Computerized risk detection towards Critical Infrastructure Protection: An Introduction of CockpitCI Project

Jianmin Jiang
Department of Computing, University of Surrey,
Guildford, GU2 7XH, United Kingdom.
Jianmin.jiang@surrey.ac.uk

Lasith Yasakethu
Department of Computing, University of Surrey,
Guildford, GU2 7XH, United Kingdom.
s.l.yasakethu@surrey.ac.uk

Abstract— **In this article we describe a new European Framework-7 (FP7) funded research project, CockpicCI, and introduce the concepts of intelligent risk detection, analysis and protection techniques for Critical Infrastructure (CI) Protections. Typical attacks could be performed blocking communication from central Supervisory Control and Data Acquisition (SCADA) to local equipment or inserting fake commands/measurements in the SCADA-field equipment communications. The paradox is that CIs massively rely on the newest interconnected and vulnerable, Information and Communication Technology (ICT), while the control equipment, legacy software/hardware, is typically old. To overcome such threats, the CockpitCI project combines machine learning techniques with ICT technologies to produce advance intrusion detection and reaction tools to provide intelligence to field equipment. This will allow the field equipment to perform local decisions in order to self-identify and self-react to abnormal situations introduced by cyber-attacks.**

Keywords- Critical infrastructures, Cyber-attacks, Risk protection.

I. Introduction

CockpitCI will focus on cyber-attacks to SCADA systems [1, 2, 3] of energy grids, that are typically interconnected with even public Telco networks and we report an actual failure scenario of a Power grid due to the unavailability of its SCADA system. Power grids and Telco networks have a large impact on daily life and are typically referred as Critical Infrastructures since their correct operation is essential for the everyday life of our modern society [4, 5, 6]. There are bi-directional dependent relationships and reciprocal influences among CIs, named interdependencies. That is especially true because CIs are more and more reliant on information and communication technology and, mainly through this reliance, they have become more and more interdependent. The successful delivery of any essential CI service depends upon the operating status not only of the CI which is intended to deliver such a service but also on the operating status of any interdependent CIs. Initial disturbances in (or even destruction of) parts of one CI, may result in cascading effects in the infrastructure itself or/and in the other interdependent CIs. A general framework for interdependencies between a Power CI and a Telco CI could be described as follows. Power CI is composed by the Electrical grid and the Telco of the Electrical grid. Telco of Electrical grid provides the communication facilities needed to control the Electrical grid and may be partly operated by the Power operator and partly by a Telco operator. Telco CI which mainly consists of the Telco network, which in turn consists of several networks that provide different Telco services (voice, mobile voice, data, mobile data, etc.). In addition Telco networks require reliable power supply that is provided by the Power CI (Power to Telco services). However, it is rather common that Telco operators have their own emergency power supply systems for each (or at least in the most important) offices, for back-up reasons in case of main power supply outages. The set of these power back-up systems, owned by the Telco operator, is called Telco emergency power supply. Telco network and Electrical grid are managed and controlled through a management and control system (SCADA system in case of a Power distribution grid).

In this article we will introduce advance intrusion detection, analysis and protection techniques which will be developed as a part of the CockpitCI project to protect CI from such cyber-attacks. In section 2 cyber-security for SCADA systems is discussed. Section 3 presents a machine learning approach for different intrusion detection techniques. Also, an information fusion based threat assessment model is presented in this section. Finally, section 4 concludes the article.

II. CYBER-SECURITY FOR SCADA SYSTEMS

SCADA systems have always been susceptible to cyber-attacks. Different types of cyber-attacks could occur depending on the architecture and configurations used in the SCADA system. These attacks fall into one of four below categorize:

1. Internal/Non-malicious - employees or contractors causing unintentional damage

2. Internal/Malicious - system users with extensive internal knowledge of the system who intentionally cause damage

3. External/Opportunistic - hackers seeking a challenge

4. External/Deliberate - malicious, well-funded political activists, organized crime groups, or nation states

All classifications of attacks can result in serious consequences. To protect cyber infrastructure from above attacks a growing collaborative effort between cyber security professionals and researchers from private and academia has involved in designing variety of intelligent cyber defence systems.

