Evolving Systems
https://doi.org/10.1007/s12530-020-09335-4

ORIGINAL PAPER

Handling of advanced persistent threats and complex incidents in healthcare, transportation and energy ICT infrastructures

Spyridon Papastergiou[1] · **Haralambos Mouratidis**[2,3] · **Eleni-Maria Kalogeraki**[1]

Received: 9 October 2019 / Accepted: 8 March 2020

Abstract
In recent years, the use of information technologies in Critical Infrastructures is gradually increasing. Although this brings benefits, it also increases the possibility of security attacks. Despite the availability of various advanced incident handling techniques and tools, there is still no easy, structured, standardized and trusted way to manage and forecast interrelated cybersecurity incidents. This paper introduces CyberSANE, a novel dynamic and collaborative, warning and response system, which supports security officers and operators to recognize, identify, dynamically analyse, forecast, treat and respond to security threats and risks and and it guides them to handle effectively cyber incidents. The components of CyberSANE are described along with a description of the CyberSANE data flow. The main novelty of the CyberSANE system is the fact that it enables the combination of active incident handling approaches with reactive approaches to support incidents of compound, highly dependent Critical Information Infrastructures. The benefits and added value of using CyberSANE is described with the aid of a set of cyber-attack scenarios.

Keywords Incident handling · Web mining · Data fusion · Risk assessment

1 Introduction

Over the past decade, critical infrastructures are operating upon robust and reliable ICT implementations of emerging technologies (e.g. IoT, Cloud Computing, Big Data), which are interconnected through complex, heterogeneous networks, providing a high level of flexibility, scalability, and efficiency on the provided services and the supported processes. The increased usage of information technology in such scenario increases the benefits for Critical Information Infrastructures (CIIs) but it also makes them more susceptible to malicious activities and cyber attacks from various types of adversaries (e.g. hackers, terrorist groups, criminal gangs and rogue states).

Nowadays, barriers to entry for would-be cyber criminals are falling rapidly because the attackers have a range of (technical) capabilities and substantial resources at their disposal, since malware and malware-as-a service have become more easily and cheaply available through various means and sources (such as Dark Web, Deep Web). As a result, a variety of advanced techniques and tools (e.g. social engineering techniques and zero-day exploits programs) are available and can be used by cyber criminals to initiate advanced targeted attacks. Such attacks employ multiple technologies and are deployed in multiple stages in order to penetrate an organization's defenses. The attack vectors vary significantly from Application-Layer, Social Engineering Unauthorized Access, Malicious Code, and Reconnaissance and Networking-based service attacks that target applications, host and client operating systems and even networking equipment. In such context, attackers use these techniques to get valuable data assets, such as financial transaction information, user credentials, and insider information. Moreover, in many cases, attackers take advantage of the interconnected nature of modern systems to

✉ Haralambos Mouratidis
H.Mouratidis@brighton.ac.uk

Spyridon Papastergiou
paps@unipi.gr

Eleni-Maria Kalogeraki
elmaklg@unipi.gr

[1] Department of Informatics, University of Piraeus, 80 Karaoli and Dimitriou Str., 18534 Piraeus, Greece

[2] Centre for Secure, Intelligent and Usable Systems, University of Brighton, Brighton BN2 4GJ, UK

[3] Department of Computer and Systems Sciences, Stockholm University, 164 07 Kista, Sweden

Published online: 04 April 2020

attack third parties. Novel multi-stage attacks can be used to exploit vulnerabilities of interconnected ICT systems beyond organizational boundaries, enabling attackers to move across multiple businesses and systems. In such scenario, attackers can use an infrastructure as the stepping stone to attack and compromise the security of a dependent organization.

Over the last few years, industrial societies are bombarded daily with headlines and reports on major cyber-attacks, new malware strains or insidious social engineering techniques that adversaries use to attack ICTs. In particular, CIIs have become lately targets for cyberattacks attracting the attention of security researchers, cyber-criminals, hacktivists (e.g. Anonymous, LulzSec) and other such role-players (e.g. cyber-spies). These cyber actors have evolved their tactics, techniques and procedures to an unruly extent, making next-generation malware toolkits available in various locations through the internet (e.g. Deep Web, Dark Web) and adopting new data exfiltration methods that give them an asymmetric quantum leap in capability. In the past years, critical infrastructures have been affected by numerous cybersecurity meltdowns and high-profile breaches.

In this context, appropriate incident handling solutions are necessary to support and facilitate the detection and analysis of cyber-attacks and threats on Critical Information Infrastructures, and to increase operators knowledge on cyber risks. Nevertheless, existing digital approaches are unlikely to stop an attack in real time, as they fail to capture and correlate, in an appropriate degree, events and information associated with cyber attacks in complex CIIs. Furthermore, existing incident handling approaches are too slow, ineffective, not taking into consideration the heterogeneity and complexity of ICT systems and the underlying infrastructures, and not providing enough insight into the real causes of the cybersecurity incidents.

