MITRE ATT&CK® for Industrial Control Systems (ICS) taxonomy, which includes tactics and techniques that may be used by the adversary. The observables provide evidence for adversary behavior through the intermediary technique and tactic nodes. One challenge in constructing this model is a lack of open-source data from cyber-attacks on OT systems. This paper demonstrates the use of both historical data and expert knowledge to construct the Bayesian network. The historical data was obtained from open-source reporting of 27 cyber-attacks affecting OT systems. The expert knowledge was obtained from a panel of subject matter experts with experience in a variety of OT cybersecurity roles and responsibilities. Finally, the Bayesian network is demonstrated using two historical case studies: the Darkside ransomware attack on the Colonial Pipeline and the destructive cyber-attack targeting the Thyssenkrupp blast furnace. By using this approach, OT cybersecurity professionals can better identify and characterize adversarial behavior in their systems to enable risk-informed investigations and interruptions before impact occurs.**

***Index Terms*—Cybersecurity, industrial control systems, operational technology.**

I. Introduction

OPERATIONAL technology (OT) systems face increasing cybersecurity risks from adversarial behavior. Detecting adversarial cyber-activity in a timely manner is critical to prevent impacts to the availability or safety of critical systems. Timely detection of adversary activity in OT systems is challenging for many reasons including the vast amounts of data generated by the systems, the system-specific knowledge required to properly interpret the data, and the complexity of the interconnections between systems and supporting infrastructure. To address these challenges, we developed a tool for cybersecurity teams to aggregate anomalous cyber-activity and correlate those anomalies to potential adversary activity.

A methodology developed by the U.S. Department of Energy program, Cybersecurity for the Operation Technology Environment (CyOTE$^{\text{TM}}$), formalizes the approach to addressing these challenges [1]. To support cybersecurity decision-making, operators and cybersecurity professionals are encouraged to focus on capturing and classifying anomalous events that are Indicators of Attack (IoAs), as opposed to Indicators of Compromise (IoCs). IoCs are only available after a known compromise has been analyzed, and an adversary can easily change the details associated with an IoC after it is well-known. In contrast, IoAs represent observable events associated with Tactics, Techniques, and Procedures (TTPs) that will be consistent across many cyber-physical environments and intrusion events. Professionals who leverage this approach will be able to comprehend and prioritize events that occur prior to high-consequence impact events without being dependent on IoCs. This approach supports decision-making during cyber-physical events that have not been previously reported, for example, the first use of certain malware or zero-day exploitation of vulnerabilities within an environment.

In this work, we present the development of a Bayesian network to enhance the comprehension of observable cyber-events caused by malicious activity in OT environments. We leverage the MITRE ATT&CK® for ICS framework as a common lexicon for describing adversarial activity in OT systems. Given observable evidence, we use the Bayesian network to make inferences about the likelihood of specific MITRE ATT&CK® for ICS techniques being used by the adversary, and infer the overall progress of the attack in terms of general adversary behavior phases. This model was developed to support human-decision makers in characterizing the nature of adversarial activity in their systems. By implementing this approach, an OT cybersecurity team can better comprehend adversarial activity in their systems

Received 16 May 2024; revised 31 January 2025 and 9 May 2025; accepted 18 May 2025. Date of publication 8 September 2025; date of current version 1 October 2025. This work was supported by the U.S. Department of Energy, Office of Cybersecurity, Energy Security, and Emergency Response (CESER). The associate editor coordinating the review of this article and approving it for publication was Dr. Mohammad Ashiqur Rahman. *(Corresponding author: Lee T. Maccarone.)*

Lee T. Maccarone is with Sandia National Laboratories, Albuquerque, NM 87185 USA (e-mail: lmaccar@sandia.gov; jcgrady@sandia.gov).

Dennis M. Buede and Charles D. Burdick are with ITA International, Newport News, VA 23606 USA (e-mail: dbuede@ita-intl.com; cburdick@ita-intl.com).

Scott T. Bowman and Shaw X. Wen are with Idaho National Laboratory, Idaho Falls, ID 83415 USA (e-mail: Scott.Bowman@inl.gov; Shaw.Wen@inl.gov).

Pawel Ambrozewicz was with ITA International, Newport News, VA 23606 USA. He is now with the Thomas Jefferson National Accelerator Facility, Newport News, VA 23606 USA (e-mail: pawel@jlab.org).

J. Connor Grady was with Sandia National Laboratories, Albuquerque, NM 87185 USA. He is now with the University of Illinois Urbana-Champaign, Champaign, IL 61801 USA (e-mail: jcgrady2@illinois.edu).

Digital Object Identifier 10.1109/TIFS.2025.3607241

© 2025 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

and conduct a risk-informed investigation and response. This model has been developed for offline analysis, and future work will focus on the application of the model for real-time analysis.

This paper is an extension of a previously published extended abstract [2]. The new contributions of this work are:

- An updated structure of the Bayesian network
- Use of additional data in model development
- Detailed discussion about model refinement
- Sensitivity and scalability analyses
- Use of two new case studies to demonstrate the approach
- Expanded analysis and interpretation of case studies

The format of the paper is as follows. First we provide a background of the challenges facing OT cybersecurity, the current state-of-the-art in cybersecurity practice, and the fundamentals of our approach. Next, we describe the development of our model using both subject matter expert (SME) elicitations and data from a set of 27 historical case studies of cyber-attacks affecting OT systems. Finally, we demonstrate the application of the Bayesian model for two case studies: the Darkside ransomware attack on the Colonial Pipeline (2021) and the cyber-attack on the Thyssenkrupp blast furnace (2014).

II. Background

This section provides an overview of challenges facing the field of OT cybersecurity, state-of-the-art OT cybersecurity approaches, and the theoretical foundation for the Bayesian network approach.

A. Challenge

Event analysis for cyber-attack identification in OT environments has several challenges. OT systems generate vast amounts of data, the efficient processing and analysis of which is a significant challenge even for cybersecurity-informed operators. The data's complexity, often stemming from heterogeneous sources, further complicates the task for operators and cybersecurity professionals attempting to differentiate between normal operations and potential adversarial behavior [3].

OT environments have increasing interconnections with supporting infrastructure, like enterprise information technology (IT) networks, adding sources of complexity to the cyber-event analysis landscape [4]. System connectivity expansion not only introduces new vulnerabilities but also makes it more challenging to detect anomalies that could be indicative of a cyber-attack. This challenge can delay decision-making and prevent an effective cybersecurity response before the cyber-attack causes consequences in the OT environment. The intricacy of these interconnected systems necessitates advanced and nuanced approaches to cybersecurity.

Real-time or near-real-time event analysis is a pivotal requirement in OT environments to maintain operational continuity [5]. OT cyber-event analysis under these time constraints puts considerable pressure on the personnel and systems designed to identify adversarial behavior with a high degree of certainty. Cyber-event analysis and investigation must be conducted in a cost-effective manner that does not affect the availability of critical systems (or offsets the cost of system downtime if downtime occurs).

There are various standards in cyber-event analysis and interpretation within OT environments which further complicates matters [6]. Decision-makers face implementation challenges when considering the diversity of these environments, data collection, and interpretation strategies. Most OT environments require high adaptability and customization. This lack of standardization increases the complexity involved in identifying adversarial behavior, necessitating a tailored approach for each unique OT environment.

Finally, the variety and velocity of cyber-attacks pose a continuous and evolving challenge for threat management. Cyber-adversaries are continuously refining their tactics to ensure the success of their objectives. Many of these tactics leverage native system capabilities which blur the line between regular network anomalies and those that are indicative of a cyber-attack [7]. This evolving nature of threats requires an equally dynamic approach to cybersecurity in OT systems that balances known threat behavior models with those that are unknown.

B. Cybersecurity State-of-the-Art

Several state-of-the-art tools and approaches exist to address the challenges identified in the previous section. Many of these tools and approaches are either complementary or enabling to other tools and approaches. The approaches include the following:

- Integrated IT-OT Security Platforms: New security platforms are emerging that integrate IT and OT data, providing a unified view for better anomaly detection [8]. These platforms often use advanced correlation techniques to identify potential threats across the entire digital landscape.
- Real-Time Data Analytics Tools: Solutions that offer real-time analysis of OT data are crucial [5]. These tools can process and analyze data as it is generated, enabling immediate detection and response to potential threats.
- Customized Security Frameworks: Recognizing the diversity in OT environments, state-of-the-art approaches involve developing customized frameworks for data collection and interpretation, tailored to the specific needs of each OT system [3].
- Behavioral Analysis Techniques: Advanced behavioral analysis techniques are used to establish a baseline of normal activities in OT systems [9]. Deviations from this baseline are scrutinized for potential cyber-attack indicators.
- Advanced Machine Learning Algorithms: Machine learning models, especially those based on deep learning, are being developed to better understand the normal patterns of OT systems [10], [11]. These models can more accurately identify deviations that may indicate a cyber-attack.
- OT Cybersecurity Taxonomies: Several taxonomies exist to provide cybersecurity professionals with a common lexicon for discussing security strategies and adversarial

tactics. Examples include the MITRE ATT&CK® and D3FEND™ matrices [12], [13], [14], [15]. The MITRE ATT&CK® for ICS matrix will be discussed in greater detail in the following subsection as it is foundational for the approach in this paper.

C. MITRE ATT&CK® for ICS

The MITRE ATT&CK® for ICS matrix was developed to provide a common lexicon for cybersecurity professionals to describe adversarial activity in ICS/OT systems [12]. The ICS matrix consists of two types of entities: tactics and techniques.

Tactics describe the reason for the adversary to take action. In other words, a tactic represents the goal of the adversary's action. The 12 tactics in the MITRE ATT&CK® for ICS matrix are: Initial Access, Execution, Persistence, Privilege Escalation, Evasion, Discovery, Lateral Movement, Collection, Command and Control, Inhibit Response Function, Impair Process Control, and Impact. Although the tactics are typically shown starting with Initial Access and ending with Impact, the tactics do not need to be completed sequentially, and all tactics do not need to be completed in order for a cyber-attack to cause an impact.

Techniques describe the actions taken by the adversary to achieve a tactical goal. For example, the Spearphishing Attachment technique is a means of accomplishing the Initial Access tactic, and the Indicator Removal on Host technique is used to accomplish the Evasion tactic. There are 81 techniques in the MITRE ATT&CK® for ICS matrix, ranging from two to 14 techniques corresponding to a given tactic. Most techniques correspond to a single tactic, but there are 11 techniques that correspond to two tactics. The MITRE ATT&CK® Enterprise matrix has sub-techniques that provide additional details about the techniques, but the ICS matrix does not.

The MITRE ATT&CK® for ICS matrix is sufficiently general to capture the breadth of cyber-attacks affecting OT systems, including many attacks that exploit zero-day vulnerabilities. In other words, although a zero-day attack includes novel technical exploits, the attack can still be described using the entities in the MITRE ATT&CK® for ICS taxonomy. Techniques are rarely added to the taxonomy, and this approach can easily accommodate new techniques if needed.

