A Bayesian Network Approach for the Interpretation of Cyber Attacks to Power Systems*

Davide Cerotti[1], Daniele Codetta-Raiteri[1], Giovanna Dondossola[2], Lavinia Egidi[1], Giuliana Franceschinis[1], Luigi Portinale[1] and Roberta Terruggia[2]

[1] Computer Science Institute, DiSIT, Univ. of Piemonte Orientale, Alessandria, Italy
{davide.cerotti, daniele.codetta, lavinia.egidi, giuliana.franceschinis, luigi.portinale}@uniupo.it
[2] RSE: Ricerca Sistema Energetico, Milano, Italy
{giovanna.dondossola, roberta.terruggia}@rse-web.it

Abstract. The focus of this paper is on the analysis of the cyber security resilience of digital infrastructures deployed by power grids, internationally recognized as a priority since several recent cyber attacks targeted energy systems and in particular the power service. In response to the regulatory framework, this paper presents an analysis approach based on the Bayesian Networks formalism and on real world threat scenarios. Our approach enables analyses oriented to planning of security measures and monitoring, and to forecasting of adversarial behaviours.

1 Introduction

One of the key challenges for cyber security experts in charge of defending digital infrastructures from malicious activities is the capability to anticipate the behaviour of their offensive counterparts. This permits to reinforce the defense measures before the occurrence of the attack, and to increase the resilience of cyber-physical systems such as those used for energy operations. In order to provide the desired capability, current research programs are developing new tools for forecasting the attack probability and for the early detection of malicious activities.

The focus of this paper is on the analysis of the cyber security resilience of digital infrastructures deployed by power grids [22]. This issue is internationally recognized as a priority since in the last years several cyber attacks targeted energy systems and in particular the power service. Stuxnet, Dragonfly, Black Energy and Industroyer are among the threats that affected industrial control systems. The European member states have recently implemented the EU Network and Information Security (NIS) Directive 2016/1148 regulating the cyber security space of essential service operators.

*This work is original and has a financial support of the Università del Piemonte Orientale and of the Research Fund for the Italian Electrical System in compliance with the Decree of Minister of Economical Development April 16, 2018

In response to the regulatory framework, this paper presents a methodology developed in a just started research project that can be used both for selecting the meaningful evidences of malicious actions and for investigating the causes of attack effects. A probabilistic approach based on the Bayesian Networks formalism has been selected as the best tool able to capture the various dependencies of a very flexible cyber security model.

Our proposal enables different kinds of analyses. It can be used to spot vulnerabilities in the network that are most likely to lead to a compromise of significant impact, in order to correctly plan security measures. Also, it can be used in evaluating the most appropriate monitoring configuration for detection of intrusions, by evaluating the sensors that are most likely to lead to the exposure of significant threats. Finally, it is a tool that helps to forecast adversarial moves, being able to deduce from observed behaviour the most likely threats.

The focus of this paper is on methodology but we also contribute an analysis of attack patterns in the context of smart grids, with specific attention to the way an intruder moves from the corporate part of the network to the more sensible operational area. We also analyze a possible attack pattern directed to disruption of grid functionality.

Our analysis combines its theoretical value with real-world data. Our attack models are founded on a knowledge base of attack techniques [27], derived breaking down publicly reported incidents. Moreover probabilities of frequencies for each attack step are derived from the same project [27] and from the the US CERT-ORG vulnerabilities bulletin [9].

We assign a great value to early detection of attackers' movements and therefore we integrate in our model monitoring aspects, once again derived from a real-world project, [26].

The structure of the paper is the following: Section 2 presents the context of the power system digital architecture, then Section 3 explains the modelling methodology, and details the attack model. Section 4 illustrates some examples of analyses that can be carried out. Finally Section 5 presents the conclusions and some future works.

1.1 Related work

In the last twenty years several probabilistic approaches have been used for developing cyber security assessment tools. This session discusses the approaches related to the proposed methodology and clarifies the novelty of our contribution.

In our analysis we based estimates of the frequencies of attacks on data extracted from the "Adversarial Tactics, Techniques and Common Knowledge" (ATT&CK) enterprise database from a MITRE project [21, 27], a knowledge base compiled analyzing reports on real attacks.

A new knowledge base with data specific to Industrial Control System (ICS) security breaches has been recently announced by MITRE [1] but has not been made publicly available yet. Therefore, for estimates of frequencies of specific power grid attacks we resorted to the Common Vulnerability Scoring System

(CVSS) [5], that provides a way to express on a 0-10 scale the severity of vulnerabilities. See [4] and [11] for analyses and proposals on the use of CVSS based metrics.

