CIGRE-319 **2020 CIGRE Canada Conference**
Toronto, Ontario, October 19-22, 2020

Communications Modeling: a new approach to Intrusion Detection, Performance and Diagnostics

EUGENIO CARVALHEIRA
OMICRON
USA
ANDREAS KLIEN
OMICRON
AUSTRIA
STEEL McCREERY
OMICRON
CANADA

SUMMARY

Several high-profile cyber-attacks on critical infrastructure have encouraged utilities to enhance their security posture. The substation network represents a critical attack vector within the power grid and faces many unique challenges not found within the traditional IT or control center networks. The inherent transient nature of the work force and equipment used within substation LAN for commissioning and routine maintenance lends the substation LAN to be at a much higher risk of infection from viruses. The potential introduction of viruses and malware designed specially to perform tasks such as network and traffic reconnaissance to lay the ground work for more sophisticated attacks is real. Several recent attacks of this nature, which having gain notoriety, are a testament to the legitimacy of such threats. Gateways and firewalls located at the substation LAN perimeter cannot prevent transient technologies introduction of viruses well within that defense perimeter. The one main stay characteristic of these more sophisticated attacks is that the attacker is within the perimeter for a significant time crafting the attack. During these intrusions, the generation of additional network traffic or a slight change in behavior of a device may be the only indicators of the presence of an intrusion. For several years the IT world has countered with the deployment of Intrusion Detection Systems or IDS's. The traditional IDS's that are deployed are typically one of two major variants: the signature-based (or black list-based) IDS or the learning-based IDS, sometimes referred to as an artificial intelligence-based IDS. Signature-based IDS's rely on a black list built from prior knowledge of key characteristics or, if you prefer, signatures of the viruses. The problem with this approach is that there have been only a few known attacks on substations and so there is limited data to discern a signature or black list from. Further it is likely that a new attack presenting a new signature not known by the IDS would have a high probability of going undetected. The learning or so-called artificial intelligence variant requires a learning phase of several weeks to learn the typical traffic patterns of a healthy system to develop rules. Once this learning phase has been completed the IDS monitors the traffic. Traffic observed to contradicts these rules would generate an alarm. The main issue with this approach is that valid but infrequent traffic scenarios may not take place during

Steel.mccreery@omicronenergy.com

the learning phase such as a real protection trip or maintenance traffic both of which are valid network traffic. Inevitably when this valid traffic is observed by the IDS false alarms are generated. Another problem with this approach is the typical nature of how such alarms are reported. To discern the cause of alarms generated by this type of IDS, when applied to an OT network, requires the combined skills of an IT professional and an OT professional with a good understanding of the system and its operation. As a result, the decoding is inevitably labor intensive and difficult. Once the cause is known, additional time and skills must be employed to create rules to "tune the IDS" which over time proves to be a daunting task.

This paper presents a new approach to intrusion detection within the IEC 61850 substation. A precise communications model is derived from the IEC 61850 System Configuration Language (SCL) files: there is no learning phase. The characteristics and rules for all real-time communications are built from the SCL file. The model, in addition to precise timing, allows the IDS to use the real time resources within the communications themselves to discern legitimate from illegitimate communications. Further, this modeling approach allows the IDS to measure performance characteristics to detect hardware failures or malfunctions such as the IED's loss of time synchronization.

KEYWORDS

IEC 61850, Cyber Security, Intrusion Detection System, Digital Substation, Functional Monitoring

Introduction

For this paper we will define a cyber-attack on a substation as an event where an attacker modifies, degrades, or disables a service of at least one protection, automation, or control device within the substation. This paper will first examine the contrast in cyber security priorities between IT and OT networks, the concept of defense in-depth, the NIST Cyber security framework and the role of an IDS. Typical IDS installation and a brief overview of the principles of operation of both traditional signature and learning based IDS will then be examined prior to the review of an IDS based on communication modeling and the advantages of this approach for substation IDS applications.

