Secure and Private Smart Grid: The SPEAR Architecture

Panagiotis Radoglou Grammatikis[†], Panagiotis Sarigiannidis[†], Eider Iturbe[‡], Erkuden Rios[‡],
Antonios Sarigiannidis[§], Odysseas Nikolis[¶], Dimosthenis Ioannidis[¶], Vasileios Machamint[‖],
Michalis Tzifas[‖], Alkiviadis Giannakoulias[**], Michail Angelopoulos[††],
Anastasios Papadopoulos[‡‡], Francisco Ramos[x]

***Abstract*—Information and Communication Technology (ICT) is an integral part of Critical Infrastructures (CIs), bringing both significant pros and cons. Focusing our attention on the energy sector, ICT converts the conventional electrical grid into a new paradigm called Smart Grid (SG), providing crucial benefits such as pervasive control, better utilisation of the existing resources, self-healing, etc. However, in parallel, ICT increases the attack surface of this domain, generating new potential cyberthreats. In this paper, we present the Secure and PrivatE smArt gRid (SPEAR) architecture which constitutes an overall solution aiming at protecting SG, by enhancing situational awareness, detecting timely cyberattacks, collecting appropriate forensic evidence and providing an anonymous cybersecurity information-sharing mechanism. Operational characteristics and technical specifications details are analysed for each component, while also the communication interfaces among them are described in detail.**

***Index Terms*—Anomaly Detection, Anonymity, Cybersecurity, Forensics, Honeypots, Intrusion Detection, Privacy, Smart Grid**

I. Introduction

In the era of hyper-connected digital economies, Information and Communication Technology (ICT) does not constitute only an extension or service of the Critical Infrastructures (CIs) but plays a significant role at the core of their operation. In particular, regarding the energy sector, the traditional electrical grid is evolving into a new paradigm called Smart Grid (SG) where ICT introduces new capabilities, such as two-way communication (both electricity and information), distributed generation, pervasive control, self-monitoring, self-healing, etc. According to S. Tan et al. [1], SG will constitute the largest Internet of Things (IoT) application, where appropriate services will optimise the typical phases of the existing electrical grid, namely a) generation, b) transmission and c) distribution. However, this new reality raises in parallel serious cybersecurity challenges, thus exposing worldwide governments and businesses into new risks. More specifically, SG is a primary target of cybercriminals since it operates as the backbone of each CI; therefore, potential cyberattacks against SG can cause devastating consequences and cascading effects, thus affecting the overall economy and even worse causing fatal accidents.

Focusing our attention on the cybersecurity issues of SG, it inherits the vulnerabilities of the involved technologies. In particular, a vital ingredient of the electrical grid is the Supervisory Control and Data Acquisition (SCADA) systems that monitor and control critical automation processes. However, SCADA systems are characterised by severe cybersecurity issues since their operation relies on insecure communication protocols that do not comprise any authentication and access control mechanism, thus allowing cyberattacks threatening the confidentiality, integrity and authenticity of the exchanged information. Characteristic examples are Man-in-the-Middle (MiTM) attacks [2], unauthorised access [2] and false data injection attacks. On the other side, the advent and adoption of IoT affect the overall security status of SG. First, the ability of objects like sensors and actuators to communicate with each other without any human intervention creates both significant security and privacy concerns [3]. For instance, the constrained nature of IoT devices concerning the computing resources allow the execution of successful Denial of Service (DoS) attacks. Moreover, since IoT is based on the insecure Internet model, the corresponding threats and vulnerabilities should be taken into account. Finally, the new characteristics introduced by IoT bring also several cyberattacks.

It is obvious that sufficient countermeasures should be adopted for protecting efficiently SG. In this paper, we analyse the Secure and PrivatE smArt gRid (SPEAR) architecture which aims to provide an entire solution regarding the timely

[*] This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 787011 (SPEAR).

[†] P. Radoglou Grammatikis and P. Sarigiannidis are with the Department of Electrical and Computer Engineering, University of Western Macedonia, Kozani 50100, Greece - E-Mail: {pradoglou, psarigiannidis}@uowm.gr

[‡] E. Iturbe and E. Rios are with TECNALIA, Basque Research and Technology Alliance (BRTA), Derio, Spain - E-Mail: {Eider.Iturbe, Erkuden.Rios}@tecnalia.com

[§] A. Sarigiannidis is with Sidroco Holdings Ltd, 3113, Limassol, Cyprus - E-Mail: asarigia@sidroco.com

[¶] O. Nikolis and D. Ioannidis are with the Center for Research and Technology Hellas / Information Technologies Institute, 6th km Charilaou-Thermi Road, Thessaloniki, Greece - E-Mail: {odynik, djoannid}@iti.gr

[‖] V. Machamint and M. Tzifas are with Eight Bells Ltd, Agias paraskevis 23, P.C. 2002, Strovolos, Nicosia, Cyprus - E-Mail: {vasilis.machamint, tzifas}@8bellsresearch.gr

[**] A. Giannakoulias is with European Dynamics, 12, Jean Engling str. L-1466 Luxembourg - E-Mail: alkiviadis.giannakoulias@eurodyn.com