In parallel to the ATT&CK adversarial model, and based on its tactics and analytics, MITRE has compiled the Cyber Analytics Repository (CAR) [26], a knowledge base focused on detection of ongoing attacks.

Bayesian Networks (BNs) are a widely used formalism for representing uncertain knowledge in probabilistic systems, applied to a variety of real-world and complex problems [15]. The adoption of BNs for security modeling has been advocated by several researchers [13, 16, 29, 2, 30, 7]. Such approaches start from attack graph models to show how BNs can be derived, stressing the evidence-based analysis allowed by such a formalism. In particular, in [7], attack graphs and BNs are used to derive an overall security measure of a network. Attack graphs are used to represent the causal relationships among attacks. Attack probabilities are obtained using directly CVSS base scores, and in addition cumulative probabilities are defined as a way to reflect in the metrics the causal relationship between vulnerabilities. From the attack graph the authors derive BNs in order to have probabilities that reflect interdependencies of events. In contrast, we base most of our quantitative data on real world data from the MITRE ATT&CK project, and model also detection mechanisms. We show that our model allows different kinds of security analyses of the network to tackle both planning and detection issues.

In [23] the authors build a BN on top of a graph that captures the security dependencies of objects in the system, taking into account also evolution in time. The aim of this research is security analysis of the network, and their specific interest is identifying attack paths that make use of zero-day vulnerabilities. In [19] the authors present a methodology for the cybersecurity analysis of Information and Communication Technology (ICT) infrastructures. A predefined modelling language is used to model ICT components, attack steps and defences. From the model a large number of attack/defence graphs in the form of Bayesian-like networks are automatically generated and the success rate of a given attack step associated to a given architectural asset is evaluated as a function of the Time to Compromise. Although the automatic generation of attack probabilities may simplify the work of the analysts, the built-in attack steps and their logical dependencies make it difficult to validate the underlying model and the analysis results.

The new research [17] is related to our work since, as our approach, its object is the security of a power grid network and it is based on the MITRE ICS ATT&CK matrix. Their model is based on MITRE techniques and analytics, and their goal is to determine the security level of the network in a time interval.

Other research targets in particular security of the smart grid or of ICS, with an approach different from ours: in [12] the authors propose to use ensemble classifiers to learn the (low entropy) benign traffic in a smart grid in order to later detect anomalies; [8] introduces an attack chain specifically designed for IIoT environments, that takes into account the multilevel architecture of an ICS, and propose the use of machine learning classification techniques to map security alerts to different phases of the attacks; in [28] graphs are used to model cybersecurity properties on the smart grid and evaluate the effectiveness of security mechanisms. In [18], the attention is restricted to attack detection in the Advanced Metering Infrastructure; the study is based on the ADVISE tool (see references in [18]) which allows to determine, based on a mechanism of costs and rewards for the attacker, the most likely attack path in a given time interval, along with probability of detection of the attack, and distinguishes between different kinds of attackers. The approach is different, since it considers detailed information on different classes of adversarial profiles, but it might in future be integrated in our methodology.

2 IT/OT Architecture Description: IT/OT covergence

The evolution of the power grid towards a digitalized infrastructure requires to address the cyber security aspects with paramount attention. The power system architecture evolves from isolated and proprietary systems with physical separation of the areas to a layered architecture where components placed in different domains have to interact. It is composed by several areas covering different domains coming from the business and corporate area moving to control

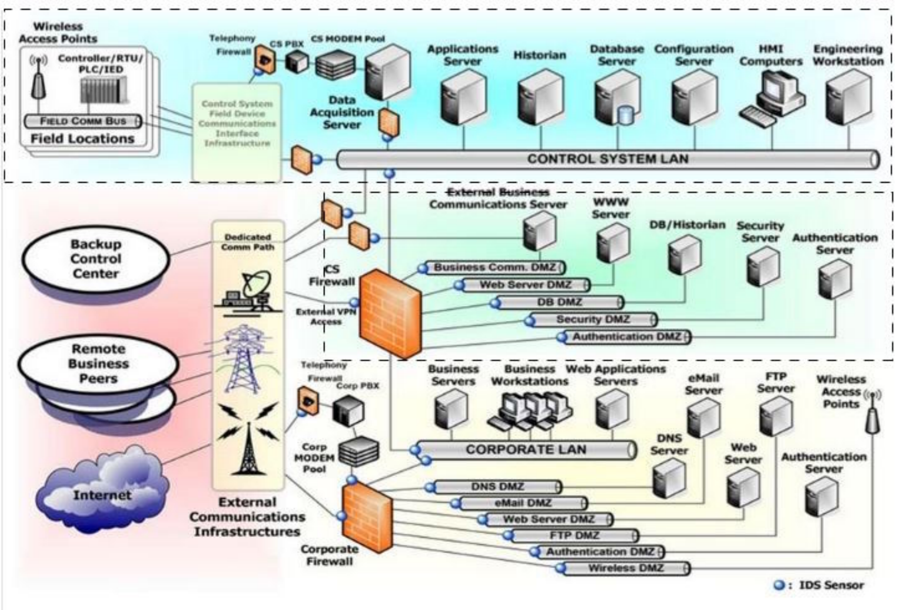

Fig. 1. IT/OT architecture NIST 800-82 [20]

