Enabling Convergence of Physical and Logical Security Through Intelligent Event Correlation

Gianfranco Cerullo, Luigi Coppolino, Salvatore D'Antonio, Valerio Formicola, Gaetano Papale and Bruno Ragucci

Abstract Until now, in most organizations, physical access systems and logical security systems have operated as two independent elements, and have been managed by completely separate departments. The lack of interoperability between the two sectors often resulted in a security hole of the overall infrastructure. An attacker who has physical access can not only steal a PC or confidential data, but can also compromise network security. Therefore, a combination of physical and logical security definitively allows for a more effective protection of the organization. In this work we present a correlation system which aims at bringing a significant advancement in the convergence of physical and logical security technologies. By "convergence" we mean effective cooperation (i.e. a coordinated and results-oriented effort to work together) among previously disjointed functions. The holistic approach and enhanced awareness technology of our solution allows dependable (i.e. accurate, timely, and trustworthy) detection and diagnosis of attacks. This ultimately results in the achievement of two goals of paramount importance, and precisely guaranteeing the protection of citizens and assets, and improving the perception of security by citizens. The effectiveness of the proposed solution is demonstrated in a scenario that deals with the protection of a real Critical Infrastructure. Three misuse cases have been implemented in a simulation environment in order to show how the correlation system allows for the detection of different attack patterns.

G. Cerullo (✉) · L. Coppolino · S. D'Antonio · V. Formicola · G. Papale · B. Ragucci
University of Naples "Parthenope", Napoli, Italy
e-mail: gianfranco.cerullo@uniparthenope.it

L. Coppolino
e-mail: luigi.coppolino@uniparthenope.it

S. D'Antonio
e-mail: salvatore.dantonio@uniparthenope.it

V. Formicola
e-mail: valerio.formicola@uniparthenope.it

G. Papale
e-mail: gaetano.papale@uniparthenope.it

B. Ragucci
e-mail: bruno.ragucci@uniparthenope.it

P. Novais et al. (eds.), *Intelligent Distributed Computing IX*,
Studies in Computational Intelligence 616,
DOI 10.1007/978-3-319-25017-5_40

1 Introduction

Technologies for implementing security services in the physical and in the logical domain are both stable and mature, but they have been developed independently of each other. Some of them have recently merged, but real convergence of physical and logical security technologies is still a faraway target. By "convergence" we mean: effective cooperation (i.e. a coordinated and results-oriented effort to work together) among previously disjointed functions. In the recent years, some achievements have been made, but much is yet to be done. As an example, Security Operation Centers (SOC) technology has improved significantly, but SOC solutions have typically been developed using vertical approaches, i.e. based on custom specific needs. In this paper we focus on SOC technology as key tool for increasing security of critical infrastructures through the convergence of physical and logical security. Specifically, a SOC aims to effectively detect and diagnose cyber-attacks and, in order to do so, it collects and analyses activity reports (e.g. system logs, notification and alert messages, traps, etc.)—also known as "events"—provided automatically by electronic and computer systems or manually by the personnel operating on the infrastructure. In order to be effective, the analysis performed by a SOC must be dependable, i.e. accurate, timely and trustworthy.

- Accurate—The detection rate will be high (i.e. a very high percentage—higher than what is currently achieved by state-of-the-art products—of real attacks will be detected) and the false positive rate will be low (i.e. a very low percentage—lower than what is currently achieved by state-of-the-art products—of real attacks will go undetected). It is worth emphasizing that in contexts such as highly available systems and applications (e.g. Critical Infrastructures) and crowded places (e.g. a stadium or an airport), false alarms can be as dangerous and harmful as false negatives [1]. Accuracy will be achieved by performing sophisticated correlation of the multitude of diverse events which will be collected in the two domains (namely: logical and physical). Evidence is demonstrating that this approach is effective [2–4].
- Timely—The aforementioned sophisticated correlation will be done in near real-time. This is a challenging task, since the amount of data that the system will have to process is massive and highly heterogeneous (both from the format and from the semantics point of view).
- Trustworthy—A largely overlooked issue in the design and development of security products is "who defends the defender" [5–7]. This means that the SOC platform has to be designed and implemented using fault- and intrusion-tolerant techniques. The platform will thus be resilient to fault and attacks, i.e. it will be able to perform its tasks correctly even in the presence of faults and/or if it will be under attack.

