Towards insider threats detection in smart grid communication systems

Beibei Li1,2, Rongxing Lu2,3, Gaoxi Xiao1, Haiyong Bao4, Ali A. Ghorbani2

1School of Electrical and Electronic Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore
2Faculty of Computer Science, University of New Brunswick, ITC Building, 550 Windsor Street, Fredericton, NB, Canada
3School of Computer Science and Information Engineering, Zhejiang Gongshang University, Hangzhou, Zhejiang, People’s Republic of China
4E-mail: ru1@unb.ca

Abstract: In today’s communication systems, the most damaging security threats are not originating from the outsiders but from the trusted insiders – both malicious insiders and negligent insiders. Always endowed with high privileges, insiders are arguably the largest national critical engineering infrastructure – is suffering from a multitude of security threats initiated from both outsiders and insiders. Without security guarantee, the promising benefits of achieving an efficient, green, and reliable power grid would not be a success. In this study, the authors investigate the insider threats and summarise the existing threats detection solutions in smart grid communication systems. In addition, a novel hybrid insider threats modelling, analysis, and detection framework, which is based on stochastic Petri net and behaviour rule specifications, is proposed to contain insider threats in smart grid communication systems.

1 Introduction

The rapid advances in information and communications technologies (ICTs) and the exponential proliferation of intelligent electronic devices (IEDs) have considerably facilitated the emergence of the smart grid, which has been widely believed as a revolutionary alternative for the existing power grid. To meet the ever-increasing demands of electricity in the 21st century and overcome the shortages of the existing power grid, the concept of smart grid has been rolled out with the expectation of achieving an efficient, green, and reliable power generation, transmission, and distribution system [1]. Smart grid is a primary representative of cyber-physical systems (CPSs). A CPS [2] is defined as a fusion system incorporating physical infrastructure (physical space) with communication and control system (cyber space) – see the conceptual architecture of a CPS in Fig. 1. Through close interactions between the cyber space and physical space, i.e. data sensing and commands execution via communication networks, CPSs are envisioned to enable real-time monitoring, accurate coordination, and automatic control objectives over the physical infrastructure.

However, despite these promising benefits, contemporary development of the smart grid (like any other CPSs) is still in its infancy stage facing many critical challenges. Particularly, cyber security issues have been rapidly emerged as primary concerns of the smart grid communication systems over the years [3–7]. The increasing interconnectivity and openness to the outside world inevitably expose the potential vulnerabilities of not only the conventional communication system in existing power grids but also the newly-introduced ICTs. Without security guarantee, the smart grid is susceptible to malicious penetrations and can easily suffer from dire consequences, e.g. catastrophic physical damages, enormous revenue losses, and immeasurable social impacts. The potential adversaries in contemporary communication systems, who can launch malignant cyber assaults by fully exploiting the system vulnerabilities, can be classified into two categories: insiders and outsiders [8]. Although in many industries the absolute numbers of outsiders are still higher than those of insiders, the percentage of insiders is experiencing a rapid rise in recent years [9]. More importantly, the insiders are contributing to increasingly damaging impacts during security incidents, compared to outsiders. According to IBM X-Force Threat Intelligence Index 2018 [10], misconfigured cloud servers and networked backup incidents unintentionally exposed more than two billion records, which is approximately 70% of the total number of compromised records being tracked. There were 424% more records compromised as a result of these types of incidents in 2017 than those in the previous years. This implies that an increasing number of attackers may have shifted their attack paradigms to employing the insiders over the recent years. Insiders always have overwhelming superiorities over the untrusted external adversaries in terms of conducting malignant acts. Specifically, insiders are well aware of the existing security mechanisms (e.g. firewalls, cyber access controls, physical access controls etc.), thus they can fully utilise such knowledge to easily circumvent the defenses in place. This is, however, intractable for untrusted outsiders.

In smart grid communication systems, intensive research efforts in recent years have been devoted towards detecting the outsider threats, e.g. denial-of-service (DoS) attacks, man-in-the-middle attacks, and eavesdropping, but seldom on insider threats [11]. As a result, there must exist many limitations when applying existing threats detection solutions directly to insider threats in smart grid communication systems, because insider threats may have completely distinct attack strategies and behaviours. To address these limitations and provide useful insights for future research studies, this paper focuses on insider threats detection solutions in...
smart grid communication systems. Our research contributions are in three-fold:

- We carefully investigate the insider threats in smart grid communication systems and summarise the existing threats detection solutions.
- Most importantly, we propose a novel hybrid insider threats modelling, analysis, and detection framework for smart grid communication systems, which is based on stochastic Petri net and behaviour rule specifications. The effectiveness of the proposed model is validated by PIPE2 [12].
- A learning algorithm of transition rates in the proposed stochastic Petri net-based model is designed to achieve near real-time evaluation of the system risk degree.

