Behavior Based Anomaly Detection Model in SCADA System

Xiaojun Zhou^{1,2}, Zhen Xu¹, Liming Wang¹, Kai Chen¹, Cong Chen^{1,2}, Wei Zhang^{1,2}

¹State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, 100195 E-park C1 North, No. 80 Xingshikou Road, Haidian District, Beijing, China

²School of Cyber Security, University of Chinese Academy of Sciences, 100049 No.19(A) Yuquan Road, Shijingshan District, Beijing, P.R.China

Abstract. With the arrival of Industry 4.0, more and more industrial control systems are connected with the outside world, which brings tremendous convenience to industrial production and control, and also introduces many potential security hazards. After a large number of attack cases analysis, we found that attacks in SCADA systems can be divided into internal attacks and external attacks. Both types of attacks are inevitable. Traditional firewalls, IDSs and IPSs are no longer suitable for industrial control systems. Therefore, we propose behavior-based anomaly detection and build three baselines of normal behaviors. Experiments show that using our proposed detection model, we can quickly detect a variety of attacks on SCADA (Supervisory Control And Data Acquisition) systems.

1 Introduction

With the continuous development of the industrial control system and the introduction of the concepts of "Industry 4.0" and "Internet +", the industrial control system is no longer an isolated and closed operating environment. Instead, the control system has been a combination of communication technology, computer network technology and industrial control technology. The industrial control system has gradually evolved into a standard system of openness, intelligence and interaction to enhance the production efficiency and support large scale of production. However, the increasing risk of information security is accompanied by it. In 2010, the earthquake-stricken network "STUXNET" attacked the SIMATIC WinCC monitoring system and SCADA system at the Iranian nuclear power station^[1], breaking the myth of the absolute security of a "closed" industrial control system. Very similar to the Stuxnet, duqu trojan mainly targets industrial control systems for stealing private information^[2]. Havex^[3], a malware specifically targeting the ICS / SCADA system in 2014, has the capability of disabling hydroelectric dams and overloading nuclear power plants; hackers have used it to attack industrial systems in Europe and the United States. On December 23, 2015, the Ukrainian power network was attacked by BlackEnergy's APT and eventually the system crashed, resulting in a massive power outage^[4].

SCADA system, as the core control system of ICS (industrial control system), faces the most serious security threats. After a lot of analysis, we found that SCADA systems mainly face two types of attacks, i.e. internal attacks and external attacks. However, at present, the security protection for SCADA lacks a very effective

security protection method. Since SCADA system attacks are inevitable, we need to combine the SCADA system features within itself to develop security methods. As a result, this paper presents a method based on the entity behavior for security protection, the results of experiments proved its efficiency and effectiveness.

The structure of this paper is as follows: Section 2 depicts the structure and characteristics of the SCADA system, analyzes the security threats faced by the SCADA system, and proposes a mechanism of security protection based on entity behavior. Part 3 details the security framework based on entity behavior and the role of each module, and analyzes the operating principles of each module. Section 4 uses experiments to validate our proposed framework for entity-based behavior. The 5th part is the conclusion of the paper and 6th part gives an introduction of future work. Finally, we make our acknowledgements in Section 7.

2 SCADA system model

The National Institute of Standards and Technology (NIST) defines and describes an industrial control system as such: Industrial Control Systems (ICSs) are a collective term for a class of control systems for industrial production that includes a supervisory control and data acquisition system (SCADA), Distributed Control Systems (DCS), and other small control systems commonly found in the industrial sector and critical infrastructure such as programmable logic controllers (PLCs). Now, let's have a close look at the SCADA system components.

* Corresponding author: author@e-mail.org

2.1. SCADA system structure

A typical industrial network^[5] shown in Figure 1, which is the reference model of Purdue University.

Fig. 1. The architecture of SCADA system.

The complete industrial control system consists of five parts: enterprise system, business planning and logistics system, site manufacturing operations, area supervisory controls, basic monitoring and control, and physical processes. Among them, enterprise system, business planning and logistics system are traditional IT systems that form the business information system of an enterprise. The remaining layers make up the on-site control system. SCADA system is the core control system of the entire ICS. On the one hand, it controls and dispatches the underlying on-site control equipment according to the production instructions of the upper level; on the other hand, it real-time monitors the production status of the industrial site and collects statistics to provide feedback for the upper control.

2.2 Threat analysis of SCADA system

Industrial control system is a complex system involving a variety of computer technologies and network technologies. It has both computer operating system level of industrial control software, monitoring programs, database systems, but also network systems involved in the network protocol and data packet processing mechanism. In addition, with the industrial control system docking with the external network, to access open Internet, it has also become an integral part of the Internet. The following describes major security threats faced by the SCADA system.