Typical systems that provide SOC with data are physical access control systems with real-time data processing features, service monitoring systems, infrastructure performance monitors, logical security systems. A number of facilities is also available to

enforce controls for safeguarding the operation of the system, as well as for protecting the surrounding environment. In this paper we illustrate how a SOC can be used to detect attacks that are perpetrated by company employees and are usually referred to as "internal" attacks. In particular, the SOC is required to understand whether an outage is due to a misconfiguration caused by a legitimate maintenance operation or it is the effect of a malicious attack. In addition to information about regular operations, a SOC analyzes the information contained in the maintenance reports in order to distinguish the following cases:

- Events representing planned maintenance operations;
- Events representing failures of specific system components due to non-malicious faults;
- Events representing failures of specific system components due to malicious faults (attacks).

The paper is structured as follows. Section 2 presents the general architecture of a Security Operation Center. Section 3 illustrates the proposed correlation system, which allows to improve the capability of a Security Operation Center to combine physical and logical security technologies, thus achieving a higher attack detection performance. In Sect. 4 the correlation system is validated in three misuse cases where the attack strategy consists in exploiting both physical and logical security vulnerabilities. Finally, Sect. 5 provides some concluding remarks.

2 Architecture of a Security Operation Center

A Security Operation Center monitors and manages several types of security events to perform Real Time Device Monitoring (RTDM), Network Fault Management, Security Incident Management, Policy Management and Enforcement, Vulnerability Assessment and Policy Compliance Verification.

The following subsections present the tasks performed by a SOC that contribute to the implementation of the security scenarios addressed by this paper.

2.1 Real Time Device Monitoring

Real Time Device Monitoring is a continuous activity for real time monitoring of security-related events. It manages events generated by network devices (routers, switches), security devices (firewalls, IDS/IPS, antivirus, etc.), servers, and applications (e.g., web servers, application servers, proxies, etc.), physical access control systems (badges, intrusion detection systems, door and window alarms, etc.). RTDM systems are used for accurate analysis of alarms or events generated by monitored devices and initialization of corrective actions for alarms that exceed specific security

thresholds; implementation of appropriate alerting mechanisms in accordance with defined procedures and escalation strategies; definition of new correlation rules to identify new threats; and tuning of existing correlation rules to avoid/reduce false positives. Security Information and Event Management (SIEM) systems are largely used to perform Real Time Device Monitoring activities.

A SIEM system is responsible for collection and correlation of all the events coming from the operational domain context and from the corporate areas.

2.2 Video Surveillance

In remote sites, the exterior zones of critical infrastructures are exposed to (sometime) extreme weather conditions and the use of video surveillance is critical and complex. In this case video surveillance is the complement to burglar alarms, and is used to minimize false alarms of physical violations. Indeed, borders have peculiarity for which intrusion can occur and must be accepted at a certain degree. Since guards are not everywhere, an attacker could tag along to a car in transit hiding himself from view. In these circumstances, video surveillance is a support for forensic and investigations, not being possible a continuous view and detection over all the facilities disseminated in the sites.

2.3 Physical Access Control

Physical Access Control systems must provide identification, authentication and authorization of people entering and exiting each zone of the infrastructure. Authentication can be single or double factor based. The single factor authentication—typically badge control—is less strong, but is the most commonly implemented, because there is no need for acquiring additional (e.g. biometric) parameters from the user. Also, in some countries there are laws limiting the usage of biometric data for physical security. One of the most important requirements for a physical access control is the analysis of authentication attempts. Indeed, physical access attempts must be recorded in order to discover when suspicious thresholds are exceeded. In order to properly supervise physical alarms, the SOC must be correctly tuned through severity or priority values. Indeed, in some operational contexts many physical access events are not so relevant, i.e. must be associated to low priority warnings. In other contexts, such as access to very critical rooms, attempts must trigger highest priority alerts.