This paper proposes CyberSANE, a novel approach, which aims to improve the detection and analysis of cyber-attacks and Advanced Persistent Threats (APTs) on CIIs and increase the knowledge of the current cyber threat landscape. In particular, the CyberSANE system supports organisations to raise their preparedness, improve their cooperation with each other, and adopt appropriate steps to manage security risks, and to report and handle security incidents. The rest of the paper is structured as follows. Section 2 discusses the current cybersecurity status in CIIs of three critical sectors of the European and global economy: Transportation, Healthcare and Energy. Section 3 presents related work in Security Incident Handling, Threat Intelligence and Information Sharing. Section 4 introduces the CyberSANE approach, while Sect. 5 goes more in depth and describes the CyberSANE system and its key components. Section 6 analyses the data flows of the CyberSANE components and the overall system operation. The application of the CyberSANE system to a number of demonstration scenarios alongside a discussion of the benefits and added value of CyberSANE are presented in Sect. 7. Section 8 reflects the innovative aspects of the system, while Sect. 9 draws the conclusions of this work.

2 Critical information infrastructures cybersecurity status

This section provides a summary on the cybersecurity status of three critical industrial sectors of the European and global economy: Transportation, Health and Energy.

2.1 Transportation cybersecurity status

Although transportation stakeholders do not publish unified figures about their cybersecurity status, data provided by the reports of independent cybersecurity firms attribute the cyber attacks rise to the lack of adequate training, best practices and cybersecurity standards with compulsory implementation. Despite the lack of consolidated transportation cybersecurity figures, the qualitative information available confirms that though transportation stakeholders are handling constantly increasing volumes of information, many of them are still working with obsolete legacy ICT systems, which raises their cyber vulnerability. In fact, the overall low sense of urgency and cybersecurity awareness in combination with the constantly increasing volumes of information exchanged and the ICT complexity makes transportation stakeholders, particularly, vulnerable to cyber-attacks, which can result in severe transportation services disruptions. Notwithstanding, the transport industry increasingly relies on ICTs to enable essential transportation operations (i.e. navigation, propulsion, freight management, traffic control, etc.) transportation regulations and policies, considering only the physical aspects of security and safety. This might be partially explained by the low number of known cybersecurity incidents incurred in the transportation sector, which did not create sufficient media exposure to trigger bold reaction from the stakeholders involved. Looking at the bigger picture, when it comes to the global cybersecurity market, the estimated cost of cyber-crime in 2017 amounts to $100bn, whereas in 2014, 81% of large organisations suffered from cyber attacks, the average cost of which varied from $900 K to $2.3 M (MTI Network 2015). Considering these facts, ENISA has published the Cybersecurity Act on 13.09.2017 attempting to pave the way for the rollout of new products and services for cyber protection across the Member States. In light of this, the European Commision (EC) has proposed the "European Cybersecurity Certification Framework for ICT products and services" aiming to reduce the cost of the cybersecurity certification, improve its effectiveness and make it more commercially attractive

for European organisations. Concerning the low maturity of cybersecurity awareness among the transportation world, most of the stakeholders invest less than the expected in fortifying their digital assets. Only transportation operators, such as the Port of Rotterdam, are able to demonstrate specific cybersecurity investments. However, the effort of the EC to raise cybersecurity awareness paves the way for engaging more transportation stakeholders to devote budgets on cybersecurity solutions.

2.2 Healthcare cybersecurity status

According to SANS reports (Filkins 2016), given the highly sensitive nature of data processed, healthcare is among the top 5 vertical industries with the highest spending on cybersecurity with standard annual budgets up to $600 k. Taking into account the recent impose of GDPR, it is expected that these investments will be increased in order to meet the stricter requirements of the regulation. Despite these investments, it has been reported that several medical IT systems are severely outdated (Mukherjee 2017) rendering health CIIs extremely vulnerable to a new class of hackers who target sensitive patients' information (i.e. medical records, credit cards, etc.). Additionally, reports show that apart from external factors, insiders are responsible for a great number of security incidents, which confirms that insiders are the biggest threat in healthcare (Widup 2018).

2.3 Energy cybersecurity status

Similarly to the transportation sector, cybersecurity awareness is also low in the energy industry. According to (Scott 2018), the booming expansion of IoT technologies entails the interconnection of the energy grids with a constantly increasing number of smart devices (Scott 2018). Therefore, the possibility of a cyberattack firing physical damages is higher than ever before. Concerning the abovementioned and that Energy is an asset-intensive industry with many remote and hazardous sites, this sector adopted digitisation rather late. In further detail, although most energy industry executives (76%) have cited that business interruption due to cyber-attacks is the most impactful loss of their organisation, more than half of them could not quantify or did not even know what their worst possible loss exposures could be (Marsh report 2018).

