2016 International Conference on Software Security and Assurance

Detection of Man-in-the-Middle Attacks on Industrial Control Networks

Oliver Eigner*, Philipp Kreimel* and Paul Tavolato*
*Institute of IT Security Research
University of Applied Sciences St. Pölten,
St. Pölten, Austria, {oliver.eigner,philipp.kreimel,paul.tavolato}@fhstp.ac.at

***Abstract*—In this paper we present a method to detect Man-in-the-Middle attacks on industrial control systems. The approach uses anomaly detection by developing a model of normal behaviour of the industrial control system network. To come as close as possible to reality a simple industrial system, a conveyor belt with sensors and actuators, was set up with controllers widely used in industry. A machine learning approach based on the k-Nearest Neighbors algorithm with Bregman divergence was used to define a model of normal (valid) behaviour. Afterwards Man-in-the-Middle attacks were launched against the system and its behaviour during the attack was compared to the valid behaviour model. The results show that the approach taken was able to detect such attacks with satisfactory accuracy.**

***Keywords*-Anomaly Detection; Intrusion Detection System; Industrial Control Systems; Machine Learning;**

I. Introduction

The combination of information processing components with physical processes and objects, as well as the ongoing networking of automation components has been present in automation systems since the 1970s. The new essential aspect, called Cyber-Physical System (CPS), expands this combination via connection to the global Internet. This results in considerable consequences - systems are randomly linked, their connections during operations are changed, closed and re-established. In addition, all gathered data, information and services within a CPS are used and are also available for retrieval. [1] Frequently mentioned examples presenting these characteristics are listed below:

- Product and production systems connected via components and system boundaries.
- Smart grids, in which local units are networked for distributing and generating energy through IT networks.
- eHealth systems, in which devices and sensors are connected to a medical information system.
- Autonomous automotive systems, which connect vehicles to the transport infrastructure.
- And many more ...

This paper focuses on industrial control systems, particularly on systems used in production environments. The continuous availability of services and data as well as the high degree of cross-linking of automation systems do not only provide new and seminal prospects, but also involve risks such as manipulation, sabotage, theft and many other attack vectors. [2]

This is reflected by the growing concerns about attacks on industrial control systems that were recognized at the latest with the detection of attacks such as Stuxnet [3], the 2014 German steel mill incident [4] and Dragonfly/Havex [5]. The possibility of such sophisticated attacks on industrial systems was a wake-up call for many and increased the awareness on security. Besides the increased attack surface of CPSs, the capabilities, sophistication and motivations of threat actors are increasing, too. Furthermore, update and patch mechanisms as used in traditional information systems often are not applicable in industrial control systems: updates may rather not be available because of the age of systems used, or downtimes for update procedures are not acceptable in 24/7 systems, or the risk of service interruption after an update is rated too high.

To succeed at protecting these environments, the communications with the control systems must be secured and monitored in order to identify any cyber-attacks on the cyber-physical system. There are solutions available that monitor network activity of control systems, however, the system operation of the CPS, which can be modified by an adversary, is usually not monitored. Thus, attacks that disrupt normal system operation may not be identified in time and could cause damage to equipment or even people. In this paper, an anomaly-based attack detection approach, which uses normal system behaviour as training data and compares the current system activity with the behaviour model, is presented. The system data was obtained from a custom-built industrial CPS scenario, a conveyor belt system, which uses industrial hardware in order to provide authentic data and ensure practicability of the approach.

II. Existing Approaches

The area of attack detection by analysing network traffic in industrial control systems is rather new. Some work about anomaly-based approaches can be found, though they are mainly used in connection with common IT security devices such as firewalls. They are generally restricted to perimeter security measures and do not analyse machine behaviour

978-1-5090-4388-0/16 $31.00 © 2016 IEEE
DOI 10.1109/ICSSA.2016.19

itself.
In [6], an industrial firewall system that applies abnormal traffic filtering mechanism for industrial protocols, is presented. A real-time intrusion tolerance system, which is based on anomaly based intrusion detection, is presented in [7]. The effectiveness of the approach was tested in a simulated environment and various attacks could be detected. In [8], a machine learning based approach to anomaly detection, using data from a live running industrial process control network, is presented.

