Bayesian Network Models in Cyber Security: A Systematic Review

Chockalingam, Saba; Pieters, Wolter; Herdeiro Teixeira, André; van Gelder, Pieter

DOI
10.1007/978-3-319-70290-2_7

Publication date
2017

Document Version
Accepted author manuscript

Published in
Proceedings of the Nordic Conference on Secure IT Systems (Nordic 2017)

This is a preprint of an article accepted for publication in the 'Proceedings of the Nordic Conference on Secure IT Systems (Nordic 2017)', Springer under DOI https://doi.org/10.1007/978-3-319-70290-2_7

Bayesian Network Models in Cyber Security: A Systematic Review

Sabarathinam Chockalingam (✉), Wolter Pieters, André Teixeira, and Pieter van Gelder

Faculty of Technology, Policy and Management, Delft University of Technology, The Netherlands
{S.Chockalingam, W.Pieters, Andre.Teixeira, P.H.A.J.M.vanGelder}@tudelft.nl

Abstract. Bayesian Networks (BNs) are an increasingly popular modelling technique in cyber security especially due to their capability to overcome data limitations. This is also exemplified by the growth of BN models development in cyber security. However, a comprehensive comparison and analysis of these models is missing. In this paper, we conduct a systematic review of the scientific literature and identify 17 standard BN models in cyber security. We analyse these models based on 8 different criteria and identify important patterns in the use of these models. A key outcome is that standard BNs are noticeably used for problems especially associated with malicious insiders. This study points out the core range of problems that were tackled using standard BN models in cyber security, and illuminates key research gaps.

Keywords: Bayesian attack graph · Bayesian network · Cyber security · Information security · Insider threat

1 Introduction

The lack of data, especially historical data on cyber security breaches, incidents, and threats, hinders the development of realistic models in cyber security [**?**,**?**]. However, standard (or classical) Bayesian Networks (BNs) possess the potential to address this challenge. In particular, the capability to combine different sources of knowledge would help to overcome the scarcity of historical data in cyber security modeling.

Standard BNs belong to the family of probabilistic graphical models [**?**]. A standard BN consists of two components: qualitative, and quantitative [**?**]. The qualitative part is a Directed Acyclic Graph (DAG) consisting of nodes and edges. Specifically, each node represents a random variable, whereas the edges between the nodes represent the conditional dependencies among the corresponding random variables. The quantitative part takes the form of conditional probabilities, which quantify the dependencies between connected nodes in the DAG by specifying a conditional probability distribution for each node. A toy example of a standard BN model, representing the probabilistic relationships between cyber-attacks ("Denial of Service Attack" and "Malware Attack") and

symptoms ("Internet Connection" and "Pop-ups"), is shown in Figure 1. Given symptom(s), the BN can be used to compute the posterior probabilities of various cyber-attacks as shown in Figure 1. In this case, the user sets evidence for the "Pop-ups" node as "True", and "Internet Connection" node as "Normal" in the BN model based on his/her observations. Based on these evidences, the BN computes the posterior probabilities of the other nodes "Denial of Service Attack" and "Malware Attack" using Bayes rule. The BN model shown in Figure 1 determines that the presence of pop-ups and normal internet connection are more likely due to a Denial of Service attack rather than to a Malware attack.

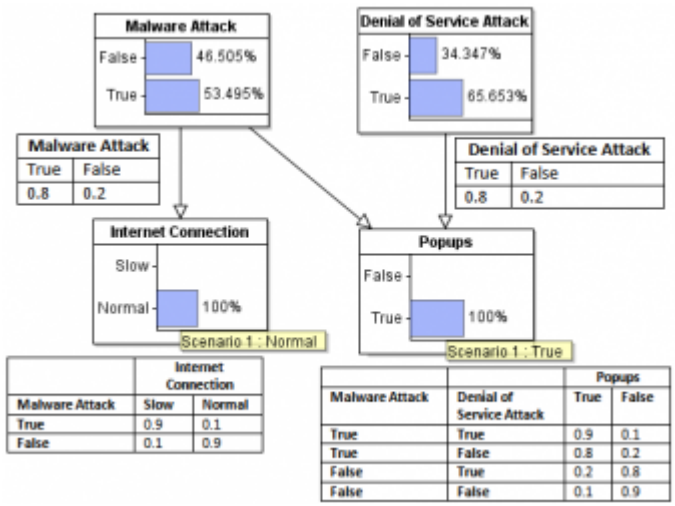

Fig. 1. Standard BN Model - Example

The major advantages of standard BNs include: the ability to combine different sources of knowledge, the capacity to handle small and incomplete datasets, and the availability of a broad range of validation approaches apart from data-driven validation approaches [**?**,**?**]. Some notable real-world applications of standard BNs include medical diagnosis [**?**] and fault diagnosis [**?**]. In addition, the advantages lead to the predominant use of standard BNs in domains where there is a limited availability of data, notably in Ecosystem Services (ESS)[**?**], water resource management [**?**], and security [**?**]. Similarly, we have seen the use of standard BNs in cyber (or information) security in recent years [**?**,**?**]. However, an overarching comparison and analysis of standard BN models in cyber security which could help to identify important usage patterns is currently lacking. Kordy et al. give a broader overview of modeling approaches based on DAGs, and thus only briefly mention BNs [**?**]. In contrast to Kordy et al., we specifically focus on BN models with the aim of performing comparison and analysis of these models to identify important usage patterns and key research gaps. This review would benefit the practical application of BN models in cyber security by providing important usage patterns and key research gaps. Therefore, this research aims to fill this gap by addressing the research question: "What are

the important patterns in the use of standard Bayesian Network (BN) models in cyber security?". The research objectives are:

- **RO 1.** To identify standard BN models in cyber security literature.
- **RO 2.** To identify the important patterns in the use of standard BN models in cyber security based on the analysis of identified models.