IT Cyber Security versus OT Cyber Security

The data of operational technology networks or OT networks is used to do something physically such as the control of lighting within a building or to activate control system I/O of an industrial process while information technology or IT LANs move data. Cyber security for utilities consists of both IT and OT LAN security. In an IT network, the confidentiality of data is the highest priority as corporations are legally obligated to ensure private information does not become public, data integrity being the second priority while network availability has the lowest priority: a short duration communications outage within a small area of a large entities' communications infrastructure is a concern but the lowest of the three in terms of priority. For the OT network within substations, the highest priority is network availability: communication between all devices is critical to control the process. Integrity and authenticity of the message is the next priority: nobody shall be able to inject a trip GOOSE while confidentiality is not important within the substation as the content of a trip GOOSE is well known. Given the contrast in priority, encryption while widely accepted as the silver bullet for IT network security, does not carry the same importance for the substation OT network. The IEC 62351 standard provides encryption for GOOSE and MMS. However, in the substation environment there are few applications where confidentiality of messages is important. If messages cannot be tampered with (integrity) and the originator can be verified (authentication) – which is fulfilled by using authentication in GOOSE and MMS (IEC 62351-6), it is not necessary to encrypt the messages. One example where encryption might be required is if routable GOOSE (R-GOOSE) messages are transmitted over an unencrypted communication path. Encryption requires additional CPU processing, increases GOOSE transmission time and impedes testing scenarios, and in most

cases does not provide additional security than that provided by authentication codes. It is also worth mentioning that if encryption within the substation is chosen, a key management infrastructure is also needed within the substation to deliver and maintain authentication keys for each IED. This is a major undertaking and as a result, these GOOSE security mechanisms have not gained widespread use. Encryption also makes a later analysis of traffic recordings and monitoring strategies such as the ones described later in this paper more difficult.

Defense In-Depth, the NIST Security Framework v1.1 and the Role of an IDS

There are different regulations in place throughout the world, for example, the NIS directive is applicable to many EU countries while in North America NERC CIP compliance is required. The main difference is freedom of choice with respect to implementation of security measures with some being a process based on risk while others are compliance based. One of the common threads or principle running through these standards is the concept of defense in-depth where there are layers of security. The substation is an excellent example of this principle. The fences, gates, and locks provide physical security while network segmentation, and role-based access provide layers of cyber security.

A second common thread is the adoption of the NIST Cybersecurity Framework which is a set of guidelines for private sector companies to follow to be better prepared in identifying, detecting, and responding to cyber-attacks. This framework was developed by the National Institute of Standards and Technology under the United States Commerce Department and is not only used in North America but has been adopted by many other countries. The core assumption of this security process is that there is no 100% protection. Attacks can always come through your layers of defense. It proposes that cyber security is seen as a process that you continuously improve. This process has 5 steps: 'Identify', 'Protect', 'Detect', 'Respond' and 'Recover'. The first step is the identification of attack vectors as will be presented in the next section. The second step is to protect against the vectors with highest risk by implementing countermeasures. If an attacker is able to break through these barriers, the third step is the detection of the attack as it occurs, and the fourth step, to immediately respond or act upon the attack to allow the recovery to normal status (which is the final step) as quickly as possible, to minimize damage in addition to learn from the attack. With the lessons learned in the detect and response steps, new attack vectors are identified, and new countermeasures implemented allowing the cyber security posture to evolve. This process continues to repeat itself. The role of the IDS within this process is the detection of attacks or intrusions as they occur to trigger the response. The inability to detect intrusions was a key factor in the success of several attacks mentioned in the next section, substation attack vectors.

Substation Attack Vectors

Figure 1 depicts a typical substation LAN and actual cyber-attack vectors (marked with a number) that have been successfully used in past attacks. An attacker could enter through the control center connection (1), as happened in the first of two high profile cyber-attacks in Ukraine, where the firmware of gateway devices was modified (causing their destruction). A second entry point is through engineering PCs (2) connected directly to the substation. When a protection engineer connects his PC to a relay to modify (protection) settings, malware on the PC could install malware on the relay comparable to what happened with the PLCs in the Stuxnet cyber-attack. Laptops used for testing the IEC 61850 system are often directly connected to the station bus which is also a potential way to infect IEDs (3). For this reason, new IEC 61850 testing tools are available which provide a cyber-secure separation between Test PC and substation network. This leaves the testing device itself (4) as a potential entry path. For this reason, it is important that test set vendors invest in hardening their devices to eliminate this as a possible entry path for an attacker to exploit.