III. Cyber-Physical Systems

Cyber-physical systems are integrations of physical processes with networks and computation. Embedded computing devices sense, monitor, and control the physical processes through networks, usually with feedback loops in which physical processes affect computations and vice versa [9]. CPSs interact with the physical world and with users through Human-Machine Interfaces (HMI), work stations, smart phones, etc. [10]
A cyber-physical system connects various components of the cyber world and the physical world together.

A. Control Components

Control components are used to control and monitor processes. Key components include controller hardware, remote diagnostics and maintenance utilities, and I/O devices. The following is a list of major control components used in CPSs [11]:

- Industrial Control System (ICS): ICS is a general term that encompasses several types of control systems often found in the industrial sector and critical infrastructures.
- Programmable Logic Controller (PLC): A PLC is a computer-based solid-state device that is used for controlling or regulating industrial equipment and processes.
- Intelligent Electronic Device (IED): An IED is a microprocessor-based controller that can acquire data from sensors and issue control commands.
- Human-Machine Interface (HMI): The HMI is a component of certain devices that are capable of handling human-machine interactions.
- Data Historian: The data historian is a centralized database that logs all process information within the control system.
- Data Acquisition Server: The data acquisition server uses industrial protocols to connect to field devices such as PLCs, RTUs and IEDs.
- Sensor: Sensors are components for measuring physical quantities and converting it into a corresponding output signal.
- Actuator: An actuator converts a source of energy (electric or pneumatic signal) into motion in order to move or control a mechanism or system.

B. Network Components

Network topologies across different CPS implementations vary, as control networks have merged with corporate networks to allow controlling and monitoring of the systems from the outside. The following is a list of the key network components used in a CPS network, regardless of the network topology [11]:

- Fieldbus Network: The fieldbus network links sensors and other devices to control components.
- Control Network: The control network connects the supervisory control level to lower-level control modules.

C. Control System Architecture

Modern cyber-physical systems that implement an industrial process control network consist of two control layers:

- the physical layer, which contains devices such as sensors, actuators and IEDs; and
- the cyber layer, which contains all the communication and information devices and software that acquire data and control processes.

Both of those layers are connected over the network fabric [12], which provides the mechanisms for the computing devices to communicate. A variety of protocols is being used for communications. Together, both layers form a cyber-physical system.

D. State of Security in Industrial CPS

Historically, industrial control systems were physically isolated and based on proprietary hardware and communication protocols. However, industrial CPSs are evolving: implementation of standard protocols and interconnection with corporate networks. Consequently, the security characteristics of CPSs have changed. Most industrial CPSs were designed to meet availability and reliability requirements. In most cases they were physically isolated from outside networks and were only controlled through proprietary systems. Therefore, cyber security measures were often not implemented. However, the introduction of IP-based technology and standard computing devices into operational environments introduces new vulnerabilities and opens new points of exposure [13].
The growing concerns about attacks on industrial control systems became clear at the latest in 2010 with the detection of the Stuxnet virus [3], which was targeting highly specialized industrial systems in critical infrastructures. The virus attacked Iranian nuclear facilities and took control of the normal operation of the centrifuge system using operating system vulnerabilities. The possibility of such sophisticated attacks on industrial systems was a wake-up call for many and increased the awareness on security. [14]
In 2015, the SANS Institute conducted a survey on ICS security [13], which was taken by 314 participants that actively maintain, operate or provide consulting services

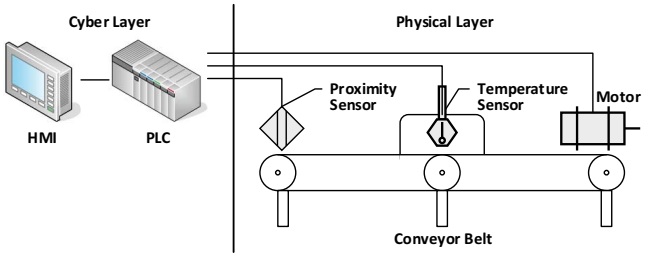

Figure 1. Conveyor belt schematic

to facilities with industrial control systems. The results show that organizations are concerned about maintaining the availability and reliability of their ICS operations. They also show that lowering risk/improving security is a pressing concern for organizations. However, 32% of the participants indicated that their control system networks or assets had been infiltrated or infected at some point, and 44% were unable to identify the source of the attack.