In this paper, we focus on comparison and analysis of standard BN models [?,?] which also include Bayesian Attack Graphs (BAGs) [?,?,?] as they possess more comparable features. This would help to identify consistent patterns in the use of standard BN models in cyber security. However, the approaches in cyber security modeling that extend BN such as Bayesian Decision Network (BDN) [?], Causal event graph [?], Dynamic BN [?,?,?], Extended influence diagram [?,?], and Multi-entity BN [?,?] are beyond the scope of this paper as they are incomparable especially based on their structure development. For instance, decision and utility nodes are specific to BDN/Influence Diagram which would allow decision making under uncertainty. In contrast, these types of nodes are not applicable to standard BN.

The scope of this comparison and analysis is the structured development, application and validation of the existing standard BN models in cyber security. The comparison and analysis of identified models is performed using the characteristics that were chosen based on related literature and domain-specific objectives as described in Section 2. The key contributions of this work are: important patterns in the use of standard BN models in cyber security, and key research gaps in the use of standard BN models in cyber security.

The remainder of this paper is structured as follows. Section 2 describes the review methodology. In Section 3, we perform the comparison and analysis of identified BN models using the characteristics that we chose, followed by a discussion on the key findings in Section 4. Finally, we highlight important patterns in the use of standard BN models in cyber security followed by future work directions in Section 5.

2 Review Methodology

We perform the systematic literature review based on the guidelines provided by Okoli et al. [?]. The methodology which we used to select the standard BN models in cyber security literature and the appropriate characteristics to perform the comparison and analysis of the selected BN models is described below.

The selection of standard BN models in cyber security literature consists of two stages:

- Searches were performed on ACM Digital Library, DBLP, Google Scholar, IEEE Xplore Digital Library, Scopus, and Web of Science – All Databases. Search-strings were constructed from keywords "Bayesian", "Bayesian Belief Network", "Bayesian Network", "BBN", "BN", "Cyber*", "Information*", and "Security". The wildcard "*" was used for "Cyber" and "Information" to match all words around these two keywords.

- Models were selected from the search results according to the listed criteria:
  - The model should employ standard BN.
  - The model should address problem(s) associated with cyber (or information) security.
  - The literature should have basic information about both DAGs and Conditional Probability Tables (CPTs). This criterion is important taking into account the scope of our comparison and analysis which is the structured development, application and validation of the existing standard BN models in cyber security.
  - The literature should be in English language.

Once a standard BN model in cyber security was selected, the scientific literature that cited it was also traced.

The characteristics used to perform the analysis of the selected BN models were chosen based on related literature and domain-specific objectives as described in Section 2 and 3. Landuyt et al. presented 47 BN models in ESS published from 2000 to 2012 [**?**]. In addition, they analysed these models based on 9 characteristics. Similarly, Phan et al. presented 111 BN models in water resource management [**?**]. Moreover, they analysed these models based on 10 characteristics. We adopted the characteristics from Landuyt et al. and Phan et al. that are generic and relevant to the scope of our analysis, as shown in Table 1. Also, we adapted and used the characteristic: *Citation details* provided by Phan et al. to perform the analysis of BN models in cyber security as described in Section 3.

Table 1. Adopted Characteristics from Landuyt et al. and Phan et al.

Characteristics used in our Analysis	**Adopted from Landuyt et al.**	**Adopted from Phan et al.**
I. Citation details		✓
II. Data sources used to construct DAGs and populate CPTs	✓	✓
III. The number of nodes used in the model	✓	
IV. Type of threat actor		
V. Application and Application sector		
VI. Scope of variables		
VII. The approach(es) used to validate models	✓	✓
VIII. Model purpose and Type of purpose		

3 Analysis of Standard Bayesian Network Models in Cyber Security

This section aims to address *RO 1. To identify standard BN models in cyber security literature*, and *RO 2. To identify the important patterns in the use of standard BN models in cyber security based on the analysis of identified models.* Based on the methodology described in Section 2, we identified 17 standard BN models in cyber security. The corresponding article titles are listed in Table 2. Furthermore, this section performs the analysis of identified BN models based on the following characteristics.

- Citation details

- Data sources used to construct DAGs and populate CPTs
- The number of nodes used in the model
- Type of threat actor
- Application and Application Sector
- Scope of variables
- The approach(es) used to validate models
- Model purpose and Type of purpose

3.1 Citation Details

We adapted and used the components of the characteristic *"Citation details"* provided by Phan et al. Specifically, we used an additional component citations in our definition of *"Citation details"* because this will help us to assess the research impact/quality of each BN model [**?**]. In Table 2, citations is the number of citations of the article according to Google Scholar Citation Index as on 15th September 2017. The number of articles covering standard BN model in cyber security varies between 0 and 3 per year. No noticeable increase in the number of papers over time is encountered. The largest number of citations (247) is acquired by Poolsappasit et al. [**?**] published in 2012. The second most cited paper, among analysed, with 136 citations, is Frigault et al. [**?**] which is published in 2008. Interestingly, BAG-based standard BN models [**?**,**?**,**?**] are extensively used compared to the other standard BN models [**?**,**?**,**?**,**?**,**?**,**?**,**?**,**?**,**?**,**?**,**?**,**?**,**?**,**?**,**?**,**?**] in cyber security based on the number of citations.

