Anomaly Detection Methods for IIoT Networks

Luying Zhou, and Huaqun Guo
Institute for Infocomm Research, A*STAR
Singapore
{lzhou, guohq}@i2r.a-star.edu.sg

Abstract—IIoT networks are different from general IT networks such as office or business networks where multiple various types of applications, protocols and traffic profiles are presented, and the cyber security challenges are more on protecting data confidentiality and integrity than on network availability. IIoT networks have special features and face unique challenges in defending against cyber-attacks. This paper briefly describes the requirements and challenges in IIoT network security, and presents an overview of the existing network anomaly detection methods. The paper further presents other anomaly detection methods that are specifically applicable to IIoT networks, as those methods exploit the deterministic features of the physical world in detecting the anomalies in the observed data of digital world.

Keywords - Anomaly Detection, IIoT Networks, Specification, Modeling

I. INTRODUCTION

An interpretation of Internet of Things (IoT) is that it represents the interconnection and communication among heterogeneous entities, where the entity refers to a human, an intelligent electronic device, or potentially anything that may request/provide a service [1, 2]. In the industrial environment, IoT is termed Industrial Internet of Things (IIoT), where the things can be specific industrial devices such as controllers, sensors, processors, actuators, and mechatronics that are interconnected. IIoT greatly enhances the automation and productivity in critical industrial sectors, e.g., manufacturing, processing, and transportation. IoT opens a completely new dimension to security, as the increased connectivity exposed the once closed/isolated system to cyber attackers [3, 4]. Industrial systems have strict performance and availability requirements, and maintaining uninterrupted and safe operation is often the top priority even when the system is experiencing cyber-attack.

IIoT faces the attack threats, such as spoofing, sniffing, replay, packet injection, man-in-the-middle attacks, and others similar to the attacks in the IT systems, but the consequences could be very different. In the IIoT, the attack threat moves from manipulating information to controlling actuation, with the serious consequences of disrupting operation or endangering safety.

Existing security solutions may not be adequate for addressing the challenges in IIoT. For example, the perimeter based protections (i.e., devices are under protection of firewalls), and cloud based defenses (i.e., computation resource intensive functions are moved to the cloud) are not suitable for the IIoT with the real time response and intolerable disruptions’ requirements.

Anomaly detection is critical in defending systems and networks against malicious activities. A number of detection mechanisms have developed to tackle the network attacks. There are many good survey work on anomaly detection methods, e.g., [5, 6, 7, 8], but few considered the applicability to the IIoT networks. IIoT systems have the restrictions that should be considered, and the features that could be utilized in anomaly detection.

This paper presents an overview of the well-known anomaly detection techniques, and discuss the anomaly detection techniques specifically suitable for the IIoT systems.

II. IIOT NETWORK SECURITY CHALLENGES

IIoT network security vulnerabilities are inherited from the general IT networks, and are at risk to a variety of attacks. Attacks can be grouped as passive and active attacks. Passive attacks are hidden and generally undetectable, such as eavesdropping and traffic analysis. Active attacks involve packet dropping, packet injection, and interfere with the network normal operation. The active attacks, e.g., malware infection, denial of server (DoS), unauthorized access and fraudulent packet injection attacks, are generally detectable. In the paper the detection methods for detecting active attacks are discussed.

In the following, the characters and aims of some active attacks are briefly summarized [3, 5].

- Malware (virus, worms, Trojan) attacks: Infect the computer systems and damage system files or compromise the system by controlling it. Malware could appear like a safe software and be downloaded to the system, or be injected to the system by an attacker who exploits the security vulnerabilities in the operation system or programs.
- DoS attacks: Consume system or network resources resulting in resource unavailability. In the distributed DoS (DDoS) attack, a large number of devices are compromised and employed to send high volume of traffic to networks or servers to render the resource unavailable to legitimate users.
- Unauthorized access attacks: Probe computers or network for finding vulnerability; sniff or intercept packets for information collection.
- Fraudulent packet injection attacks: replay captured packet, send forged packet or manipulated packet to interference or sabotage the operation.

To protect the system from the cyber attacks, various countermeasures have been developed and deployed. Some of the countermeasures are discussed in the following.
- **Embedded security**: Only authentic software should be used and installed in devices, and devices must be authenticated and authorized to be connected into the network to perform assigned functions.