In summary, such systems address various cyber-security threats, including trojans, spam, viruses and worms. The system protects the cyber-infrastructure at two levels and combat threats at network and host based levels. Network based defense system control the network traffic by network-fire wall, antivirus, spam filters and network intrusion detection techniques, where as the host based defense system control the data flow in a workstation by host firewall, antivirus and host intrusion detection techniques. Intrusion detection process can be divided into several components to include: Information sources; Data acquisition tools; Data pre-processing; Data analysis and intrusion detection; and Threat assessment and response.

Data acquisition tools capture events from network and host based information sources. If an event originates from the network traffic, it is categorized as a network based event where as if an event originates with log files, it is categorized as a host based event. Host based event is a collection of system calls traces. These intrusions are in the form of anomalous subsequence of traces. Network based event is a collection of network traffic data, such as IP (Internet Protocol) or TCP (Transmission Control Protocol) network packets. These intrusions typically occur as anomalous patterns. Data pre-processing stage deals with data cleaning, fusion, selection techniques, feature extraction and transformation techniques to support the data analysis. Numerous intrusion detection mechanisms are employed to investigate the behaviour of the cyber-infrastructure by analysing the input data. This is considered as the principal component of the intrusion detection system. More details on intrusion detection methodologies will be discussed in section 3. Once a cyber-attack is identified a threat assessment is deployed and a decision is taken accordingly.

III. INTRUSION DETECTION VIA MACHINE LEARNING APPROACHES

Intrusion detection techniques can be classified into three main modules: Signature detection (misuse detection), Anomaly detection and Hybrid detection. Detection principles behind each module are discussed in the following subsections.

Signature detection (misuse detection): Signature detection also known as misuse detection generates alarms when a known cyber-attack occurs. In this technique the behaviour of the system is compared with unique patterns and characteristics of known attacks, called signatures. This is typically done by measuring the similarity between the input events and signatures of known attacks. If a match is found, an alarm is triggered. As a result, known cyber-attacks can be detected immediately with low false-positive rate. However, signature detection can only detect known attacks, which also heavily rely on the prior knowledge of attack signatures. Thus the effectiveness of the detection mechanism rely on frequent updating of the signature database.

Due to the availability of prior knowledge on attack signatures, hence the availability of labelled data, supervised machine learning techniques are generally used for signature based intrusion detection.

Anomaly Detection: Anomaly detection is an IDS triggering method that generates alarms when an event behaves different from the normal behaviour patterns. Thus this can be defined as a problem of finding patterns in data that do not confirm to expected behaviour of a system. Figure 1 illustrates the anomalous data patterns in a simple 2-dimentional data set.

In this example the data has two normal regions, N1 and N2. Data that sufficiently deviate from these regions, i.e. point A1, point A2 and region A3 are considered as anomalies.

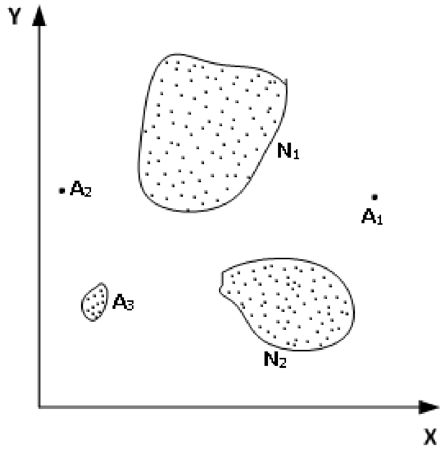

Figure 1. Anomalies in a simple 2-dimentional data set

The anomaly detection approach has two main steps: training and detection. In the training step, machine learning techniques are used to generate a profile of normal behaviours that define the healthy cyber-infrastructure. In the detection step, an event is classified as an attack if the event records deviates sufficiently from the normal profiles. Unlike signature detection, anomaly detection has the potential to detect novel attacks. However, anomaly detection typically has a high false-positive rate. This is because in anomaly detection any sufficient deviation from the base line is flagged as an intrusion. Thus is it likely that non-intrusive behavior that falls outside the normal region generate an alarm, resulting in a false-positive.