3.2 Data Sources Used to Construct DAGs and Populate CPTs

We used the characteristic *"Data sources used to construct DAGs and populate CPTs"* to identify the type of data sources utilised in the reviewed BN models. We employed the coding scheme provided by Phan et al. in Table 2 where "Expert Knowledge (K)" refers to domain expert(s) and/or article's author(s) knowledge, "Empirical Data (D)" refers to observational or experimental evidence or data, either available directly to the authors or derived from the literature [**?**]. From Table 2, we observe that 5 out of 17 BN models used only expert knowledge to construct DAGs, whereas 5 out of 17 BN models employed only empirical data to construct DAGs. 7 out 17 BN models made use of both expert knowledge and empirical data to construct DAGs. In particular, 10 out of 12 BN models which utilised empirical data to construct DAGs relied on the literature. In contrast, 2 out of 12 BN models which utilised empirical data to construct DAGs relied on the inputs from vulnerability scanner [**?**] and incidents data [**?**].

From Table 2, we infer that 11 out of 17 BN models utilised only expert knowledge to populate CPTs, whereas 3 out of 17 BN models used only empirical data to populate CPTs. On the other hand, there were 3 out of 17 BN models which employed both expert knowledge and empirical data to populate CPTs. Specifically, the sources of empirical data includes literature, incidents data, National Vulnerability Database (NVD), Open Source Vulnerability Database

Table 2. List of Bayesian Network Models in Cyber Security (Ordered by the number of citations)

Article Title (Year)	**Citations**	**Data Source (DAG)**	**Data Source (CPT)**	**Application**	**Application Sector**
Dynamic Security Risk Management Using Bayesian Attack Graphs [**?**] (2012)	247	D	K	Risk Management	Non-specific
Measuring Network Security Using Bayesian Network-Based Attack Graphs [**?**] (2008)	136	K	K	Risk Management	Non-specific
Network Vulnerability Assessment Using Bayesian Networks [**?**] (2005)	106	K	K	Risk Management	Non-specific
Reasoning about Evidence using Bayesian Networks [**?**] (2008)	39	K	K	Forensic Investigation	Law Enforcement
A Bayesian Network Model for Predicting Insider Threats [**?**] (2013)	35	D,K	D,K	Threat Hunting (Insider Threat)	Non-specific
Identifying at-risk Employees: Modeling Psychosocial Precursors of Potential Insider Threats [**?**,**?**] (2012,2010)	31,24	D,K	K	Threat Hunting (Insider Threat)	Non-specific
Identifying Compromised Users in Shared Computing Infrastructures: A Data-Driven Bayesian Network Approach [**?**] (2011)	23	D	D	Forensic Investigation	University
Development of Cyber Security Risk Model using Bayesian Networks [**?**] (2015)	21	D,K	K	Risk Management	Nuclear
Studying Interrelationships of Safety and Security for Software Assurance in Cyber Physical Systems: Approach Based on Bayesian Belief Networks [**?**] (2013)	20	K	K	Risk Management	Petroleum (Oil)
Vulnerability Categorization using Bayesian Networks [**?**] (2009)	10	D	D	Vulnerability Management (Classification)	Software
Quantitative Assessment of Cyber Security Risk using Bayesian Network-based Model [**?**] (2009)	8	D	D,K	Risk Management	Non-specific
A Bayesian Network Model for Likelihood Estimations of Acquirement of Critical Software Vulnerabilities and Exploits [**?**] (2015)	7	D,K	D,K	Governance	Software
Analysis of the Digital Evidence Presented in the Yahoo! Case [**?**] (2009)	2	K	K	Forensic Investigation	Law Enforcement
Modeling Information System Availability by using Bayesian Belief Network Approach [**?**] (2016)	1	D,K	K	Risk Management	Non-specific
A Bayesian Network Model for Predicting Data Breaches [**?**] (2016)	0	D,K	K	Risk Management	Health Care
Information Security Risk Assessment of Smartphones using Bayesian Networks [**?**,**?**] (2016,2015)	0,0	D,K	K	Risk Management	Smartphone (In Finland)
Bayesian Network Modelling for Analysis of Data Breach in a Bank [**?**] (2011)	0	D	D	Risk Management	Banking

(OSVDB), and Exploithub to populate CPTs. Notably, the review of BN models in water resource management and ESS pointed out *model simulations* as another data source used to construct DAGs and populate CPTs [**?**,**?**]. *Model simulations* refers to outputs of other empirical, deterministic or stochastic models [**?**]. Interestingly, there was no standard BN model in cyber security that used *model simulation* as the data source to construct DAGs and populate CPTs.

3.3 The Number of Nodes used in the Model

The number of nodes can be used to describe the model complexity [**?**]. A high number of nodes often lead to a lot of intermediary layers between the layer of input nodes and the layer of output nodes. This could weaken the relation between input and output nodes. Marcot et al. recommended to limit the number of node layers or sequential relationships to less than five to prevent this dilution of interactions [**?**].