The remainder of this paper is organised as follows. Section 2 gives a high-level overview of the smart grid architecture with its specific security requirements. Section 3 showcases a taxonomy of existing insider attacks in smart grid communication systems. Following this, existing threats detection solutions in smart grid communication systems are categorised in Section 4. Our proposed framework for insider threats modelling, analysis, and detection is elaborated in Section 5, followed by the numerical results in Section 6. Finally, Section 7 concludes this paper with future research directions.

2 Smart grid architecture and security requirements

In this section, we formalise the architecture of a smart grid and identify its specific security requirements.

2.1 Architecture of smart grid

Smart grid is a fusion system comprising a physical space of power grid as well as a cyber space of communication and control system (as shown in Fig. 2). The physical space incorporates power generation plants, transmission substations, distribution substations, and customer premises (e.g. residential houses, commercial buildings, and industrial factories). Phasor measurement units (PMUs) are deployed at each bus or substation along the power grid to measure the real-time status of the power grid. In addition, smart meters are positioned at each customer site to collect detailed electricity consumption data. Note that in smart grid, not only the information flow but also the power delivery is no longer restricted to a single direction [13]. Customers with capabilities of producing renewable energies, e.g. solar, wind, hydro etc., are allowed to supply their surplus energy back to the power grid for other customers.

In the cyber space, a variety of wired and wireless communication networks, such as wireless sensor networks, 3G/4G cellular networks, optical fibre networks etc., are integrated to support diverse communication requirements in different parts of the smart grid. These communication channels are responsible for timely and efficient information transmission across the entire power grid. Specifically, it includes transporting real-time measurement and sensing data from PMUs and smart meters to the system control centre, and delivering control commands from the control centre to IEDs and actuators. The control centre is the core component of the whole power grid, wherein reported data are carefully analysed for multiple applications, e.g. state estimation, event diagnostics, optimal power flow analysis etc. These applications are planned to make optimal feedback decisions to coordinate the demand-supply relations and optimise system operations.

Strictly speaking, the smart grid cannot be distinctly divided into cyber space and physical space, as the heavy interdependence between the two spaces has fused them together as a whole system. Sensing data analysis and control commands execution tightly interconnect the power grid and communication networks, forming into a strict and efficient close-loop control framework. This framework enables automatic real-time supervision and control of the power grid, and thereby boosts the efficiency, reliability, and robustness of the system.

2.2 Security requirements of smart grid

The high reliability, efficiency, and robustness of a smart grid heavily rely on the security guarantee of the entire system. Therefore, a line of high-level security requirements must be satisfied to ensure system security. Here, we discuss a few of these security requirements, and show a few scenarios where inside attackers can breach or misuse these requirements.

- **Integrity**: The integrity in smart grid refers to preventing malicious modification, delay or drop of communication information, such as sensing data, metering readings, and control commands, by unauthorised parties. Inside attackers can conduct false data injection attacks, data replay attacks, data delay or drop attacks over the communication networks to breach this requirement.
- **Availability**: The availability is concerned with ensuring that legitimate access or use of information and/or communication resources for authorised users cannot be delayed or denied by unauthorised parties. For the smart grid, these resources include all the IT elements of the system, like databases, operator
workstations, management systems, defense mechanisms, as well as the communication systems among these elements. Successful implementation of attacks against this requirement may cause significant safety issues and catastrophic physical damages, considering that system operators may lose control of the power grid over the disabled communication channels. This requirement mainly includes defense against DoS attacks, distributed DoS (DDoS) attacks, and attacks of service attacks.

- **Authentication:** Authentication is a process of determining whether an identity is declared to be a legitimate system user by matching the credentials with those stored in the database. Authentication is required before each communication session starts, and between any two parties that have communication needs. For example, the control centre has the need to authenticate PMUs before accepting the measurement data, or system actuators have the need to authenticate the control centre before executing given commands. Without effective authentication mechanisms in place, malicious attackers, for example, may disguise as a legitimate PMU who can report false measurement data to mislead the control centre into issuing damaging commands. Authentication is of paramount significance to distinguish between legitimate and illegitimate users in smart grid. However, inside attackers can easily misuse their legitimate identities to launch malicious attacks.