D. Bayesian Networks

A Bayesian network (BN) is both a joint probability distribution (JPD) over N discrete variables as well as a message-passing inference algorithm that propagates evidential uncertainty throughout the joint distribution as new evidence is received. The JPD is represented as a graph with a node for each variable and directional arcs representing probabilistic dependence between pairs of variables. The absence of arcs between variable pairs is a statement of probabilistic independence that is often overlooked. While probabilistic dependence is not directional, directional arcs are used in BNs to define which conditional probability distributions are being used in the JPD. An arc from X to Y means the conditional probability $p(Y|X)$ is defined. Aligning the arcs with causality (when causality is known) is recommended for two reasons: it is easier to specify the conditional probability distributions for causal relationships, and there will be fewer arcs in the BN.

Special properties of the directed graph include the requirement to be acyclic, the graph's modularity, and the existence of operations such as arc reversal and node absorption. Since a BN is a JPD, a BN must follow the requirements of probability theory. One such requirement is that the JPD can be factored using the chain rule of probability into N factors, one for each variable. The factors are either marginal (no predecessors) or conditional probability distributions (one or more predecessors). Such a factorization restricts the graph from having one or more cycles (i.e., a sequence of arcs that returns to where it started). This is the only global requirement placed on a BN. A BN is modular in the sense that a change to any variable in the JPD will only affect that node and nodes that are direct predecessors or successors of it. There are strict rules for reversing the arc in a BN; such a reversal is possible as long as it does not create a cycle. Similarly, a variable can be absorbed by summing over its states. Operations for reversal and absorption can require the addition of new arcs between the remaining nodes.

For evidential propagation each piece of evidence is added to a specific node of the BN. An evidence node can be either a successor or a predecessor of an existing BN node. Predecessor evidence changes the prior while successor evidence works via a likelihood function, which is a Bayes' rule operation. Evidence patterns are described in [16]. The evidence propagation process of BNs is an efficient and effective approach to updating risk as a cyber-attack proceeds.

Numerous inference algorithms were developed from 1986 through 2000. Even though such inference can be NP-complete for BNs when many nodes are highly connected [17], most models can be solved nearly instantaneously on laptop computers.

Bayesian networks can be built by using expert judgment, learning from data, or both. The learning process must specify both the structure and the probability distributions. References for building BNs include [18], [19], [20], [21].

There has been substantial work over the past 40 years on how to model the progress of cyber-attacks and the associated risk to owners and users of a network. This review deals with a focus on network attack models [22], [23] with an emphasis on BNs as a subset of attack graphs/networks.

Numerous papers and book chapters have been published on the use of attack trees, graphs, and nets [24], [25], [26]. In early work these attack structures are developed in advance, while in later cases they are developed as the attack progresses. Four of the better discussions for the use of BNs for modeling the risk of cyber-attacks are [27], [28], [29], [30].

Dynamic BNs (DBNs) were also considered as a potential approach for this work. A DBN is an extension of a BN that describes the relationship between variables over time. Examples of DBNs applied for security applications are [31], [32], [33]. DBNs were not selected for this work because of their increased computational complexity relative to BNs. The information provided by a BN was deemed to be sufficient for our objective of enhancing human decision-making and characterization of adversary behaviors.

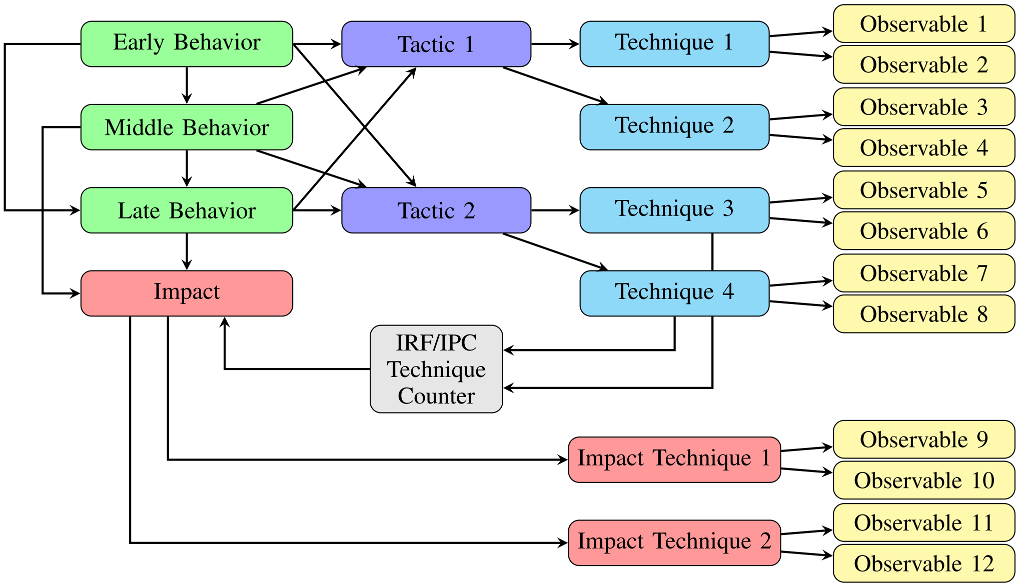


Fig. 1. The structure of the Bayesian network.

The MITRE ATT&CK® framework was trademarked in 2015 so it was not available for most of these efforts. Our approach uses the MITRE ATT&CK® for ICS framework, so the elements of the BN are known [2]. The behavioral nodes and tactic nodes are established in advance while the technique nodes and observable evidence nodes of the techniques are added as the network owners and users perceive observables. This approach to attack graph/network development has not been described by others in the literature to the best of our knowledge.

III. Model Development

This section describes the development of the Bayesian model. The model was developed using a combination of expert elicitation and data from 27 historical case studies of cyber-attacks affecting OT systems. The MITRE ATT&CK® for ICS matrix is the foundation of the Bayesian network and enables evidence propagation from observables to adversary behavior phases. Each subsection provides the rationale for the initial conditional probability tables (CPTs) used in the model and describes how they were improved to their final values.

A. Operational Technology Cyber-Attack Studies

The Bayesian model was developed using case studies of 27 highly publicized cyber-attacks that impacted OT systems from 2000 to 2022. For each case study, a team of control system cybersecurity analysts, cybersecurity researchers, and control system engineers leveraged open-source reporting to identify the sequence of MITRE ATT&CK® for ICS techniques and tactics used by the adversary, and the corresponding observables of the adversary's activity. A summary of the case studies is provided in Table I. Ten of the case studies were ransomware attacks affecting OT systems.

TABLE I
Cyber-Attack Case Studies Summary

Attribute	Min.	Max.	Mean	Std. Dev.
Attack Duration (Days)	0.2	1,562	254	320
Num. Techniques	6	39	19	8
Num. Tactics	4	12	9	2

TABLE II
Bayesian Network Node Descriptions

Node Type	Parents	Children	Count
Adversary Behavior	Adv. Behavior	Adv. Behavior & Tactic	3
Tactic	Adv. Behavior	Technique	11
Impact	Adv. Behavior & Counter	Technique	1
Technique	Tactic	Observable	92
Observable	Technique	–	14,730
Counter	Technique	Impact	1

B. Bayesian Network Structure

A simplified version of the Bayesian model is shown in Figure 1. There are six different types of nodes in the network, which are summarized in Table II. The core of the Bayesian network is a process model that describes the stages of adversary behavior. The remainder of the model is based on the MITRE ATT&CK® for ICS matrix, which includes tactics and techniques [12]. Tactics are objectives that an adversary may seek to accomplish, and techniques are the means by which a tactic is achieved. We initially postulated that tactics may be connected to one, two, or all three phases of adversary behavior. The final version of the Bayesian model connects every tactic to each of the behavior phases, as will be discussed in the following sections. The Impact node is a special case in this model because it acts as both an adversary behavior node and a tactic node. The observables provide evidence for adversary behavior through the intermediary technique and tactic nodes. Finally, a deterministic node is used to count the number of techniques belonging to the Inhibit Response Function (IRF) or Impair Process Control (IPC) techniques.

In the Early Phase, the adversary obtains limited privileges and access to the network, and has partial visibility of the network with a basic user presence. In the Middle Phase, the adversary attempts to escalate privilege and access to the network and expand visibility of the network with capabilities common to power users. In the Late Phase, the adversary often obtains elevated privileges on the network and is able to cause an impact to the asset. This model was constructed with separate nodes for early, middle, and late behavior because these behaviors are not mutually exclusive. Three states characterize each phase of adversary behavior: (1) "None" corresponds to no adversary activity in that phase, (2) "Ongoing" corresponds to active adversary activity in that phase, and (3) "Complete" corresponds to completed adversary activity in that phase.

The direction of the arcs from adversary behaviors to tactics, to techniques, to observables was selected because of the cause-and-effect relationships between these nodes. The cause-and-effect relationships were deduced given the definitions of the elements in the MITRE ATT&CK® for ICS matrix. Given that tactics are defined as broad tactical goals and a technique is defined as an action by which an adversary achieves a tactic, it is reasonable to state that the adversary's goals drive their actions, and therefore tactic nodes ought to be parents of technique nodes. Similar logic can be applied to the adversary behavior nodes and tactic nodes. Finally, technique nodes are parents of observable nodes because observables must be caused by actions. This arc directionality has the added benefit of greatly reducing the complexity of the CPTs in comparison to those in a BN where the arcs are reversed.

C. Adversary Behavior

The adversary behavior phases were defined to allow the Bayesian model to track the progress of the cyber-attack. These phases are crucial for the eventual practical application of the model because they will provide the user with a comprehensive situational awareness to guide cybersecurity response, rather than the likelihoods of individual techniques and/or tactics. We initially created three adversary behavior phases: Early, Middle, and Late. Our rationale was that some attacks take six or more months to finish, and the literature is replete with attack descriptions that stress specific behaviors for early, middle, and late attack strategies [34]. Our case studies demonstrate a wide range of attack duration. For example:

- 12+ months: 6 attacks (e.g., Night Dragon and Norsk Hydro)
- 6-12 months: 6 attacks (e.g., EKANS and Conti)
- 2-6 months: 6 attacks (e.g., Triton and JBS Foods)
- 0-2 months: 8 attacks (e.g., Colonial Pipeline and WannaCry)
- Note: the duration of one case study was not defined from the open-source reporting

Since the three phases are being used to track the attack progress, we decided that three states (None, Ongoing, and Complete) were most appropriate for each of the behavior nodes. Note, it is important to stress that the probabilities (beliefs) associated with each state in each phase represent the knowledge of the operators of the network being attacked about the status of the adversary. An attack may be underway, and the operators may not have observed events related to the behavior. As a result, the probability for "None" for each phase will be greater than zero if there is no observed evidence of an attack. Another important point is that this Bayesian model does not require that one phase is complete before the subsequent phase can begin (e.g., the Early phase does not need to be complete before the Middle phase begins).