Landuyt et al. indicate that BN models with nodes lower than 40 can safeguard the functionalities of BNs [**?**]. Based on our analysis, we conclude that the amount of nodes is relatively kept low in the identified BN models in cyber security as 16 out of 17 BN models have a node number lower than 40. On the other hand, the BN model developed by Shin et al. exceeds the node number 40 [**?**]. However, the BN model developed by Shin et al. is a combination of two networks. If it is not possible to keep the model structure shallow, Marcot et al. suggested to break up the model into two or more networks [**?**]. Shin et al. utilised this idea to prevent the dilution of interactions between the input and output nodes.

3.4 Type of Threat Actor

We used the characteristic *"Type of threat actor"* because this will allow us to understand whether the BN model in cyber security was developed with a focus on particular type of threat actor(s). We classified threat actors as insider versus outsider [**?**]. Furthermore, we also considered their intentions, which could be either malicious/deliberate or accidental [**?**]. Figure 2 shows the general distribution of the BN models reviewed according to the type of threat actors and their intent.

From Figure 2, we infer that 4 out of 17 BN models are used only for problems associated with insiders [**?**,**?**,**?**,**?**]. In particular, we observe that 4 out of these 4 BN models are appropriate for malicious insiders [**?**,**?**,**?**,**?**], and only 1 out of these 4 BN models is relevant for accidental insiders in addition to malicious insiders [**?**]. Holm et al. developed a BN model with a focus on malicious outsider (professional penetration tester) [**?**].

Importantly, there was no integrated BN model that considers problem(s) associated with both insider and outsider type of threat actors, and their interactions. This type of BN models would help to combat especially social engineering attacks, and outsider collusion attacks [**?**]. Finally, there were 12 out of 17 BN models which did not focus on any specific type of threat actor [**?**,**?**,**?**,**?**,**?**,**?**,**?**,**?**,**?**,**?**,**?**,**?**]. For instance, the BN model developed by Pecchia et al.

is used to identify compromised users in shared computing infrastructures based on alerts [**?**]. This model did not focus on any specific type of threat actor, but rather focused on alerts which could be appropriate to any type of threat actor. Therefore, we categorized it as 'non-specific'.

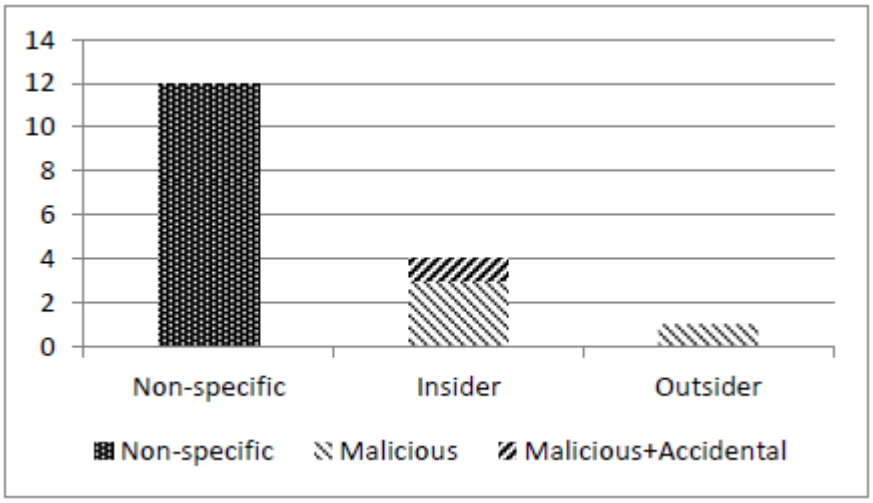

Fig. 2. Characterization of Threat Actors in the BN Models Reviewed

3.5 Application and Application Sector

We used the characteristic *"Application"* to understand the type of applications that partially or completely benefit from these BN models. We used the Chief Information Security Officer mind map as the basis to classify the reviewed BN models based on their application [**?**]. In addition, we used the characteristic *"Application Sector"* to identify the type of application sectors in which these BN models were demonstrated. From Table 2, we infer that 10 out of 17 BN models in cyber security completely or partially benefit Risk management. In addition, Forensic investigation, Governance, Threat hunting, and Vulnerability management were the other applications which completely or partially benefit from these BN models. From Table 2, we observe that the application sectors were quite diverse. However, 15 out of 17 BN models focused on the cyber security of Information Technology (IT) environment. In contrast, 2 out of 17 BN models focused on the cyber security of Industrial Control Systems (ICS) environment [**?**,**?**].

3.6 Scope of Variables

We used the characteristic *"Scope of Variables"* to identify the entities to which the variables used in the reviewed BN models are related. In addition, we classify the variables used in the reviewed BN models based on the key elements of cyber

security. Cyber security is a combination of three key elements: People, Process and Technology [?].

From Table 3, we observe that the variables used in the BN models that focus on the cyber security of ICS environment did not consider the 'people' element of cyber security [?,?]. Importantly, the variables used in these BN models are mainly related to the technological components of ICS ('Technology' focussed) [?,?]. In addition, we infer that the variables used in 2 out of 4 BN models employed for the problems associated with insiders consider the three key elements of cyber security [?,?] which are application-specific, whereas the variables used in 2 out of 4 BN models employed for the problems associated with insiders take into account only the 'people' element of cyber security [?,?] which might be applicable to different organizations.