The increased targeting of control systems can attributed to multiple factors:

- There are significant differences in the security objectives between CPSs and traditional IT systems.
- The interconnection of CPSs with corporate networks and the Internet increases the attack surface.
- There is a lack of security testing technology for CPSs, as traditional intrusion defence strategies for IT systems are not effective enough in operational environments [15].
- Industrial control systems that monitor and control critical infrastructures present an attractive target for adversaries for the potentially large impacts they might have. At the same time, the capabilities, sophistication and motivations of threat actors are also increasing [13].

E. The Project Setup: A CPS Scenario

In order to check the possibility of detection of Man-in-the-Middle attacks on a real CPS, a scenario of a simple industrial system, a conveyor belt, was created as a testing environment. Here we present a brief overview of the system architecture of the CPS scenario created to analyse attacks. To test vulnerabilities of CPSs under real conditions, a conveyor belt system, using Siemens SIMATIC hardware, was built. Figure 1 shows a schematic overview of the system architecture. The cyber layer consists of a PLC and an HMI and the physical layer contains sensors and actuators. The PLC is connected to the physical layer and controls the actuators to power the conveyor belt and uses the sensors to gather process data. The system was created on physical hardware to come as close to reality as possible. The system contains a heating chamber, in which the products are heated to a specific temperature. However, due to hardware limitations the heating process is only simulated.

The conveyor belt system uses an industrial controller, in order to represent the operation of a real industrial system. For this scenario, Siemens SIMATIC hardware was used. The following components were used in the setup:

- PLC: SIMATIC S7-1214C AC/DC/Rly [16]
- HMI: SIMATIC KTP600 Basic colour PN [17]
- Engineering Software: SIMATIC STEP 7 Professional (TIA Portal)

The PLC is implemented as the primary control component in this control system configuration. It is connected to an engineering workstation over LAN. The sensors and actuators are connected over a fieldbus network.

The hardware setup of the system contains multiple infrared obstacle avoidance sensors [18], which determine the position of the product, and actuators to control the motor. Based on the position, a specific instruction is performed on the PLC.

IV. Normal System Behaviour

In order to detect anomalies in system behaviour, first a model of normal or valid behaviour must be extracted from reference information collected by various means. The anomaly-based detection process later compares this model with the current activity. There is vast amount of literature on this topic; see e.g. the surveys in [19], [20].

A. Data Acquisition

To log near real-time data of the systems operation, the data acquisition process should be performed on the PLC. However, PLCs generally use industrial-grade memory cards and therefore offer no additional storage for saving log data. Thus, a Modbus TCP server was implemented on the PLC to allow communication with devices on the same network. In order to develop a model of normal system behaviour, the sensor and actuator data of the conveyor belt during normal operation was used as reference information. The operation was performed repeatedly and logged. On average, the time span of one process life cycle was 2 minutes and 20 seconds. Within this time frame approximately 2800 log entries (samples) were collected.

B. Feature Extraction

Feature extraction was used to reduce the dimensionality of the data, while still retaining sufficient accuracy. The following features were extracted:

- Minimum value; extracted only from non-Boolean variables
- Maximum value; extracted only from non-Boolean variables
- Arithmetic mean; extracted from all variables
- Standard deviation; extracted from all variables

The result of the feature extraction is a set of 32 features for each data file. Those examples were used to represent valid system behaviour.

Table I
EXTRACTED FEATURES

cntt_stdev	cntt_avg	temp_min	temp_max	temp_stdev	temp_avg
3.30167	4.82656	20.00000	60.00000	12.87631	53.30429
3.31753	4.77255	20.00000	60.00000	13.08286	53.13698
3.28911	4.75833	20.00000	60.00000	12.66858	53.60363
3.28840	4.80021	20.00000	60.00000	12.61219	53.63203
3.29452	4.76375	20.00000	60.00000	12.77102	53.50218
3.32203	4.71702	20.00000	60.00000	13.24500	53.04589

The set of features shows that during normal system operation deviations of the extracted features lie in a fairly small range as can be seen in Table I. The accuracy of the valid behaviour model highly depends on precise training data. However, some fluctuations are to be expected, as no system operation is identical.

C. Model of Valid Behaviour

To come up with a behaviour model we used machine learning to find the thresholds of the example data which show the range in which valid behaviour occurs. To this end we used RapidMiner [21], a software platform providing an Anomaly Detection Extension [22], which comprises the most well-known anomaly detection algorithms.

