POSTER: Obfuscation of Critical Infrastructure Network Traffic using Fake Communication

Sungho Jeon, Jeong-Han Yun, and Woo-Nyon Kim
The Attached Institute of ETRI
Daejeon, South Korea
sdeva14@gmail.com, {dolgam, wnkim}@ensec.re.kr

1. INTRODUCTION

Motivation. Attacks on critical infrastructure such as power plants, nuclear reactors, and other highly significant service institute have rapidly increased recently. One of the reasons these accidents on critical infrastructure occur again and again – despite the considerable attention and effort toward cyber security – is the absence of security device in Supervisory Control and Data Acquisition (SCADA) [2], a system used for controlling and monitoring critical infrastructure. Many critical infrastructures fully depend on the security device in an access point such as firewall. In this case, penetration or bypassing of the firewall means the security system of SCADA has been neutralized. Unfortunately, many critical infrastructures are facing this situation.

To address this problem, anomaly-based intrusion detection system (A-IDS) on SCADA [4] has been proposed instead of attack signature-based IDS (AS-IDS). A-IDS on SCADA creates a normal profile containing characteristics of periodic communications generated in SCADA and subsequently regards other communications as intrusion. Note, however, that A-IDS will be useless if the normal profile of A-IDS is revealed to the attacker through sniffing and analysis of communications on SCADA traffic. In the IT network, this attempt for IDS, mimicry attack, polymorphic blending attack [5] posing as normal communication so as not to be detected by A-IDS, and defense against such attack have been studied continuously.

Although research to defeat A-IDS through information extracting from SCADA traffic has yet to be conducted, such is the aim of this research. To our knowledge, there has been no research on the defense against an information extraction for malicious activity on SCADA traffic. Information extraction on SCADA traffic can make security device – even proposed recently – useless. We propose novel obfuscation method for SCADA network traffic using fake communication. Our method interferes with information extraction by attacker, such as revealing normal profile of A-IDS on SCADA.

Challenge. There were other studies that attempted to reveal information from network flow and traffic and defense against such threat on the IT network. They focus on revealing information and protecting information – such as which Internet sites are visited by the users and what they send and receive – to protect user privacy. These research studies can be categorized as "Web Fingerprinting Attack" [3] and "Network Obfuscation" [8]. Note, however, that these are not helpful to SCADA since there are two major differences between the IT network and SCADA. Futhermore, although next generation SCADA [6] provides more advance security function, it can not solve this problem.

- First, the content of communication in SCADA can be analyzed by the attacker. A research studied the IT network based on the hypothesis that the content of communication in the IT network is encrypted during transmission. Note, however, that there is no encryption for the content of communication in SCADA; thus, the method of packet padding, a dummy packet consisting of meaningless value, is useless, because such methods are easily identified and evaded by the attacker upon analysis. Even next generation SCADA which provides encryption for SCADA communication, attacker still can guess information from network traffic as "Web Fingerprinting Attack" [3].

- Second, after installation, a system in SCADA rarely changes because of the availability and stability of SCADA that most important properties of critical infrastructures. While belief of many researchers that the latest research result and system will be installed soon, to fully introduce next generation SCADA and high performance system takes at least several years in a real world. Because, most of the administrators of critical infrastructures hesitate to introduce new system. It needs more than few years to convince and to prove that new system works on existing system without any other availability and stability problem. Therefore, a method requiring little calculation, at the same time, does not have influence on existing system is required for the present to apply on SCADA.

Based on these observations, our goal is to interfere information extraction on SCADA traffic for malicious activity without modification or interference of normal communication in SCADA.

Contribution. We do not propose advance IDS on SCADA; instead, we suggest a obfuscation method of SCADA network traffic to interfere with information extraction. Our method can protect secret information that can be extracted from SCADA setwork traffic. Especially, our method can be used to hid the real normal profile of A-IDS even if attackers sniff and analyze network traffic. No matter how outstanding the proposed security device, such as A-IDS on SCADA is, it will be useless if the normal profile of A-IDS is revealed to the attacker. So far, there is no protection method for revealing information from SCADA traffic; thus, we propose a novel obfuscation method of SCADA network traffic.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
ACSAC '13 New Orleans, Louisiana USA
Copyright 20XX ACM X-XXXXX-XX-X/XX/XX ...$15.00.

2. PROPOSED APPROACH

Threat Model. In our threat model, the attacker cannot perform malicious activity despite compromising a certain host or system in SCADA due to A-IDS. Note, however, that the attacker has prior knowledge of A-IDS, i.e., which features it uses and how normal profile is created; Futhermore, attacker can sniff whole network traffic from certain compromised host, because SCADA is a connected network. Hence, the attacker can extract normal profile of the A-IDS through monitoring SCADA network traffic.

Key Idea. As mentioned earlier, the system in SCADA rarely changes; on the other hand, the network environment is upgraded continuously by an advance Ethernet device to replace an outdated device so as not to cause a network problem. For that reason, the actual amount of network bandwidth used in SCADA is much smaller than the maximum network bandwidth of current SCADAs. Thus, our approach is based on the idea of utilizing surplus network bandwidth instead of the overhead of system resource, i.e., CPU or memory, for each host in SCADA. Specifically, our approach is to make fake communications similar to normal ones so that they are not identified by the attacker. In our threat model, attacker can identify fake communication in host which they compromised, but not in other host. Our purpose is to protect DNP3, which one of the most popular open protocols used between components in cyber physical systems [7], over TCP communication; as such, we have to consider creating fake communication at a transaction level and which includes responses from request and acknowledge beyond the packet level, so as not to be identified by the attacker.