Table 3. Scope of Variables used in the BN Models Reviewed

Authors	**Variables - Entities**	**Variables - Key Element(s) of Cyber Security**
Poolsappasit et al. [?]	Mail server, DNS server, SQL server, NAT Gateway server, Web server, Administrator machine, Local desktops	Technology
Frigault, Wang [?]	N/A	N/A
Liu, Man [?]	Network hosts	Technology
Kwan et al. [?]	Seized computer	Technology
Axelrad et al. [?]	Employee	People
Grietzer et al. [?,?]	Employee	People
Pecchia et al. [?]	User profile, Shared computing infrastructure	People, Technology
Shin et al. [?]	Organization (Management) checklist, Reactor Protector System (RPS) components	Process, Technology
Kornecki et al. [?]	Components of ICS used to control oil pipeline flow	Technology
Wang, Guo [?]	Software	Technology
Mo et al. [?]	Organization (Management), Attack pathway	Process, Technology
Holm et al. [?]	Software	Technology
Kwan et al. [?]	Suspect, Seized computer, Yahoo! email account, Internet service provider	People, Technology
Ibrahimovic, Bajgoric [?]	Organization (Management)	Process
Wilde [?]	Employee, Organization (Management), Mobile Device	People, Process, Technology
Herland et al. [?,?]	Smartphone	Technology
Apukhtin [?]	Employee, Organization (Management), Security controls	People, Process, Technology

3.7 The Approach(es) Used to Validate Models

We used the characteristic *"The approach(es) used to validate models"* to identify the type of validation approaches used in the reviewed BN models. Based on our analysis, we observe that real-world case study [?,?], cross-validation [?,?], goodness of fit [?], Monte-Carlo simulation [?], expert evaluation [?,?], and sensitivity analysis [?,?] were the approaches used to validate the reviewed BN models. Importantly, there was no validation performed in 8 out 17 BN models [?,?,?,?,?,?,?,?,?]. Finally, there was only one BN model which utilized several approaches such as sensitivity analysis, and expert evaluation to perform the validation [?]. However, the reviewed BN models validated different aspects

depending on their objectives. For instance, Wilde [**?**] validated the usefulness of their model in practice, whereas Herland et al.[**?**,**?**] validated the accuracy and completeness of the qualitative BN model.

3.8 Model Purpose and Type of Purpose

We used the characteristic *"Model Purpose"* to point out the problems that were tackled using BN models in cyber security. In addition, we used the characteristic *"Type of Purpose"* to identify the corresponding category of model purpose. Table 4 highlights the authors of the BN model, the corresponding purpose of the BN model, and the corresponding type based on the model purpose.

Table 4. BN Model Purpose and Type of Purpose

Authors	**Model Purpose**	**Type of Purpose**
Poolsappasit et al. [**?**]	To quantify the chances of network compromise at various levels	Predictive
Frigault, Wang [**?**]	To determine the likelihood of attaining the goal state by exploiting vulnerabilities in a network	Predictive
Liu, Man [**?**]	To perform quantitative vulnerability assessment of a network of hosts	Predictive
Kwan et al. [**?**]	To reason about digital evidence in the BitTorrent case	Diagnostic
Axelrad et al. [**?**]	To predict degree of interest in a potentially malicious insider	Predictive
Greitzer et al. [**?**,**?**]	To predict the psychosocial risk level of an individual	Predictive
Pecchia et al. [**?**]	To detect compromised users in shared computing infrastructures	Diagnostic
Shin et al. [**?**]	To evaluate the cyber security risk of the reactor protection system	Predictive
Kornecki et al. [**?**]	To jointly assess safety and security of a SCADA system used to control oil pipeline flow	Predictive
Wang, Guo [**?**]	To categorise software security vulnerabilities	Diagnostic
Mo et al. [**?**]	To evaluate the security readiness of organizations	Predictive
Holm et al. [**?**]	To estimate the likelihood that a penetration tester is able to obtain information about critical vulnerabilities and exploits for these vulnerabilities corresponding to a desired software and under different circumstances	Predictive
Kwan et al. [**?**]	To reason about digital evidence in the Yahoo! Case	Diagnostic
Ibrahimovic, Bajgoric [**?**]	To predict information system availability	Predictive
Wilde [**?**]	I. To predict the probability of a data breach caused by a group of insiders who lose employee- and employer-owned mobile devices or misuse the employer-owned mobile devices, II. To help health care organizations determine which additional measures they should take to protect themselves against data breaches caused by insiders.	Predictive, Diagnostic
Herland et al. [**?**,**?**]	To assess information security risks related to smartphone use in Finland	Predictive
Apukhtin [**?**]	To predict the probability of a data breach in a bank caused by a malicious insider	Predictive

From Table 4, we observe that the reviewed BN models in cyber security were mainly used for two types of purposes based on their model purpose: I. Diagnostic: To reason from effects to causes, and II. Predictive: To reason from causes to effects. Importantly, 13 out of 17 BN models in cyber security were used for predictive purposes.

4 Discussion

In the previous section, we identified key usage patterns of BNs in cyber security. This section discusses potential reasons for the key findings and suggests future research directions.

There is an emphasis on problems associated with insiders compared to outsiders in the use of standard BN models in cyber security. In general, this emphasis could be due to the most significant threat posed by insiders. This was elucidated by IBM's cyber security intelligence index which concluded that 60% of all attacks were carried out by insiders [**?**]. In connection with the use of standard BNs, the availability of characteristics associated with insiders in the literature provided a good starting point to determine appropriate variables and their relationships which form an integral part of a standard BN. In addition, the variables and their relationships determined from the literature were fine-tuned and/or complemented with other suitable variables based on expert knowledge in a few instances. This is one of the major advantages of standard BNs described in Section 1 which is the ability to combine different sources of knowledge. This could be the rationale behind the predominant use of standard BNs for problems associated with the insiders.