and process areas. Security characteristics and requirements vary across environments, thus calling for a segmentation in different trust levels, called security domains.

A reference model commonly used to represent the main layers of an ICS is the Purdue model [14]. The components within each layer communicate by means of local networks, and the communications between layers exploit possibly heterogeneous technologies implementing local as well as wide area networks. In some cases the latter are public connections deploying heterogeneous levels of cyber security measures. Networks linked to other networks may introduce common vulnerabilities spanning over multiple domains.

This paper addresses the analysis of cyber attack processes targeting an IT/OT power system ICT architecture as presented in Figure 1. The architecture comprises Information Technology (IT) and Operational Technology (OT) components. It is not possible to clearly split the energy architecture in IT and OT partitions since in the real world IT devices must interact with OT components. In the same way, the cybersecurity analysis and implementations need to address both IT and OT aspects in a comprehensive evaluation. So, starting from the architecture in Figure 1, our analysis addresses a cyberattack process exploiting the corporate area as entry point, passing through the different networks and reaching the control and process environments. Indeed most ICS kill chains of real cyber attacks occurred in the last years, targeting energy organizations, and in particular electrical utilities, initiated from the IT network, more specifically from the office/corporate network, and through the inter-domain connections reached the OT control network.

The focus of our analysis is on attack processes exploiting the security breaches of the demilitarized zone (DMZ) between the enterprise network (IT) and the control network (OT). The DMZ is a basic cybersecurity measure necessary in order to segregate two environments with very different cyber security requirements and implemented security measures, but this is also a weak point used by the attackers to reach the operational area. Our analysis covers the areas enclosed in dotted rectangles in Fig. 1 (the DMZ being marked by the lower rectangle in the figure and the control area by the upper one) hosting the IT and OT components employed in power monitoring and control activities.

3 Modeling

3.1 Methodology and tools

The information of the ATT&CK knowledge base [21, 27] is presented in a matrix whose columns are labelled by *tactics* whereas the entries in each column are *techniques*. A tactic can be viewed as a goal that the attacker is trying to attain (e.g. Initial Access, Privilege Escalation, Lateral Movement, etc.); a technique defines the means by which the attacker tries to attain its goal: for instance Initial Access can be attained using Valid Accounts or a Spearphishing Link; on the other hand Valid Accounts can be used also to attain Privilege Escalation.

Each technique is documented with a description, recommendations for mitigation and detection, and a list of references. In addition, for each technique MITRE provides a list of actor groups that are known to have applied the technique, and a list of software (both malware and pen-testing tools) that make use of it. MITRE also proposes a way of scoring the "popularity" of techniques in the ATT&CK matrix [3], that can be visualized loading a *json* layer [24] on the ATT&CK navigator [25]. The score is based on how often it has been reported (i.e. the number of references in which it appears), how many groups have used it and how many different software tools apply it. Notwithstanding the limitations due to the lack of public accurate data on all attacks and to the error-prone procedure of analyzing by hand existing reports, it is very attractive being able to base the analysis on real-world data. Therefore we adopted the ATT&CK scores as probabilities of technique occurrence, normalizing the scores in 0-1.

Since our focus is on the way a security breach can propagate to the OT environment of a smart grid and there wreak havoc, we also used in our attack graph techniques that are proper to the ICS world. Since the ICS ATT&CK matrix is not yet available, in this case we derived our scores from the US Department of Homeland Security's ICS-CERT Advisories [9], selecting advisories on software that is used in the electricity sector, limiting the analysis to advisories released or revised in year 2015 or later and to those for which a Common Vulnerability Scoring System (CVSS) [5] score was provided. We processed the scores in order to obtain metrics that reflect more accurately occurrence probabilities, i.e. focusing on their *Exploitability* components.