By experimenting with different algorithms the k-NN Global Anomaly Score operator showed the best results; it calculates the anomaly score according to the selected measure type. The operator offers a wide range of measure types including Euclidean distance, Kernel Euclidean distance, Manhattan distance, dynamic time warping distance and Bregman divergence. The calculations were performed using various measure types. For each measure type the process was conducted multiple times with different k values. A comparison of the results using different measure types is shown in Table II.

The process calculates the outlier score for each example in the training set. All examples are compared with each other and deviating data is flagged as an outlier. The higher the outlier the more anomalous the instance. From the results one can see that the examples have a high similarity, as the average outlier is fairly low.

The Table shows that the Bregman divergence measure type with $k = 3$ provides highly precise results, with an outlier average of 0.0287. Based on this result a threshold for the outlier score can be set to identify valid or anomalous behaviour from unlabelled data. The threshold was calculated from the highest outlier of the example sets, 0.0535, and multiplied by three in order to specify the range of normal behaviour. Thus, the threshold was set to 0.15. Using the operators Filter Examples, Map, and Append we produced a labelled dataset. Once the data is labelled it is added to the example set. Thus, the model is continuously built up and known anomalous behaviour is easily identifiable.

Table II
RESULTS OF DIFFERENT MEASURE TYPES

	Euclidean Distance			Kernel Euclidean Distance			Bregman Divergences		
k	3	5	7	3	5	7	3	5	7
Example 1	0.1555	0.1956	0.2493	0.0535	0.0829	0.1374	0.0273	0.0426	0.0725
Example 2	0.1501	0.1991	0.2623	0.0487	0.0865	0.1555	0.0248	0.0446	0.0834
Example 3	0.1645	0.2469	0.3479	0.0644	0.1413	0.2740	0.0331	0.0748	0.1566
Example 4	0.2074	0.2955	0.2649	0.1024	0.1952	0.3420	0.0535	0.1053	0.2004
Example 5	0.1599	0.1883	0.3811	0.0530	0.0733	0.1601	0.0269	0.0375	0.0865
Example 6	0.1814	0.2923	0.2321	0.0721	0.1950	0.3140	0.0369	0.1063	0.1809
Example 7	0.1193	0.1870	0.2417	0.0283	0.0812	0.1222	0.0143	0.0420	0.0641
Example 8	0.1563	0.1903	0.2374	0.0528	0.0769	0.1295	0.0269	0.0394	0.0684
Example 9	0.1430	0.1859	0.3160	0.0447	0.0757	0.1258	0.0227	0.0389	0.0661
Example 10	0.1386	0.2349	0.4001	0.0400	0.1312	0.2294	0.0202	0.0697	0.1278
Outlier Statistics									
Minimum	0.1193	0.1859	0.2321	0.0283	0.0733	0.1222	0.0143	0.0375	0.0641
Maximum	0.2074	0.2955	0.4001	0.1024	0.1952	0.3420	0.0535	0.1063	0.2004
Average	0.1576	0.2216	0.2933	0.0560	0.1139	0.1990	0.0287	0.0601	0.1107

V. ATTACK EXECUTION AND DETECTION

We assume that an attacker has already gained access to the control network using common attack routes and now is able to launch Man-in-the-Middle attacks. The system behaviour is logged during the attacks and compared with the valid behaviour model. Comparison is done by calculating the anomaly score and checking whether it lies within the thresholds defined for anomalies – in this case +- 0.15 of the highest outlier in the training data set.

The question is whether such cyber-attacks can be detected by looking at the deviations in the measured data from the valid behaviour model.

A Man-in-the-Middle attack (MitM) [23] is an attack in which the attacker secretly intercepts and relays messages between two parties who believe they are communicating directly with each other. In theory, any device that uses communication protocols that do not implement security features, such as encryption or advanced authorization algorithms, can be attacked.

A Man-in-the-Middle attack against two components of the conveyor belt system, the PLC and the Modbus logging client, was performed. This attack point is rather reasonable as logging will occur in any realistic industrial setting and will deliver all information an attacker would be interested in. The schematic of the MitM attack is shown in Figure 2. It shows the normal connection path between the Modbus logging client and the PLC, which runs the Modbus server. During normal operation the client requests data and the PLC sends a response containing the requested data. This data stream from the PLC to the Modbus logging client is intercepted by the attack.