- **Access control**: Policies are set for resource access control, and for user authentication and authorization. The privileges of users, and applications are defined by the policies to gain access to services and resources.

- **Cryptographic approach**: Message confidentiality and integrity are protected by encryption.

- **Firewall**: Inspect the packets for traffic control. The firewall filters the packets based on predefined rules or attack signatures such that any known network attacks are able to be detected.

- **Updates and patches**: the firewall rules will constantly be updated once the signature of newly discovered attack is identified; and the system or network software is continuously patched to fix a newly discovered program flaw.

The timing and opportunities for updating the hardware and software in the case of IIoT systems can be severely limited, and the attack defense solutions that simply shut down a potentially compromised system, or replace its components and subsystems, will not be suitable for protecting the IIoT systems considering the non-interruptive operation restriction.

The various security measures, such as the cryptographic methods, i.e., encryption key enabled data confidentiality, data and user authentication, and data integrity, and access control and firewall approaches, could protect the system from many types of attacks. However, an attacker could still manage to launch attacks successfully against the system. DDoS attack and insider attack are some notable examples of such attacks even if these security measures are in place but could fail to prevent them.

Network anomaly detection will be a second defense line to protect the systems, and we will present it in the following.

### III. ANOMALY DETECTION METHODS

As the network attacks aim compromising the system’s information confidentiality, integrity and resource availability, they will in some way deviate from normal network operation and exhibit abnormal behaviors. There are surveys on the developed network anomaly detection methods, e.g., in [5, 6, 7, 8]. The work in [8] provided a comprehensive survey of the research on anomaly detection, where the detection methods are categorized based on their underlying techniques. The work in [5, 6] presented an overview of network anomaly detection methods, and classified methods according to computation techniques used. Machine learning and data mining methods for cyber security anomaly detection were surveyed in [7].

In general, anomaly detection aims to find patterns in data that do not conform to expected behavior, and there are three types of anomalies classified as below in [8].

**Point anomaly**: An data instance of an individual is anomalous with respect to the rest of data [8]. This is the simplest type of anomaly. For instance, assume that the data of a control command is defined using only one feature: amount of the value, then a command with a data value out of the predefined range will be a point anomaly.

**Contextual anomaly**: A data instance which is only anomalous in a specific context is termed as a contextual anomaly [8]. The context is induced by the structure in the dataset and has two sets of attributes: (a) contextual and (b) behavioral attributes. The anomalous behavior is identified within a specific context. For example, a command with data value within the predefined range could be contextual anomaly if the command is issued during predefined no-show time period.

**Collective anomaly**: A collection of related data instances which is anomalous with respect to the entire data set is termed as a collective anomaly [8]. The individual events are normal when they occur alone, but the collection of events is an anomaly. For example, a single TCP connection request is normal, but multiple such requests received continuously from the same source could be a DoS attack, which is anomaly.

Network anomaly detection deals with detecting anomalies in network traffic data, using devices or software applications to monitor and analyze network traffic to detect malicious activities. Network anomaly detection methods are grouped into two broad categories: Misuse-based detection vs. anomaly-based detection.

**Misuse-based (or Signature-based)** detection attempts to identify known patterns of attacks (attack signatures). A misuse-based detection system detects attacks by comparing the gathered information with a set of attack signatures to search for a specific attack that has already been known. It can be certain to detect an attack if a signature of the attack has been specified. But, the method is unable to detect previously unknown attacks, as it cannot define attack signatures for the undiscovered attacks. Pre-defined malicious activity patterns or attack signatures are needed to detect attacks. An example of applications of signature based approach is the firewall where rules describe the characteristics of a cyber attack are predefined and any packets that match with these characteristics are blocked.

**Anomaly-based detection** relies on the assumption that attacks will result in behaviors that do not conform to the previously observed normal behaviors. An anomaly-based detection method models the normal network behavior, and identifies anomalies by comparing the current behavior with the modeled behavior for notable deviations from the modeled behavior. Anomaly detection is capable of detecting previously unseen attacks which exploit unaware security flaws. However, the drawback of this method is its high false positive detection rate due to the difficult to set up the right thresholds for identifying the anomalies.