Initially, we adopted a Markov approach to connect the adversary behavior nodes: the Late phase was only influenced by the Middle phase and the Middle phase was only influenced by the Early phase. This is the simplest set of assumptions. We were not sure this Markov approach would be sufficient but thought it was best to start here. We also decided to treat the Impact tactic as a fourth adversary behavior phase.

We started by defining common-sense principles for the CPTs. Note, these principles conflict to some degree so not all of them can be true. We desired to find a good set of CPTs so that the principles are mostly met. Note some of these principles pertain to the beliefs (marginal probabilities) in the four phase nodes and other principles pertain to the CPTs that are used to compute the marginals.

1) In no case should the presence/absence of a state be certain given the state(s) of the parent node(s) (i.e., no zeros or 100s in the CPTs).
2) The CPT for the Middle phase given the Early phase should be similar to the CPT for the Late phase given the Middle phase.
3) Without evidence in the Bayesian model, the probability of "None" for each behavior phase should increase as the adversary behavior phases progress (i.e., $p(\text{Early} = \text{None}) < p(\text{Middle} = \text{None}) < p(\text{Late} = \text{None}) < p(\text{Impact} = \text{None})$).
4) Without evidence in the Bayesian model, the probability of "Complete" for each behavior phase should decrease as the adversary behavior phases progress (i.e., $p(\text{Early} = \text{Complete}) > p(\text{Middle} = \text{Complete}) > p(\text{Late} = \text{Complete}) > p(\text{Impact} = \text{Complete})$).
5) The probability that the Impact phase is "None" given the Late phase states should be greater than the probabilities that the Late phase is "None" given the Middle phase states and that the Middle phase is "None" given the Early phase states (i.e., $p(\text{Impact} = \text{None}|\text{Late}) > p(\text{Late} = \text{None}|\text{Middle}) > p(\text{Middle} = \text{None}|\text{Early})$).
6) The probability that the Impact phase is "Complete" given the Late phase states should be less than the probabilities that the Late phase is "Complete" given the Middle phase states and that the Middle phase is "Complete" given the Early phase states (i.e., $p(\text{Impact} = \text{Complete}|\text{Late}) < p(\text{Late} = \text{Complete}|\text{Middle}) < p(\text{Middle} = \text{Complete}|\text{Early})$).

We next did a series of "what-if" tests such as entering findings of "Complete" for the Middle, Late, and Impact phases one at a time and observing what the marginals were for the other phases (e.g., observing the Early states when the Middle

TABLE III
ADVERSARY BEHAVIOR PHASE CPTS

(a) Early Behavior CPT

None	Ongoing	Complete
60	37	3.0

(b) Middle Behavior CPT

Early	None	Ongoing	Complete
None	85	14.9	0.10
Ongoing	60	36	4.0
Complete	2.0	48	50

(c) Late Behavior CPT

Early	Middle	None	Ongoing	Complete
None	None	85	14.9	0.10
None	Ongoing	60	36	4.0
None	Complete	5.0	55	40
Ongoing	None	85	14.9	0.10
Ongoing	Ongoing	60	36	4.0
Ongoing	Complete	3.0	47	50
Complete	None	83	16	1.0
Complete	Ongoing	58	36	6.0
Complete	Complete	2.0	38	60

phase was set to “Complete”, observing the Early and Middle states when the Late was set to “Complete”). These “what-if” exercises revealed our initial CPTs should be revised. Some additional changes were made to the adversary behavior phase CPTs as a result of additional “what-if” analyses described in the following section. These tactic-level analyses also led us to reject our initial assumption of a Markov relationship between the phases. This decision was made due to insufficient changes in the probabilities of earlier adversary behavior phases given evidence for later adversary behavior phases. For example, evidence for the Late phase did not result in expected outcomes for the Early phase when the Markov connections were implemented. The Early phase was connected to the Late phase and the Middle phase was connected to the Impact phase. We chose not to connect the Early phase to the Impact phase because there was little evidence from our case studies or expert justification for connecting those phases. These structural modifications to the network also have the benefit of increasing the model’s robustness to cyber-attack patterns that do not follow linear progressions from one behavior phase to the next. The final CPTs for the Early, Middle, and Late phases are given in Table III. The numbers in the CPTs are percentages (e.g., 0.1 is shown as 10). The CPT for the Impact phase/tactic will be discussed in the following section.

D. Tactics

The tactics used here are from the MITRE ATT&CK® for ICS matrix, though we are treating the Impact tactic as a special case as described above. To begin, we asked a panel of five SMEs to evaluate how frequently they would expect to see each tactic (besides Impact) used in association with each of the adversary behavior phases. The SMEs were asked to evaluate each tactic-phase pair as “common”, “occasional”, “rare”, or “never”. SMEs were instructed not to evaluate a given tactic-phase pair if they were not confident in their assessment. The experts first completed the assessments individually. The tabulated results were then discussed as a group to arrive at a consensus decision. The results are shown in Table IV. The highlighted cells in the table indicate the consensus decision.

TABLE IV
ASSIGNMENT OF TACTICS TO BEHAVIORAL PHASES BY EXPECTED FREQUENCIES OF COMMON (C), OCCASIONAL (O), RARE (R), AND NEVER (N)

	Behavioral Phase											
	Early				Middle				Late			
Tactic	C	O	R	N	C	O	R	N	C	O	R	N
Initial Access	**5**				1	**3**		1		1	1	**2**
Execution	3	**2**			**5**				**3**		1	
Persistence	2	1		**2**	**5**				1		**3**	
Privilege Escalation	1	2		**1**	**4**	1			1	1	**2**	
Evasion	**4**				**4**	1			1	**3**	1	
Discovery	**4**				**3**	2				1	**3**	
Lateral Movement	1	2	▨	1	**5**				1	1	**3**	
Collection	**2**	1		1	**4**	1			1	1	**2**	
Command and Control	1	**3**		1	**4**	1			**2**		2	
Inhibit Response Function			**1**	3	2	**1**			**5**			
Impair Process Control		1	**1**	2	2	**1**			**5**			

Our SMEs told us that none of the evaluated tactics would be associated with only one of the behavioral phases. After our first elicitation session with the SMEs, it became clear that most tactics were associated with all three of the phases. As can be seen by the numbers in Table IV, there was some disagreement among the SMEs. There are only two or three cases in each phase that saw agreement among all the SMEs that voted. After a round of discussion among the SMEs we achieved consensus for the assignments shown with a blue background. Note there is one case where the consensus assignment had not received any individual votes (Lateral Movement was assigned Rare for Early). Based on these consensus positions, Initial Access is not connected to the Late phase, and Persistence and Privilege Escalation are not connected to the Early phase. All other tactics have links to all three phases.

On the basis of the consensus shown in Table IV and the votes associated with each phase for a given tactic, we developed CPTs for the tactics. We identified several tactics that were assigned identical CPTs as a starting point:

- Persistence and Privilege Escalation
- Inhibit Response Function and Impair Process Control
- Command & Control and Execution
- Evasion, Discovery, and Collection

Once the CPTs for the Tactics were created, we performed additional “what-if” exercises by setting the value of particular phases to individual states with certainty and observing how the marginals for the tactics changed. The “what-if” exercises and tactics are too numerous to describe in detail, but they resulted in many changes to the tactic CPTs based on discussions of the results with the SMEs.

The next phase was a round of “what-if” exercises in which the marginal probability of “Complete” or “None” for a specific tactic (or small groups of tactics) was set to 100. This “what-if” case would create changes via Bayes’ rule for the behavioral nodes and probability propagation for the other tactics. Again, we observed numerous changes that should be fixed (e.g., $p(\text{Early} = \text{None})$ increased when it was expected to decrease). We identified a list of these "incorrect" reactions to "what-if" scenarios and developed reasonable changes to the CPTs for the behavioral phase and Impact nodes that adhered to the principles identified earlier. The CPT for the Middle phase did not need to change much. We made a range of changes to the CPTs for the Late phase and Impact nodes. We also decided to add three major changes to the structure of the nodes and arcs:

1) Arcs were added from the Early phase to the Late phase and from the Middle phase to Impact. This rejection of our initial Markov approach addressed some major issues even though they were unlikely to happen.
2) A new node was added to count the number of techniques associated with primary tactics for causing impacts (Inhibit Response Function and Impair Process Control). This change was implemented to reduce the probability that Impact was "None" when multiple impact-causing techniques were being employed.
3) All tactics (besides Impact) were connected to all three adversary behavior phases. This change was implemented to counteract anomalous decreases in the probability of Early adversary behavior.

The Impact CPT is now much larger and is shown in Table V. For brevity, it is not feasible to show the CPTs for all tactics.

At the conclusion of the 27 case studies, we revisited our initial survey of the SMEs and our updates to the tactic CPTs while considering the frequencies at which each tactic was used in each phase of adversary behavior in the case studies. All of the initial SME assessments were reasonably accurate relative to the data obtained from the case studies. Unsurprisingly, the most accurate assessments corresponded to the tactics that typically start or end a consequential cyber-attack (e.g., Initial Access, Inhibit Response Function, Impact). The SME assessments that were least accurate relative to the case study data were addressed in the previously discussed "what-if" analyses.

E. Techniques

The techniques used here are directly from the MITRE ATT&CK® for ICS matrix. To develop the technique CPTs, we started with a prior and updated it based on the frequency of the technique over the 27 case studies. The prior CPT is given in Table VI. For ease of explanation of our updating method, the CPT is shown using counts instead of percentages.

For simplicity, the counts corresponding to the "None" row were not modified. The probability corresponding to the "Yes" state for the technique were calculated for the "Ongoing" and "Complete" rows using the formula

$$p(\text{Yes}) = \frac{\text{Count}_{\text{Yes}} + n_t}{\text{Count}_{\text{Yes}} + \text{Count}_{\text{No}} + n_c} \quad (1)$$

where the Counts are given by Table VI, n_t is the number of times the technique occurred in the case studies, and n_c is the number of case studies. We updated the technique CPTs periodically as we progressed through the case studies, but $n_c = 27$ for the final CPT calculations. The probabilities of the "No" state for the "Ongoing" and "Complete" rows are the complement of the probabilities of the corresponding "Yes" state.