The MitM attack targeting the logging client and the PLC was performed using Ettercap [24]. ARP poisoning was used to intercept all data between the components, as shown in the Figure. When using industrial protocols, this type of attack is virtually undetectable by the targets, as these protocols generally do not offer security features.

The MitM attacks were performed during the processing of the conveyor belt. The operation was captured with the

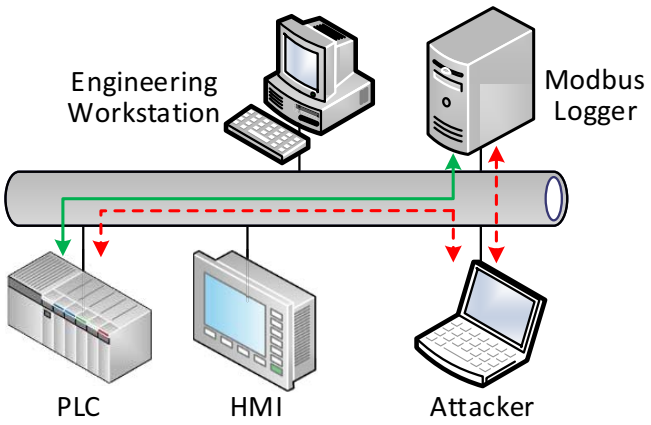

Figure 2. Schematic of the Man-in-the-Middle attack

Table III
RESULT OF OUTLIER DETECTION

Type	Label	Outlier	Prediction
Average of valid behavior	valid	0.028	valid
Man-in-the-Middle attacks	unknown	1.231	anomalous

logging client, then the features were extracted, and lastly the unlabelled data was compared with the valid behaviour model.

The outlier detection was done in RapidMiner using the same settings as before. The result of the RapidMiner process is shown in Table III. An outlier score of 1.231 was calculated as mean value for the logged data of the MitM attacks. Thus, it was classified as an anomaly. This can be attributed to the fact that even though the data in itself is not anomalous, the higher transmission time of the data, which is caused by the interception of the attacker, resulted in slight deviations of the values.

VI. CONCLUSION AND FURTHER WORK

To show the viability of anomaly detection in industrial control systems a simple project setup with industrial components (PLC, sensors, actuators, HMI) was set up and its network data generated during normal operations was used to train a behavioural model.

Then MitM attacks were launched against the system and the data generated was matched to the normal behaviour model. The results showed outlier values that were far beyond the thresholds defined for anomalies and hence made it possible reasonably large deviations to clearly conclude that an attack had taken place. The deviance was in all cases high enough and there were no false-positives during the experiment.

The project showed clearly that MitM attacks can be detected using anomaly detection. However, it should be noted that the valid system behaviour was captured in a lab environment exhibiting ideal network conditions, i.e. no unnecessary protocols or other data or noise were transmitted over the network. But this is not a very big shortcoming as real industrial implementations are often showing the same setting. Another difference to realistic settings is the number of sensors and actuators involved: usually a larger number of such elements is present in real industrial environments, which would make the process of feature selection during the machine learning a more challenging task.

Further work will comprise two activities: First, we will widen the spectrum of anomaly detection to other types of attacks, such as denial-of-service attacks or replay attacks. Second, we will try to apply the anomaly detection method to data gained from real industrial facilities, thus delivering even more realistic training sets for the machine learning process.

ACKNOWLEDGMENT

Our project is funded by the KIRAS program of the Austrian Research Promotion Agency (FFG). KIRAS funds projects in the field of security, with IT security being a subcategory in this context.

REFERENCES

[1] A. Hellinger and H. Seeger, "Cyber-physical systems. driving force for innovation in mobility, health, energy and production," *(acatech POSITION PAPER)*, 2011.

[2] G. Loukas, *Cyber-physical attacks: A growing invisible threat*. Butterworth-Heinemann, 2015.

[3] N. Falliere, L. O. Murchu, and E. Chien, "W32. stuxnet dossier," *White paper, Symantec Corp., Security Response*, vol. 5, 2011. [Online]. Available: {https://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf}

[4] G. R. Clarke, D. Reynders, and E. Wright, *Practical modern SCADA protocols: DNP3, 60870.5 and related systems*. Newnes, 2004.

[5] Symantec Security Response, "Dragonfly: Cyberespionage Attacks Against Energy Suppliers," https://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/Dragonfly_Threat_Against_Western_Energy_Suppliers.pdf, 2014, [Online; Status: Jun 16^{th} 2016].