[††] M. Angelopoulos is both with the Department of Economics, University of Piraeus and the Testing Research & Standards Center / Public Power Corporation SA, Leontariou 9, Kantza, 15351, Athens, Attica - E-Mail: m.angelopoulos@dei.com.gr

[‡‡] A. Papadopoulos is with the Testing Research & Standards Center / Public Power Corporation SA, Leontariou 9, Kantza, 15351, Athens, Attica - E-Mail: a.papadopoulos@dei.com.gr

[x] F. Ramos is with the Schneider Electric, Charles Darwin s/n, Edificio Bogaris, 41092 Sevilla, Spain - E-Mail: francisco.ramos@schneider-electric.com

978-1-7281-5684-2/20/$31.00 ©2020 IEEE

detection of possible cyberattacks against SG, considering in parallel privacy-related issues and the collection of the appropriate forensic-related elements. Moreover, SPEAR intends to enhance situational awareness of energy-related stakeholders, by establishing an anonymous repository of incidents where the involved members will be able to share with each other technical details about the various cybersecurity incidents without endangering their reputation. It should be noted that SPEAR is a research Horizon 2020 programme cofounded by the European Union (EU) which will be validated against four real use cases.

Based on the aforementioned remarks the contribution of this paper is summarised in the following sentences.

- Providing an entire, secure architecture concerning the efficient protection of SG, taking into account the relevant privacy issues and the EU legislation.
- Defining the technical details and specifications of the proposed architecture.
- Determining the communications among the involved components, specifying the corresponding technical details.

The rest of this paper is organised as follows: Section II describes previous works related to the research areas that SPEAR deals with. In section III, we provide the methodological framework utilised to form the SPEAR architecture, while section IV analyses its components. Finally, section V concludes this paper.

II. Related Work

This section describes previous works related to the research areas that SPEAR deals with, namely: a) intrusion detection using big data and visual analytics, b) honeypots, c) network forensics, d) cybersecurity information sharing and e) cybersecurity training. Based on the insights of these works, Section IV presents the SPEAR architecture and its novelties. It should be noted that the last research area (cybersecurity training) is not analysed here since this paper focuses mainly on the technical aspects of SPEAR.

A. SIEM and IDS Systems

Security Information and Event Management (SIEM) tools constitute an emerging technology capable of collecting, normalising and analysing data from various sources, thus generating security events. In [4], R. Leszczyna and M. Wróbel assess three open-source SIEM tools regarding their efficacy in protecting SG. According to the evaluation results, AlienVault Open Source SIEM (OSSIM) presents the most beneficial characteristics. Accordingly, in [5], K. Kavanagh and T. Bussa examine multiple SIEM systems, documenting their strong points and weaknesses. Finally, in [6], the H2020 DiSIEM project analyses seven SIEMs, namely a) HP ArchSight, b) IBM QRadar, c) Intel McAfee Enterprise Security Manager, d) Alienvault OSSIM and USM, e) ATOS XL-SIEM, e) Splunk and f) Elastic Stack in terms of various criteria, including a) data sources supported, b) data storage capabilities, c) processing capabilities, d) flexibility in security directives, f) behavioural analysis at application-level, g) risk analysis capacity, h) exposed APIs, i) resilience, j) security event management and visualisation capabilities, k) reaction capabilities, l) simplicity of deployment and support provided and m) licensing.

In [7], the authors provide a detailed analysis of various Intrusion Detection Systems (IDS) for SG and specific directions for future work. Similarly, in [8], Vasilomanolakis et al. investigate various Collaborative Intrusion Detection Systems (CIDS) devoted to protecting large Information Technology (IT) and critical infrastructures. In particular, the authors document first the main requirements of such CIDS, including a) accuracy, b) minimal overhead, c) scalability, d) resilience, e) privacy, f) self-configuration and g) interoperability. Next, they describe their basic building blocks and discuss disclosure and evasion techniques against them. Finally, R. Mitchell and I. Chen in [9] present a noteworthy survey in which multiple IDS dedicated to protecting Cyber-Physical Systems (CPS) are investigated, providing also directions for future work.

B. Honeypots

Honeypots are entities emulating the behaviour of real assets, but they do not have real production value [10]. Their purpose is to mislead possible cybercriminals to execute attacks against them instead of the real assets [10]. In [11], C. Dalamagkas et al. present a survey exclusively related to the honeypot applications on SG. In particular, they describe and compare many honeypots and honeynets such as Honeyd, HoneydV6, Conpot, CryPLH, ShaPe, CockpitCI, DiPot, etc. in terms of the services supported. On the other side, in [12], M. Nawrocki et al. provide a more extensive study where a plethora of honeypots is analysed, by classifying them into four main categories: a) Low Interaction Server Honeypots, b) High Interaction Server Honeypots, c) Low Interaction Client Honeypots, d) High Interaction Client Honeypots. Next, the authors focus on data analytics related to the information collected by honeypots, such as attack profile, attack target, attack frequency, attack propagation and attack patterns. Finally, in [13], W. Fan et al. provide a taxonomy about various honeypots, considering multiple characteristics, including a) fidelity, b) physicality/virtuality, c) scalability, d) adaptability, f) role, g) deployment strategy, h) resource type, i) attack monitoring, j) attack prevention, k) attack detection, l) attack response and m) attack profiling.