In order to model aspects related to detection, we integrated in our model also analytics from the Cyber Analytics Repository (CAR) [26]. An analytic describes events whose observation is significant to the security analyst. Among all analytics listed, some are strong indications that a certain technique is being deployed by an attacker (type "Tactic, Technique, Procedure" (TPP) analytics), some are more general symptoms of a behavior not observed normally (type "Anomaly"), others can be seen as indicators of the general health of the network (type "Situational Awareness"). Monitoring these events enables to raise alerts of varying levels of criticality. One single analytic is of type "Forensic" which means that it is specifically useful as an investigation tool rather than to raise alerts (not of interest to us and thus not used in the present context). Each analytic is connected to the techniques (with a single exception) that the analytic helps to expose.

3.2 Attack graphs

The typical attacks targeting energy and utility networks, comprise multistage processes where the first phases rely on stealing administrative credentials. The foothold in enterprise networks and the theft of administrative credentials allow the intruders to exploit administrative protocols and connections, to move laterally and obtain information about the process environment. Then the remote attacker is able to traverse the network and to perform malicious actions targeting most critical areas of the power digital infrastructure.

The attack scenario under exam assumes that a workstation in the corporate network has been already compromised. From there, the intruder attacks a host located in the DMZ and then concentrates its efforts on a second host, in the OT section. Once the latter host is compromised, the target of the attack may be a service of the ICS subsystem. The attack graphs in Figs. 2 and 3 depict possible attack paths: nodes represent states and edges transitions between states that are enabled by the exploitation of some technique; the label of each edge is a pair *tactic/technique* describing the specific ATT&CK *technique* that enables the transition, as a vehicle for realizing the specified *tactic*.

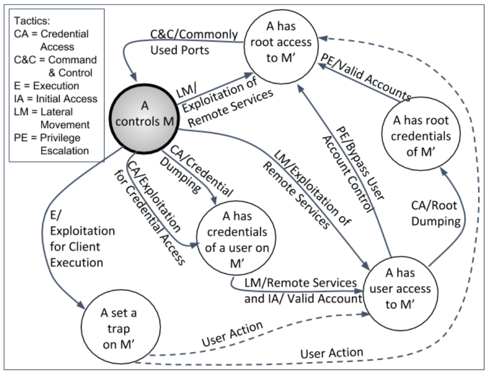
Tactics:
CA = Credential Access
C&C = Command & Control
E = Execution
IA = Initial Access
LM = Lateral Movement
PE = Privilege Escalation
C&C/Commonly Used Ports
A has root access to M′
A controls M
LM/ Exploitation of Remote Services
PE/Valid Accounts
A has root credentials of M′
CA/Credential Dumping
CA/Exploitation for Credential Access
LM/Exploitation of Remote Services
PE/Bypass User Account Control
CA/Root Dumping
E/ Exploitation for Client Execution
A has credentials of a user on M′
A has user access to M′
LM/Remote Services and IA/ Valid Account
A set a trap on M′
User Action
User Action

Fig. 2. Lateral movement attack graph.

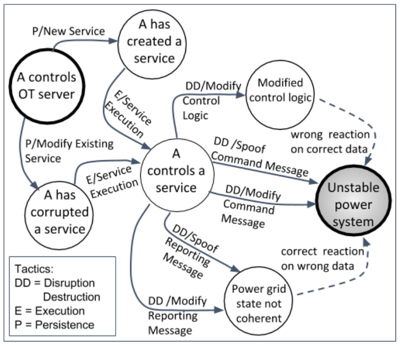
P/New Service
A has created a service
A controls OT server
E/Service Execution
DD/Modify Control Logic
Modified control logic
wrong reaction on correct data
P/Modify Existing Service
A controls a service
DD /Spoof Command Message
Unstable power system
E/Service Execution
A has corrupted a service
DD/Modify Command Message
DD/Spoof Reporting Message
correct reaction on wrong data
Tactics:
DD = Disruption Destruction
E = Execution
P = Persistence
DD /Modify Reporting Message
Power grid state not coherent

Fig. 3. ICS service attack graph.

The multigraph of Fig. 2 is cyclic and describes the lateral movement from one host to the next one. The shaded state represents both the initial state in which the attacker has control of a generic host in the network and the goal when, having followed one of the available paths, the attacker remotely controls another machine in the network. The intruder may directly obtain root privileges through the *Lateral Movement* tactic by means of the *Exploitation of Remote Services* technique, and then can establish its *Command and Control* channel using the *Commonly Used Ports* technique. Otherwise, she may gain first a normal user access, and then escalate privileges. In this case, normal user access can be achieved with three different attack patterns: one is via a different usage of *Lateral Movement/Exploitation of Remote Services*. The other two patterns both achieve *Credential Access* either via *Credential Dumping* or *Exploitation for Credential Access*; then the *Lateral Movement* is completed using the *Remote Services* technique combined with the use of a *Valid Account* that provides *Initial Access* to the victim. Finally, from the initial state, if the foothold of the attacker is a server, the attacker can exploit a vulnerability on the server (*Execution/Exploitation for Client Execution*) and wait for a user (or superuser) to fall in the trap and offer her access to the victim as user (or root), as represented by dotted arcs in the figure. If the attacker has gained user access,