According to the sources, the security threats to industrial control systems mainly include external threats and internal threats (Figure 2). External threats are: APT attacks, Trojans, malware, viruses, etc., which may be politically relevant or may be related to industrial espionage.

Internal threats include: annoying employees (including internal attacks and data breaches), employee

misuse, software contractors, third-party integrators, and more. Due to the lack of authentication and encryption mechanisms that limit user activity in most industrial control systems, users have unfettered access to devices on the network, even modifying device configuration and operating parameters. Typical attacks include the incident of Maluqi, Australia Vitek Boden, former engineer of the technology service provider of the plant, was deliberately repaying for being dissatisfied with the renewal of his contract of work. A total of 1 million liters of untreated sewage were drained directly into the natural water system through storm drains. The United States Davis-Besse nuclear power plant was attacked by the Slammer worm^[6].

Fig. 2. Threats faced by SCADA system.

Fig. 3. Commands and operations will be reflected on entity behaviors.

A supplier provided application software to the server and established an unprotected T1 link at the back end of the nuclear plant's network firewall, through which the virus entered the nuclear power plant network. The United States Hatch nuclear power plant automatic shutdown event^[7], an engineer operates a computer of the plant business network (used to collect diagnostic data in the control network) for software updates to synchronize the business network and control data in the network, when the engineer restarts the computer, the synchronization program resets the data of the control network, the control system experiences such a sudden drop in reactor water storage reservoir that it automatically shuts down the entire unit. In view of the many threats to industrial control, it is necessary to take effective security measures to ensure the safety, security

and stable operation of industrial control systems. And we come to the conclusion that all these threats and attacks will be reflected on the behavior of devices (Figure 3).

Since SCADA system attacks are inevitable, there is a need for a method that can detect attacks in time and avoid attacks such as Stuxnet^[8]. Therefore, we propose a behavioral based anomaly detection mechanism.

3 Framework for behavior-based anomaly detection mechanism

In this section, we will describe the model in detail. The framework is shown in Figure 4.

Fig. 4. The framework of behavior-based anomaly detection.

The basic anomaly detection steps include: information collection, uniquely entity determination, constructing three kinds of normal behavior baseline from different dimensions and using the baseline for anomaly detection. Below we describe each step in detail.

3.1. Information collection

We employ passive information collection to avoid any possible system interference. The best solution is to have a transparent network snoop on the ICS system components. Passive recognition methods utilize PCAP files generated by tools such as Wireshark or direct on-line sniffers for data analysis. It does not inject network traffic and does not respond to upcoming messages, thus ensuring that ICS system operations are not interrupted. In addition, not all network flow data is valuable. The pre-process step filters ICS network session-independent data and dirty data (such as TCP retransmissions, duplicate ACK packets, etc.). Five basic conversational features were extracted and the eigenvalues were rated. The five basic conversational features include:

I. Source IP (S-IP)

II. Source Port (S-Port)

III. Target IP (D-IP)

IV. The target port (D-Port)

V. Unit interval length (1s) (SegSize)

Then we use machine learning method to process all the information.

3.2 Entity determination

This step we want to uniquely identify an entity. Device fingerprint is a series of device-related data that uniquely depicts a device. This information includes: device operating system, configuration information, operational behavior features, and more.

Generally, industrial control protocols (such as Modbus) provide the query function, which can be used to collect information. Moreover, the conversations in industrial control systems have significant stability and periodicity. Researchers can make fingerprints of industrial control systems by using data sources such as network traffic characteristics and interaction modes based on time dimension.

Then we will determine the relationship between entities and infer the topology. Industrial control systems have their inherent characteristics and drawbacks over conventional Internet and corporate LANs. First of all, compared with the traditional IT systems, industrial control devices in industrial control systems generally have a longer life cycle; second, the industrial control system has a stable network topology; Finally, in industrial control systems, the role of a single device is usually unique, with a fixed communication object.

Therefore, the network topology of the SCADA system can be reconstructed based on the traffic in the SCADA system.

3.3 Baseline construction

After each communication object is uniquely identified, combining the topological relations among different entities, we can construct a normal behavior baseline for each entity in the SCADA system.

The establishment of normal behavior baseline is divided into three aspects:

(1) Historical Baseline. It is based on the notion that a device's role and function are relatively fixed, and therefore today's behavior and historical behavior should have obvious similarities. If there is inconsistency between the two, you can be judged as abnormal.