[6] B. K. Kim, D. H. Kang, J. C. Na, and T. M. Chung, "Abnormal traffic filtering mechanism for protecting ics networks," in *2016 18th International Conference on Advanced Communication Technology (ICACT)*. IEEE, Jan 2016, pp. 436–440.

[7] S. Huang, C.-J. Zhou, S.-H. Yang, and Y.-Q. Qin, "Cyber-physical system security for networked industrial processes," *International Journal of Automation and Computing*, vol. 12, no. 6, pp. 567–578, 2015. [Online]. Available: http://dx.doi.org/10.1007/s11633-015-0923-9

[8] M. Mantere, M. Sailio, and S. Noponen, "Network Traffic Features for Anomaly Detection in Specific Industrial Control System Network," *Future Internet*, vol. 5, no. 4, pp. 460–473, 2013.

[9] E. A. Lee, "Cyber Physical Systems: Design Challenges," EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2008-8, Jan 2008. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-8.html

[10] S. Adepu, A. Mathur, J. Gunda, and S. Djokic, *Algorithms and Architectures for Parallel Processing: 15th International Conference, ICA3PP 2015, Zhangjiajie, China, November 18-20, 2015, Proceedings, Part III*. Cham: Springer International Publishing, 2015, ch. An Agent-Based Framework for Simulating and Analysing Attacks on Cyber Physical Systems, pp. 785–798. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-27137-8_57

[11] K. Stouffer, J. Falco, and K. Scarfone, "Guide to industrial control systems (ICS) security," *NIST special publication*, vol. 800, no. 82, 2011.

[12] E. A. Lee and S. A. Seshia, *Introduction to embedded systems: A cyber-physical systems approach*. LeeSeshia.org, 2011.

[13] D. Harp and B. Gregory-Brown, "The State of Security in Control Systems Today," SANS Institute, Tech. Rep., Jun 2015. [Online]. Available: https://www.sans.org/reading-room/whitepapers/analyst/state-security-control-systems-today-36042

[14] S. Karnouskos, "Stuxnet worm impact on industrial cyber-physical system security," in *IECON 2011 - 37th Annual Conference on IEEE Industrial Electronics Society*. IEEE, Nov 2011, pp. 4490–4494.

[15] X. Fan, K. Fan, Y. Wang, and R. Zhou, "Overview of cyber-security of industrial control system," in *2015 International Conference on Cyber Security of Smart Cities, Industrial Control System and Communications (SSIC)*. IEEE, Aug 2015, pp. 1–7.

[16] Siemens AG, "Product Details 6ES7214-1BG40-0XB0," https://mall.industry.siemens.com/mall/en/WW/Catalog/Product/6ES7214-1BG40-0XB0, 2016, [Online; Status: May 11^{th} 2016].

[17] ——, "Product Details 6AV6647-0AD11-3AX0," https://mall.industry.siemens.com/mall/en/WW/Catalog/Product/6AV6647-0AD11-3AX0, 2016, [Online; Status: May 11^{th} 2016].

[18] Art of Circuits, "Infrared Obstacle Avoidance Proximity Sensors Module FC-51," http://artofcircuits.com/product/infrared-obstacle-avoidance-proximity-sensors-module-fc-51, 2016, [Online; Status: May 11^{th} 2016].

[19] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 15:1–15:58, Jul. 2009. [Online]. Available: http://doi.acm.org/10.1145/1541880.1541882

[20] V. J. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artificial Intelligence Review*, vol. 22, no. 2, pp. 85–126, 2004.

[21] RapidMiner Inc., "RapidMiner Official Website," https://rapidminer.com/, 2016, [Online; Status: Jun 16^{th} 2016].

[22] M. Goldstein, M. Amer, J. Gebhardt, P. Kalka, and A. Elsawy, "RapidMiner Anomaly Detection Extension," https://github.com/Markus-Go/rapidminer-anomalydetection, 2015, [Online; Status: Jun 16^{th} 2016].

[23] A. Ornaghi and M. Valleri, "Man in the middle attacks," in *Blackhat Conference Europe*, 2003.

[24] E. Project, "Ettercap Home Page," https://ettercap.github.io/ettercap/, 2016, [Online; Status: Jun 16^{th} 2016].