- **Confidentiality and privacy:** This requirement in smart grid is concerned with preventing disclosure of secret information, such as system topology and configurations, measurement data, and customer power consumption data etc., to unauthorised parties. Breach of this requirement may contribute to major security issues. On the one hand, leakage of power grid status information (i.e. topology and configurations, measurement data) may cause significant security threats to the system, as malicious attackers can benefit from such information to launch false data injection attacks or more complicated cyber-physical attacks. On the other hand, in case that the detailed customer consumption data is exposed to the outside world, adversaries with vicious intentions can gain insights into users' daily life behaviours by analysing these data with a certain side information. However, insiders with malicious intentions are able to easily break this requirement by stealing and selling such sensitive information to others.

- **Authorisation:** Also known as access control, authorisation in smart grid refers to giving official permissions for specific users or IEDs with full or limited access to certain system resources, say databases, management systems etc. There are many different roles and user groups in the smart grid, and they all have diverse privileges and responsibilities, thereby an effective authorisation mechanism is required to prevent system resources from illegitimate access. Note that, since authorisation is able to distinguish between legitimate and illegitimate users for specific resources, it can also, to a certain extent, protect the integrity and confidentiality. However, similar to that for authentication requirement, insiders with legitimate privileges can easily carry out malicious activities.

- **Non-repudiability:** In the context of smart grid, non-repudiability deals with preserving the undeniable proofs of whom initiated a specific action in the system, say issuing a command or visiting a database. This security requirement expects to establish accountability and liability in smart grid, by which the actors can easily be identified with irreproachable proof of their previous actions. This requirement provides effective proofs to identify insider threats. Unfortunately, in some cases, insiders may be able to ruin these proofs before being disclosed.

- **Auditability:** Auditability allows auditors to reconstruct the complete operation history of a specific user or device based on historical records, e.g. log files, duty schedules, of all its own or relevant performed acts. This security requirement aims to identify the reasons for any malfunction or security incident so that perpetrators can easily be discovered, and the potential vulnerabilities of the system can be identified and further fixed. Such a requirement can be breached by malignant insiders who may access these historical records and delete or ruin them permanently.

### 3 Insider threats in smart grid

To clearly identify the attack strategies, objectives, and potential impacts of insider threats, here we develop a taxonomy of existing insider threats in smart grid. We then introduce three types of attack behaviours of these insider threats. Note that insiders considered in this paper are those who can play an authorised role in the whole smart grid communication systems. In other words, insiders have authorised entities with legitimate credentials to access the communication system including system operators, utility employees, third-party service providers, end users, and particularly, all IEDs joining in this system.

#### 3.1 Taxonomy of existing insider threats

As summarised in Table 1, we present a taxonomy of existing insider threats in smart grid. In this taxonomy, there are four types of insider threats: false data injection, false command injection, delay of service, and insider data theft. The attack strategies, objectives, and impacts on the system are shown as follows:

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>False data injection</td>
<td>Compromised meter devices (e.g. PMUs and smart meters) report fake data to the control centre</td>
</tr>
<tr>
<td>False command injection</td>
<td>System operators or local control agents issue fake commands to actuators, say line breakers and generators</td>
</tr>
<tr>
<td>Delay of service</td>
<td>IEDs create delays in latency-sensitive services, e.g. delayed report of real-time measurement data</td>
</tr>
<tr>
<td>Confidential data theft</td>
<td>Utilities employees or compromised devices steal sensitive information or intellectual property</td>
</tr>
</tbody>
</table>

where $H$ is an $m \times n$ measurement Jacobian matrix, implying the knowledge of grid topology and configurations. $\eta = (\eta_1, \eta_2, \ldots, \eta_m) \sim \mathcal{N}(0, W)$ is an independent measurement error vector with zero-mean and covariance $W$, a diagonal matrix. Statistical estimation method, like maximum likelihood estimation, is exploited to estimate the real system state variables by

$$\hat{x} = [HW^{-1}H]^{-1}HW^{-1}z.$$ (2)
compromised PMUs to collaboratively construct a set of bad measurement data $z_{bad}$ as

$$z_{bad} = z + Hc,$$  \hspace{1cm} (3)

where $c$ is the desired offset to be injected to $x$. As such, fake system states are estimated by

$$x_{hat} = \left( H^T W^T H \right)^{-1} H^T W^T (z + Hc)$$  \hspace{1cm} (4)

With such a set of fake data being undetected, the system control centre is much likely to issue false feedback commands to the power grid, resulting in serious damages.

### 3.1.2 False command injection

A false command injection attack attempts to inject illicit control commands to the system actuators to mislead the normal physical processes [4]. This kind of attack can be implemented by not only inside system operators (say disgruntled employees) but also IEDs (say decentralised control agents). Inside system operators, who oversee the entire power grid with high privileges, can easily inject illicit commands. Likewise, IEDs controlled by remote hackers can also directly inject fake commands, e.g. abnormally opening a circuit breaker to cause power outage, or exceptionally turning on a generator to mess up normal demand-supply relations.