The storage of settings (2a) and test documents (3a) could also be a source of malware making this storage server part of the Electronic Security Perimeter (ESP). It therefore makes sense to remove this server from the ESP and introduce a separate, isolated and protected data management solution for such data.

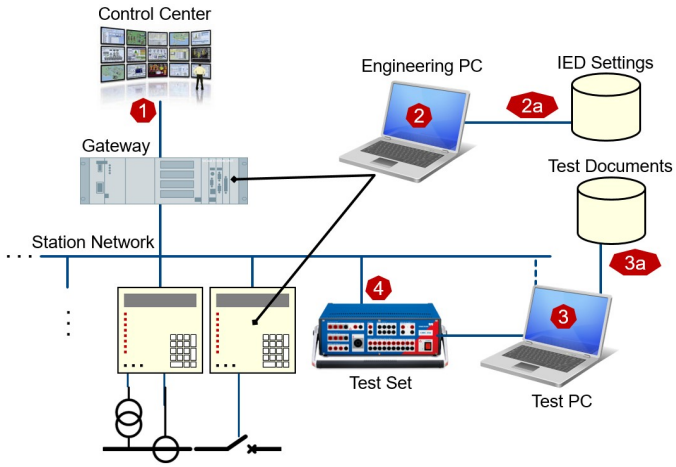

Figure 1. Attack vectors of a Substation

Intrusion Detection System Installation within Substations

The Intrusion Detection System would be connected as depicted in Figure 2. Switch ports that have been configured as “mirror ports” forward a copy of all network traffic appearing on other ports of the switch to the IDS. The IDS inspect all network traffic received over the mirror ports. The most important traffic to monitor is that of the traffic between the gateway and the IED’s and so at a minimum, the traffic of the switches connected to the gateway, in addition to traffic from any other network entry points should be sent to the IDS via mirror ports. Having said this, there is no way to predetermine which device could be infected and so it is ideal to monitor network traffic from all devices on the substation LAN. It should be noted that if switch chips inside the IED are used to interconnect IED’s it is likely that it would not possible to monitor the traffic between IED’s using these internal switch chips.

While classical IT security is concerned with high-performance servers with millions of connections at the same time, substation Operational Technology (OT) security deals with devices with limited resources, custom operating systems, real-time demands, and specialized redundancy protocols and so we will see that intrusion detection systems from classical Information Technology (IT) world are not ideally candidates for the substation environment.

Until very recently, there were only two main approaches for IDS: “signature-based” IDS and “learning-based” IDS.

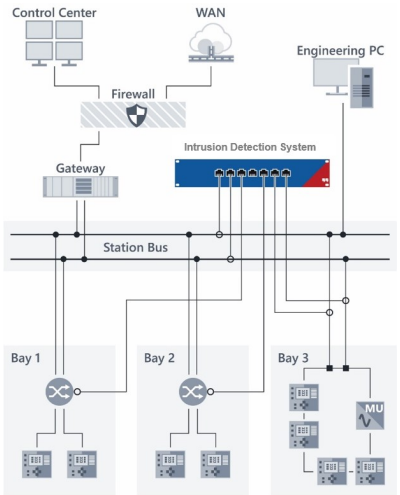

Figure 2. Connection of IDS to the substation network

Signature-based IDS

Signature-based IDS's rely on a black list built from prior knowledge of key characteristics or, if you prefer, signatures of the viruses. The IDS scan the data stream for known patterns from this list which is similar to how a PC virus scanner scans the PC's memory and drives for virus signatures. The problem with this approach is that there have been only a few known attacks on substations and so there is limited data to discern a signature or black list from. Further, it is likely that a new attack presenting a new signature not known by the IDS would have a high probability of going undetected. Given that the first occurrence of this new attack could have severe consequences, a substation IDS must be able to detect attacks without any previous knowledge of what the attack might look like, which is a very different approach then that employed by the signature-based IDS.