Special importance is given to problems associated with malicious insiders compared to accidental insiders in the use of standard BN models in cyber security. In general, this could be due to the fact that malicious insiders are more natural than accidental insiders in security contexts, as malicious insiders have a clear intent of compromising security, while accidental insiders do not. Moreover, malicious insiders have been shown to be the cause of more incidents than accidental insiders, as it was demonstrated by IBM's cyber security intelligence index which concluded that 44.5% of attacks were carried out by malicious insiders, and accidental insiders were responsible for 15.5% of attacks [**?**]. In order to use standard BNs for problems associated with accidental insiders compared to malicious insiders, it is important to identify features associated with accidental insiders in the literature to determine appropriate variables and their relationships, which form an essential part of a standard BN. There are studies which identify features associated with accidental insiders in the literature [**?**,**?**]. Once the appropriate variables and their relationships are determined for problems associated with accidental insiders, this could always be updated based on expert knowledge. It would also be useful to explore variables and their relationships in the reviewed BN models that focus on problems associated with malicious insiders, as some of the indicators might also apply for problems associated with accidental insiders [**?**].

The focus on insiders may also explain why there is little research on applications in the ICS domain. The reviewed BN models that focus on problems associated with the insiders might not be suitable for ICS environments, especially for control rooms with an operator. This is prevalent in control rooms that are used to operate sluices in the Netherlands. Not accepting feedback, Anger management issues, Confrontational issues, Counterproductive behaviour towards individuals (CPB-I), Counterproductive behaviour towards the orga-

nization (CPB-O) were some of the variables used in the reviewed BN models [?,?]. Most of these variables might be measured/observed based on interactions of the particular individual with the co-workers. However, this would not be possible in the control rooms where there would be no co-worker. It would be interesting to explore in the future whether the variables and their relationships in the reviewed BN models focused on problems associated with the insiders are suitable for ICS environment, and also whether the size of the organization in which the BN model would be applied have an effect on these variables and their relationships. In general, the limited use of standard BN models in cyber security on problems associated with ICS environment could be due to the shortage of ICS security expertise [?] as majority of the reviewed BN models relied on expert knowledge especially to construct DAGs and populate CPTs.

There is no integrated BN model which takes into account the problem(s) associated with both insiders and outsiders, and their interactions. The German steel mill incident is a typical example of a cyber-attack which involves both accidental insiders and malicious outsiders, and their interactions [?]. As an initial step, the adversaries used both the targeted email and social engineering techniques to acquire credentials for the plant's office network. Later, once they acquired credentials for the plant's office network, they worked their way into the plant's control system network and caused damage to the blast furnace. Standard BNs would help to tackle problem(s) associated with both insiders and outsiders, and their interactions, for instance a standard BN model that could predict the probability of an individual being deceived by outsider(s) to cause a cyber-attack in an organization, given certain risk factors and symptoms. This BN model would especially help to identify vulnerable individuals in an organization against social engineering attacks, and effective measures which could reduce the likelihood of an individual deceived by outsiders to cause a cyber-attack in an organization.

It is evident that the initial attempts in the use of standard BN models in cyber security were using BAG-based standard BN models [?,?,?]. BAG-based standard BN model combines acyclic attack graph which acts as the DAG with computational procedures of BN. Attack graph is one of the extensively used approaches in security modeling which was introduced in 1998 [?,?]. The use of BAG-based standard BN models in the intial attempts could be due to practicality. It could be practical to build attack graphs first which had been extensively studied in this domain and use BN computational procedures for quantification during the early stages in the use of standard BN models in cyber security. Similarly, there were attempts in the safety domain which mapped fault tree to BN [?,?]. Importantly, BAG-based standard BN models model static systems. Therefore, they are not directly applicable to multi-step attacks.

Risk management, forensic investigation, governance, threat hunting, and vulnerability management were the applications of standard BNs in cyber security. However, it would also be useful to investigate the potential of standard BNs to benefit other applications. Chockalingam et al. highlighted the importance of integrating safety and security especially in the context of modern ICS [?]. BNs

possess the potential to develop an integrated BN model that could diagnose the root cause of an abnormal behavior in the ICS especially whether the abnormal behavior is caused by an attack (security-related) or technical failure (safety-related) by taking into account certain risk factors and symptoms. This would allow the operator(s) to point out the best possible response strategy. For instance, the process of routing traffic through a scrubbing center would be a suitable response strategy for a Distributed Denial of Service (DDoS) attack whereas this may not be an appropriate response strategy for a sensor failure.

The sources of empirical data used to construct DAGs and populate CPTs include: literature, incidents data, NVD, OSVDB, and exploithub. It is important to identify other domain-specific empirical data sources which would help to develop realistic models in cyber security. For instance, Capture-The-Flag (CTF) events like SWaT security showdown (S3) [**?**] could be a potential data source to construct DAGs and populate CPTs. CTF events could generate datasets that are realistic in nature [**?**]. However, this could have been overlooked because the data generated in these events would be in most cases specific to that particular system, and the quality of data generated could depend on the participants.

5 Conclusions and Future Work

In this paper, we have identified 17 standard BN models in cyber security. Based on the analysis, we identified important patterns in the use of standard BN models in cyber security.