C. Network Forensics

Unfortunately, although there are many studies related to network forensics [14], [15], only a bit of them focus on SG. In particular, in [16] M. Kantarci and H. Mouftah introduce SG forensics, by identifying appropriate ingredients of SG that can assist in carrying out forensic-related processes efficiently. More detailed, the role of smart meters, SCADA and Phasor Measurement Units (PMUs) is discussed since their data can constitute significant evidence in order to prove the presence and cause of a relevant cybercrime.

D. Information Sharing and Anonymous Repository of Incidents

In [17], A. Triantafyllou et al. present the requirements of an anonymous repository of incidents exclusively related to SG. In the same study, similar information-sharing paradigms, such as the Trusted Automated Exchange of Indicator Information (TAXII) are discussed, while proposed technologies are also proposed. In particular, the authors suggest the utilisation of group signature and k-anonymity techniques. On the one side Group signature is adopted in order to hide the identity of the organisation uploading the security events, while k-anonymity is applied in order to anonymise possible information inside in the security event.

III. Methodology

The SPEAR architecture has been designed by using the ARCADE methodological framework [18]. ARCADE, influenced by IEEE 1471-2000, Recommended Practice for Architecture Description of Software-Intensive Systems is an open architecture description framework consisting of five primary viewpoints, namely a) *Context Viewpoint*, b) *Requirement Viewpoint*, c) *Component Viewpoint*, d) *Distribution Viewpoint* and e) *Realisation Viewpoint*. In particular, the *Context Viewpoint* is responsible for describing all aspects of the SPEAR platform related to the external SG ecosystem, thus including all interfaces of SPEAR with each SG related asset as well as with the energy-related stakeholders. On the other hand, the goal of the *Requirement Viewpoint* is to identify all functional and non-functional requirements of SPEAR based on the end-users' needs and regulatory framework analysis. Accordingly, the *Component Viewpoint* focuses on the components of SPEAR, while the *Distribution Viewpoint* describes the logical distribution of the software and hardware components by depicting how the components are logically placed and separated from each other. Finally, the *Realisation Viewpoint* aims at documenting how the final system's components should be implemented and deployed into a real-life environment.

In this paper, we concentrate on the *Component Viewpoint* and, more specifically, on the decomposition model, by describing the SPEAR components as well as their interfaces and technical specifications. Each component is described in detail by analysing its subcomponents and technologies.

IV. SPEAR Architecture

Fig. 1 illustrates the architecture of SPEAR, which is composed of three main components, namely a) *SPEAR Security Information and Event Management (SPEAR SIEM)*, b) *SPEAR Forensic Readiness Framework (SPEAR FRF)* and *c) Anonymous SPEAR Repository of Incidents (SPEAR RI)*. *SPEAR SIEM* is devoted to collecting, normalising and analysing both network traffic and operational data, thus detecting possible cyberattacks and calculating the reputation values of each asset. *SPEAR SIEM* consists of multiple subcomponents that collaborate efficiently with each other, taking into account the requirements defined in the *Requirement Viewpoint*. On the other side, *SPEAR FRF* is responsible for the procedures related to forensics, giving emphasis to the forensic readiness level. An important part of *SPEAR-FRF* is also the *AMI honeypots* and the *Honeypot Manager*. *AMI honeypots* emulate the behaviour of real industrial devices, emphasising on the operational data generated by themselves while the *Honeypot Manager* is responsible for managing and controlling the *AMI honeypots*. Finally, *SPEAR RI* establishes the anonymous repository of incidents where the various energy-related organisations will be able to exchange information regarding the various incidents detected by *SPEAR SIEM*. The following subsections analyse in detail each of the aforementioned components.

A. SPEAR SIEM

As depicted inf Fig. 2, *SPEAR SIEM* consists of six subcomponents namely a) *OSSIM (both OSSIM Server and OSSIM Sensor)*, b) *SPEAR SIEM Basis*, c) *Message Bus*, d) *Big Data Analytics Component (BDAC)*, e) *Visual-based Intrusion Detection System (VIDS)* and f) *Grid Trusted Module (GTM)*. In particular, both *SPEAR and OSSIM sensors* are distributed in the various subnets of an environment, thus monitoring and capturing the corresponding network traffic generated by the individual assets. *SPEAR Sensors* also collect operational data, such as electricity measurements (e.g., voltage, current, battery time, etc.). Then *OSSIM and SPEAR Sensors* transmit their collected and parsed data to the *OSSIM Server* and *Data Acquisition, Parsing and Storage (DAPS)* respectively, where the other subcomponents can receive and analyse them in order to extract possible security events. *OSSIM Server* includes many software intrusion detection tools that are based on signature-based techniques, while *BDAC* and *VIDS* receive the various data from *DAPS* and apply anomaly-based detection techniques based on machine learning and visual analytics, respectively. All security events are transmitted to the *Message Bus* component which constitutes an intermediate node where *GTM* and *VIDS* can receive the various security events in near-real-time in order to calculate the reputation of each asset and visualise the cybersecurity incidents. Moreover, it is worth noting that *Message Bus* acts as a message producer for *SPEAR FRF* and *SPEAR RI*. Next, we describe the functionality of each subcomponent with more details.