Learning-based IDS

The learning or so-called artificial intelligence variant requires a learning phase of several weeks to learn the typical traffic patterns of a healthy system to develop rules. Such systems look at frequency and timing of certain protocol markers to attempt to learn the usual behavior of the system. Once this learning phase has been completed, the IDS monitors the traffic and an alarm will be raised if one of the markers is significantly outside the expected range. The main issue with this approach to intrusion detection is that valid but infrequent traffic scenarios may not take place during the learning phase such as a real protection trip, uncommon switching or automation actions, or routine maintenance and testing traffic all of which are valid network traffic. Another problem with this approach is the way in which such alarms are reported. These systems do not understand the semantics of the protocols, and so the alarm messages are expressed in terms of technical protocol details. Hence, the root cause of alarms can only be determined by the combined efforts of engineers skilled in IEC 61850 protocol details and those familiar with IT network security. The engineers examining the alarm also must know about the operational situation to judge if certain IEC 61850 protocol events correspond to valid behaviors. As a result, the decoding is inevitably labor intensive and difficult. Once the cause is known, additional time and skills must be employed to create rules to "tune the IDS" which over time proves to be a daunting task as there is a tendency to be a high number of these false alarms generated within a substation and there are many substations all of which requiring the same highly skilled personnel to root cause and tune the IDS. This often leads to alarms being ignored or alarms discarded without an investigation of the cause, ultimately leading to the IDS being ignored or disabled.

Communications Modeling

For IEC 61850 substations, the whole automation system, including all devices, their data models, and their communication patterns are described in a standardized format – the System Configuration Language (SCL). System Configuration Description (SCD) files can be generated by engineering configuration tools and will contain this information in addition to information about primary assets and potentially the single-line diagram.

Communications Modeling is a term that describes a new approach to the detection of intrusions by deriving a precise communications model of the automation and power system communications from this information. This precise model allows the IDS to compare each packet received from the network against the communications model. Further, the communications model allows the IDS to perform deep packet inspection: the IDS can compare the variables contained within the communicated (GOOSE, MMS, SV) messages (or frames) against the expected content derived from the communications model. This level of inspection is made possible without the need for a learning phase. The IDS configuration process is initiated by the import of the SCL file into the IDS. The information contained in the SCL file is used by the IDS to create the system communications model and whitelist the messages that were described within the SCL file for the system. With the comparison of the monitored network traffic against the system communications model, zero (or minimum) false alarms are expected.

Functional Security Monitoring

As discussed, the communications modeling approach described above enables the IDS to perform very detailed monitoring of the substation traffic to detect traffic that should not be there. The ability of the IDS to inspect the content of the message (GOOSE, MMS, or SV) and compare it to what is expected bring the monitoring to a new level: the IDS is capable of what could be described as "functional security monitoring'. Let's use GOOSE messaging to provide a clear understanding of what functional security monitoring is and how it is accomplished. In addition to the actual data, GOOSE messages contain additional codes that allows the IDS to detect errors. For example, monitoring of the state number (stNum) and sequence number (sqNum) of each GOOSE message received allows the IDS to detect glitches in the sequence of GOOSE messages being received which indicates intermittent network communications issues to the respective IED. The EntryTime timestamp (time that the GOOSE was sent) can be compared to the time the GOOSE message was received at the IDS port. The difference between these two times is the network transmission time. If this time is significantly longer than 3 ms for a "protection" GOOSE (referring to the performance specified in IEC 61850-5), it is an indication that the network has excessive latency due either to a hardware failure or a time synchronization issue.

An understanding of MMS affords the IDS additional functional security monitoring capabilities. Drawing from the system model (contained within the SCL file) the IDS understands which Logical Nodes in the IEDs control which primary switchgear. Thus, the IDS can distinguish between correct/incorrect, and critical/noncritical actions. With reference to figure 1, if the PC has been defined within the IDS as a Test PC, an example of a legitimate Test PC command would be to put a relay into test mode, but the operation of a breaker by the Test PC is most probably not allowed and if this were to occur, an IDS with functional security monitoring capability would alarm on this type of command originating from the Test PC. Let's take a closer look at this example in the following sections of this paper.