TABLE V
IMPACT BEHAVIOR/TACTIC CPT

Middle	Late	IRF/IPC Count	None	Ongoing	Complete
None	None	$0 \leq c < 2$	92	7.7	0.3
None	None	$2 \leq c < 5$	77	22	1.0
None	None	$5 \leq c < 9$	50	43	7.0
None	None	$9 \leq c < 20$	30	50	20
None	Ongoing	$0 \leq c < 2$	70	26	4.0
None	Ongoing	$2 \leq c < 5$	55	35	10
None	Ongoing	$5 \leq c < 9$	30	40	30
None	Ongoing	$9 \leq c < 20$	10	20	70
None	Complete	$0 \leq c < 2$	12	61	27
None	Complete	$2 \leq c < 5$	6.0	45	49
None	Complete	$5 \leq c < 9$	2.0	28	70
None	Complete	$9 \leq c < 20$	1.0	4.0	95
Ongoing	None	$0 \leq c < 2$	75	24.7	0.3
Ongoing	None	$2 \leq c < 5$	62	37	1.0
Ongoing	None	$5 \leq c < 9$	40	53	7.0
Ongoing	None	$9 \leq c < 20$	30	50	20
Ongoing	Ongoing	$0 \leq c < 2$	60	36	4.0
Ongoing	Ongoing	$2 \leq c < 5$	40	50	10
Ongoing	Ongoing	$5 \leq c < 9$	20	50	30
Ongoing	Ongoing	$9 \leq c < 20$	10	20	70
Ongoing	Complete	$0 \leq c < 2$	8.0	46	46
Ongoing	Complete	$2 \leq c < 5$	5.0	48	47
Ongoing	Complete	$5 \leq c < 9$	3.0	28	69
Ongoing	Complete	$9 \leq c < 20$	1.0	4.0	95
Complete	None	$0 \leq c < 2$	85	14.7	0.3
Complete	None	$2 \leq c < 5$	62	37	1.0
Complete	None	$5 \leq c < 9$	40	53	7.0
Complete	None	$9 \leq c < 20$	30	50	20
Complete	Ongoing	$0 \leq c < 2$	60	36	4.0
Complete	Ongoing	$2 \leq c < 5$	40	50	10
Complete	Ongoing	$5 \leq c < 9$	20	50	30
Complete	Ongoing	$9 \leq c < 20$	10	20	70
Complete	Complete	$0 \leq c < 2$	4.0	65	31
Complete	Complete	$2 \leq c < 5$	1.0	50	49
Complete	Complete	$5 \leq c < 9$	1.0	29	70
Complete	Complete	$9 \leq c < 20$	1.0	4.0	95

TABLE VI
PRIOR TECHNIQUE CPT
(COUNT FORMAT)

Tactic	No	Yes
None	19	1
Ongoing	2	1
Complete	1	2

TABLE VII
SPEARPHISHING ATTACHMENT CPT

Tactic	No	Yes
None	95	5.0
Ongoing	63	37
Complete	60	40

An example of a final technique CPT is shown in Table VII for the Spearphishing Attachment technique. The Spearphishing Attachment technique occurred in 10 of the case studies. For brevity, it is not feasible to show the CPTs for all techniques here.

TABLE VIII

NORMAL OPERATIONS LABELS, LOGIC, AND PROBABILITIES

Label	Logic for "Yes"	p(No) as %	p(Yes) as %
Yearly	1 of 500	99.8	0.2
Quarterly	4 of 400	99	1
Monthly	12 of 400	97	3
Weekly	50 of 350	86	14
Persistent	490 of 500	2	98

The lack of subtechniques in the MITRE ATT&CK® for ICS matrix does not affect the model's primary objective of inferring the behavior phases demonstrated by the adversary. If a Bayesian network was constructed with a similar structure for the MITRE ATT&CK® Enterprise matrix, then subtechnique nodes would be children of technique nodes and parents of observable nodes.

F. Observables

An observable is "an event (benign or malicious) on a network or system" [35]. Observables are the source artifacts, or evidence, for developing hypotheses about the likelihood of adversarial behavior. Observables may be cyber or physical events, and are often characterized as IoCs or IoAs [36]. An IoC is digital evidence that a network has been breached (e.g., signatures, IP addresses). IoAs are actions that an adversary must conduct to successfully complete an intrusion, and are less specific than IoCs (e.g., code execution, lateral movement). A cyber-attack consisting entirely of zero-day exploits would not produce a recognized IoC, but may produce recognized IoAs. The value of this model is most significant when considering IoAs as observable evidence, due to their robustness for identifying both known and zero-day attacks.

The CPT for each observable contains two rows. The first row is for normal operations when the technique generating the observable is not in use. The second row addresses when the technique is being used, which may cause the observable to appear. The CPT also has two columns: "No" the observable is not present and "Yes" the observable is present. Initially, the risk analysts and the SMEs agreed upon four labels for the "Yes" in the first row: yearly, quarterly, monthly, and weekly. We experimented with daily and weekly but found that there was little impact on the numerical results when one was used instead of the other, and weekly seemed to be the next logical step from monthly. Later, we decided that something more common than weekly/daily was needed. Persistent was a much more common assessment by the SMEs than daily, so we adopted persistent (skipping daily). A similar approach was defined for the "Yes" in the second row (observable detected when the technique is in use): critical (or very high), high, medium, and low.

The following tables provide the rationale for assigning probabilities for the first and second rows, respectively. The first table addresses the probability of No/Yes, the Observable is not/is present for an observer to perceive given normal operations. The second table addresses the probability of No/Yes, the Observable is not/is present given that the technique is in use.

TABLE IX

TECHNIQUE USAGE LABELS, LOGIC, AND PROBABILITIES

Label	Logic for "Yes"	p(No) as %	p(Yes) as %
Critical	99 of 100	1	99
High	4 of 5	20	80
Medium	1 of 2	50	50
Low	1 of 5	80	20

TABLE X

DIAGNOSTICITY VALUES

Technique	Normal Operation				
Usage	Persistent	Weekly	Monthly	Quarterly	Yearly
Critical	1.01	7.07	33.0	99.0	495
High	0.816	5.71	26.7	80.0	400
Medium	0.510	3.57	16.7	50.0	250
Low	0.204	1.43	6.67	20.0	100

"Yearly" in Table VIII includes time well beyond a year. The SMEs often said "hopefully never" when choosing yearly as the label. As a result, we chose 500 days as the representative time for yearly. For quarterly we rounded 365 days in a year to 400 and selected four as the number of quarters in a year. We took a similar approach for monthly. For weekly, which included hours and days within it, we used 350 as the comparison and chose 50. The last three columns are calculated from the "Logic for Yes" column. For persistent, we went back to 500 as the benchmark and selected 490 as a representation of persistent.

The "Logic for Yes" in the second table represents an assessment of how likely the observable will be present for the observer to perceive if the technique is used. This is designed so that if an observable scores Persistent and Critical, the resulting likelihood ratio (99 divided by 98) will be greater than one for "yes" and lower (one divided by two) for "no".

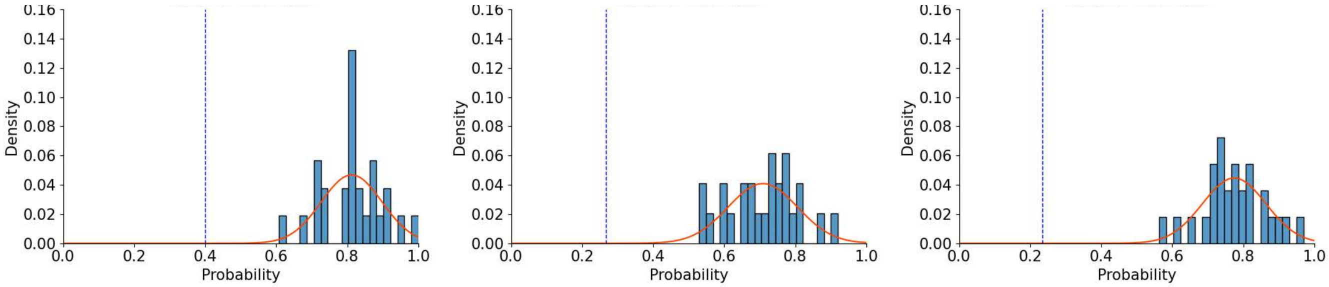
The value of the observable for diagnosing a particular technique, or "diagnosticity", can be defined as a likelihood ratio between the two rows of the observable's CPT. Diagnosticity values can range from 0.2 to 495 based on our CPT definitions. Diagnosticity values of less than one mean that the observable decreases the probability of the technique and values greater than one mean that the observable increases the probability of the technique. All possible diagnosticity values are given in Table X.

IV. MODEL EVALUATION

This section details several analyses evaluating the performance of the model, including its accuracy, scalability, and sensitivity to the case studies used to construct the model.

A. Accuracy

To evaluate the accuracy of the model, the probability of adversary behavior was calculated at the conclusion of each adversary behavior phase for each of the 27 case studies. For each phase, the probability of ongoing or complete behavior was examined given cumulative evidence up to the conclusion of that phase (e.g., the probability of ongoing or complete early behavior given evidence up to the conclusion of the early phase).



(a) Early phase behavior

(b) Middle phase behavior

(c) Late phase behavior

Fig. 2. The probability of ongoing or complete adversary behavior given the cumulative evidence at the completion of the corresponding adversary behavior phase. Data shown is from analysis of all 27 case studies [37].

TABLE XI
PARAMETERS OF THE FITTED NORMAL DISTRIBUTIONS FOR EACH ADVERSARY BEHAVIOR PHASE PROBABILITY GIVEN CORRESPONDING CUMULATIVE EVIDENCE [37]

Phase	Mean	Std. Dev.	r
Early	0.812	0.0857	0.982
Middle	0.709	0.0977	0.994
Late	0.774	0.0894	0.989

It was required that the probability of a given adversary behavior phase given corresponding evidence ought to be significantly different from the prior value (i.e., probabilities without any observable evidence). It is not required for these probabilities to consistently be close to one, however, due to two reasons:

1) Some case studies have observables with greater diagnosticity values than those in other case studies. If the observables for a given case have low diagnosticity values (i.e., large normal frequency and small probability given the technique), then it is more difficult to be certain of the corresponding Technique and therefore more difficult to be certain of adversary behavior phases.
2) Some case studies include tactic-technique pairs that are more strongly correlated with specific adversary behavior phases than those included in other case studies. If a tactic-technique pair is not strongly correlated with a particular phase, it is more difficult to be certain of that phase.

The results are shown in Figure 2. The prior values for each phase are shown by the vertical dashed blue lines. Normal distributions were fit to each data set and the means and standard deviations are given in Table XI. Probability-quantile plots were used to verify that the normal distributions adequately fit the data. A linear relationship between the probabilities and theoretical quantiles indicates that the normal distribution is a good fit for the data. The correlation coefficients, r, of the probability-quantile data are also given in Table XI. The correlation coefficients were very close to one for all three phases, indicating that the normal distributions fit the data well.

Given these results, the model is deemed to be sufficiently accurate. The probabilities are consistently significantly greater than the priors and often approach one for cases that contain observables with high diagnosticity and tactic-technique pairs that are strongly correlated with adversary behavior phases. In cases where the phase probabilities are not as close to one, the user may still determine that investigation or response is appropriate, depending on the organization's risk tolerance [37].

B. Scalability

One consideration for implementation of this model is scalability. The structure of the network is well-defined, and most categories of nodes are limited in number based on the structure of the MITRE ATT&CK® for ICS matrix (e.g., the maximum number of tactic nodes is 12). The most significant factor potentially affecting scalability is the number of observables. The observable nodes are simply connected to the rest of the network via the technique nodes, therefore including large numbers of observables in the model was not expected to be a challenge.