(2) Peer Baseline. It is based on the behavior of peer devices for analysis. Multiple devices in a SCADA system will perform the same functions. If there is a large difference in behavior between devices of the same type that perform the same function, they can be determined to be abnormal.

(3) Partner Baseline. According to the feedback from the communication partner, if a certain machine suddenly bursts out frequently sending query information packets, it is considered as abnormal.

3.4 Anomaly detection

By building three behavioral baselines, you can quickly discover anomalous behavior. Moreover, based on historical data, three baseline information can be verified against each other horizontally and vertically, so as to ensure the authenticity and accuracy of alarms generated, and to reduce the proportion of false alarms and false negatives. The whole anomaly detection process is shown below in Figure 5.

Fig. 5. Anomaly detection process.

4 Experimental verification

Let's experiment with the behavior-based anomaly detection model we built.

4.1. Fake attack

Due to the lack of authentication of communication entities in SCADA systems, there is a large number of counterfeit attacks. We have installed a PLC simulation software in an experimental notebook, the laptop disguised as PLC and communicated with host computer. In this way, anomaly detection system installed on the host computer will quickly detect abnormalities. Because according to the uniqueness of the entity, you can determine that PLC is fake which is now communicating with the host computer.

4.2. Tampering attack

Tampering with packet attack is an attack that many attackers prefer. The main means of this kind of attack is to modify the instruction of the data packet, tamper with the data packet load, modify the measured value of the data packet and so on. According to the historical behavior, we can determine that such behavior is abnormal.

4.3. Logical disorder attack

This kind of attack is more difficult to defend because all the packets are valid and the communicating entity is legal, but we can detect abnormalities in time by recording the historical behavior of the communicating partners.

5 Conclusion

This paper first analyzes the structure of SCADA system and the security threats faced by SCADA system. Then it introduces the framework of behavior-based anomaly detection mechanism and introduces each module in the framework in detail. Based on the collected information, we construct three different normal behavior baseline from multiple dimensions and use these to detect the attack. Experiments show that the proposed detection model can find fake attacks, data packet tampering attacks and logical sequential attacks well.

6 Future work

Future work we intend to focus on the following two research directions:

(1). Establish a defense-in-depth system^[9] suitable for industrial control systems. According to the characteristics of industrial control system, the depth defense system will be improved so that it can be better applied to the security of industrial control systems^[10].

(2). Establish a kill-chain model^[11] specially for industrial safety. The current kill chain model is rough, ignoring a lot of details of attack. We intend to create a kill chain model that is specific to industrial control systems.

Acknowledgements

We thank our shepherds—Zhen Xu, Liming Wang in our research group, for providing insightful feedback of the draft that helped improve the final paper. We would also like to thank Kai Chen, Zelong Chen and Zhenbo Yan for their help in early discussions and providing insightful comments. This work was supported by Security Services for Informatization Applications Program, Institute of Information Engineering, Chinese Academy of Sciences, under grant No. XXH13505-02, for which we are grateful.

References

1. Langner R. Stuxnet: Dissecting a cyberwarfare weapon[J]. IEEE Security & Privacy, 2011, 9(3): 49-51.
2. Bencsáth B, Pék G, Buttyán L, et al. Duqu: A Stuxnet-like malware found in the wild[J]. CrySyS Lab Technical Report, 2011, 14: 1-60.
3. HENTUNEN D T. A: Havex Hunts for ICS/SCADA Systems [on-line][J]. 2014.
4. Case D U. Analysis of the cyber attack on the Ukrainian power grid[J]. Electricity Information Sharing and Analysis Center (E-ISAC), 2016.
5. Knapp E D, Langill J T. Industrial Network Security: Securing critical infrastructure networks for smart grid, SCADA, and other Industrial Control Systems[M]. Syngress, 2014.
6. Moore D, Paxson V, Savage S, et al. Inside the slammer worm[J]. IEEE Security & Privacy, 2003, 99(4): 33-39.
7. Krebs B. Cyber incident blamed for nuclear power plant shutdown[J]. Washington Post, June, 2008, 5: 2008.
8. Byres E, Ginter A, Langill J. How Stuxnet spreads—A study of infection paths in best practice systems[J]. Tofino Security, white paper, 2011.
9. Kuipers D, Fabro M. Control systems cyber security: Defense in depth strategies[R]. Idaho National Laboratory (INL), 2006.
10. Stouffer K, Falco J, Scarfone K. Guide to industrial control systems (ICS) security[J]. NIST special publication, 2011, 800(82): 16-16.
11. Greenert J, Welsh M. Breaking the kill chain[J]. Foreign Policy, 2013, 16.