False command injection attacks usually lead to more severe damages for CPSs than false data injection attacks, as attackers can directly manipulate the control units and all its operations. For example, in 2000 an Australia sewage control system was invaded by an ex-contractor, with a motive of revenge, who injected fake commands to open the flood gates. These malignant acts resulted in more than one million litres of raw sewage being poured into two local fresh water streams, which polluted the water for the subsequent three months. Unfortunately, only a few research studies have paid attention to this kind of attacks in the smart grid. One rare example is that Gao et al. in 2010 investigated the false command injection attacks in smart grid supervisory control and data acquisition (SCADA) system and the relevant detection techniques [15]. More research efforts are demanded towards preventing and detecting false command injection attacks in smart grid.

### 3.1.3 Delay of service

Instead of denying services (e.g. dropping data) or offering bad services (e.g. falsifying data), delay of service attacks prefer to compromise the quality of services by creating delays when offering latency-sensitive services. In other words, similar to DoS attacks, delay of service attacks target at the availability requirement in smart grid [16]. This kind of inside attackers can employ distinct strategies in real-world applications, as various services may have different latency requirements. For example, the power grid can tolerate up to 4 ms for protective relaying service, a few seconds for SCADA data transmission, several minutes for real-time market pricing information, and some hours for smart meter readings etc.

### 3.1.4 Confidential data theft

A confidential data theft attack in smart grid is concerned with illicit insider acts of stealing sensitive data stored on database servers, desktop computers, or other internal devices. Specifically, disgruntled employees or compromised IEDs with legitimate access to data resources are able to launch data theft attacks by stealing information such as the system topology and configurations data, user power consumption data, intellectual property etc. Once these data are leaked or sold to the third parties, e.g. marketers, maintenance utilities, household appliance enterprises, or even criminals, they can be exploited for different purposes. For example, customer electricity consumption data may reveal the daily lifestyle of an individual house, as each appliance has a unique electricity usage profile. Criminals, if having such data, may figure out the daily routines of a family and commit burglary when people are absent, or cause personal injuries when someone is at home alone.

Confidential data theft attack has been a growing critical problem for corporations and organisations over the years. According to the 2018 Data Breach Investigations Report [17] by Verizon, the year 2017 has confirmed 2216 data breaches across 65 countries. In particular, the healthcare, education, professional, and public domains have experienced an ever-growing number of data breaches.

### 3.2 Attack prototypes

In this subsection, we show three attack prototypes in terms of malicious insiders' attack strategies. With different illegitimate intentions and under different circumstances, inside attackers may behave in distinct manners. Some may prefer to lurk behind the power grid as long time as possible by performing less malicious acts, while others may prefer to cause as extensive damages as possible by carrying out malicious acts whenever having an opportunity. In this paper, insider attacks are categorised into three prototypes: reckless, random, and opportunistic attacks [8].

- **Reckless attacks** are carried out without hesitation whenever the attackers have the capability and have not been detected yet. The attacking probability of a reckless attack can be defined as $P_{att} = 1$, wherein these attacks behave in a rude manner with the intention to cause as extensive damages as possible. Reckless attacks are, therefore, the easiest type of attacks to be identified by existing detection techniques.
- **Random attacks** are conducted with a certain attacking probability, i.e. $P_{att} = \rho$, where $\rho \in [0, 1]$. Unlike reckless attacks, these attacks are performed with consideration to balance between the damages they may cause and the chances they could be identified. Hence, such attacks are carried out in a relatively moderate mode. Usually, it will cost the detection system a certain efforts to detect these random attacks.
- **Opportunistic attacks** are the most cunning attacks, which are usually implemented based on the system noise level with an attacking probability of

$$P_{att} = f(\alpha, \beta, \gamma) \cdot P_{n},$$  \hspace{1cm} (5)

where $f(\alpha, \beta, \gamma)$ is a coefficient function, $\alpha$, $\beta$, and $\gamma$ denote the difficulty level of the attack mission, operating time of the insider (employee or device), and the latent period since the insider has been compromised, respectively. In addition, $\epsilon$ denotes a scalar of the system noise level $P_{n}$, where $P_{n} \in [0, 1]$. In particular, $\epsilon > 1$ represents conservative opportunistic attacks, while $\epsilon < 1$ represents aggressive ones. From (5), we see that the higher the system noise level $P_{n}$ is, the larger the attacking probability $P_{att}$ is, and vice versa. This reveals that opportunistic attackers are good at hiding themselves based on the system noise level. As a result, it is usually hard to detect these attacks.