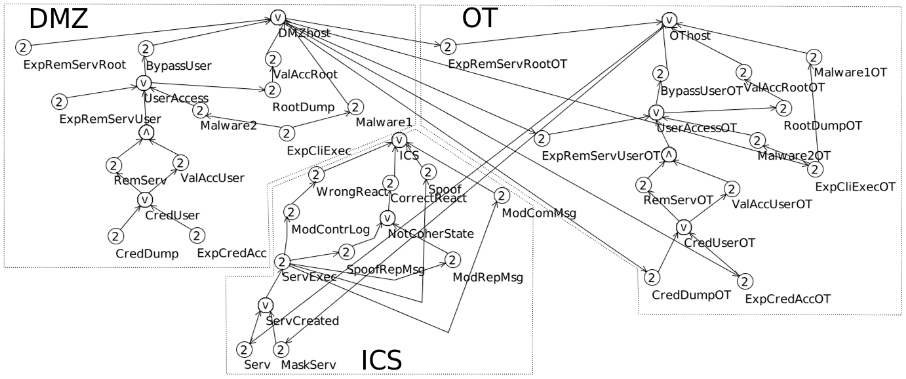

Fig. 4. BN model of the attack paths

then she will escalate privileges using *Privilege Escalation/Bypass User Account Control* or via a *Credential Access/Credential Dumping* followed by a *Privilege Escalation/Valid Accounts*. In all cases, the attacker concludes this phase of the attack establishing *Command and Control*. So the attack graph of Fig. 2 models a generic lateral movement and depending on the services possibly running on the initial host or on the next victim, and on the overall setting, the attacker can follow one or more attack paths. We will model an attack pattern in which the attacker performs two lateral movements (two complete cycles in this attack graph) from its initial foothold on a workstation in the corporate network, to a server on the DMZ, and then to a server in the OT network.

The multigraph of Fig. 3 assumes that the attacker has control of a relevant server on the OT network ("A controls OT server"). The goal is making the system unstable (shaded node in the graph). In the scenario depicted, she can take advantage of her position and either install a *New Service* or *Modify an Existing Service* (*Persistence* tactic) and then execute it (*Execution/Service Execution* technique). The new service can be used to obtain *Disruption* or *Destruction*, using one of the techniques: *Spoof Command Message* or *Spoof Reporting Message*; an existing service can be used for the same tactic using either *Modify Command Message*, *Modify Control Logic* or *Modify Reporting Message*, depending on the specific setting. These techniques can lead directly to an "Unstable power system", the final goal of the attacker. Or they can lead to a "Power grid state not coherent" in which case a correct reaction of the system to wrong data leads to the final goal. When the technique *Modify Control Logic* is used, then data are still correct, but the reaction of the system is wrong, again leading to the final attack goal.

3.3 Bayesian Network model

A Bayesian Network (BN) is a probabilistic graphical model where a Directed Acyclic Graph is used to connect discrete random variables which are dependent

[10, 15]. The joint probability of the model can be obtained through a set of conditional distributions which are locally specified at each variable node, with respect to the set of parent nodes (i.e. the joint distribution is factorized by means of the product of such local distributions). Because of that, any desired conditional probability in the model can be computed, meaning that we can ask for the probability of any variable (or set of variables), given that other variables are observed (set to a specific state). We call the observed variable states the *evidence*. From the attack scenarios depicted in Figs. 2 and 3, a BN can be derived, as shown in Fig. 4 where nodes correspond to the involved techniques and the states of the system, while arcs correspond to dependency relations that take the form of cause-effect relations (if A causes or produces B, then an arc is drawn from node A to node B). In our model, each node is binary: value 0 (false) indicates that the state or the technique has not occurred; value 1 (true) indicates the occurrence of the state or the technique. In Fig. 4 three subnets can be identified and represent the attack paths to the DMZ host, to the OT host, and to the ICS service, respectively. The first two represent instances of attacks that follow paths from the attack graph in Fig. 2. The third one represents an instance of the attack graph in Fig. 3. We explain in detail how the first subnet is obtained from the attack graph.