1) OSSIM: AlienVault OSSIM [19] is a widely known open-source SIEM tool that constitutes a basic ingredient of SPEAR SIEM, providing multiple operations such as log management, asset discovery, vulnerability assessment, network flow analysis and signature-based intrusion detection. Its architecture is divided into two main components, namely a) *OSSIM Sensors* and b) *OSSIM Server*.

OSSIM Sensors are deployed throughout the infrastructure, thus monitoring each subnet and collecting information related to the various assets. Next, this information is normalised into a specific format and transmitted to the *OSSIM Server*. In particular, the data collection process is carried out via agents

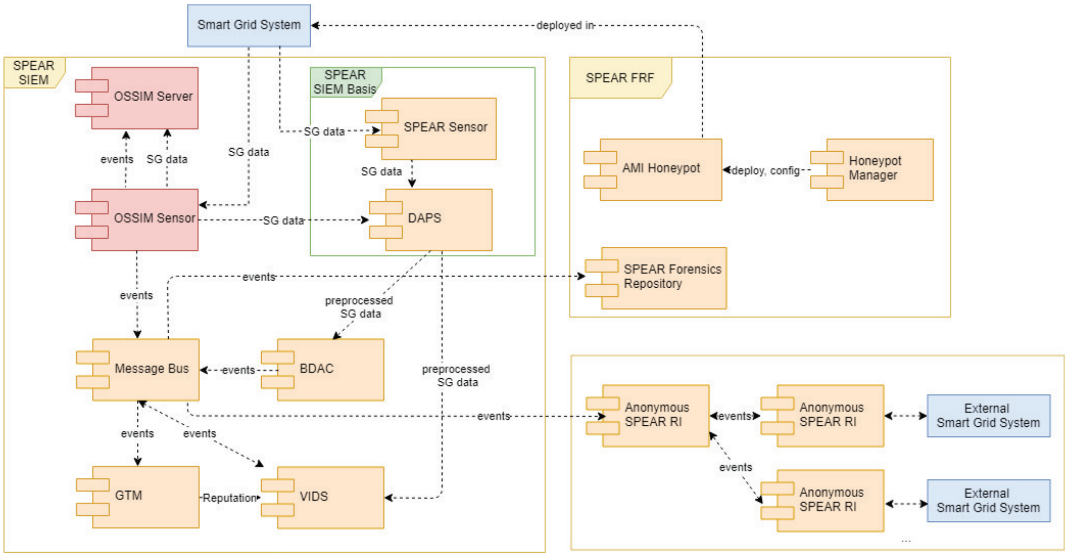

Fig. 1: SPEAR Architecture

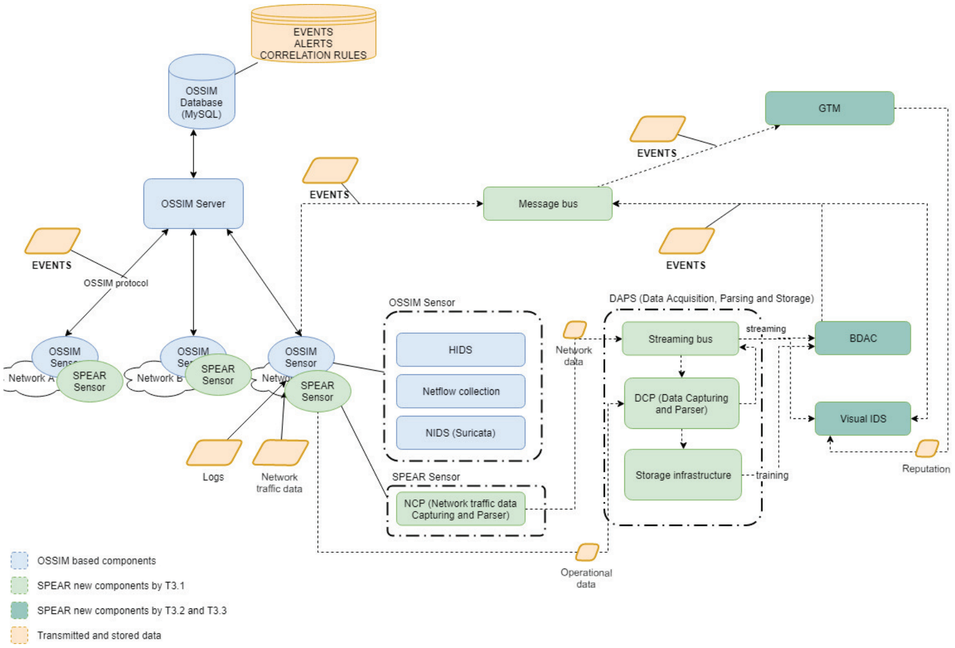

Fig. 2: SPEAR SIEM

running on the *OSSIM Sensors*. These agents include software modules called plugins that define how to collect, process and convert information into security events.