The key challenge for anomaly detection in intrusion detection is the analysis of huge amount of data with high dimensional feature space. It requires computationally efficient data mining techniques to handle large amount of input data. Furthermore, the data typically comes in a streaming fashion, thus requires online analysis. As the data amounts to millions, even a few present of false alarms can be overwhelming when comes to decision making.

In anomaly detection, labelled data corresponding to normal system behaviour are usually available, while the labelled data for intrusions are not. As a result, unsupervised and semi supervised machine learning techniques are preferred for anomaly detection.

Hybrid detection: From the above discussions it is understood that both signature and anomaly detection techniques have advantages as well as disadvantages. Most signature detection techniques have high detection rate and a low false alarm rate. But they cannot detect novel attacks. Whereas anomaly detection techniques are capable of detecting novel attacks. However, anomaly detection suffers from a high false alarm rate. Since signature and anomaly detection techniques have compensational capabilities and functions, hybrid detection methods have been proposed to integrate the accuracy and reliability of signature detection techniques and the intelligence and flexibility of anomaly detection techniques [7, 8]. However, a simple integration of the two systems cannot assure a better performance than a single intrusion detection (signature or anomaly detection system) system. For a successful hybrid system, two systems should incorporate in an effective manner such that both systems will benefit from the positive features of each other. This requires two important issues to be considered in implementation: 1) best candidates for signature and anomaly detection techniques should be selected, 2) optimum method of integration should be determined such that two systems will complement one another. However, this would require research as the above criteria's depend on input data types, type of intrusion/application and contextual information (i.e. intelligence of the systems, adversary, network conditions, etc.). In general following approaches could be used to integrate the two IDS to obtain a hybrid system.

a) anomaly detection followed by signature detection

b) anomaly detection and signature detection in parallel

c) signature detection followed by anomaly detection

Information fusion for threat assessment: For each individual alarm triggered by the IDS, the decision making process needs to understand how likely it is that the alarm corresponds to an actual attack. Using Bayes's theorem, the probability of a sensor alarm meaning an actual attack could be expressed as follows [9]:

$$P(attack|alarm) = \frac{P_d P_a}{P_d P_a + P_{fa}[1 - P_a]}$$

Where P_d is the probability of detecting an attack, P_a is the probability of an attack occurring and P_{fa} is the probability of a false alarm. For the context of this paper we define the term $P(attack|alarm)$ as the positive predictive value. Figure 2 shows the relationship between the positive predictive value and probability of a false alarm (P_{fa}) for different P_a values. For this illustration we have used a detection probability (P_d) of 0.9.

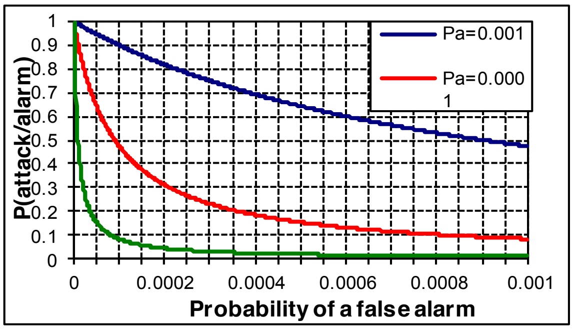

Figure 2. Relationship between positive predictive value and probability of false alarm