- The standard BN models in cyber security were significantly used for problems associated with malicious insiders.
- There is an emphasis on the use of standard BN models in cyber security for problems associated with IT environment compared to ICS environment. In addition, the standard BN models that focus on the cyber security of ICS environment did not consider the 'people' element of cyber security. This implies that there is no standard BN model which deal with problem associated with insiders in ICS environment.
- There is a lack of standard BN models usage for problems associated with insiders and outsiders, and their interactions.
- Expert knowledge, and empirical data predominantly from literature were the data sources utilised to construct DAGs and populate CPTs.
- The standard BN models in cyber security completely or partially benefited risk management, forensic investigation, governance, threat hunting, and vulnerability management.
- The approaches used to validate standard BN models in cyber security were real-world case study, cross-validation, goodness of fit, monte-carlo simulation, expert evaluation, and sensitivity analysis.

These patterns in the use of standard BN models in cyber security would help to make full use of standard BNs in cyber security in the future especially by pointing out the current trends, limitations and future research gaps.

In the future, it is important to investigate whether the BN models used for problems associated with insiders are applicable for ICS environments, especially for a control room with an operator. It would be useful to demonstrate the capacity of standard BNs to tackle problems associated with both insiders and outsiders, and their interactions like social engineering attacks, collusion attacks. It would be intriguing to investigate how to deal with multi-step attacks using standard BNs. The potential of alternative data sources like model simulations, CTF events to construct DAGs and populate CPTs in cyber security also needs to be explored, as well as the capability of standard BNs to completely or partially benefit the other applications in cyber security.

Acknowledgement

This research received funding from the Netherlands Organisation for Scientific Research (NWO) in the framework of the Cyber Security research program under the project *"Secure Our Safety: Building Cyber Security for Flood Management (SOS4Flood)"*.

References

1. WEF: Partnering for Cyber Resilience: Towards the Quantification of Cyber Threats. (2015)
2. Yu, S., Wang, G., Zhou, W.: Modeling malicious activities in cyber space. IEEE network 29, 83-87. (2015)
3. Ben-Gal, I.: Bayesian Networks. Encyclopedia of Statistics in Quality and Reliability. John Wiley & Sons, Ltd. (2008)
4. Darwiche, A.: Bayesian networks. Foundations of Artificial Intelligence 3. (2008)
5. Landuyt, D., et al.: A review of Bayesian belief networks in ecosystem service modelling. Environmental Modelling & Software 46, 1-11. (2013)
6. Uusitalo, L.: Advantages and challenges of Bayesian networks in environmental modelling. Ecological modelling 203, 312-318. (2007)
7. Nikovski, D.: Constructing Bayesian Networks for Medical Diagnosis from Incomplete and Partially Correct Statistics. IEEE Transactions on Knowledge and Data Engineering 12(4), pp.509-516. (2000)
8. Nakatsu, R.T.: Reasoning with Diagrams: Decision-Making and Problem-Solving with Diagrams. John Wiley & Sons. (2009)
9. Phan, T.D., et al.: Applications of Bayesian belief networks in water resource management: A systematic review. Environmental Modelling & Software 85, 98-111. (2016)
10. Kordy, B., Pièvre-Cambacédès, L., Schweitzer, P.: DAG-based attack and defense modeling: Don't miss the forest for the attack trees. Computer science review 13, 1-38. (2014)
11. Poolsappasit, N., Dewri, R., Ray, I.: Dynamic security risk management using bayesian attack graphs. IEEE Transactions on Dependable and Secure Computing 9, 61-74. (2012)
12. Frigault, M., Wang, L.: Measuring network security using bayesian network-based attack graphs. IEEE. (2008)

13. Liu, Y., Man, H.: Network vulnerability assessment using Bayesian networks. In: Proc. SPIE, pp. 61-71. (2005)
14. Kwan, M., Chow, K.-P., Law, F., Lai, P.: Reasoning about evidence using Bayesian networks. In: IFIP International Conference on Digital Forensics, pp. 275-289. (2008)
15. Axelrad, E.T., Sticha, P.J., Brdiczka, O., Shen, J.: A Bayesian network model for predicting insider threats. In: Security and Privacy Workshops, pp. 82-89. (2013)
16. Greitzer, F.L., et al.: Identifying at-risk employees: Modeling psychosocial precursors of potential insider threats. In: System Science (HICSS), Hawaii International Conference on, pp. 2392-2401. (2012)
17. Greitzer, F.L., et al.: Identifying at-risk employees: A behavioral model for predicting potential insider threats. Pacific Northwest National Laboratory. (2010)
18. Pecchia, A., et al.: Identifying compromised users in shared computing infrastructures: A data-driven bayesian network approach. In: Reliable Distributed Systems (SRDS), 2011 30th IEEE Symposium on, pp. 127-136. IEEE. (2011)
19. Shin, J., Son, H., Heo, G.: Development of a cyber security risk model using Bayesian networks. Reliability Engineering & System Safety 134, 208-217. (2015)
20. Kornecki, A.J., Subramanian, N., Zalewski, J.: Studying interrelationships of safety and security for software assurance in cyber-physical systems: Approach based on bayesian belief networks. In: Computer Science and Information Systems (FedCSIS), 2013 Federated Conference on, pp. 1393-1399. IEEE. (2013)
21. Wang, J.A., Guo, M.: Vulnerability categorization using Bayesian networks. In: Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research, pp. 29. ACM. (2010)
22. Mo, S.Y.K., Beling, P.A., Crowther, K.G.: Quantitative assessment of cyber security risk using Bayesian Network-based model. In: Systems and Information Engineering Design Symposium, 2009. SIEDS'09., pp. 183-187. IEEE. (2009)
23. Holm, H., Korman, M., Ekstedt, M.: A bayesian network model for likelihood estimations of acquirement of critical software vulnerabilities and exploits. Information and Software Technology 58, 304-318. (2015)
24. Kwan, M., et al.: Analysis of the digital evidence presented in the Yahoo! case. In: IFIP International Conference on Digital Forensics, pp. 241-252. Springer. (2009)
25. Ibrahimović, S., Bajgorić, N.: Modeling information system availability by using bayesian belief network approach. Interdisciplinary Description of Complex Systems 14, 125-138. (2016)
26. Wilde, L.: A Bayesian Network Model for predicting data breaches caused by insiders of a health care organization. University of Twente. (2016)
27. Herland, K., Hammainen, H., Kekolahti, P.: Information Security Risk Assessment of Smartphones using Bayesian Networks. Journal of Cyber Security and Mobility 4, 65 - 85. (2016)
28. Herland, K.: Information security risk assessment of smartphones using Bayesian networks. (2015)
29. Apukhtin, V.: Bayesian network modeling for analysis of data breach in a bank. University of Stavanger, Norway. (2011)
30. Khosravi-Farmad, M., Rezaee, R., Harati, A., Bafghi, A.G.: Network security risk mitigation using Bayesian decision networks. In: Computer and Knowledge Engineering (ICCKE), 4th International eConference on, pp. 267-272. IEEE. (2014)
31. Pan, S., Morris, T.H., Adhikari, U., Madani, V.: Causal event graphs cyber-physical system intrusion detection system. In: Proceedings of the Eighth Annual Cyber Security and Information Intelligence Research Workshop, pp. 40. ACM. (2013)
32. Frigault, M., et al.: Measuring network security using dynamic bayesian network. In: Proceedings of the 4th ACM workshop on Quality of protection, pp. 23-30. (2008)