OSSIM Server aggregates and correlates all information sent by *OSSIM Sensors*, thus generating additional security events. In particular, OSSIM supports mainly two types of correlation: a) Logical Correlation and b) Cross-Correlation. Logical Correlation is accomplished via the security directives supported by OSSIM by combining the various security events. Security directives are eXtensible Markup Language (XML) files organised into specific categories (e.g., security directives for DoS attacks). Each security directive is composed of several security rules that are organised in a hierarchical manner, including different correlation levels. Specifically, the fields of the security events are compared with the conditions of a security rule and whether the number of occurrences reaches a specific threshold, then the next and deeper security rule is checked. On the other hand, Cross-Correlation checks the destination IP address field of the security events and whether this IP address is characterised by vulnerabilities stored in the database of the *OSSIM Server* via OpenVAS, then the reliability of these events is increased to the maximum value. Finally, *OSSIM Server* includes a web-based Graphical User Interface (GUI) where the user can monitor and control the overall security status of the infrastructure.

It is worth mentioning that OSSIM does not include User and Entity Behavior Analytics (UEBA) [6]. SPEAR will enhance OSSIM with machine learning capabilities and advanced visual analytics via *BDAC* and *VIDS*, respectively.

2) SPEAR SIEM Basis: The main role of the *SPEAR SIEM Basis* is to capture, parse and store both network traffic and operational data that are transmitted to detect possible cyberattacks or anomalies. Also, it operates as an asset discovery mechanism, collecting the necessary information for calculating the reputation value of each asset. More concretely, the functionality of the *SPEAR SIEM Basis* relies on two subcomponents, namely *SPEAR Sensors* and *DAPS*. *SPEAR Sensors* undertake to capture and parse data from an SG environment, specifically collect and parse network traffic data such as Transmission Control Protocol/Internet Protocol (TCP/IP) network flows as well as the payload of the network packets of industrial protocols. To this end, CICflowmeter [20] and T-shark [21] are used. Once all network traffic data is collected and parsed, the appropriate shipper is activated to transmit this data to *DAPS*. The operational data (e.g., electricity measurements) is collected either by 1) asynchronous agents deployed in the operational infrastructure of the SG that sends the data as it is generated in production, or by b) synchronous agents that periodically access the SG asset who hosts the operational data, and transmitting them to *DAPS*. Finally, SPEAR sensors include a third agent, called Asset Discovery (AD) that utilises nmap in order to discover and enumerate periodically the available assets in each subnet.

DAPS is a centralised subcomponent capable of handling a vast amount of data, thus storing and parsing all data coming from *SPEAR Sensors*. Subsequently, it distributes this preprocessed data to the other components of *SPEAR SIEM* to detect directly possible cyberattacks and anomalies (prediction phase) in near real-time or to train the machine learning-based intrusion/anomaly detection models (training phase). To this end, Apache Kafka [22] and ELK (Elasticsearch, Logstash, Kibana) stack [23] are used, respectively.

3) BDAC: BDAC performs machine learning and deep learning models in order to detect potential cyberattacks and anomalies, thus generating the corresponding security events. It constitutes an anomaly-based IDS which perfectly complements the signature-based Host-based IDS (HIDS) and Network-based IDS (NIDS) of *OSSIM*. In particular, *BDAC* analyses TCP/IP network flows, network packets related to industrial application layer protocols, honeypots data, as well as operational data (e.g., electricity measurements). The architecture of *BDAC* consists of five modules, namely: a) the *Data Receiving Module*, b) the *Data Pre-processing Module*, c) the *Training Module*, d) the *Analysis Engine Module* and e) the *Security Event Extraction Module*. The first one is responsible for receiving data from *DAPS*. The second module undertakes to pre-process this data before executing the machine learning-based detection models. The *Training Module* is responsible for generating the various intrusion and anomaly detection models by running offline the training process required by the machine learning and deep learning models based on the historical data of *DAPS*. Four categories of intrusion/anomaly detection models are defined: a) *TCP/IP Network Flow-Based Anomaly Detection Models*, b) *Packet-Based Anomaly Detection Models*, c) *Operational Data-Based Anomaly Detection Models* and e) *Honeypot-Based Anomaly Detection Models*. It is worth mentioning that the *Training Module* periodically generates new intrusion/anomaly detection models that replace the previous ones, whether their detection performance is considered better in terms of accuracy [24] and F1 measure [24]. All the aforementioned models form the *Analysis Engine Module*. Finally, the *Security Event Extraction Module* receives the security events and forwards them to the *Message Bus* component. The development of *BDAC* relies on various technologies, including [25], Pyspark [26], Scikit-learn [27], Keras [28] and PyOD [29].

4) Message Bus: This component provides a communication system to the SPEAR components/subcomponents (OSSIM, BDAC, VIDS, GTM, SPEAR RI) that either generate security events or need to receive them for implementing other functions. Message Bus uses Apache Kafka [22] in order to handle efficiently the asynchronous nature of the security events.