### 4 Existing detection countermeasures for security threats in smart grid

In this section, we give an overview of state-of-the-art threats detection approaches in smart grid, including signature-, anomaly-, and specification-based approaches. These detection approaches are all grounded on data forensics and network surveillance techniques, for example by comparing the collected measurement data with the predefined rule specifications, or matching the traffic flow patterns with those stored in databases. Features of different detection approaches usually vary a lot, resulting in a wide variety of superiorities and limitations while solving threats detection problems in smart grid. We summarise and compare the mechanisms, advantages, and disadvantages of the these approaches in Table 2.


Table 2: Comparison of existing threat detection approaches

<table>
<thead>
<tr>
<th>Detection approach</th>
<th>Signature-based</th>
<th>Anomaly-based</th>
<th>Specification-based</th>
</tr>
</thead>
<tbody>
<tr>
<td>mechanism</td>
<td>data mining</td>
<td>statistical</td>
<td>normal behaviour</td>
</tr>
<tr>
<td>methods are utilised to detect possible attacks based on known attack patterns (with a pre-established blacklist)</td>
<td>statistical techniques are utilised to differentiate between the normal and abnormal behaviours (with a pre-established whitelist)</td>
<td>model is built based on either man-made or mined rule specifications to detect non-known-pattern attacks</td>
<td>normal behaviour is modelled based on the historical observations.</td>
</tr>
<tr>
<td>advantage</td>
<td>extremely efficient for known-pattern attacks; low false positive rate</td>
<td>capable of detecting both known-pattern and false positive rate</td>
<td>able to effectively detect not only known-pattern but also new-pattern attacks; low false alarms</td>
</tr>
<tr>
<td>disadvantage</td>
<td>inefficient for new-pattern attacks; high false negative rate</td>
<td>suffering from computational burden; high false positive rate</td>
<td>costly efforts to manually develop a complete and detailed set of legitimate rule specifications for complicated networks</td>
</tr>
</tbody>
</table>

4.1 Signature-based detection

Signature-based detection approaches, also known as misuse detection approaches, aim to look for patterns of malicious behaviours by matching the observed runtime features with known-attack signatures from a pre-established database. These type of approaches have been widely used in existing CPS communication networks. For example, in order to detect known suspicious or malicious communication activities in smart grid, Yang et al. [18] utilised a signature-based detection scheme by specifying a line of attack behaviours (e.g. spontaneous messages storm, unauthorised interrogation commands to a server, reset process command from unauthorised client etc.), and matching user profiles with these attack behaviours to identify potential cyber attacks.

The main advantage of signature-based detection approach is the high efficiency for known-pattern attacks detection, as it can quickly respond to known attacks with the blacklist database in hand. As a result, this type of approach usually maintains significantly low false positive rates. However, since the pre-established database cannot accommodate all the attack signatures, the major disadvantage is obviously the inefficiency for new-pattern attacks detection, usually yielding high false negative rates.

4.2 Anomaly-based detection

Anomaly-based detection approaches look for deviations from a pre-established database of normal behaviour profiles via statistical methods. These normal behaviour profiles are established based on a collection of historical observations (unsupervised) or a set of training data (semi-supervised). Anomaly-based detection technique is valuable to thwart malicious activities in CPS communication networks. For instance, Faisal et al. [19] employed an anomaly-based detection system using data mining techniques to secure the advanced metering infrastructure of the smart grid. Unlike legacy static data mining techniques, they opted to stream data mining techniques, which they believe is a more realistic approach in real-world applications.

The key advantage of anomaly-based approaches is that they are capable of identifying new-pattern attacks, as they aim to only differentiate normal and abnormal behaviours regardless of new- or known-pattern attacks. Therefore, this approach eliminates the efforts for collecting as many known attack patterns as those for signature-based approach to keep the database up-to-date. However, the major disadvantage of this category is the susceptibility to false positives. Since the user's normal behaviours may change over time under different circumstances, thus some of these changed behaviours may easily be considered as abnormal. Moreover, this type of approaches may also suffer from heavy computational burden while matching the observed behaviours with nearly all the normal behaviour patterns in the database.

4.3 Specification-based detection

Specification-based detection approaches look for abnormal behaviours by comparing the actual behaviours with the pre-built normal behaviour model, which is based on a set of either man-made or mined rule specifications. Unlike anomaly-based detection approaches which compare to specific limited normal user profiles, specification-based detection approaches formally define legitimate specification model and identify potential attacks when deviations from this behaviour model is observed. One key advantage of specification-based detection approaches is that they can not only be utilised to effectively detect known-pattern attacks but also new-pattern attacks, with low false alarms. Its major disadvantage, however, is the efforts cost to manually develop a complete and detailed set of legitimate specifications in a complicated network.