Performance testing was conducted to confirm the model's capability to include large numbers of observables. Synthetic case studies were constructed using various numbers of techniques and observables. The number of tactics was held constant, and the techniques were evenly assigned to the tactics. Similarly, the observables were evenly assigned to the techniques. Inference was timed and the results are shown in Figure 3.

These results indicate that observable count does in fact drive inference time. Although there appears to be some effect due to technique count, it is outweighed by the observable count. Surprisingly, at least for large numbers of observables, inference time scales inversely with technique count. We speculate that this effect is due to the selected Bayesian inference algorithm [38]. It appears that in border cases of few nodes and many edges, the effects of the latter are dominant. In any case, in regions of reasonable use we see that the inference times range from near instantaneous at the low end to a few minutes at the high end, and are sufficient for the model's current purpose of offline use supporting human decision-makers. Approximate inference methods may be considered for real-time use of the model, depending on the expected number of observables for real-time use.

C. Model Sensitivity to CPTs

A sensitivity analysis was conducted to determine the model's sensitivity to perturbations of the values of the CPTs.

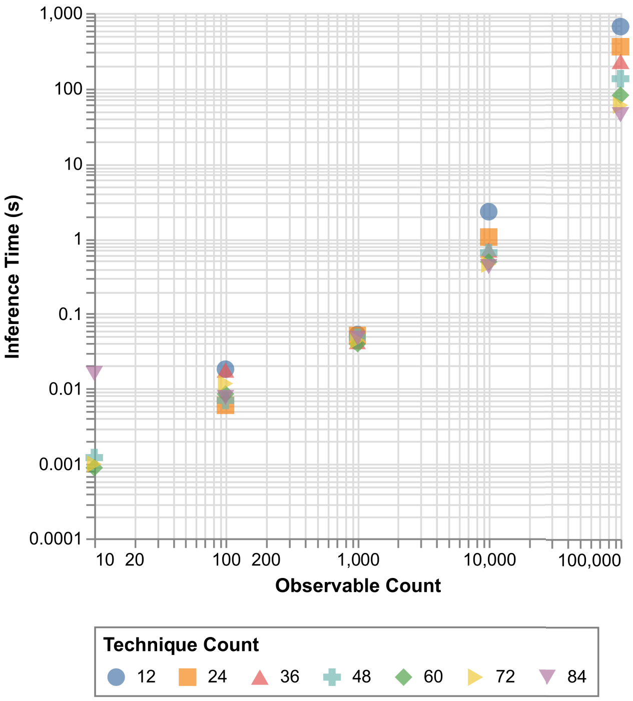


Fig. 3. Scalability testing results. Inference times on the scale of nanoseconds are omitted.

The greatest focus was placed on the CPTs for the observable nodes because the CPTs for the adversary behavior, tactic, and technique nodes were either updated with data collected from the 27 case studies, or iteratively modified based on expert analysis of the case study results as discussed in Section III. Specifically, we focused on the sensitivity to changes in the row of the observable CPT corresponding to the probability of the observable given the use of the corresponding technique. Although we have implemented five possibilities for the normal frequency of the observable, this value could be tailored to the user's systems using system-specific data.

Our sensitivity analyses indicated that small perturbations to the probability of the observable given the technique did not significantly affect the model's output. This is because many observables from multiple techniques are generally necessary to significantly affect the adversary behavior nodes. To demonstrate this, we show the results from analyzing the CPT for an observable corresponding with a normal frequency of "Monthly" and a technique usage probability of "High". The probability of the observable given the technique is defined to be 0.8, and we examined perturbations to values between 0.7 and 0.9. The effects on inference of the technique usage are shown in Figure 4. The range of the probability of the technique is small (0.043) when findings are entered for one observable and the difference is negligible when findings are added for a second observable. Differences of this magnitude have a negligible effect on inferences for the adversary behavior nodes.

D. Model Sensitivity to Case Study Data

The model's output was analyzed for sensitivity to the 27 case studies using leave-one-out cross validation (LOOCV). In this approach, the case studies are analyzed to determine the sensitivity of the model's output to including each case study in the training set. For each case study, the model's results are compared when all 27 cases are used as the training set to when the individual case study is excluded from the training set (i.e., the remaining 26 cases are the training set). This approach was selected rather than dividing the 27 cases into training and testing sets because of the limited number of case studies.

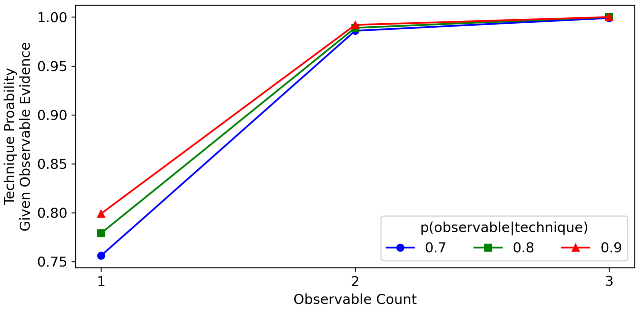


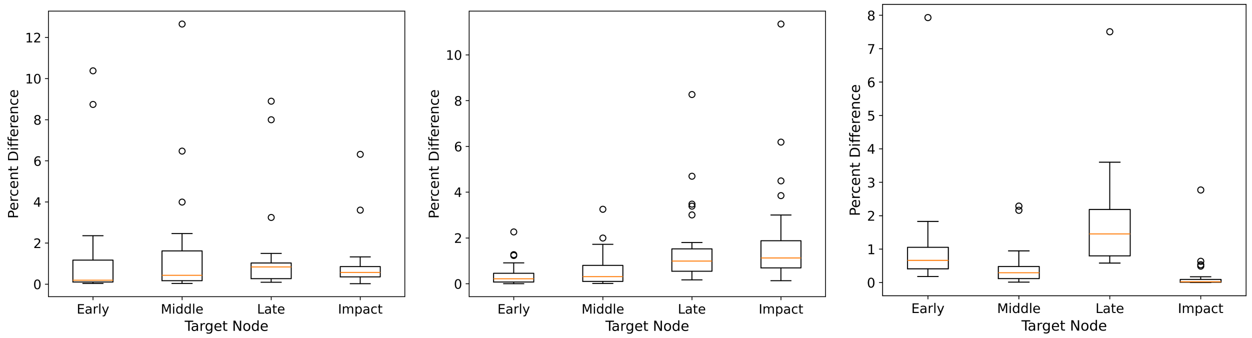
Fig. 4. Probability of a technique given cumulative observable evidence for perturbations of the observable CPT.

The target nodes of early, middle, late, and impact were analyzed at three time points for each case study: at the conclusion of all early phase adversary behavior, at the conclusion of all middle phase adversary behavior, and at the conclusion of all late phase adversary behavior. The time points were selected in this manner to ensure consistency across the case studies which varied significantly in both duration and number of adversarial techniques. The outputs considered in the analysis are the sum of the probabilities of the Ongoing and Complete states for the adversary behavior nodes.

The results of the cross validation are summarized in Figure 5. The percent difference between the full case study set and partial case study set was small for all target nodes at the conclusion of each phase of adversary behavior. The outliers in Figure 5 often correspond to case studies with irregular progressions of adversary behavior phases, but are small nevertheless. These small percent differences indicate that the model is not significantly sensitive to any one case study. Given the variations in the adversary behavior sequences of the 27 case studies, the results also indicate that the model is robust to variations in the sequence of adversary behavior phases. This is particularly useful for detection of cyber-attacks with an unusual structure (e.g., some zero-day attacks).

V. Cyber-Attack Analyses

This section contains two historical case studies to demonstrate the application of the Bayesian network to identify adversarial activity. The first case study is the Darkside ransomware attack on the Colonial Pipeline occurring in 2021 and the second case study is the destructive cyber-attack on the Thyssenskrupp Blast Furnace occurring in 2014. Both case studies leverage publicly available information to identify observables of the attack and approximate adversary technique timelines. For each case study, the probability of each



(a) Percent difference at conclusion of early phase behavior

(b) Percent difference at conclusion of middle phase behavior

(c) Percent difference at conclusion of late phase behavior

Fig. 5. Percent difference of probability of ongoing or complete adversary behavior given cumulative evidence at the conclusion of each phase of adversary behavior.

TABLE XII
Techniques and Corresponding Observables Used in the Darkside Ransomware Attack on the Colonial Pipeline

Tech. ID	Technique	Tac. ID	Tactic	Num. of Obs.
T0822	External Remote Services	TA0108	Initial Access	4
T0859	Valid Accounts	TA0110	Persistence	4
T0890	Exploitation for Privilege Escalation	TA0111	Privilege Escalation	7
T0869	Standard Application Layer Protocol	TA0101	Command and Control	6
T0884	Connection Proxy	TA0101	Command and Control	5
T0811	Data from Information Repositories	TA0100	Collection	3
T0882	Theft of Operational Information	TA0105	Impact	7
T0809	Data Destruction	TA0107	Inhibit Response Function	11
T0881	Service Stop	TA0107	Inhibit Response Function	3
T0826	Loss of Availability	TA0105	Impact	8
T0828	Loss of Productivity and Revenue	TA0105	Impact	4

phase of adversary behavior is calculated given the reported observables.

A. Darkside Ransomware Attack on the Colonial Pipeline

On 7 May 2021, Colonial Pipeline Co., a Houston, Texas, based refined fuel pipeline operator, experienced a ransomware attack on their enterprise network, resulting in the company shutting down pipeline operations. Colonial's CEO, Joseph Blount, reported that an employee in the control center identified a ransom note on a system in the enterprise network that demanded payment to regain access to the system [39]. The employee immediately notified the operations supervisor at the control center [39]. The supervisor decided to halt operations to isolate the OT network from the attack [39]. This resulted in the shutdown of 8850 km of pipeline that delivered refined fuel products to approximately 260 points across 13 states [40]. The outage impacted availability of commercial gasoline services, resulting in higher prices and consumer-driven shortages [40]. Shortly after discovering the ransomware note, Colonial paid a ransom of US$4.4 million to the adversaries for a decryption tool [40]. The company restarted its pipeline operations on 12 May 2021, five days after the initial shutdown [41].

The cyber-attack began with the adversaries exploiting a vulnerable virtual private network (VPN) account lacking multi-factor authentication, enabling them to gain initial access to the IT network [42]. They then executed the DarkSide ransomware to encrypt files [43], likely using techniques to maintain persistence such as modifying user accounts [44]. The adversaries likely escalated their privileges and evaded detection by disabling security tools and obfuscating their activities [45]. Credential access and reconnaissance efforts were likely undertaken to navigate the network and identify targets [46]. Although the ransomware primarily impacted the IT network, there was potential for lateral movement to the OT network [47]. The attack's primary observable impact was the file encryption leading to the operational shutdown of the pipeline [48].

The reported observables are summarized in Table XII and Figure 6. Figure 6 is a heat map of the diagnosticity of the observables of each technique shown in sequential order. In Figure 6, the "D-n" notation indicates the number of days before the triggering event (i.e., the event that prompted the victim to initiate an investigation) and "D+n" notation indicates the number of days after the triggering event. The same convention is used with the letter "H" to indicate the number of hours before and after the triggering event. Many of the observables corresponding to the earliest techniques had moderate diagnosticity scores, and the greatest number of highly diagnostic observables occurred in the Impact phase when the Loss of Availability technique was used.