It is noted that when the probability of the attack decreases (as we hope in the case of cyber-attacks) the positive predictive value will also decrease for a given false alarm rate. In other words, the confidence that a sensor alarm actually reflecting a true attack will reduce. Furthermore, the relations indicate that for a high positive predictive value it requires that P_{fa} be much lower than P_a. For instance, when P_a =0.0001, P_{fa} needs to in the range of $1x10^{-5}$, in order to have a 90% chance of that alarm corresponding to a true attack. However, developing an intrusion detection sensor that has such a low false-positive rate is an extremely difficult task. Thus, it is highly likely that the decision making process will have to rely on moderate positive predictive values to make decisions. However, this does not imply that no action should be taken. It implies that the decision to react and the type of reaction should take into account factors such as cost of the reaction, associated risk level and the frequency of the alerts. For instance, a high cost action should not be taken if the positive predictive value is not close to 1, unless the associated risk level is extremely high. If a high cost action is taken with a low positive prediction level and later on if the action proves to be unnecessary, then the confidence in the cyber-security infrastructure could be questioned. As a result, for an efficient cyber-security infrastructure it requires an automatic threat assessment module to be incorporated to the IDS.

The objective of the threat assessment module is to quantify the risk(s) associated with the attack and the cost for the action(s) to be taken. This will provide intelligence to the field equipment to reliably identify the threat and to take correct actions to prevent it automatically. However, from the above discussion it is understood that due to the moderate positive prediction value of an abnormal event, additional information is required by the threat assessment module in order to take correct reactions to the alerts raised. This raises the necessity of an information fusion framework. Information fusion is the process of intelligently combining information from different sources to enhance understanding on the data and its implications in order to provide an outcome that is superior to any provided by an individual source. Following a triggered alarm indicating a potential attack, there are number of information that the decision making process would like to be aware of. In addition to the positive prediction value of the attack, information such as the time of attack, extent and the time of contamination, whether it was intentional or not and the cost of the reaction to prevent it are valuable information in order to make a reliable decision. This information will come from different sources in different formats. Moreover, this information could arrive through uncertain sources and may conflict with one another. Also, information may be collected at different times and locations. Thus the information fusion framework should have an approach to analyse the different types of data and to merge the information in order to present a reliable and informative decision support to the operator. This would require huge amount of data processing which will mainly involve the following tasks:

- data cleaning: noise and irrelevant data will be removed
- data selection techniques: only useful features are extracted from the data to obtain a reduce data set while keeping the integrity of the original information
- data transformation: different data is transformed to a suitable format(s) in order to combine the knowledge of each data source
- pattern recognition: useful patterns of the data are identified

Due to the amount of data mining involved in the above tasks, machine learning techniques are necessary to address the challengers of the information fusion framework. Pattern recognition, artificial intelligence and statistics could be used to analyse, group and extract features from the entities to perform the above tasks. Thus, the processes will exploit analysis tools from machine learning methods (both supervised and unsupervised depending on the nature of the information), mathematical algorithms and statistical tools to discover and merge the relationships among different information. Figure 6 illustrates an information fusion driven automatic threat assessment architecture base on the above principles discussed. The output of the threat assessment module will provide intelligence to the field equipment to take correct actions to prevent cyber-attacks.

CockpitCI shall introduce a highly-innovative, unseen before functionality of machine learning based solutions for CI protection. The project aims to investigate and develop machine learning algorithms to support different types of intrusion detection techniques discussed in section 3. These algorithms will be tested and validated on real equipment and scenarios provided by the Israel Electric Corporation. Furthermore, as discussed in section 4, an information fusion based automatic threat assessment module will be developed and integrated to the cyber-security system, to react to abnormal situations introduced by cyber-attacks. With the developments of the above techniques CockpitCI will be able to:

- deploy smart detection agents to monitor the potential cyber threats according to the types of ICT based networks (e.g. SCADA) and types of devices that belong to such networks.
- identify, in real time, the CI functionalities impacted by the cyber-attacks and assesses the degradation of CI delivered services.
- broadcast an alerting message through an improved Secure Mediation Gateway at different security levels (low and high level).
- manage a strategy of containment of the possible consequences of cyber-attacks at short, medium and long term.