Fig. 1 Main functional blocks of a real-time intelligent event correlation system

3 Correlation System

The main contribution of this paper is the development and validation of a correlation engine that has been implemented in order to enhance the capability of a Security Operation Center to protect a critical infrastructure from sophisticated attacks involving both the physical and logical domain. Real time correlation allows to combine huge amounts of micro-data generated by heterogeneous information sources, and obtain semantically richer macro description of faults in real time. This process is a key building block for a Situation Aware Security Operation Center since it enables timely and accurate detection and diagnosis of (both maliciously induced and not) faults on complex critical systems. The real time correlation process involves three main tasks, which are: collection and preprocessing of events at the edge of the SOC framework and in proximity of the data sources; distribution of these events from the edge to the core processing systems; data processing, i.e. correlation of information belonging to multiple layers of the infrastructure; semantic fusion of the information and final generation of ranked evidence. This processing chain is represented in the Fig. 1.

3.1 Data Collection

The collection task is in charge of gathering data generated from heterogeneous information sources, and to output these messages in a format which is processable by the centralized correlation engine. The main sub-tasks performed by the collection system can be summered in: data gathering, i.e. collection of data based on different transfer protocols; message format parsing, i.e. tokenization of fields from variously structured messages; message filtering, i.e. dropping of irrelevant messages; message pre-aggregation, i.e. coalescence of similar entries; format normalization, i.e. generation of fields in a standard format; forwarding, i.e. propagation of the events to the core processing. The parsing step extracts tokens from streams of events represented in syntactically and semantically heterogeneous data formats. In order to identify the input format, an "Event Id" is typically configured on the parsing system and is associated with the specific information source. For instance, IP address of the source, data transfer protocol used, collection port, session-id and tags can be used to identify the input format. Filtering of events is typically based on Regular Expressions (RegEx), so that it is possible to associate the specific filter to the required class of

events. Filters wait for new Parsed Events to operate, or can be optimized to work during the parsing process. In the latter case, events matching the dropping RegEx rule are discarded and the parsing consumes the next message. Pre-aggregation step performs a first level of aggregation when similar entries are coalesced into a single message. In this case semantic reasoning is still required and is delegated to the core processing, as we see below.

3.2 Data Distribution

The message forwarding model is demanded to the requirements of the processing systems, i.e. how many processors will take charge of evaluating and processing the collected events. An effective solution comes from the Publish/Subscribe mechanism, which enables the publisher (i.e. the data collector in this case) to publish a single message that can be consumed by multiple subscribers (i.e. the correlation system and others): the Publisher only takes care of publishing its message on a "topic" hosted by some broker; the latter provides the message to the consumers subscribed to that topic. Finally, this messaging model supports asynchronous communication, and improves the scalability of the system. The messaging system must ensure reliability by guaranteeing the delivery of messages, through a trade-off between throughput and reliability. It is possible to introduce different levels of reliability, which assign to messages a greater or lesser relevance. Furthermore the messaging systems are usually supported by persistent storage systems that preserve messages and protect them from attacks aiming at violating data confidentiality and integrity. Another aspect is the robustness of the system. Actually publishers, subscribers and network can have failures, and redundancy can mitigate this issue. In addition to the common message brokering systems, such as Java Message Systems (JMS), a very effective technology is provided by Apache Kafka, that combines the model of messaging system with log aggregators. Also, it implements scalable and distributed processing of queues and topics.

3.3 Data Processing

Data processing task concerns with the centralized correlation of events coming from the distributed collectors at the edge. The centralized processing enables global view of the infrastructure state, and can take advantage of high computing resources available at the core systems. The correlation process aims at finding a relation among the data fields composing the normalized events, and eventually at producing an aggregated message. The aggregated output contains fields extracted from the input events and metrics obtained by combining each evidence. Whatever the metrics and the data fusion model—one will be discussed in the next section—, the aggregated message outputs the security risk level of aggregated events. In order to have effective sit-

uational awareness of the system state, it is of paramount importance refining the huge amount of information obtained from the monitoring systems. This ultimately means correlating both the information coming from the system operations (e.g. system logs, security incidents) with information related to the operational context (e.g. maintenance plans, physical events). One of the most widely and effective solutions to correlate streams of events is Complex Event Processing (CEP). We use CEP for defining relations among the events, i.e. to describe the correlation model (e.g. a simple matching rule or more sophisticated inter-event analysis), for combining metrics useful to rank the alerts, and for defining the structure of the output alert message. An example of CEP is EsperTech Esper [19]. The computational load of this processing is distributed by means of high performance and dependable computing technologies. An effective solution to merge the semantics of Esper with distributed, scalable and fault-tolerant processing is Apache Storm [18].