The adversary behavior phase probabilities are shown in Figure 7. The value "D-∞" corresponds to the output of the network without evidence and the value D0 corresponds to the triggering event. For clarity, some techniques occurring

TABLE XIII
TECHNIQUES AND CORRESPONDING OBSERVABLES USED IN THE CYBER-ATTACK ON THE THYSSENKRUPP BLAST FURNACE

Tech. ID	Technique	Tac. ID	Tactic	Num. of Obs.
T0865	Spearphishing Attachment	TA0108	Initial Access	11
T0817	Drive-by Compromise	TA0108	Initial Access	33
T0863	User Execution	TA0104	Execution	36
T0869	Standard Application Layer Protocol	TA0101	Command and Control	69
T0859	Valid Accounts	TA0110	Persistence	10
T0834	Native API	TA0104	Execution	33
T0846	Remote System Discovery	TA0102	Discovery	28
T0888	Remote System Information Discovery	TA0102	Discovery	42
T0859	Valid Accounts	TA0109	Lateral Movement	19
T0861	Point and Tag Identification	TA0100	Collection	16
T0802	Automated Collection	TA0100	Collection	21
T0814	Denial of Service	TA0107	Inhibit Response Function	18
T0885	Commonly Used Port	TA0101	Command and Control	16
T0882	Theft of Operational Information	TA0105	Impact	13
T0872	Indicator Removal on Host	TA0103	Evasion	24
T0813	Denial of Control	TA0105	Impact	21
T0827	Loss of Control	TA0105	Impact	21
T0880	Loss of Safety	TA0105	Impact	8
T0879	Damage to Property	TA0105	Impact	4
T0828	Loss of Productivity and Revenue	TA0105	Impact	4

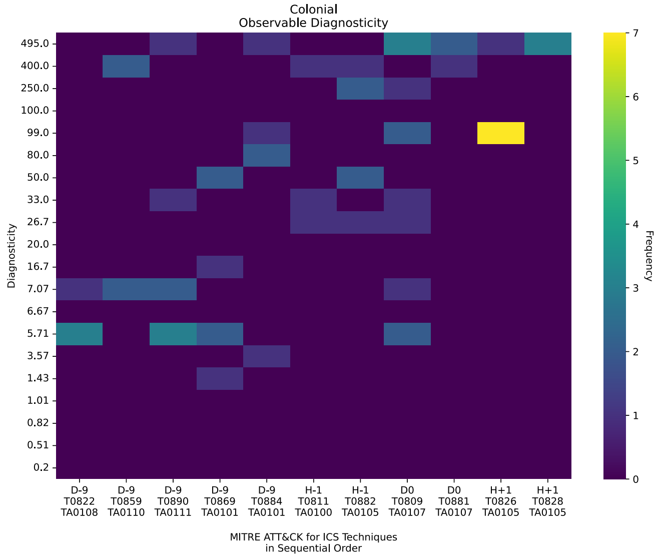


Fig. 6. Diagnosticity of the observables of the DarkSide ransomware attack on the Colonial Pipeline.

at the same time step are spread over multiple ticks on the horizontal axis. The probability of each phase is the sum of the probability of ongoing and complete behavior for that phase, and the probabilities are cumulative over all preceding observables.

For the first tick corresponding to nine days before the triggering event in Figure 7, there are significant increases in the probabilities of early and middle adversary behavior corresponding to the use of External Remote Services for Initial Access and Valid Accounts for Persistence. At the second D-9 tick, probability of middle adversary behavior increases further with the use of Exploitation for Privilege Escalation while the probability of early does not change much. The opposite is true at the third and final D-9 tick, where the probability of early adversary behavior increases for Standard Application Layer Protocol for Command and Control, while the probability of middle adversary behavior does not change much. There are significant increases in the probability of late and impact adversary behavior one hour before the triggering event when Theft of Operational Information occurs. These probabilities increase further when Loss of Availability and Loss of Productivity and Revenue occur one hour after the triggering event.

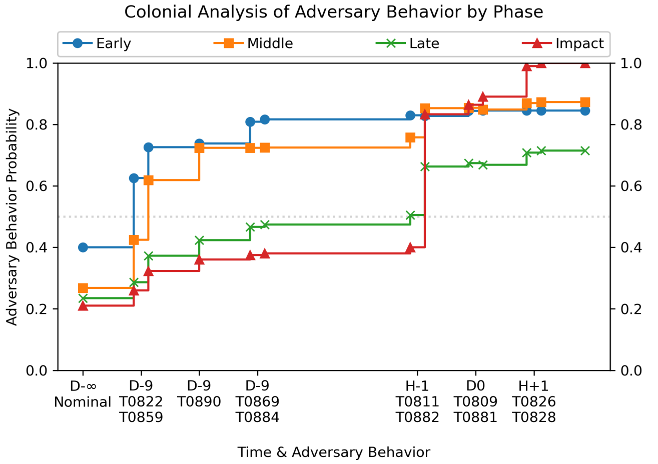


Fig. 7. Probability of each phase of adversary behavior given the observables of the DarkSide ransomware attack on the Colonial Pipeline.

B. Cyber-Attack on Thyssenkrupp Blast Furnace

In December 2014, the German Government's Federal Office for Information Security (BSI) released a report detailing a cyber-attack on a German steel mill that occurred earlier that year, though exact dates and details of the attack were not revealed [49]. While the report did not specify the name of the company, multiple sources identified the victim as one of Europe's largest steel manufacturers, Thyssenkrupp AG [50], [51]. Further, Thyssenkrupp announced on 16 May of that year that Europe's largest blast furnace, "Schwelgern 2," located at its facility in Duisburg, Germany, would be offline for several weeks for repairs and upgrades [52], [53], suggesting Schwelgern 2 was likely the target of the attack.

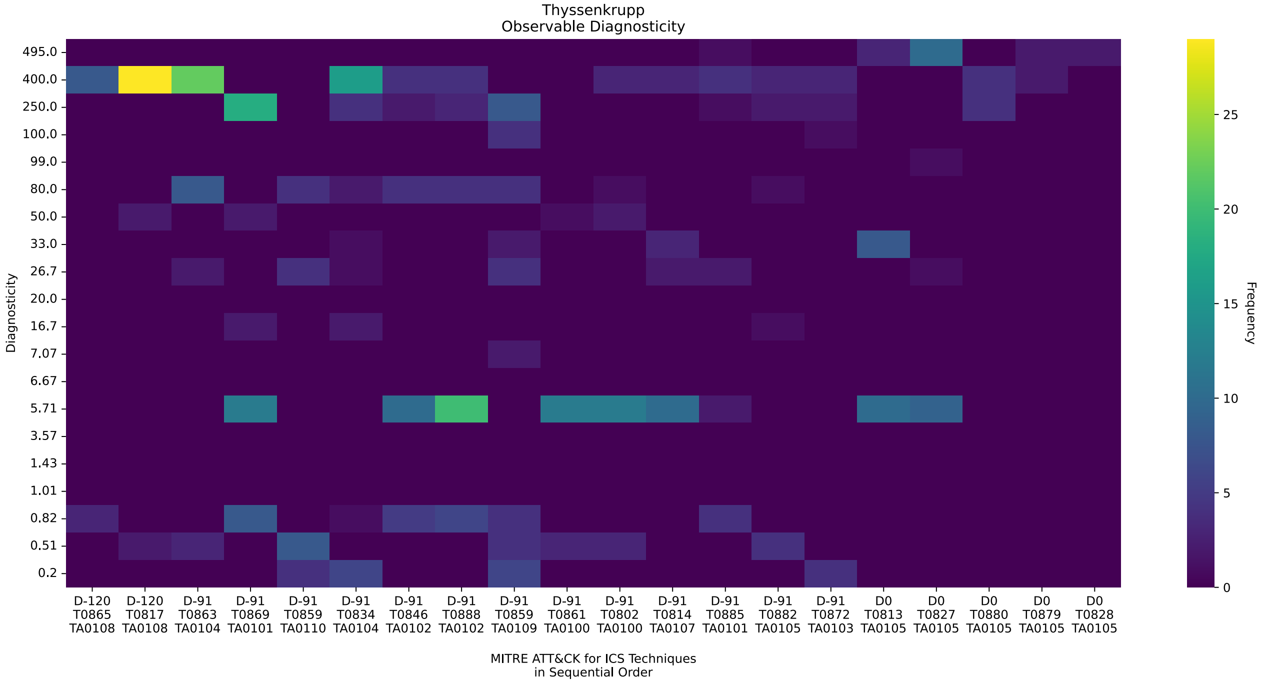


Fig. 8. Diagnosticity of the observables of the cyber-attack on the Thyssenkrupp blast furnace.

The attack began in early 2014, when adversaries infiltrated the victim's IT network via a spearphishing campaign, then worked their way into the OT environment, where they executed software that caused denial of service, denial of control, and eventually a loss of control. This led to the blast furnace shutting down without proper safety procedures, resulting in catastrophic physical damage. No lives were lost in the incident, but Thyssenkrupp suffered US$4 million in damage to the blast furnace and an additional US$6 million in lost revenue [54].

The adversaries required specialized knowledge and expertise in steel production, which enabled them to compromise a variety of internal systems and components across both IT and OT networks. The attack also demonstrated detailed knowledge of the ICS and production processes being used. This combination resulted in one of the earliest known publicly reported cybersecurity incidents resulting in physical damage to ICS equipment.

The reported observables are summarized in Table XIII and Figure 8, and the adversary behavior phase probabilities are shown in Figure 9. Figures 8 and 9 have the same format as Figures 6 and 7.

Unlike the previous case study, several of the techniques used in the early phase had observables with large diagnosticity values. These observables corresponded to anomalous links in spearphishing emails. There were many observables with moderate diagnosticity scores throughout the case study.

The cyber-attack on Thyssenkrupp begins 120 days before the triggering event with External Remote Services for Initial Access and Valid Accounts for Persistence, resulting in large increases in the probabilities of early and middle adversary behavior. The probability of early adversary behavior further increases with the User Execution technique and the probability of middle adversary behavior increases with Valid Accounts for Persistence at 91 days before the triggering event. The probabilities of all phases of adversary behavior continue to increases slightly as additional techniques are used 91 days before the triggering event, until the Theft of Operational Information technique causes a large increase in the probabilities of late and impact adversary behavior. The probability of impact adversary behavior approaches one as five impact techniques occur at the time of the triggering event.