BLITHE [8] is a representative work of specification-based detection approaches for smart grid communication systems. In BLITHE, PMUs are assumed to be the possible compromised IEDs that can launch false data injection attacks. To identify the potential compromised PMUs, BLITHE first employs a set of manually developed normal rule specifications for the measurement data, which are then used to build up a normal behaviour model. With this model, BLITHE is able to assess the compliance degree of each PMU. By comparing the compliance degree with a pre-defined threshold for benign PMUs, BLITHE can then identify whether a PMU is compromised or not and further whether a false data injection attack is launched.

5 Proposed framework

In this section, we elaborate our proposed insider threats modelling, analysis, and detection framework. This framework is designed based on stochastic Petri net and behaviour rule specifications (see Fig. 3 for the proposed stochastic Petri net model). A stochastic Petri net [20] is an extended Petri net, where the time intervals of state transitions are characterised by random distributions and usually exponential distributions are used. Stochastic Petri nets have the advantage of being easily transformed to discrete time Markov processes, allowing further steady-state analysis. In our proposed framework, the stochastic Petri net is used to model the behaviours of insiders, analyse the steady-state risk degree of the whole system, and detect the malicious insiders including both intentional and unintentional ones.

5.1 Introduction of the proposed stochastic Petri net model

A basic Petri net can be described as a 4-tuple \((P, T, F, M)\), where \(P\) is a finite set of places, \(T\) is a finite set of transitions (or actions), \(F \subseteq (P \times T) \cup (T \times P)\) is a finite set of input and output arcs, and \(M\) is a finite set of markings denoting the number of tokens in each place at a time instant. Here, a token in a place represents an object that holds a specific condition or the occurrence of a specific event that the specific place defines. Tokens can be transferred from one place to another when a transition is fired once a specific condition holds or the occurrence of a specific event. In our proposed stochastic Petri net model, there are three places and five transitions, which are summarised in Tables 3 and 4.

5.2 Behaviour rule specifications

The behaviours of insiders are the essential evidences to assess their operating status. In our framework, we consider two categories of insider behaviours: suspicious and malicious. Specifically, suspicious behaviours are unauthorised behaviours or
Since there are many existing studies introducing the construction behaviours or prohibited acts that can cause great damages or have high risks of contributing to significant impacts. Correspondingly, performing a malicious act as $R_{MI}$ describes leads to firing of transition $T_{BM}$. Then, a token from place $P_{BI}$ is transferred to place $P_{MI}$. The marking now is updated from $M_t = \{N-1,1,0\}$ to $M_t = \{N-2,1,1\}$.

- **The third event:** In the third event, we consider the utility employee in place $P_{SI}$ visiting an unsafe or prohibited website after the second event. In this case, $T_{SM}$ is triggered, and one token from place $P_{SI}$ is transferred to place $P_{MI}$. The current marking is updated from $M_t = \{N-2,1,1\}$ to $M_t = \{N-2,0,2\}$.

- **The fourth event:** In the fourth event, we consider that a detected malicious insider in place $P_{MI}$, e.g. the system operator in the second event, is fired by the company and a new operator is hired. In this case, ‘resetting’ an insider fires transition $T_{R}$, pulling one token from place $P_{MI}$ to place $P_{BI}$. The current marking is updated from $M_t = \{N-2,0,2\}$ to $M_t = \{N-1,1,0\}$.

- **The fifth event:** In the fifth event, we consider that a suspicious insider in place $P_{SI}$ is observed with long time benign behaviours. In this case, it is reasonable to conclude that this insider was considered being suspicious because of an unintentional suspicious act and now it can be cleared of suspicion. In this way, transition $T_{C}$ is fired and a token from place $P_{SI}$ is transferred to $P_{BI}$. Assume that the last time instant marking is $M_{t-1} = \{N-3,3,0\}$, then the current marking is updated to $M_t = \{N-2,2,0\}$.

As we can see, the proposed stochastic Petri net model can be used to model and analyse the insider behaviours, and detect suspicious or malicious insiders. Note that, a threshold $T_{th}$ can be defined based on expert knowledge or history experience as the monitoring time window of suspicious insiders. If no more suspicious or malicious behaviours of these suspicious insiders are observed, they can be considered as benign insiders. In addition, insiders in place $P_{MI}$ can be naturally regarded as being malicious. Close investigations or other further actions are recommended to clear such insider threats.