3.4 Data Fusion

Data Fusion [8] is the process of combining information from a number of different sources to provide a robust and complete description of an environment or process of interest. Data Fusion process is applied where a large amounts of data must be combined and fused to obtain information of appropriate quality and integrity on which decisions can be made. In any data fusion problem, there is an environment, process or quantity whose true value, situation or state is unknown. The sources provide some parameters, imperfect and incomplete knowledge, that are processed and then transformed in decisions, that provides effective support for human or automated decision making. Data fusion is the process of combining data to refine estimates and predictions of the state that is observed. Joint Directors of Laboratories (JDL) defines data fusion as a "multilevel, multifaceted process handling the automatic detection, association, correlation, estimation, and combination of data and information from several sources". The proposed correlation engine exploits the features provided by the Dempster-Shafer Theory that allows to combine multiple pieces of evidence for detecting an ongoing attack.

4 Misuse Cases

The Correlation Engine (CE) aims to correlate relevant information from the physical and the electronic domain in an effective way to ensure the security and the detection of potential attacks and threats. The information, coming from a huge amount of sources, is aggregated in real-time fashion using correlation rules. A correlation rule aggregates symptoms based on a set of parameters, such as the attack type, the target component and the temporal proximity. Alerts are not generated as results of all the monitored symptoms, but only when the correlation among such symptoms

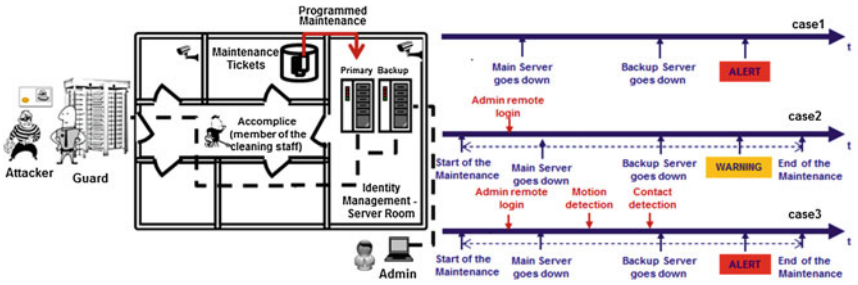

Fig. 2 Attack scenario storyboard

indicates a potential attack, thus reducing the number of false positives and improving the detection capability of the overall system. To explain the functioning of our solution a storyboard with three possible cases of attack is presented. The actors involved in the selected misuse cases are the Maintenance Scheduling Programming (MSP), the Network Management System (MS) and the Videosurveillance System (VS). An accomplice of an attacker wants to take advantage of a scheduled maintenance service on the server which manages the identity of users, named Primary Server, in order to allow the attacker to enter the building. When the Primary Server is down a Backup Server replaces it. The accomplice enters the Identity Management Server Room and opens the rack, in which the servers are placed. After that, he unplugs the Ethernet cord of the Backup Server, so that the system for identity management fails down, and the attacker can enter the building with a fake badge, as shown in Fig. 2.

- **Case 1**

  In this case we consider only the MS. The Primary Server goes down and the Backup Server replaces it. After a while also the Backup Server goes down and the CE is aware of the anomalous events occurring within the infrastructure and raises an alert. The CE observing an outage of the Primary Server and a non-operation of the Backup Server considers this situations as a possible symptom of an attack since the identity management system is out of service.

- **Case 2**

  The actors involved in this scenario are MSP and MS. The first provides a maintenance ticket in which it warns that the Primary Server stops the services offered, due to a maintenance job performed by the system administrator. The administrator logs in remotely and the Primary Server stops working. At the same time the Backup Server replaces it, afterwards also the Backup Server goes down. The CE correlates and aggregates the information provided by the two sources and raises a warning because it considers this event as a possible attack. It is aware that Primary Server is down for a maintenance operation. As soon as the Backup Server goes down, the CE detects this situation as a malfunction or an ongoing attack.