Anomaly detection methods can be classified based on their underlying computation techniques used:

- statistical based
- knowledge based (rule, expert system, and logic based)
- machine learning based (Classification based, clustering and outlier based)
In statistical based method, from the captured network traffic an activity model of the network traffic is created that characterizes the stochastic behavior of the monitored network parameters. This model serves as a description of normal activity, and is used to test the currently observed instances to decide if an instance fits in with this model. Bayesian networks presented in [9] are capable of detecting anomalies in the circumstance of multiple classes. The work in [10] introduced an anomaly detection method that applies Bayesian networks, and the Bayesian decision process enhances the detection accuracy and greatly reduces false alarm rate. In [11], the activity profiles are represented by a number of probability density functions, and the test to determine the similarity of the monitored variables is based on Kolmogrov-Smirnov (K-S) statistic test. The important property of the K-S is that it is distribution-independent and thus a normal distribution could be assumed for the traffic since the underlying traffic distribution is usually not well known.

Knowledge based methods examine the observed data against a set of predefined rules. The methods detect the attacks based on available prior knowledge so that any occurrence of those known attacks could be detected. The methods examine activities of known attacks, by comparing with pre-identified attack representations. Expert systems and rule-based [12, 13] approaches are some examples of knowledge-based methods. Snort [14] is a popular rule-based intrusion detection system. It examines the observed data against the defined rules to detect the anomalies.

Machine learning based methods are to validate the collected data patterns against an establish model. The methods generally involve three phases, i.e., training, validation, and testing (see, e.g., [5, 6] and references therein). With machine learning approach, an observed behavior model of the system is first learned and established by training it with reference data (previously observed data), then the model is utilized to verify the collected data to find any deviations. Machine learning is mainly for classification. Classification method establishes an explicit or implicit model and categorizes network traffic patterns into several classes. SVM (support vector machines) is a classification based detection method. The method could be applied to detect instances that do not belong to the learned class [15]. Clustering method assigns instances into groups such that the instances in the same group have similar properties (see, e.g., in [16]).

There are practical issues in applying machine learning methods for network attack detection. As discussed in [17], machine-learning methods perform fundamentally much better at finding activities that fit to the established model than at identifying activities that do not comply with the normal system behavior. In other words, machine-learning methods are more suitable for finding activity that is similar to something previously observed, rather than for finding novel activities, as the established model may not fully cover all normal activities. This is why machine learning finds wide and successful applications in commercial world, e.g. in product recommendation systems [18] and natural language translation [19], but not for anomaly detection in operational systems.

IV. ANOMALY DETECTION METHODS SPECIFIC TO THE IIOT NETWORKS

IIoT networks are likely to run a limited number of applications and protocols, and have static network configurations and regular traffic patterns, which are very different from general IT networks having a large and changing number of applications, protocols and users. And furthermore, IIoT systems have a scheduled operation procedure and predictable operation process states governed by the underlining physical reality. Those unique deterministic features in the IIoT networks could be exploited for assisting in detecting malicious network activities that change the process states. An anomaly detection model could be built on the deterministic features and applied to examine the execution procedure of the protocols, the pattern of communication, or the states of the physical operation, to discover the causes that result in any deviations from the model. Two types of such methods, specification based anomaly detection and physical state dynamic estimation are discussed here.

A. Anomaly detection methods based on specifications

Specification based anomaly detection method constructs a model according to the specifications, e.g. program execution specification, protocol specification, and operation procedures, which describes the legitimate (inherent rather than observed) system behavior, and detects attacks as deviations from the modeled behavior. It avoids high false positive alarm rate, as in the machine learning approach, by wrongly examining legitimate but unobserved behavior.

The idea of specification-based anomaly detection was first presented in [20], where the detection is based on program specification description in which the intended behavior of security-critical programs is defined. In the specification based approach, allowed operation sequences of the programs are described, and the actual execution of the programs are monitored and compared to detect deviations from the specification description. This specification based approach detects attacks through detecting anomaly behavior resulted from the attacks. Such approach has the capability of detecting attacks which are unknown previously.

It is time consuming to develop detailed specification description to fully cover the inherent behavior in order to achieve low false positive alarm rate. The work in [21] described a specification description language, an extended finite state automata, and developed the description according to protocol specifications. In the work the protocol specifications are described via the developed state machine, such as for network layer and transport layer protocols, and further, based on the state machine, the statistical properties of state transitions for normal traffic are also presented. With such state machine models, the system behaviors are mapped to transitions of the state machine, and based on the transitions frequency the unusual behaviors can be detected.