### 5.4 Steady-state analysis of system risk degree

As aforementioned, the stochastic Petri net is employed to achieve steady-state analysis.

#### 5.4.1 Definition of risk degree

In steady states of a discrete time Markov model transferred from the stochastic Petri net, the steady state probabilities for each state (i.e. marking here) can be obtained. In this way, the steady state numbers of tokens $N_0, N_1$, and $N_2$ in place of $P_{BI}$, $P_{SI}$, and $P_{MI}$, respectively, can also be easily harvested. Based on these data, we define a metric risk degree to assess the overall operating status of the whole communication system, which is given by

\[ \text{Risk Degree} = \sum_{i=0}^{2} \frac{N_i}{\sum N_i} \]
\[ R = 1 - \frac{w_b N_b + w_s N_s + w_m N_m}{N_b + N_s + N_m}, \]  
(6)

where \( w_b, w_s, \) and \( w_m \) are the impact factors denoting the contributions that each type of insiders (benign, suspicious, and malicious) can create on the overall system operating status, respectively.

To obtain the values of \( N_b, N_s, \) and \( N_m \), we need to calculate the steady-state probabilities of the underlying discrete time Markov model. Prior to that, we introduce the concept of firing rates first. The firing rate of a transition refers to the average number of tokens transferred through this transition in a unit time. We use \( \lambda = (\lambda_b, \lambda_s, \lambda_m, \lambda_{bs}, \lambda_{bm}) \) to denote the firing rates for transitions \( T_{BS}, T_{SB}, T_{MB}, T_{BM}, \) and \( T_{SM} \), respectively.

\[ \text{Let } A(\lambda) \text{ be the transition matrix of the Markov model obtained from the reachability graph and the firing rate vector } \lambda. \text{ Then, the steady-state probabilities } \pi = (\pi_1, \pi_2, \ldots, \pi_K) \text{ can be solved by} \]
\[ \sum_{i=1}^{K} \pi_i = 1, \quad \text{and} \quad \pi_i = \frac{M_i(P_{\text{BI}}) \cdot \pi_j}{\lambda_i}, \quad \text{for } i = 1, 2, \ldots, K, \quad \text{where } K \text{ is the total number of states in the Markov model.} \]
(7a)
(7b)

where \( M_i(P_{\text{BI}}), M_i(P_{\text{SI}}), \) and \( M_i(P_{\text{MI}}) \) denote the number of tokens in places \( P_{\text{BI}}, P_{\text{SI}}, \) and \( P_{\text{MI}} \) in state \( M_i \), respectively.

### 5.4.2 Learning of the transition matrix

To achieve steady-state analysis and obtain the risk degree, system defenders need to know the transition matrix \( A(\lambda) \). An initial transition matrix \( A_0(\lambda) \) can be learned based on historical data by Algorithm 1 (see, Fig. 4). Then, given new observations of insiders’ behaviours, the real-time transition matrix \( A(\lambda) \) can be obtained also by using Algorithm 1 with an updating strategy, such as periodically updating it in a fixed unit time \( T_0 \).

To help learn the transition matrix, each piece of an insider’s behaviours of interest can be recorded in the following format in chronological order (see Fig. 5):

- Each insider is assigned with an Insider_ID. Behaviours of interest (denoted by Behaviour_of_Interest) includes proactive behaviours breaking suspicious or malicious rules and passive behaviours being cleared of suspicion or reset to benign. Place_of_Origin and Place_at_Present denotes the current place this insider should stay. Insider_Score evaluates the overall anomalous level of an insider, taking values from 0 to 2.
- The value of Risk_degree is defined as the increment of Insider_Score (e.g. 0.2) for insiders in place P_SI breaking suspicious rules.

### 6 Numerical results

In this section, we conduct simulations to evaluate the effectiveness of the proposed framework using PIPE2 and MATLAB. Specifically, we conduct two groups of simulations, where the first group is to evaluate the relationship between the risk degree \( R \) and the firing rates (i.e. \( \lambda_b, \lambda_s, \lambda_{bs}, \lambda_{bm}, \lambda_m \)) and the second is for the relationship between the risk degree \( R \) and the insiders handling factors (i.e. \( w_b, w_s, \) and \( w_m \)).

