MEDUSA: Mining Events to Detect Undesirable uSer Actions in SCADA

Dina Hadžiosmanović, Damiano Bolzoni and Pieter Hartel

Distributed and Embedded Security, University of Twente
{dina.hadziosmanovic,damiano.bolzoni,pieter.hartel}@utwente.nl

Abstract. Standard approaches for detecting malicious behaviors, e.g. monitoring network traffic, cannot address process-related threats in SCADA(Supervisory Control And Data Acquisition) systems. These threats take place when an attacker gains user access rights and performs actions which look legitimate, but which can disrupt the industrial process. We believe that it is possible to detect such behavior by analysing SCADA system logs. We present MEDUSA, an anomaly-based tool for detecting user actions that may negatively impact the system.

1 Problem

There is a wide range of possible attacks that can be carried out against SCADA environments [1, 2]. We classify possible threats in two groups: system- and process-related. System-related threats are typical of "regular" computer networks, e.g., malware or Denial of Service attacks. Attackers leverage vulnerabilities in networked systems and programmable logic controllers (PLCs) to alter or disrupt the industrial process. Process-related threats imply that an attacker gains user access rights (e.g., through social engineering) and performs legitimate SCADA commands which will negatively affect the industrial processes. Typical security countermeasures, e.g., antivirus or network IDSes, can hardly detect process-related threats, as they lack process semantic.

In this work, we focus on the detection of process-related threats. Based on interviews with stakeholders, we distinguish two types of threat scenarios, namely 1) an attacker impersonates a system user or 2) a legitimate system user makes an operational mistake. A SCADA engineer manages object libraries and user interfaces, sets working ranges for devices, etc. If an attacker succeeds in acquiring the access rights of an engineer, she is then able to perform actions such as altering a device parameter (e.g., change capacity of a tank) or altering the system topology (e.g. some devices become "invisible", and thus inaccessible). A SCADA operator monitors the system status and reacts to events, such as alarms, so that the process runs correctly. An attacker, impersonating an operator or an engineer, can generate a sequence of actions where each action is legitimate, but the combination (or even a single action) can damage the process.

We argue that to detect process-related attacks one needs to analyse data passed through the system (Bigham et al. [1]) and include a semantical understanding of the process and user actions. This can be achieved either by employing a tool such Bro, which requires the network protocol specifications (but those could be hard to obtain due to the closeness of SCADA systems), or by analysing system logs.

2 Solution

Typically, SCADA system logs provide detailed information about industrial processes. However, based on interviews with stakeholders, logs are not normally processed. The reason for this is that system engineers lack time, skills and specific tools for performing a thorough analysis. The size and high dimensionality of the logs make manual inspection infeasible. For instance, a SCADA system for a water treatment process in a medium-size city, depending on daily activities, records between 5.000 and 15.000 events per day.

We believe that system logs can be used to detect process-related threats and user mistakes automatically. We propose a visualization tool, MEDUSA(Mining Events to Detect Undesirable uSer Actions in SCADA), whose engine is based on anomaly detection. MEDUSA automatically analyses system logs, detects and alerts users about situations in which the system behaves unconsistently with past behavior. As a result, the number of security-related alarms that operators have to look at is decreased. The anomaly detection models in MEDUSA are built using two data mining techniques. First, we use algorithms for mining outliers to detect individual actions and events that are significantly different from previous entries. Secondly, we analyse sequences of events and actions in the logs to provide a better view on the situation context. Once we train our model on history logs of a particular system, we plan to use the model in real-time analysis a SCADA system.

Preliminary results show that our approach is feasible. The initial dataset consists of 100.000 entries which correspond to approximatively 15 days of process work. The attributes are mostly categorical. The goal of our initial analysis was to transform the dataset in such a way that anomalous entries are highlighted. We managed to extract several events that may semantically represent suspicious behavior (eg., a previously unseen engineer activity in late night hours, user expression errors when connecting to critical communication backbones).

References

1. John Bigham, David Gamez, and Ning Lu. Safeguarding scada systems with anomaly detection. In *MMMACNS '03: Proc. 2nd International Workshop on Mathematical Methods, Models and Architectures for Computer Network Security*, LNCS, pages 171–182. Springer Verlag, 2003.
2. Clyde G. Chittester and Yacov Y. Haimes. Risks of terrorism to information technology and to critical interdependent infrastructures. In *Journal of Homeland Security and Emergency Management: Vol. 1 : Iss. 4, Article 402*, pages 341–348, 2004.