Functional Security Monitoring Alarm Messages

By using the substation section in the SCL file, an overview diagram of the substation (similar to a single line diagram) can be created automatically as shown in Figure 3. In this example, an alarm is shown as an arrow from the active participant (Test PC) performing a prohibited action (sending a control command), and the "victim" of the action – a bay controller in bay Q01. In addition to the avoidance of the generation of false alarms, it is of vital importance that the textual alarm messages that are generated are easy to comprehend by those engineers who are responsible for the operation of the protection, automation and control (PAC) functions within the substation. This new approach uses information contained within the system model to generate alarms that use PAC engineer terminology rather than terminology associated with a protocol analyzer as would be the case if using a traditional IDS. Figure 4 is an example of such an alarm.

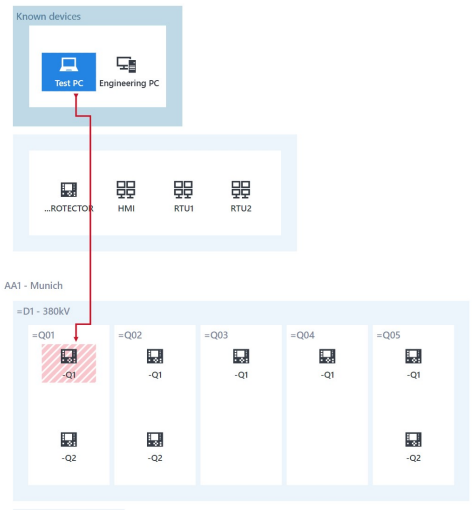

Figure 3. Graphical alarm display

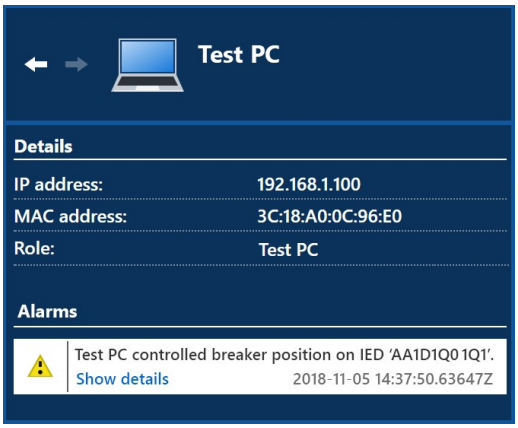

Figure 4. Alarm details: Test PC attempting unauthorized control of circuit breaker

The graphic and PAC language based textual alarms allows security engineers and PAC engineers to collaborate when tracing alarms within a substation. To support the alarm response process even further, it should be possible to associate the IDS alarms with sequence of events (SOE) and event logs in the substation SCADA / HMI systems.

Maintenance Mode

To avoid false alarms, routine testing and maintenance conditions can be included in the substation system model. The testing and engineering equipment, including protection test sets, can be added into the system. When in maintenance mode, a Test PC could perform certain commands on the IEDs of this bay such as changing the IEC 61850 test or simulation mode of IED -Q1 without the generation of an alarm. However, if the Test PC operates a breaker within the bay a similar alarm as discussed previously will be generated.

IDS are Passive

The IDS is usually passive. If an action is “not allowed”, an alarm is triggered as the IDS does not interfere with the substation communications The alarms can be communicated via syslog to a Security Incident Event Management (SIEM) system or potentially support user-definable binary outputs that can be connected directly to an RTU allowing alarms to be integrated into the SCADA signals of the substation.

Conclusion

Substations cyber-attack vectors must be determined and effectively protected to prevent potentially severe consequences for the grid. For IEC 61850 substations a new approach is available for intrusion detection which takes advantage of the information contained within the SCL configuration files to detect malicious traffic in addition to functional communications issues and failure with a minimal configuration effort and a low probability of the generation of false alarms. Further, this approach allows the display of detected events in the language of protection, automation and control engineers allowing a more efficient root cause analysis.