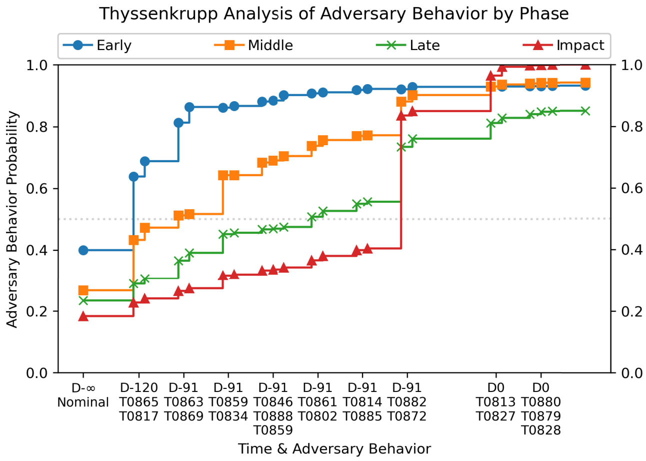


Fig. 9. Probability of each phase of adversary behavior given the observables of the cyber-attack on the Thyssenkrupp blast furnace.

VI. Conclusion

In this work, we have presented the development of a Bayesian network to calculate the likelihood of adversary behavior given common observables of a cyber-attack on an OT system. This model was developed using a combination of SME elicitations and data gathered through the analysis of 27 historical case studies of cyber-attacks affecting OT systems. The MITRE ATT&CK$^{®}$ for ICS framework provides the foundation for the model. When findings are assigned to an observable node, the evidence propagates to the corresponding technique node, then to a tactic node, and finally to the adversary behavior nodes. By implementing this approach, an OT cybersecurity team can better comprehend adversarial activity in their systems and conduct a risk-informed investigation and response.

Benefits of the approach include:

1) Proactive Estimation Capabilities: By modeling Bayesian networks according to the generic phases of a cyber-attack (Early, Middle, Late, and Impact), this approach allows for consistent adversarial behavior estimation. This proactive stance is crucial for early detection and mitigation of threats.
2) Comprehensive Threat Modeling: The integration of MITRE ATT&CK$^{®}$ for ICS Tactics, Techniques, and Observable Events with historical OT cyber-attack data provides a comprehensive framework. This rich, multi-layered model offers a more detailed understanding of potential attack vectors and their manifestations in an OT environment. In addition, this model makes no assumptions about which tactics (and techniques) will be used, and in what order they will be used.
3) Improved Anomaly Detection: The use of Bayesian networks aids in distinguishing between normal operational anomalies and those indicative of adversarial behavior. By considering the probabilistic relationships between different events and states, this approach reduces false positives, enhancing the reliability of threat detection.
4) Dynamic Risk Assessment: The approach allows for dynamic risk assessment by continuously updating the probability estimates as new evidence is collected. This real-time evaluation of the security posture ensures that the most current data is used in decision-making, enabling a more responsive defense strategy.
5) Customization and Adaptability: The proposed model can be customized to the specific characteristics of different OT environments. By incorporating observable events from historical attacks relevant to the particular setup, the model becomes more adept at recognizing threats pertinent to that environment.
6) Facilitates Strategic Decision Making: By estimating the likelihood of different stages of adversarial behavior, the approach aids in strategic decision-making. It enables security teams to prioritize resources and responses effectively, focusing on the most probable threats at any given time.
7) Enhanced Learning from Past Incidents: Utilizing historical data in the Bayesian Network model allows for a learning component where past incidents inform future detection. This continuous learning cycle ensures that the system evolves and adapts to new cyber-attack trends. This benefit also makes this approach useful for cybersecurity training exercises.
8) Effective Integration with Existing Systems: The approach can be effectively integrated with existing security systems and protocols. Leveraging the MITRE ATT&CK$^{®}$ framework ensures compatibility with widely-used security standards and practices.

Future work will examine the development of risk-informed thresholds for each of the adversary behavior phases to guide cybersecurity response. Predictive capabilities will also be explored to identify which technique or tactic the adversary might use next, given a sequence of observed techniques. Cyber-physical modeling and simulation will be a valuable tool for further refinement and testing of the model. Finally, the model will be adapted for real-time analysis. Additional implementation considerations such as integration with security information and event management (SIEM) tools and cybersecurity operations centers (CSOCs) will be addressed to enable real-time use of the Bayesian model. Seamless integration of the model with the people, processes, and technology in existing CSOCs will be essential to maximize the impact of the model in real-time applications.

Acknowledgment

The authors would like to thank Robert Brulles and Alexander Safonov of Sandia National Laboratories for their reviews.

This article has been authored by an employee of National Technology & Engineering Solutions of Sandia, LLC under Contract No. DE-NA0003525 with the U.S. Department of Energy (DOE). The employee owns all right, title, and interest in and to the article and is solely responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan https://www.energy.gov/downloads/doe-public-access-plan

This document was prepared by Idaho National Laboratory (INL) under an agreement with and funded by the U.S. Department of Energy. INL is a U.S. Department of Energy National Laboratory operated by Battelle Energy Alliance, LLC. INL/JOU-24-77873.

References

[1] Idaho National Laboratory.(2024). *Cybersecurity for the Operational Technology Environment (CyOTE)*. [Online]. Available: https://cyote.inl.gov/

[2] L. Maccarone et al., “Development of a Bayesian network to model malicious cyber-activity in operational technology environments,” in *Proc. 16th UAI Bayesian Model. Appl. Workshop*, Eindhoven, The Netherlands, Aug. 2022, pp. 1–4.

[3] K. Stouffer et al., “Guide to operational technology (OT) security,” U.S. National Institute of Standards and Technology, Tech. Rep. NIST SP 800-82r3, 2023, doi: 10.6028/NIST.SP.800-82r3.

[4] S. Santos, P. Costa, and A. Rocha, “IT/OT convergence in industry 4.0: Risks and analisy of the problems,” in *Proc. 18th Iberian Conf. Inf. Syst. Technol. (CISTI)*, Jun. 2023, pp. 1–6.

[5] I. A. Kandhro et al., “Detection of real-time malicious intrusions and attacks in IoT empowered cybersecurity infrastructures,” *IEEE Access*, vol. 11, pp. 9136–9148, 2023.

[6] G. Murino, M. Ribaudo, S. P. Romano, A. Tacchella, A. Armando, and M. Colajanni, “OT cyber security frameworks comparison tool (CSFCTool),” in *Proc. ITASEC*, 2021, pp. 9–22.

[7] N. G. Romps. (2021). *Protecting Critical Infrastructure From Cyber Threats*. Accessed: Mar. 1, 2022. [Online]. Available: https://www.mitre.org/news-insights/impact-story/protecting-critical-in%frastructure-cyber-threats#:(textasciitilde):text=As%20adversaries%20grow%20more%20sophistica%ted,MITRE%20released%20ATT%26CK%20for%20ICS

[8] L. Patera, A. Garbugli, A. Bujari, D. Scotece, and A. Corradi, “A layered middleware for OT/IT convergence to empower industry 5.0 applications,” *Sensors*, vol. 22, no. 1, p. 190, Dec. 2021. [Online]. Available: https://www.mdpi.com/1424-8220/22/1/190

[9] G. A. Weaver and D. Gunter, “Language-theoretic data analysis to support ICS protocol baselining,” in *Proc. 44th IEEE Symp. Secur. Privacy*, Mar. 2023, pp. 1–5.

[10] Y. Xin et al., “Machine learning and deep learning methods for cybersecurity,” *IEEE Access*, vol. 6, pp. 35365–35381, 2018.

[11] A. Handa, A. Sharma, and S. K. Shukla, “Machine learning in cybersecurity: A review,” *WIREs Data Mining Knowl. Discovery*, vol. 9, no. 4, p. 1306, 2019. [Online]. Available: https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.1306

[12] The MITRE Corporation.(2022). *ICS Matrix*. [Online]. Available: https://attack.mitre.org/matrices/ics/

[13] The MITRE Corporation.(2023). *Enterprise Matrix*. [Online]. Available: https://attack.mitre.org/matrices/enterprise/

[14] The MITRE Corporation.(2023). *Mobile Matrix*. [Online]. Available: https://attack.mitre.org/matrices/mobile/

[15] The MITRE Corporation.(2023). *D3FEND*. [Online]. Available: https://d3fend.mitre.org/

[16] S. Mahoney, D. Buede, and J. Tatman, “Patterns of report relevance,” in *Proc. 3rd UAI Bayesian Model. Appl. Workshop*, Edinburgh, U.K., 2005, pp. 1–12.

[17] D. Roth, “On the hardness of approximate reasoning,” *Artif. Intell.*, vol. 82, nos. 1–2, pp. 273–302, Apr. 1996. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0004370294000921

[18] F. V. Jensen, *Introduction to Bayesian Networks*. New York, NY, USA: Springer, 1997.

[19] K. B. Korb, *Bayesian Artificial Intelligence*, 2nd ed., Boca Raton, FL, USA: CRC Press, 2010.

[20] R. E. Neapolitan, *Learning Bayesian Networks*. London, U.K.: Pearson, 2019.

[21] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed., Upper Saddle River, NJ, USA: Prentice-Hall, 2010.

[22] H. Al-Mohannadi, Q. Mirza, A. Namanya, I. Awan, A. Cullen, and J. Disso, “Cyber-attack modeling analysis techniques: An overview,” in *Proc. IEEE Int. Conf. Future Internet Things Cloud Workshops (FiCloudW)*, Mar. 2016, pp. 69–76, doi: 10.1109/W-FICLOUD.2016.29.

[23] A. A. Ghorbani, W. Lu, and M. Tavallaee, *Network Intrusion Detection and Prevention*(Advances in Information Security), vol. 47. New York, NY, USA: Springer, 2010, doi: 10.1007/978-0-387-88771-5.

[24] B. Kordy, M. Pouly, and P. Schweitzer, “A probabilistic framework for security scenarios with dependent actions,” in *Integrated Formal Methods*, E. Albert and E. Sekerinski, Eds., Cham, Switzerland: Springer, 2014, pp. 256–271.

[25] K. Ingols, M. Chu, R. Lippmann, S. Webster, and S. Boyer, “Modeling modern network attacks and countermeasures using attack graphs,” in *Proc. Annu. Comput. Secur. Appl. Conf.*, Dec. 2009, pp. 117–126, doi: 10.1109/ACSAC.2009.21.

[26] I. Kotenko and A. Chechulin, “A cyber attack modeling and impact assessment framework,” in *Proc. 5th Int. Conf. Cyber Conflict (CYCON)*, Tallinn, Estonia, Jun. 2013, pp. 1–24.

[27] K. Huang, C. Zhou, Y.-C. Tian, W. Tu, and Y. Peng, “Application of Bayesian network to data-driven cyber-security risk assessment in SCADA networks,” in *Proc. 27th Int. Telecommun. Netw. Appl. Conf. (ITNAC)*, Nov. 2017, pp. 1–6.

[28] S. Chockalingam, W. Pieters, A. Teixeira, and P. van Gelder, “Bayesian network models in cyber security: A systematic review,” in *Secure IT Systems*, H. Lipmaa, A. Mitrokotsa, and R. Matulevičius, Eds., Cham, Switzerland: Springer, 2017, pp. 105–122.