5) VIDS: The aim of *VIDS* is twofold; first, it depicts the outcome of *BDAC* and *GTM* and secondly, it works as an anomaly-based IDS complementary to *BDAC* by providing advanced visual analytics through which the user is able to identify additional anomalies. It is noteworthy that *VIDS* cannot automatically generate security events, but via an appropriate form, the user is capable of producing them based on the corresponding visualisations. Also, *VIDS* supports different user roles that, in turn, possess different access privileges based on their cybersecurity background. In particular, three user roles are supported: a) Security Engineer, b) Facility Operator and c) Non-Technical End-User. The first one can access all visualisation mechanisms, while the second role is aimed at those users with technical knowledge regarding SG but with no cybersecurity background. Hence, they can access only some specific features and visualisations of *VIDS*. Finally, the last role is devoted to those users with no technical background; therefore, their privileges are very limited.

6) GTM: GTM calculates for each asset a particular reputation value which reflects how dangerous it is for the entire normal operation of the organisation. In particular, *GTM* takes as input: a) the security events stored in *Message Bus* and b) the outcome of AD, i.e., the unique ID for each asset, their IP addresses and the asset value. Next, it adopts fuzzy logic techniques [30] for calculating the reputation value.

B. *SPEAR FRF*

According to Fig. 1, SPEAR FRF is composed of three main subcomponents, namely a) SPEAR Forensics Repository, b) AMI Honeypots and c) Honeypot Manager. The following subsections analyse them in detail. It is worth mentioning that part of SPEAR FRF is also the SPEAR Privacy-Preserving Framework, which establishes a particular Privacy Impact Assessment (PIA) [31], taking into account the applicable EU legislation and regulatory requirements concerning privacy. To this end, the DPIA CNIL tool [32] and a SPEAR Microsoft Excel-based tool are used.

1) SPEAR Forensic Repository: Before performing the necessary forensic processes, SPEAR evaluates whether the infrastructure is forensically ready via a SPEAR tool based on Camunda BPM [33]. If not, the appropriate changes are proposed. Next, the OSCAR methodology [34] is adopted, which includes five phases, namely a) Obtain Information, b) Strategise, c) Collect evidence, d) Analyse and e) Report. These phases are conducted by combining multiple existing tools such as NfSen, Wireshark, NfDump and Xplico, while the *SPEAR Forensic Repository (SPEAR FR)* relying on ELK [23] supports the overall process by storing centrally the appropriate elements: a) syslogs and EventLogs for Linux and Microsoft Windows operating systems respectively, b) network traffic data (including Packet Capture (PCAP) files, network flows and relevant statistical data) and c) security events generated by *OSSIM*, *BDAC* and *VIDS*. It should be noted, that OSCAR raises some requirements related to a) the necessary means for collecting the appropriate forensic evidence, b) how this evidence remains unforged and c) how the user privacy is guaranteed. Concerning the first challenge, *SPEAR-FRF* will use the security events generated by *SPEAR SIEM* as well as the logs of *AMI Honeypots*, while for the other challenges, advanced encryption techniques are adopted.

2) AMI Honeypots: Honeypots aim to imitate SG assets and act as a decoy in order to a) hide the real assets and b) attract possible cyberattackers, thus gathering useful information regarding their malicious activities. In the context of SPEAR, *AMI Honeypots* are based on the various SG protocols, including IEC 60870-5-101, IEC 60870-5-103, IEC 60870-5-104, DNP3, Modbus, MMS, Goose, SSH, FTP, Telnet, Bacnet, HTTP and HTTPS. Existing honeypot implementations such as Conpot [35] and Cowrie [36] are used for this scope. Moreover, *AMI Honeypots* emulate the behaviour of the real assets by transmitting similar network traffic data via efficient Generative Adversarial Networks (GANs) that are trained with the network traffic data of the real assets. The logs of *AMI Honeypots* is recovered by DAPS through the *Honeypot Manager* in a secure way in order to be used by the *BDAC Honeypot-based Anomaly Detection Models*. Finally, the fact that all data captured by the *AMI Honeypots* is considered as malicious eases the forensic investigation processes, gathering useful information for the cyberattackers' activity.

3) Honeypot Manager: Honeypot Manager is responsible for handling the lifecycle of various *AMI Honeypots* which goal is to act as deception mechanism in the SG. More specifically, *Honeypot Manager* consists of two main parts, namely a) *Planner*, b) *Deployer*. *Planner* is a decision support system based on game theory, which supports the end user to decide the best configuration of honeypots for a given infrastructure of the SG. The honeypots' deployment will be controlled by a game theory-based module, which will indicate the optimal honeypots' deployment in terms of their characteristics and their number. Finally, *Deployer* is based on a Terraform server that implements the services provided though a Representational State Transfer (REST) Application Programming Interface (API) to deploy or destroy the infrastructure of virtual machine instances containing the different honeypots.