DMZ host. If we concentrate on the subnet modeling the attack to DMZ, we can notice that success of the attack (node *DMZhost*) may be caused by four techniques: *Exploitation of Remote Service* (node *ExpRemServRoot*), *Bypass User Account Control* (node *BypassUser*), *Valid Account* (node *ValAccRoot*), *User Action* (node *Malware1*). The first technique does not need the execution of previous attack steps in order to be exploited; therefore the node *ExpRemServRoot* has no parent nodes. The other techniques, instead, can be executed only if previous steps have been successfully performed. Let us consider the *Bypass User Account Control* (node *BypassUser*); it needs a normal user access (node *UserAccess*) which in turn may be caused by the *Exploitation of Remote Service* (node *ExpRemServUser*), by a *User Action* (node *Malware2*), or by the contemporary successful execution of two techniques: *Remote Service* (node *RemServ*) and *Valid Account* (node *ValAccUser*). Both techniques are enabled by a normal user access (node *CredUser*) which can be achieved by means of the *Credential Dumping* technique (node *CredDump*) or the *Exploitation for Credential Access* (node *ExpCredAcc*). The *Valid Account* technique (node *ValAccRoot*) requires access to root credentials (node *RootDump*), which in turn needs a normal user access (node *UserAccess*) obtained as described above. Finally, both User Actions (nodes *Malware1*, *Malware2*) are enabled by the technique *Exploitation for Client Execution* (node *ExpCliExec*).

Analytics. The BN model in Fig. 4 can be extended in order to include the analytics, events possibly signaling the occurrence of a set of techniques. For each analytic we add a node whose parent nodes correspond to the involved techniques. The analytics nodes are again binary: 0 expresses that the analytic has not occurred; 1 represents the occurrence. For example, the analytic *Remote Desktop Logon* may indicate the execution of the *Valid Account* technique at normal or root user level. Therefore the parent nodes of *RemDeskLog* are *ValAccUser* and *ValAccRoot*.

CPT. The conditional distribution of a node n is locally specified by means of the conditional probability table (CPT) which provides for each possible value of n its occurrence probability, given any possible combination of the parent nodes' values. The CPT of technique-nodes without parents, such as *ExpRemServRoot* or *ExpCredAcc*, include only independent probabilities. Such values are the *a priori* probabilities of technique occurrence and are derived from the normalized scores, as explained in Section 3.1. When a technique-node has at least one parent, the definition of the CPT is required. For instance the conditional probabilities of the node *ByPassUser* depend on the values of *UserAccess*, since the technique *Bypass User Account Control* requires (unprivileged) user access to be exploited. In particular, *ByPassUser* is surely 0 when *UserAccess* is 0; in case *UserAccess* is 1, *ByPassUser* is 1 with probability p, and $1 - p$ otherwise. In this case $p = 0.9779785432$ and it is the probability of occurrence of *Bypass User Account Control* estimated from its score. For the definition of the analytic-nodes, such as *RemDeskLog*, we make the simplifying assumption that all the analytics are "perfect" meaning that they raise an alert if and only if at least one of the monitored techniques was used. In a BN such behavior can be represented by an OR-gate that allows us to derive, for all combinations of parent nodes' values, the correspondent conditional probability according to the OR operator semantics.

4 Analysis

Several kinds of analyses (inferences) can be perfomed on a BN. In general, the probability distribution of the values of any node (queried variable) can be computed. The computation can be conditioned by the observation of the value (state) of any set of other nodes in the BN (observed variables or evidence). In a predictive analysis the observed variable represents a cause, and the queried variable an effect. In a diagnostic one, the observed variable represents an effect, and the queried variable a cause. In a mixed diagnostic/predictive analysis both causes and effects of a set of queried variables are part of the evidence. In the BN modeling our attack graph, the variables representing attack goals (*DMZhost*, *OThost*, *ICS*), and those representing analytics can be considered as effects, while those representing techniques can be considered as causes.

In our context, an interesting predictive analysis is the probability of a successful attack, given the observation of a set of analytics, since it is a guide for interpreting alerts for early detection of intrusions. Tab. 1 shows the probability of a successful attack to DMZ host, OT host, and ICS service, conditioned by five different evidence sets (ES) of analytics' occurrences. For all sets, the probability of compromise of the DMZ host is the highest because it is the first goal to be reached by the attacker, while the other goals (OT host, ICS service) require further steps in order to be achieved. We can forecast adversarial threats, after early detection of a first security breach: if the DMZ host has been compromised,

Goal	ES^1	ES^2	ES^3	ES^4	ES^5
DMZhost	0.1413388	0.1384550	0.1280185	0.1261751	1.0000000
OThost	0.0200506	0.0196415	0.0176786	0.015602	0.1236536
ICSservice	0.0001628	0.0001594	0.0001435	0.0000000	0.0000000

Table 1. Compromise probabilities given different evidence sets. (ES^1: no analytics deployed. ES^2: only DMZ analytics deployed, no alerts raised. ES^3: only DMZ and OT analytics deployed, no alerts raised. ES^4: all analytics deployed, no alerts raised. ES^5: all analytics deployed, only DMZ's *DLL inject* [6] raises an alert.)