IV. CONCLUSIONS

In today's growing “cyber world”, where a nation's vital communications and utilities infrastructure can be brought down in minutes by hostile attacks, the need for critical infrastructure protection and advanced cyber-security is at all-time high. Indeed, security failure for such systems can result in an Armageddon with consequences sprawling at different layers of society.

The article provides the CockpitCI concept and roles of intelligent computing functions and machine learning methods to prevent cyber-attacks are discussed. A discussion on this concept emphasizes the need of intelligent rick detection, analysis and protection techniques for CI [10]. With the intelligence of machine learning solutions, CockpitCI will contribute to a safer living environment for people especially by providing smart detection tools, early alerting systems and strategic security system. The distributed framework of the system will ensure an operational deployment of the security all over Europe and will improve the European Critical Information Infrastructure Protection (CIIP) strategy. The research carried out during the CockpitCI project will allow improvements to the security industry. Indeed the project will develop smart detection tools for SCADA and IT networks, new methodologies of detection and analysis likely to give a real advantage in security market in these domains.

REFERENCES

[1] A. Bobbio, G. Bonanni, E. Ciancamerla, R. Clemente, A. Iacomini, M. Minichino, A. Scarlatti, R. Terruggia, E. Zendri - Unavailability of critical SCADA communication links interconnecting apower grid and a Telco network - Reliability Engineering and System Safety Journal – Elsevier editor, December 2010.

[2] A. Bobbio, E. Ciancamerla, S. Diblasi, A. Iacomini, F. Mari, I. Melatti, M. Minichino, A. Scarlatti, R. Terruggia, E. Tronci, E. Zendri “Risk analysis via heterogeneous models of SCADA interconnecting Power Grids and Telco Networks”, The Fourth International Conference on Risks and Security of Internet and Systems, Toulouse, France, October 2009.

[3] E. Ciancamerla, C. Foglietta, D. Lefevre, M. Minichino, L. Lev and Y. Shneck - Discrete event simulation of QoS of a SCADA system interconnecting a Power grid and a Telco network - 1st IFIP International Conference on Critical Information Infrastructure Protection, World Computer Congress 2010 - Brisbane, Australia, 20-23 September 2010

[4] Y. Haimes, B. Horowitz, J. Lambert, J. Santos, C. Lian and K. Crowther, “Inoperability Input-Output Model for Interdependent Infrastructure Sectors”. I: Theory and Methodology, Journal of Infrastructure Systems, Vol. 11(2), pp. 67-79, June, 2005.

[5] S. De Porcellinis, S. Panzieri, R. Setola, and G. Ulivi, “Simulation of Heterogeneous and Interdependent Critical Infrastructures”, Int. J. Critical Infrastructures (IJCIS), vol. 4, n. 1/2, pp. 110 128, 2008.

[6] P. Capodieci & al., "Improving Resilience of Interdependent Critical Infrastructures via an On-Line Alerting System," Proc. of COMPENG 2010 - Complexity in Engineering, 2010.

[7] S. Dua and X. Du, Data Mining and Machine Learning in Cybersecurity, New York, CRC Press, 2011.

[8] J. Zhang, M. Zulkernine, and A. Haque, Random-Forests-Based Network Intrusion Detection Systems, IEEE Transaction on Systems, Man and Cybernetics, Vol.38, No.5, September, 2008

[9] T.J. Dasey and J.J. Braun, Information Fusion and Response Guidance, Lincoln Laboratory Journal, Vol.17, No.1, 2007.

[10] S. Bologna, and R. Setola, “The Need to Improve Local Self-Awareness in CIP/CIIP”, Proc. of First IEEE International Workshop on Critical Infrastructure Protection (IWCIP 2005), pp. 84-89, Darmstadt, Germany, 3-4 November 2005.