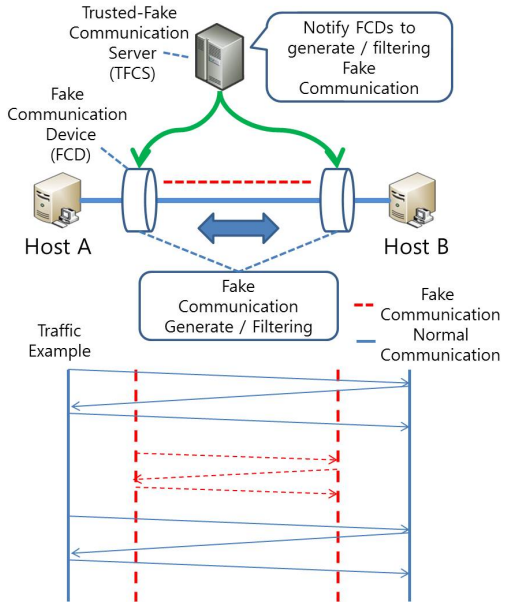

Figure 1: Fake Communication Generation and Filtering

Generation and Filtering. The "Fake Communication Device (FCD)" located in front of each host or SCADA then communicates with other FCDs in a mock normal communication. "Trusted-Fake Communication Server (TFCS)", whose concept is similar to that of the trusted-anonymization server in location privacy research [1] lets FCD located in front of host A make fake communication and gives a clue on how to filter fake communication to FCD located in front of host B (Figure 1). We assume that all FCDs are connected to TFCS safely. TFCS is located in critical infrastructure, but separately from SCADA.

Timing and Fake Traffic Control. When we design our method, we consider two important properties. First, our approach is a probabilistic method to avoid generating a distinguishable pattern. Generating many fake communications at the first or last periodic time is not effective in concealing the periodic time, because it can still provide a clue for the approximation of periodic time. Our intention is to assign high fake generation probability when time is far from the first and last periodic times or low probability near the first and last of periodic times. For example, if the periodic time of A communication on host A is 15 seconds, we want to make more fake communication near 8 seconds, but less at 0 and 15 seconds. To achieve our goal, we decided to assign fake generation probability as Probability density function of Normal distribution with mean value of periodic time divided by 2, then multiply by 2 to increase the probability. The probability of fake generation is formally described as

$$p_{fake} = \frac{2}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \quad (1)$$

, where t is the time in current time modulo periodic time.

The second important method property is that our method must be able to control the amount of fake communication traffic. To achieve this property, we introduce a generating decision parameter (*gd*) that indicates when to consider generating. For example, if *gd* is 1 second, we consider generating fake traffic every second. Note that, if *gd* is assigned a value that satisfies *periodic time* modulo *gd* = 0, then fake generation probability is 0. Thus, *gd* is selected compared with the periodic time that obtained a value from the pre-processing stage. By adjusting this parameter, we can control the amount fake communication traffic as well as the degree of fake protection. Additionally, the variance of Normal distribution used in fake generation probability can play the role of controlling fake traffic. Briefly, for every *gd* time, our method considers generating fake communication with fake communication generation probability.

Future Work. Currently, we have considered fake communication not as packet level, but transaction level, and evaluated our method through the implementation of an fake generator simulator. We have experimented with real data collected on two sites of national critical infrastructure for two weeks. The fake generator simulator added fake communication to the original packet file. Through this simulation, we will verify the effectiveness of our method.

3. REFERENCES

[1] B. Bamba, L. Liu, P. Pesti, and T. Wang. Supporting anonymous location queries in mobile environments with privacygrid. In *Proceedings of the 17th international conference on World Wide Web*, pages 237–246. ACM, 2008.

[2] S. A. Boyer. *SCADA: supervisory control and data acquisition*. International Society of Automation, 2009.

[3] X. Cai, X. C. Zhang, B. Joshi, and R. Johnson. Touching from a distance: Website fingerprinting attacks and defenses. In *Proceedings of the 2012 ACM conference on Computer and communications security*, pages 605–616. ACM, 2012.

[4] P. Düssel, C. Gehl, P. Laskov, J.-U. Bußer, C. Störmann, and J. Kästner. Cyber-critical infrastructure protection using real-time payload-based anomaly detection. In *Critical Information Infrastructures Security*, pages 85–97. Springer, 2010.

[5] P. Fogla, M. Sharif, R. Perdisci, O. Kolesnikov, and W. Lee. Polymorphic blending attacks. In *Proceedings of the 15th USENIX Security Symposium*, pages 241–256, 2006.

[6] S. Karnouskos and A. W. Colombo. Architecting the next generation of service-based scada/dcs system of systems. In *IECON 2011-37th Annual Conference on IEEE Industrial Electronics Society*, pages 359–364. IEEE, 2011.

[7] T. MicroWorks. Dnp3 overview. *Raleigh, North Carolina (www. trianglemicroworks. com/documents/DNP3 Overview. pdf)*, 2002.

[8] D. Riboni, A. Villani, D. Vitali, C. Bettini, and L. V. Mancini. Obfuscation of sensitive data in network flows. In *INFOCOM, 2012 Proceedings IEEE*, pages 2372–2380. IEEE, 2012.