Technique	Prob.	Analytic	Prob.
BypassUser	0.021764	DLLInject	0.021764
CredDump	0.0717119	EventMonitor	0.0127349
ExpCredAcc	**0.342749**	**QuickExec**	**0.0717119**
ExpRemServRoot	0.0359553	RemDeskLog	0.0127349
ExpRemServUser	0.0057617	RPCAct	0.0150849
RemServ	0.0024305	SimLogHost	0.0127349
RootDump	0.0364453	SMBCopyExec	0.0127349
ValAccRoot	0.0019609	SMBWriteReq	0.0127349
ValAccUser	0.010796	**SuspArg**	**0.0736954**
		UserLogActMon	0.0127349
		UserLogMulti	0.0127349

Table 2. Probabilities of techniques and analytics given compromise of DMZ host

the attacker will likely target the OT host, since the probability for the OT host is 0.1418621, for the ICS service 0.0011515.

The probability that a set of techniques has been applied, given the observation of the success of attacks to DMZ, OT, or ICS, can be used as a guide for implementing effective countermeasures. In Tab. 2 we show the probabilities of techniques and analytics given the observation that the DMZ host is compromised. Among all techniques, the most probable one is *ExpCredAcc*, so it may be considered as the most relevant cause, and accordingly appropriate measures should be adopted to protect the network from such an attack. Tab. 2 shows that the most probable analytics, given the observation of the success of attacks to DMZ, OT, or ICS, are *SuspArg* (*Suspicious Arguments*) and *QuickExec* (*Quick Execution of Suspicious Commands*). This means that installing sensors for these analytics should be a cost effective detection strategy.

5 Conclusions

This paper presents a promising methodology to analyze and assess the cyber security posture of a typical power system ICT infrastructure. The model allows us to perform several types of analysis that could be used to plan the development of the security measures in terms of defense countermeasures and abnormal evidence collection. Moreover the assessment of the probability of attack processes involving different areas and in particular the point of connection between them (in this case the DMZ network), could be used to forecast the status of the infrastructure and detect possible active cyber attacks.

Bayesian Networks and their extensions represent, indeed, a useful formalism to model complex environments, as the ICT power system infrastructure, to perform cyber security analyses. This methodology can be extended in order to deploy the ability of the Bayesian Network to perform predictive and diagnostic analysis. Together with data coming from a laboratory environment, the approach could be used to identify evidences helpful to detect attack steps, thus effectively enriching the knowledge base used by the model. Another possible application of this methodology is to assess the effectiveness of the implementation of specific countermeasures in order to contrast the evolution of attack processes.

Acknowledgements

We thank Andy Applebaum for interesting insights on MITRE's ATT&CK.

References

1. Otis Alexander. ICS ATT&CK, 2017. `https://www.acsac.org/2017/workshops/icss/Otis-Alexander-ICS,AdversarialTactics,Techniques.pdf` Last accessed 11/8/2018.
2. X. An, D. Jutla, and N. Cercone. Privacy intrusion detection using dynamic Bayesian networks. In *Intl. Conf. on Electronic Commerce*, pages 208–215, 2006.
3. Andy Applebaum. Personal communication, 9/13/2018.
4. Pengsu Cheng, Lingyu Wang, Sushil Jajodia, and Anoop Singhal. Refining CVSS-based network security metrics by examining the base scores. In *Network Security Metrics*, pages 25–52. Springer, 2017.
5. Common vulnerability scoring system SIG. FIRST. `https://www.first.org/cvss/` Last accessed 11/8/2018.
6. The MITRE Corporation. CAR-2013-10-002: DLL Injection via Load Library. MITRE Cyber Analytics Repository. `https://car.mitre.org/analytics/CAR-2013-10-002` Last accessed 1/9/2019.
7. Marcel Frigault, Lingyu Wang, Sushil Jajodia, and Anoop Singhal. Measuring the overall network security by combining CVSS scores based on attack graphs and bayesian networks. In *Network Security Metrics*, pages 1–23. Springer, 2017.
8. Amin Hassanzadeh and Robin Burkett. SAMIIT: Spiral attack model in IIoT mapping security alerts to attack life cycle phases. In *Intl. Symp. for ICS & SCADA Cyber Security Research 2018*. BCS Learning and Development Ltd., 2018.
9. Industrial Control Systems Cyber Emergency Response Team (ICS-CERT). ICS-CERT advisories, 2018. `https://ics-cert.us-cert.gov/advisories` Last accessed 10/24/2018.
10. F.V. Jensen and T.D. Nielsen. *Bayesian Networks and Decision Graphs (2nd ed.)*. Springer, 2007.