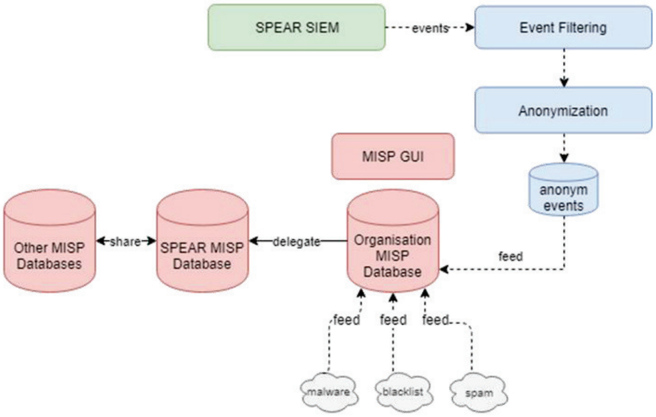

Fig. 3: SPEAR RI Architecture

C. *SPEAR RI*

SPEAR RI provides a common communication channel, where energy-related organisations across Europe are able to broadcast anonymously information related to security events without exposing the reputation of the organisation (victim of a cyberattack). Fig. 3 shows the architecture of SPEAR RI, which relies on the Malware Information Sharing Platform (MISP) platform [37]. In particular, the security events of SPEAR SIEM are filtered by the *Event Filtering module* and then are anonymised by the *Anonymisation module*, utilising pseudonymisation techniques. Then, the anonymised events feed the Organisation MISP database, which delegates its content to the SPEAR MISP database. Finally, the content of the SPEAR MISP database can be shared with the other MISP communities.

V. Conclusions

This paper presents the SPEAR architecture. SPEAR is an H2020 programme founded by the EU and aims at providing an integrated solution that will address efficiently cybersecurity issues of SG. Based on the ARCADE methodology and taking into account the state of the art cybersecurity products and the relevant research progress, SPEAR is composed of three main components, namely *SPEAR SIEM*, *SPEAR FRF* and *SPEAR RI*. *SPEAR SIEM* is devoted to recognising timely potential security events, while *SPEAR FRF* focuses on forensics. Finally, *SPEAR RI* is responsible for disseminating anonymously the security events with external relevant stakeholders. Each of the aforementioned components consists of individual subcomponents that are analysed in detail, providing also technical specifications regarding their implementation.

VI. Acknowledgement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 787011 (SPEAR).

References

[1] S. Tan, D. De, W.-Z. Song, J. Yang, and S. K. Das, "Survey of security advances in smart grid: A data driven approach," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 397–422, 2017.

[2] P. Radoglou-Grammatikis, P. Sarigiannidis, I. Giannoulakis, E. Kafetzakis, and E. Panaousis, "Attacking iec-60870-5-104 scada systems," in *2019 IEEE World Congress on Services (SERVICES)*, vol. 2642-939X, July 2019, pp. 41–46.

[3] P. I. R. Grammatikis, P. G. Sarigiannidis, and I. D. Moscholios, "Securing the internet of things: challenges, threats and solutions," *Internet of Things*, vol. 5, pp. 41–70, 2019.

[4] R. Leszczyna and M. R. Wróbel, "Evaluation of open source siem for situation awareness platform in the smart grid environment," in *2015 IEEE World Conference on Factory Communication Systems (WFCS)*. IEEE, 2015, pp. 1–4.

[5] K. M. Kavanagh, O. Rochford, and T. Bussa, "Magic quadrant for security information and event management," Gartner Magic Quadrant, Tech. Rep., 2017. [Online]. Available: https://jameskaskade.com/wp-content/uploads/2011/06/Magic-Quadrant-for-Security-Information-and-Event-Management.pdf

[6] S. G. Zarzosa, "D2.1 in-depth analysis of siems extensibility," DiSIEM Project, Tech. Rep. 1, 2017.

[7] P. I. Radoglou-Grammatikis and P. G. Sarigiannidis, "Securing the smart grid: A comprehensive compilation of intrusion detection and prevention systems," *IEEE Access*, vol. 7, pp. 46 595–46 620, 2019.

[8] E. Vasilomanolakis, S. Karuppayah, M. Mühlhäuser, and M. Fischer, "Taxonomy and survey of collaborative intrusion detection," *ACM Computing Surveys (CSUR)*, vol. 47, no. 4, p. 55, 2015.

[9] R. Mitchell and I.-R. Chen, "A survey of intrusion detection techniques for cyber-physical systems," *ACM Computing Surveys (CSUR)*, vol. 46, no. 4, p. 55, 2014.

[10] C. Dalamagkas, P. Sarigiannidis, D. Ioannidis, E. Iturbe, O. Nikolis, F. Ramos, E. Rios, A. Sarigiannidis, and D. Tzovaras, "A survey on honeypots, honeynets and their applications on smart grid," in *2019 IEEE Conference on Network Softwarization (NetSoft)*, June 2019, pp. 93–100.

[11] C. Dalamagkas, P. Sarigiannidis, D. Ioannidis, E. Iturbe, O. Nikolis, F. Ramos, E. Rios, A. Sarigiannidis, and D. Tzovaras, "A survey on honeypots, honeynets and their applications on smart grid," in *2019 IEEE Conference on Network Softwarization (NetSoft)*. IEEE, 2019, pp. 93–100.

[12] M. Nawrocki, M. Wählisch, T. C. Schmidt, C. Keil, and J. Schönfelder, "A survey on honeypot software and data analysis," *arXiv preprint arXiv:1608.06249*, 2016.

[13] W. Fan, Z. Du, D. Fernández, and V. A. Villagrá, "Enabling an anatomic view to investigate honeypot systems: A survey," *IEEE Systems Journal*, vol. 12, no. 4, pp. 3906–3919, 2017.