A specification-based detection framework was presented in [22] to analyze an industrial control system protocol semantics and verify the confirmation of network packets to the protocol specification. The specification defines valid data
fields of the packets and communication patterns between different packets. The scheme inspects the network packets to ensure the data fields dependency and the values within the valid ranges in a single packet, and the correlations between inter-packets. Communication patterns which deviate from the described specification will indicate malicious activities, such as system misconfigurations, replayed packet or DoS attacks.

The work in [23] described a specification-based intrusion detection sensor to monitor advanced metering infrastructures communication traffic. The solution relies on protocol specifications, activity constraints, and security policy to detect anomalies. A leveled structure of state machines was developed to present a legitimate activity profile for the expected behaviors from the protocol specifications, including the state machines presentations for the activities at the device level, network level and application level. Traffic features such as source and destination nodes, session duration distributions, data volume of each session, and service requests sequence are extracted. It defined security policy and rules to detect any activity deviation from the normal behaviors. A formal verification framework was adopted that consists of the network model, protocol specifications and security policy, and the monitoring rules to check the validities of the flow format, sequence of services, the flow origin, use case frequency and others, to detect the activities that violate the security policy.

Industrial control networks are likely to have fixed structure and devices, and regular communication patterns, thus a specification-based monitoring would be more suitable for such networks. [24] presented a protocol-level model for characterizing an industrial control system protocol, i.e., Modbus TCP protocol, based on the application protocol specifications. Such model is able to detect packets that contain an unspecified function code. The work also modeled the expected communication patterns among different types of devices for the Modbus TCP protocol. For example, a Modbus server’s communication patterns show that it only responds to Modbus TCP requests, and does not initiate TCP connections itself.

In [25], the unique characteristics of communications in the control system of power systems are discussed. The communication between certain senders and receivers show patterns of regular message exchanges during specific time period, and with constraints in message format and contents. Such determinism in industrial control systems is leveraged in developing anomaly detection solutions, where communication patterns, such as data value, message transmission constraints and communication behaviors, are extracted from the system specification file. The extracted information is further constructed into rules for validating the confirmation of the packets.

B. Anomaly detection methods with physical state dynamic estimation

The underlining process of the physical system is controlled generally by the operation principles, so that the process state of the physical system is predictable. A model based anomaly detection method, which models the specific normal physical operations according to the physical dynamics, can serve as complementary to the traffic information anomaly detection method, and be able to detect the cyber attacks from the anomalous states that deviate from the physical operation models.

The work [26] proposed a cyber-physical system attack resilience framework that leverages known mathematical description of physical domains, and information on forecasts and historical data, to validate the correlation between the predicted and the measured data. There are various attacks that could compromise the integrity of the measurement and control data, and the traditional methods that solely depend on cyber information process will not be effective to detect the attacks. As the security of the measurements and control commands are critical to the cyber-physical system normal operation, the authors stated the necessity to apply the model-based anomaly detection approaches that employ the underlying physical domain knowledge to assist in detecting the data integrity attack.

[27] described how to use hydrodynamic models to detect physical fault and cyber attack on a water distribution network. Based on the well known hydrodynamic principles, the water system is modeled in state and measurement equations along with unknown inputs that represent the disturbance on states and on sensor measurements. The model can reflect the effects on the system caused by abnormal events such as sensor-actuator faults or water leaks. But as the authors pointed out the cyber attack detection depends only on the physical model is not sufficient, e.g., it is difficult to detect the attack if multiple sensor measurements are compromised.

For the goal of detecting the attack by a sophisticated attacker who could exploit system vulnerabilities and inject legitimately appeared malicious control commands to sabotage the power grids, [28] proposed to combine the knowledge of physical infrastructure in power grid and cyber information to detect the attack. The attack could be detected through the estimated execution consequences of control commands. In the approach, the packets are inspected based on the protocol specification and the critical control commands are extracted to perform a simulation run through the power system operation equations. From the simulation, the system states resulted from executing the control commands are estimated and compared with trusted measurements to detect the attacks.