In the first group of simulations, we fix \( w_b = 1.0, w_s = 0.5, \) and \( w_m = 0.1 \) for the impact factors. Note that we would like to evaluate the relationship between risk degree \( R \) and the insiders behaviours using the proposed stochastic Petri net model, thus we also fix \( \lambda_{bs} = 0.15 \) and \( \lambda_{bm} = 0.2 \) which represent the risk handling capabilities of a system. As Fig. 6 shows, the numerical results of \( R \) versus \( \lambda_b \) with various values of \( \lambda_{bs} \) are plotted. It can be seen that, given \( \lambda_b = 1.0, \lambda_s = 0.5, \lambda_{bs} = 0.1, \lambda_{bm} = 0.2, \) and \( \lambda_m = 0.1, \) the risk degree \( R \) of a communication system increases as \( \lambda_b \) grows.

Fig. 4 Algorithm 1: transition rates learning algorithm

<table>
<thead>
<tr>
<th>Insider_ID</th>
<th>Behaviour_of_interest</th>
<th>Place_of_Origin</th>
<th>Insider_Score</th>
<th>Place_at_Present</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 5 Reference data format for each piece of an insider’s behaviours of interest

and considered as benign. \( \Delta_s \) is defined as the increment of Insider_Score (e.g. 0.2) for insiders in place P_SI breaking suspicious rules.
In addition, we see that given a fixed value of $\lambda_a$, a larger $\lambda_{sm}$ leads to a higher risk degree. It is natural that, the more benign insiders conducting malicious acts, the higher the system risk degree. Conducting malicious acts are increasing, there will be more insiders transferring from P_SI to P_MI, leading to an increased system risk degree. Also, we can see that a given a fixed values of $\lambda_{sm}$, the larger the $\lambda_{bm}$, the higher the system risk degree. Intuitively, if the number of benign insiders conducting malicious acts are increasing, there will be more insiders transferring from P_BI to P_MI, leading to an increased system risk degree.

When it comes to the second group, we carry out simulations to evaluate the relationship between the risk degree $R$ and the impact factors. In this group, we set $\lambda_b = 0.2, \lambda_{sm} = 0.1, \lambda_{bm} = 0.05, \lambda = 0.15$, and $\lambda = 0.2$. Hence we have $N_b = 24.52743, N_s = 3.47627$, and $N_m = 1.99631$ in steady states. Limit the values of $w_i$ and $w_{in}$ to the range of $[0, 0.5]$. The numerical results of the risk degree $R$ versus $w_i$ and $w_{in}$ are plotted in Fig. 8. It can be easily seen that large values of $w_i$ and $w_{in}$ always result in high risk degrees. This is because that, according to (7b), large values of $w_i$ and $w_{in}$ lead to small values of $w_i$, contributing to high risk degrees with given parameter settings. Particularly, if $w_i = w_{in} = 0$, then $w_i = 1$. It means that in this case, the risk degree is only determined by $N_b$, and consequently $R = 0.182$, the lowest risk degree with given parameter settings. In addition, if $w_i = w_{in} = 0.5$, then $w_i = 0$. It means that in this case, the risk degree is only determined by $N_s$ and $N_m$, and $R = 0.909$, the highest risk degree with given parameter settings.

7 Conclusion

In this paper, we investigated the insider threats in smart grid communication systems. The architecture of a smart grid and its specific security requirements were first introduced. Then, we developed a taxonomy of existing insider threats, and investigated the corresponding attack prototypes. In addition, a summary and the comparison of existing insider threat detection approaches for smart grid communication systems were presented. Most importantly, a novel hybrid insider threats modelling, analysis, and detection framework, which is based on stochastic Petri net and behaviour rule specifications, was proposed for smart grid communication systems. In future work, we aim to conduct research studies on fine-grained access control and granular data auditing to help prevent and mitigate insider threats in smart grid communication systems.

8 References


Author Queries

Q Please make sure the supplied images are correct for both online (colour) and print (black and white). If changes are required please supply corrected source files along with any other corrections needed for the paper.

Q All equations too long to fit within a single column are automatically floated to the bottom of the page and a '(see equation below)' or '(see (equation number) and (equation number)' left in their place. This is done automatically by our new XML transform, so please specify in your corrections very clearly where they should be broken as your paper is edited by non-experts.

Q1 Please check the edits made to the sentence 'Customers with ...'.

Q2 Please confirm the edits made in the sentence.

Q3 We have changed the algorithm 1 to figure 4 and figures have been renumbered accordingly as per style. Please check and approve.

Q4 As per journal style references are renumbered in the text and reference list. Please confirm.

Q5 Please provide page range in Ref. [3].

Q6 Please provide volume number and page range in Ref. [6].

Q7 Please confirm the inserted volume number in Ref. [21].

Q8 Please provide the volume number for Ref. [21] and is listed in the reference list but not cited in the text. Please cite in the text, else delete from the list.