- **Case 3**

  MSP, MS and VS are the actors involved in this misuse case. The maintenance ticket is sent to the CE and the administrator logs in remotely on the Primary Server, that goes down. Now an accomplice of the attacker enters the Identity Management Server Room, the VS detects his presence and sends a motion detection log to the CE. The CE considers the aggregated event as a normal situation because the maintenance ticket warns that the Primary Server will be inactive for a certain time window and the motion detection does not indicate an imminent attack. After this the accomplice opens the rack, in which the Primary and Backup Server are placed. At the same time the VS detects the contact and sends a log to the CE that raises a warning since it interprets the information regarding the contact with the rack as relevant for safety purposes. After few seconds the accomplice unplugs the Ethernet cord of the Backup Server, which goes down. The CE aggregates the information in a single event that is processed and through the correlation rule returns the evidence of an attack. Then an alarm is raised.

5 Related Work

The use of correlation techniques for attack and intrusion detection has been largely explored in literature. Some relevant papers dealing with this research topic are presented below. In [8] the authors propose a Simple Event Correlator relying on a rule-based correlation approach, that is used to detect and filter out relevant symptoms useful for fault diagnosis in a Supervisory Control and Data Acquisition (SCADA) infrastructure. [9] presents a Generic Intrusion Detection and Diagnosis System for detection and diagnosis of complex attack patterns in large scale Critical Infrastructures. In [10] a comprehensive analysis of the cyber-security issues concerning Smart Grids, specifically network vulnerabilities, attack countermeasures, secure communication protocols and architectures, is performed. In [11] the authors present an Intrusion Detection System (IDS) for correlating attack symptoms from diverse information sources. The presented IDS relies on an ontology to drive the correlation task and is implemented as a distributed and highly scalable system. In [12] the authors identify limits of the current SIEM systems and propose a framework to enhance services for data treatment. They also provide prototypal deployment of a case study consisting in securing a dam monitoring and control system. In [13], the authors provide a performance comparison of the most popular open source rule based correlation engines. A distributed event correlation system which performs security event detection is presented in [14]. The system detects several misuse cases, with a low false positive rate. Our solution involves Level 0—Source Pre-Processing, Level 1—Object Refinement and Level 2—Situation Refinement of the Joint Directories Laboratory (JDL) Data Fusion Process Model. A description of the JDL for cyber-security is given in [15]. Information fusion is exploited in [16] to spot frauds against a mobile money transfer service by using combination rules of the Dempster-Shafer Theory.

6 Conclusions

In this paper we addressed the need for convergence of physical and logical security in order to enhance the protection of critical assets. The convergence of these two worlds brings positive effects to the general security of an organization. Integrating the efforts of physical and logical security departments allows an organization to significantly lessen security risks while also saving time and money. This paper presented a correlation system capable of collecting and processing security relevant information and events from both physical and logical domain, thus enabling the convergence of these two security areas. The proposed system has been validated in three different scenarios, where the correlation of events generated by heterogeneous probes made it possible the detection of sophisticated attacks exploiting both physical and logical security vulnerabilities.

Acknowledgments The research leading to these results has received funding from the European Commission within the context of the Seventh Framework Programme (FP7/2007–2013) under Grant Agreement No. 313034 (Situation AWare Security Operation Center, SAWSOC Project). It has been also partially supported by the TENACE PRIN Project (n. 20103P34XC) funded by the Italian Ministry of Education, University and Research, and by the Embedded Systems in critical domains POR Project (CUP B25B09000100007) funded by the Campania region in the context of the POR Campania FSE 2007–2013, Asse IV and Asse V.