33. Sarala, R., Kayalvizhi, M., Zayaraz, G.: Information security risk assessment under uncertainty using dynamic Bayesian networks. International Journal of Research in Engineering and Technology 304-309. (2014)
34. Tang, K., Zhou, M.-T., Wang, W.-Y.: Insider cyber threat situational awareness framwork using dynamic Bayesian networks. In: Computer Science & Education, 2009. ICCSE'09. 4th International Conference on, pp. 1146-1150. IEEE. (2009)
35. Sommestad, T., Ekstedt, M., Johnson, P.: Cyber security risks assessment with bayesian defense graphs and architectural models. In: System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on, pp. 1-10. IEEE. (2009)
36. Ekstedt, M., Sommestad, T.: Enterprise architecture models for cyber security analysis. In: Power Systems Conference and Exposition, pp. 1-6. IEEE. (2009)
37. Laskey, K., et al.: Detecting threatening behavior using bayesian networks. In: Conference on Behavioral Representation in Modeling and Simulation, pp. 33. (2006)
38. AlGhamdi, G., et al.: Modeling insider behavior using multi-entity Bayesian networks. (2006)
39. Okoli, C., Schabram, K.: A guide to conducting a systematic literature review of information systems research. Sprouts Work. Pap. Inf. Syst 10. (2010)
40. Meho, L.I.: The rise and rise of citation analysis. Physics World 20, 32. (2007)
41. Marcot, B.G., Steventon, J.D., Sutherland, G.D., McCann, R.K.: Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. Canadian Journal of Forest Research 36, 3063-3074. (2006)
42. Alberts, C., Dorofee, A.: OCTAVESM Threat Profiles.
43. Bureau, F.I.P.: Unintentional Insider Threats: A Foundational Study. (2013)
44. Rehman, R.: CISO MindMap. `http://rafeeqrehman.com/wp-content/uploads/2017/07/CISO_Job_MindMap_v9.png`. (2017)
45. Andress, A.: Surviving security: how to integrate people, process, and technology. CRC Press. (2003)
46. 2016 Cyber Security Intelligence Index. IBM Security. (2016)
47. Greitzer, F.L., et al.: Unintentional insider threat: contributing factors, observables, and mitigation strategies. In: System Sciences (HICSS), 2014 47th Hawaii International Conference on, pp. 2025-2034. IEEE. (2014)
48. Antonioli, D., et al.: Gamifying Education and Research on ICS Security: Design, Implementation and Results of S3. arXiv preprint arXiv:1702.03067. (2017)
49. RISI Database.: German Steel Mill Cyber Attack. http://www.risidata.com/database/detail/german-steel-mill-cyber-attack. (2017)
50. Lippmann, R.P., Ingols, K.W.: An annotated review of past papers on attack graphs. MASSACHUSETTS INST OF TECH LEXINGTON LINCOLN LAB. (2005)
51. Bobbio, A., Portinale, L., Minichino, M., Ciancamerla, E.: Improving the analysis of dependable systems by mapping fault trees into Bayesian networks. Reliability Engineering & System Safety 71, 249-260. (2001)
52. Khakzad, N., Khan, F., Amyotte, P.: Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches. Reliability Engineering & System Safety 96, 925-932. (2011)
53. Chockalingam, S., et al.: Integrated Safety and Security Risk Assessment Methods: A Survey of Key Characteristics and Applications. In: International Conference on Critical Information Infrastructures Security (CRITIS). Paris. (2016)
54. Salem, M.B., Hershkop, S., Stolfo, S.J.: A survey of insider attack detection research. Insider Attack and Cyber Security 69-90. (2008)