[29] J. Wang, M. Neil, and N. Fenton, “A Bayesian network approach for cybersecurity risk assessment implementing and extending the FAIR model,” *Comput. Secur.*, vol. 89, Feb. 2020, Art. no. 101659.

[30] A. Zimba, H. Chen, and Z. Wang, “Bayesian network based weighted APT attack paths modeling in cloud computing,” *Future Gener. Comput. Syst.*, vol. 96, pp. 525–537, Jul. 2019.

[31] M. Frigault, L. Wang, A. Singhal, and S. Jajodia, “Measuring network security using dynamic Bayesian network,” in *Proc. 4th ACM Workshop Quality Protection*, Oct. 2008, pp. 23–30, doi: 10.1145/1456362.1456368.

[32] T.-H.-G. Vu, T.-H. Hoang, and M.-T. Nguyen, “Assessing Web security risks using dynamic Bayesian network,” in *Proc. 11th Int. Symp. Inf. Commun. Technol.*, Hanoi, Vietnam, Dec. 2022, pp. 165–172, doi: 10.1145/3568562.3568591.

[33] P. K. Vaddi et al., “Dynamic Bayesian networks based abnormal event classifier for nuclear power plants in case of cyber security threats,” *Prog. Nucl. Energy*, vol. 128, Oct. 2020, Art. no. 103479.

[34] M. J. Assante and R. M. Lee, “The industrial control system cyber kill chain,” The SANS Inst., North Bethesda, MD, USA, Tech. Rep. 36297, 2015.

[35] C. Johnson, L. Badger, D. Waltermire, J. Snyder, and C. Skorupka, “Guide to cyber threat information sharing,” U.S. National Institute of Standards and Technology, Gaithersburg, MD, USA, Tech. Rep. NIST SP 800-150, 2016, doi: 10.6028/NIST.SP.800-150.

[36] K. Baker, “IOA vs IOC,” CrowdStrike, Austin, TX, USA, 2022.

[37] J. C. Grady, S. X. Wen, L. T. Maccarone, and S. T. Bowman, “Statistical methods for developing cybersecurity response thresholds for operational technology systems using historical data,” in *Proc. IEEE 6th Int. Conf. Trust, Privacy Secur. Intell. Syst., Appl. (TPS-ISA)*, Oct. 2024, pp. 549–554, doi: 10.1109/TPS-ISA62245.2024.00075.

[38] A. L. Madsen and F. V. Jensen, “Lazy propagation: A junction tree inference algorithm based on lazy evaluation,” *Artif. Intell.*, vol. 113, nos. 1–2, pp. 203–245, Sep. 1990.

[39] The United States Senate Committee on Homeland Security & Governmental Affairs. (2021). *Testimony of Joseph Blount, President and Chief Executive Officer Colonial Pipeline Company*. Accessed: Mar. 1, 2022. [Online]. Available: https://www.hsgac.senate.gov/imo/media/doc/Testimony-Blount-2021-06-08.%pdf

[40] W. Turton and K. Mehrotra, “Hackers breached colonial pipeline using compromised password,” Bloomberg, New York, NY, USA, 2021.

[41] Colonial Pipeline.(2021). *A Message to Our Customers and Those Who Depend on Us*. [Online]. Available: https://www.colpipe.com/safe-operations/cybersecurity-response?gclid=Cj%0KCQjwheFBhDZARIsALHjIKcFyX6srFFwyDhqviGH4ViV05Ib7YG0b2ZazA1NCijn6ivM7ezYaqcaA%swHEALwwcB

[42] D. E. Sanger, C. Krauss, and N. Perlroth, “Cyberattack forces a shutdown of a top U.S. pipeline operator,” New York Times, New York, NY, USA, 2021.

[43] B. Krebs, “A closer look at the DarkSide ransomware gang,” Krebs on Security, Merrifield, VA, USA, 2021.

[44] *DarkSide Ransomware Gang Quits After Servers, Bitcoin Stash Seized*, The Hacker News, New York, NY, USA, 2021.

[45] J. Tidy, “Colonial hack: How did cyber-attackers shut off pipeline?,” BBC News, London, U.K., 2021.

[46] *Ransomware Attack Shuts Down Colonial Pipeline*, Check Point Software, Tel Aviv-Yafo, Israel, 2021.

[47] J. Goodchild, “The Colonial Pipeline cyberattack: A comprehensive timeline,” CSO Online, New York, NY, USA, 2021.

[48] J. Koetsier, “Colonial Pipeline hack explained: Everything you need to know,” CNET, San Francisco, CA, USA, 2021.

[49] *The State of IT Security in Germany 2014*, Federal Office for Information Security, Bonn, Germany, 2014.

[50] M. Riley and J. Robertson, “Cyberspace becomes second front in Russia’s clash with NATO,” Bloomberg, New York, NY, USA, 2015. Accessed: Dec. 10, 2022.

[51] C. Wiener, “Penetrate, exploit, disrupt, destroy: The rise of computer network operations as a major military innovation,” Ph.D. dissertation, George Mason Univ., Fairfax County, VA, USA, 2016.

[52] Thyssenkrupp.(2014). *First Campaign Ends After 21 Years: Europe’s Biggest Blast Furnace to Be Modernized*. [Online]. Available: https://www.thyssenkrupp.com/en/newsroom/press-releases/first-campaign-%ends-after-21-years–europe-s-biggest-blast-furnace-to-be-modernized-3303.html

[53] (2014). *Europe’s Biggest Blast Furnace Relit: ‘Schwelgern 2’ Producing Iron Again*. Accessed: Oct. 14, 2022. [Online]. Available: https://www.thyssenkrupp-steel.com/en/newsroom/press-releases/europes-biggest-blast-furnace-relit-schwelgern-2-producing-iron-again.html

[54] S. S. Buchanan, “Cyber-attacks to industrial control systems since Stuxnet: A systematic review,” Ph.D. dissertation, Capitol Technology University, Laurel, MD, USA, 2022.

Lee T. Maccarone received the B.S. and Ph.D. degrees in mechanical engineering from the University of Pittsburgh in 2016 and 2021, respectively.

He joined Sandia National Laboratories as a Post-Doctoral Appointee in 2021, and has been a Principal Cybersecurity Research and Development Engineer in the Energy Security Department since 2024. His research interests include the development of security-by-design methodologies and the development of analytical methods to enhance security decision-making.

Dr. Maccarone was a U.S. Department of Energy Nuclear Energy University Program Graduate Fellow and was the recipient of the Second Place Award in the Energy Policy Category of the Innovations in Nuclear Technology Research and Development Program sponsored by the U.S. Department of Energy Office of Nuclear Technology Research and Development. He received the Graduate Certificate in Nuclear Engineering at the University of Pittsburgh in 2021.

Dennis M. Buede received the B.S. degree in aerospace engineering from the University of Cincinnati in 1971 and the M.S. and Ph.D. degrees in engineering-economic systems from Stanford University in 1973 and 1977, respectively.

He has over 45 years of experience in both the theoretical development and engineering application of systems engineering, probabilistic analysis and forecasting, and decision support technologies. He was a Professor at the Stevens Institute of Technology and George Mason University. He is currently an Analytics Specialist at ITA International. He has authored the first two editions of *The Engineering Design of Systems: Models and Methods* (and co-authored the third edition). He has authored or co-authored numerous books and professional articles. He belongs to the Institute for Operations Research and Management Science, the Institute of Electronic and Electrical Engineers, and the International Council on Systems Engineering.

Dr. Buede is a fellow of INCOSE.

Scott T. Bowman received the B.A. degree in international relations and Russian language from the University of Oklahoma in 2007 and the M.S. degree in information management and security from Syracuse University in 2021.

From 2021 to 2023, he was a Control System Cybersecurity Analyst, focusing on analyzing adversarial cyber behavior patterns for protecting critical infrastructure from emerging cyber threats. In 2023, he transitioned to a new role as a Technical Lead for Cyber-Physical Systems Modeling and Simulation at the Idaho National Laboratory. This position allows him to leverage his extensive knowledge in cybersecurity to model complex challenges in the field of cyber-physical systems.

Mr. Bowman's work has been recognized in the academic community, notably through his contribution to an award-winning paper on Bayesian Modeling of Cyber Attack Behavior in Operational Technology, presented at the Uncertainty in AI Conference. He received the Graduate Certificates in Industrial Control System Cybersecurity and Digital Forensics and Response from the SANS Technology Institute in 2022.

Pawel Ambrozewicz received the Ph.D. in particle physics from Temple University, Philadelphia, in 2002.

He has over 20 years of experience in data analysis and analytics, Monte Carlo simulations, programming, and experimental particle physics. His experience stems from holding research positions at Florida International University, North Carolina A&T State University, and Jefferson Lab. He was also involved in medical physics research aiming at improving proton therapy precision and effectiveness, and has extensive physics and statistics teaching experience. He holds a patent that resulted from work on shielding proton therapy facilities.

Charles D. Burdick received the B.S. degree in psychology from Rensselaer Polytechnic Institute (RPI), Troy, NY, USA, in 1966, and the M.S.A. degree in operations research and systems analysis from George Washington University, Washington, DC, USA, in 1974.

Following his seven years of active military service as an Army Military Intelligence Officer, he has over 40 years of experience in modeling and simulation of military systems, operations, and communications. He has applied these simulations to analytical, training, and testing applications for the services, defense agencies, and the intelligence community. He is currently a Principal C5ISR Analyst at ITA International, and previously led simulation developments and analysis at Lockheed Martin and BDM International. He has authored or co-authored numerous articles, reports, and presentations on simulation and communications network emulations. He is an Institute for Operations Research and Management Science Certified Analytics Professional (CAP) and a member and regular contributor of presentations to the Military Operations Research Society Symposiums, and has presented in several other modeling and simulation forums.

Mr. Burdick is an honor graduate of the Army Command and General Staff College and the National Defense University.

J. Connor Grady received the B.S. degree in mathematics and physics from the University of California, Santa Barbara (UCSB) in 2018 and the M.S. degree in mathematics from the University of Illinois Urbana–Champaign (UIUC) in 2021, where he is currently pursuing the Ph.D. degree in mathematics.

He began as a Graduate Research and Development Intern at Sandia National Laboratories in 2024. His research at Sandia has been focused on applying mathematical techniques to the analysis of vulnerabilities and risks in energy and critical infrastructure systems.

Shaw X. Wen received the B.S. degree in computer science and business administration with an emphasis in finance from California State Polytechnic University in 2010 and the M.S. degree in data science from Northeastern University, Boston, in 2021.

She worked mostly in finance and accounting prior to joining the Idaho National Laboratory, as a Data Scientist. After joining the Idaho National Laboratory, as a Data Scientist, she has worked on several nuclear proliferation projects where a combination of statistical and machine learning tools was used to detect potential misuse of nuclear materials. Additionally, she has also worked on natural language processing projects, one of which involved the use of topic modeling to extract topics from facility condition reports to identify concerning patterns in commonly reported issues, while another involved the fine-tuning of pre-trained large language models to automate SQL database querying.