References

1. Tips to reduce false security alarms with proper installation, education and training. http://www.sourcesecurity.com/news/articles/co-2173-ga.4866.html
2. Repp, N., Berbner, R., Heckmann, O., Steinmetz, R.: A cross-layer approach to performance monitoring of web services. In: Proceedings of the Workshop on Emerging Web Services Technology, CEUR-WS, December 2006
3. Yu-Sung, W., Bagchi, S., Garg, S., Singh, N.: SCIDIVE: a stateful and cross protocol intrusion detection architecture for voice-over-IP environments. In: Proceedings of Dependable Systems and Networks Conference, 28 June 2004, pp. 433–442 (2004)
4. Vigna, G., Robertson, W., Vishal, K., Kemmerer, R.A.: A stateful intrusion detection system for World-Wide Web servers. In: Proceedings of the 19th Annual Computer Security Applications Conference, 8–12 December 2003, pp. 34–43 (2003)
5. Verssimo, P., Correia, M., Neves, N., Sousa, P.: Intrusion-resilient middleware design and validation. In: Information Assurance, Security and Privacy Services (Handbooks in Information Systems, vol. 4), Emerald Group Pub. Ltd., pp. 615–678 (2009)
6. Sousa, P.: Proactive Resilience. In: Proceedings of the 6th European Dependable Computing Conference (EDCC-6) Supplemental Volume, Coimbra, Portugal, October (2006)
7. Dondossola, G., Deconinck, G., Di Giandomenico, F., Donatelli, S., Kaaniche, M., Verssimo, P.: Critical utility infrastructure resilience. In: Workshop on Security and Networking in Critical Real-Time and Embedded Systems (CRTES'06), with RTAS'06, San Jose, California, April (2006)
8. Ficco, M., Daidone, A., Coppolino, L., Bondavalli, A.: An event correlation approach for fault diagnosis in SCADA infrastructures. In: Proceedings of the 13th European Workshop on Dependable Computing (EWDC 2011), Pisa, Italy, May 2011, pp. 15–20. ACM Press (2011). doi:10.1145/1978582.1978586

9. Ficco, M., Romano, L.: A generic intrusion detection and diagnoser system based on complex event processing. In: Proceedings of the 1st International Conference on Data Compression, Communications and Processing (CCP 2011), Palinuro, Italy, June 2011, pp. 275–284. IEEE CS Press (2011). doi:10.1109/CCP.2011.43
10. Wang, W., Lu, Z.: Cyber security in the smart grid: survey and challenges. Comput. Netw. **57**(5), 1344–1371 (2013). doi:10.1016/j.comnet.2012.12.017
11. Coppolino, L., D'Antonio, S., Esposito, M., Romano, L.: Exploiting diversity and correlation to improve the performance of intrusion detection systems. In: Proceedings of the International Conference on Network and Service Security, N2S'09, Paris, June 2009, pp. 24–26 (2009)
12. Coppolino, L., D'Antonio, S., Formicola, V., Romano, L.: Enhancing SIEM Technology to Protect Critical Infrastructures. Critical Information Infrastructures Security Lecture Notes in Computer Science 7722, 10–21 (2013)
13. Rosa, L., Alves, P., Cruz, T., Simes, P., Monteiro, E.: A comparative study of correlation engines for security event management. In: Proceedings of the 10th International Conference on Cyber Warfare and Security (ICCWS-2015), Kruger National Park, South Africa (2015)
14. Myers, J., Grimaila, M.R., Mills, R.F.: Log-based distributed security event detection using simple event correlator. In: System Sciences (HICSS), 2011 44th Hawaii International Conference, Kauai, 4–7 January 2011. doi:10.1109/HICSS.2011.288
15. Giacobe, N.A.: Application of the JDL data fusion process model for cyber security. In: Proceedings of the SPIE 7710, Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2010, 77100R, 28 April 2010. doi:10.1117/12.850275
16. Coppolino, L., D'Antonio, S., Formicola, V., Massei, C., Romano, L.: Use of the Dempster Shafer theory to detect account takeovers in mobile money transfer services. J. Ambient Intell. Humaniz. Comput. (April 2015). doi:10.1007/s12652-015-0276-9
17. Multi Sensor Data Fusion: Hugh Durrant-Whyte, Australian Centre for Field Robotics, The University of Sydney NSW 2006, Australia (2006)
18. Apache Storm. https://storm.apache.org/ (2015). Accessed 15 April 2015
19. EsperTech Esper: http://www.espertech.com/esper/index_redirected.php (2015). Accessed 15 April 2015