11. P. Johnson, R. Lagerström, M. Ekstedt, and U. Franke. Can the Common Vulnerability Scoring System be trusted? A bayesian analysis. *IEEE Trans. on Dependable and Secure Computing*, 15(6):1002–1015, Nov 2018.
12. Kudrat Jot Kaur and Adam Hahn. Exploring ensemble classifiers for detecting attacks in the smart grids. In *Procs. CyberSec'18*, page 13. ACM, 2018.
13. B. Kordy, L. Piètre-Cambacédès, and P. Schweitzer. DAG-based attack and defense modeling: Don't miss the forest for the attack trees. *Computer science review*, 13:1–38, 2014.
14. NSTB assessments summary report: Common industrial control system cyber security weaknesses. Technical report, Idaho National Laboratory, Gaithersburg, MD, United States, 2010.
15. L. Portinale and D. Codetta-Raiteri. *Modeling and Analysis of Dependable Systems: A Probabilistic Graphical Model Perspective*. World Scientific Pub., 2015.
16. X. Qin and W. Lee. Attack plan recognition and prediction using causal networks. In *Annual Computer Security Application Conference*, pages 370–379, 2004.
17. Armin Rahimi, Adam Hahn, and Mathew Merrick. Continuous security monitoring techniques for energy delivery systems. `https://cred-c.org/videos/continuous-security-monitoring-techniques-energy-delivery-systems` Last accessed 11/15/2018.
18. Michael J. Rausch, Brett Feddersen, Ken Keefe, and William H. Sanders. A comparison of different intrusion detection approaches in an advanced metering infrastructure network using ADVISE. In *QEST 2016*, pages 279–294, 2016.
19. T. Sommestad, M. Ekstedt, and P. Johnson. Cyber security risks assessment with bayesian defense graphs and architectural models. In *HICSS '09*, pages 1–10, 2009.
20. K. Stouffer, V. Pillitteri S. Lightman, M. Abrams, and A. Hahn. NIST sp 800-82 rev 2, guide to industrial control systems (ICS) security. Technical report, NIST, 2015.
21. Blake E. Strom, Andy Applebaum, Doug P. Miller, Kathryn C. Nickels, Adam G. Pennington, and Cody B. Thomas. MITRE ATT&CK^{TM}: Design and philosophy. Technical report, The MITRE Corporation, 2018.
22. Chih-Che Sun, Adam Hahn, and Chen-Ching Liu. Cyber security of a power grid: State-of-the-art. *Int. Jour. of Electrical Power & Energy Systems*, 99:45–56, 2018.
23. Xiaoyan Sun, Jun Dai, Peng Liu, Anoop Singhal, and John Yen. Using bayesian networks to fuse intrusion evidences and detect zero-day attack paths. In *Network Security Metrics*, pages 95–115. Springer, 2017.
24. The MITRE Corporation. ATT&CK matrix heatmap. `https://github.com/mitre/attack-navigator/blob/master/layers/data/samples/heatmap_layer.json` Last accessed 11/8/2018.
25. The MITRE Corporation. ATT&CK NAVIGATOR. `https://mitre.github.io/attack-navigator/enterprise/` Last accessed 11/8/2018.
26. The MITRE Corporation. Cyber analytics repository (CAR). `https://car.mitre.org/wiki/Main_Page`, Last accessed 11/12/2018.
27. The MITRE Corporation. Adversarial tactics, techniques and common knowledge (ATT&CK), 2015. `https://attack.mitre.org/`, Last accessed 11/8/2018.
28. Md Touhiduzzaman, Adam Hahn, and Anurag Srivastava. Arcades: Analysis of risk from cyber attack against defensive strategies for power grid. *IET Cyber-Physical Systems: Theory & Applications*, 3:119–128, 2018.
29. P. Xie, J.H. Li, X. Ou, P. Liu, and R. Levy. Using Bayesian Networks for cybersecurity analysis. In *IEEE/IFIP DSN 2010*, pages 211–220, 2010.
30. S. Zhang and S. Song. A novel attack graph posterior inference model based on bayesian network. *Journal of Information Security*, 2(1):8–27, 2011.