[14] J. Hou, Y. Li, J. Yu, and W. Shi, "A survey on digital forensics in internet of things," *IEEE Internet of Things Journal*, 2019.

[15] B. Manral, G. Somani, K.-K. R. Choo, M. Conti, and M. S. Gaur, "A systematic survey on cloud forensics challenges, solutions, and future directions," *ACM Computing Surveys (CSUR)*, vol. 52, no. 6, pp. 1–38, 2019.

[16] M. Erol-Kantarci and H. T. Mouftah, "Smart grid forensic science: applications, challenges, and open issues," *IEEE Communications Magazine*, vol. 51, no. 1, pp. 68–74, 2013.

[17] A. Triantafyllou, P. Sarigiannidis, A. Sarigiannidis, E. Rios, and E. Iturbe, "Towards an anonymous incident communication channel for electric smart grids," in *Proceedings of the 22nd Pan-Hellenic Conference on Informatics*. ACM, 2018, pp. 34–39.

[18] S. W. Erlend Stav and B. Farshchian, "An open architectural description framework," SINTEF, The address of the publisher, Tech. Rep. 2, Dec. 2013.

[19] Alienvault ossim. [Online]. Available: https://cybersecurity.att.com/products/ossim

[20] A. H. Lashkari, G. Draper-Gil, M. S. I. Mamun, and A. A. Ghorbani, "Characterization of tor traffic using time based features." in *ICISSP*, 2017, pp. 253–262.

[21] B. Merino, *Instant traffic analysis with Tshark how-to*. Packt Publishing Ltd, 2013.

[22] D. Powers, "kafka-python documentation release 1.4.7," Kafka-Python, Tech. Rep., 2019. [Online]. Available: https://buildmedia.readthedocs.org/media/pdf/kafka-python/master/kafka-python.pdf

[23] S. Chhajed, *Learning ELK Stack*. Packt Publishing Ltd, 2015.

[24] P. I. Radoglou-Grammatikis and P. G. Sarigiannidis, "Securing the smart grid: A comprehensive compilation of intrusion detection and prevention systems," *IEEE Access*, vol. 7, pp. 46 595–46 620, 2019.

[25] W. McKinney, "pandas: powerful python data analysis toolkit," Pandas, Tech. Rep., 2019. [Online]. Available: https://pandas.pydata.org/pandas-docs/stable/pandas.pdf

[26] T. Drabas and D. Lee, *Learning PySpark*. Packt Publishing Ltd, 2017.

[27] scikit learn, "scikit-learn user guide release 0.21.3," scikit-learn, Tech. Rep., 2019. [Online]. Available: https://scikit-learn.org/stable/_downloads/scikit-learn-docs.pdf

[28] Álvaro Peris, "Nmt-keras documentation release 0.2," keras, Tech. Rep., 2019. [Online]. Available: https://buildmedia.readthedocs.org/media/pdf/nmt-keras/latest/nmt-keras.pdf

[29] Y. Zhao, Z. Nasrullah, and Z. Li, "Pyod: A python toolbox for scalable outlier detection," *Journal of Machine Learning Research*, vol. 20, no. 96, pp. 1–7, 2019. [Online]. Available: http://jmlr.org/papers/v20/19-011.html

[30] A. H. Hamamoto, L. F. Carvalho, L. D. H. Sampaio, T. Abrão, and M. L. Proença Jr, "Network anomaly detection system using genetic algorithm and fuzzy logic," *Expert Systems with Applications*, vol. 92, pp. 390–402, 2018.

[31] S. G. T. Force, "Data protection impact assessment template for smart grid and smart metering systems," Smart Grid Task Force 2012-14 - Expert Group 2: Regulatory Recommendations for Privacy, Data Protection and Cyber-Security in the Smart Grid Environment, Tech. Rep. v. 2 of 13, 2018.

[32] J. Sarrat and R. Brun, "Dpia: How to carry out one of the key principles of accountability," in *Annual Privacy Forum*. Springer, 2018, pp. 172–182.

[33] Camunda bpm. [Online]. Available: https://camunda.com/

[34] ENISA, "Forensic analysis," ENISA, Tech. Rep., 2016. [Online]. Available: https://www.enisa.europa.eu/topics/trainings-for-cybersecurity-specialists/online-training-material/documents/2016-resources/exe2_forensic_analysis_ii-toolset

[35] D. Pliatsios, P. Sarigiannidis, T. Liatifis, K. Rompolos, and I. Siniosoglou, "A novel and interactive industrial control system honeypot for critical smart grid infrastructure," in *2019 IEEE 24th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Sep. 2019, pp. 1–6.

[36] M. Oosterhof, "cowrie documentation release 19.10.0," Cowrie, Tech. Rep., 2019. [Online]. Available: https://buildmedia.readthedocs.org/media/pdf/cowrie/latest/cowrie.pdf

[37] C. Wagner, A. Dulaunoy, G. Wagener, and A. Iklody, "Misp: The design and implementation of a collaborative threat intelligence sharing platform," in *Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security*. ACM, 2016, pp. 49–56.