An anomaly detection method based on a representation of the states of the protected system, and on the prediction of the control commands on the state evolution of an industrial control system was presented in [29]. The method defines a “critical state” at which the system will be unsafe or insecure, and a “predicted state” toward which the system will evolve, and by comparing the two states the method can detect whether any anomaly activities will bring its state into critical. The system operation states and the representative system model are maintained in a firewall through capturing the communicated information and specifically acquiring information from the monitored system. When receiving a control command, the firewall will evaluate the effect of the command on the system state utilizing the maintained system information. If the command is likely to cause the system state
into a “critical state,” then the command will be blocked by the firewall without forwarding it for actual execution.

The Automatic Generation Control (AGC) is a power system frequency control application that receives power flow and frequency measurements from field devices, and based on them to produce control signal to maintain the system operation and reliability. The work in [30] presented a domain-specific model-based anomaly detection algorithm that verifies the integrity of received measurements against forecasts obtained from equations that control the operation of the underlying physical system. In the algorithm the received measurements are used by AGC to calculate the area control error (ACE), and to generate control commands accordingly. The ACE is checked for anomaly by the anomaly detection engine, based on statistical characterization of forecasted ACE values. If the ACE identified an anomaly, the AGC will use forecasted ACE instead. The system is able to perform using forecasts when measurements are detected as unreliable.

As the effect of cyber attacks will likely be reflected through the system state variables, based on system model to get state estimations, the malicious activities and faults as well can be detected. Applying the model based approach alone to detect the attack has limitations, as the system model based approaches may not accurately estimate the system states as the models are built on approximations and the estimates are subject to unknown disturbances. On the other hand, pure cyber approaches face challenges in guaranteeing security of the protected system, as there are risks of insider attackers and zero day vulnerabilities. The work of [31] described a model-based method to secure smart grids. The method complements cyber security approach with physical system modeling to achieve enhanced system security. Based on the system state dynamics equations, the approach estimates the system states, and compares that with collected measurements, and can effectively detect compromised measurements.

[32] discussed and tested the control theory modelling based fault detection and cyber security based anomaly detection approaches on a water infrastructure testbed. The work demonstrated that the two approaches can effectively detect many faults and attacks, and also pointed their limitations. In an experiment under a concurrent physical fault and cyber attack, the cyber attacker could evade the detection of the control theory modeling approach. So it is crucial to combine the state estimation from the physical dynamic modeling approach with the data analysis from the cyber security approach to enhance the cyber security for the industrial control system.

V. OPEN ISSUES AND CHALLENGES

IIoT networks have specific features which are distinguishable from that of general enterprise or business networks. The underlining physical processes are governed by the physical dynamics of the system and the processes tend to be deterministic. The predictability of the states of the physical processes should be exploited in the detection of anomalies resulted from cyber attacks or faulty operation. Cyber information anomaly detection algorithms generally do not concern the physical aspects of systems, while mainly focus on the collected data analysis to detect the deviations from earlier observations, such as statistic and machine learning approaches. In the IIoT systems, which are generally relevant to physical operational processes, well known system dynamics are available for expressing the state evolution of the physical process, and could be applied for assisting in anomaly detection.

Cyber information anomaly detection approaches and system dynamic modeling approaches have their own advantages and limitation as well. Cyber security approaches, through information analysis, may not effectively detect certain attacks, e.g., insider and zero day attacks. The physical modeling approaches may not accurately estimate the real world system states as the system could be modeled with approximations and also disturbed by interference. And the specification based approach has difficulty in efficiently or precisely describing the protocol specifications or the communication patterns. Hence, a more effective way would involve and integrate multiple approaches, including the machine learning, specification description, and physical system modelling approaches, and complement the approaches based on physical system characteristics with cyber security analysis approaches to address the anomaly detection challenges.

VI. CONCLUSION

This paper presented a survey on the anomaly detection methods, and specially on the anomaly detection methods that are more suitable to the IIoT systems, which are not well addressed in existing survey work. The methods based on specification descriptions and physic process modeling take the operation dynamics of underlining physical system into consideration, and detect the anomalies from the inherent system behaviors. Those methods can be complement to cyber security methods which detect the anomalies from the observed system behaviors. An integrated cyber security and physical state estimation approach will be more effective for IIoT network anomaly detection.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation (NRF), Prime Minister’s Office, Singapore, under its National Cybersecurity R&D Programme (Award No. NRF2014NCR-NCR001-31) and administered by the National Cybersecurity R&D Directorate. The special thanks are also given to SMRT Trains Ltd for providing domain knowledge and technical support.

REFERENCES