3 Current efforts in security incident handling, threat intelligence and information sharing

3.1 Incident handling in critical information infrastructures

The main goal of a security incident handling and response process is to define the top concerns in managing a security breach/incident/event (West-Brown et al. 2003a; Wiik and Kossakowski 2005). In practice, choosing the right approach for incident handling proves to be a difficult decision. In recent years, a number of security incident response approaches and frameworks (British Standards Institution 2011; Cichonski and Scarfone 2012; ENISA 2019; Northcutt 2003; Vangelos 2011; Werlinger et al. 2010; Khurana et al. 2009; Grobauer and Schreck 2010; Monfared and Jaatun 2012; Line 2013; Cusick and Ma 2010; Connell et al. 2013) have been introduced by the research and industrial communities and various standardization bodies. Even though many of these approaches provide useful technical guidelines, aiming to enhance the security incident response capabilities of the organizations, they demonstrate significant limitations. In particular, Grimes (Grimes 2007) argues that most of the existing incident response approaches follow a linear process that is outdated and does not support the highly efficient capability that is required to handle and manage today's incidents. Therefore, a progression flaw exists in these processes, since if one phase in the linear process is not completed, the entire process cycle may stop midstream. According to another research work, current incident response processes are too focused on the containment, eradication, and recovery-related activities and usually ignore, skip or do not emphasize on other important steps of incident management, such as investigations actions (Ahmad et al. 2012). The proposal of Shedden et al. (2011) gives emphasis on proactive preparation and reactive learning to encourage security incident learning. The existing incident handling approaches do not provide adequate guidance on how to conduct effective forensic investigations (Casey 2006; Nnoli et al. 2012; Tan et al. 2003). Hence, current methods' limitation to assist and guide the investigators in the forensic evidence analysis, undermines the value of the evidence and fails to promote incident resolution.

Most of the available security information and event management solutions lack significant reactive and post-incident capabilities for managing incidents and events in the scope of the ICT-based CIIs, providing inadequate technical guidance to the incident response professionals on how to detect, investigate and reproduce attacks. As

such, despite the socioeconomic importance of tools and techniques for handling incidents, there is still no easy, structured, standardized and trusted way to manage and forecast interrelated cybersecurity incidents taking into account the heterogeneity and complexity of the CIIs and the increasingly sophisticated types of attacks. Therefore, there is a pressing need for devising novel systems for efficient CIIs incident handling and providing a thorough and common understanding of cyber-attack situations in a timely manner.

To sum up, the main limitations of the existing approaches (FireEye 2013; Grispos et al. 2014; Ab Rahman and Choo 2015) can be summarized as follows: (1) the traditional linear incident response models are too slow, ineffective and do not support the highly efficient capability that is required to handle and manage today's incidents; (2) the focus is mostly on the proactive element (i.e. provide assistance and information to help prepare, protect, and secure) of the incident management; (3) current approaches do not provide enough insight into the underlying causes of the incident; (4) poor provisions for incident planning; (5) undermine the value of forensic evidence possibly required for subsequent legal action; (6) do not consider the risk-related results produced by existing risk assessment methodologies.

In contrast, our work provides a novel integrated approach that detects malicious activities, and it provides a thorough analysis of the detected anomalies in a more efficient, elastic and scalable reasoning way. To achieve this, it combines active approaches, which detect and analyse anomaly activities and attacks in real-time, with reactive approaches, which provide a thorough analysis of the underlying infrastructure, in order to assess the reported incident.

3.2 Attack scenarios and evidence graph representation

Over the last few years, a significant amount of effort has been invested in representing evidence and attack scenarios, using a variety of graphs. These attempts are using a variety of mathematical formula and algorithms to produce and generate the attack patterns (Leucari 2012). Nevertheless, most of them have significant limitations, as they just provide a high-level, abstract view of a complex attack (Swiler et al. 2001). Representative examples of these types of investigation graphs include scenario graphs, forensics graphs, logic exploitation graphs, attack graphs, and evidence graphs (Neralla et al. 2013). In this context, Wang and Daniels (2005), in their proposed evidence graph model, seek to facilitate the evidence presentation process and provide an automated intrusion evidence analysis considering the firewall output intrusion alerts. Later, Wang and Daniels (2006) introduce diffusion and graph spectral methods that adopt high-performance computation algorithms. Gladyshev (2004) proposes a formalized approach for Event Reconstruction based on finite state machine model in order to define all possible attack scenarios in the computer network incidents. Liu et al. (2013) engage algorithms that allow integrating different evidence graphs with a probabilistic evidence graph with or without the help of a corresponding attack. Phillips and Swiler (1998) present an approach for network risk analysis based on an attack graph that defines the set of attack paths that have a high probability of success for the attacker. Sheyner et al. (2002) propose automated techniques based on a set of algorithms to establish the attack graphs associated with specific attack scenarios. Eventually, Bruschi et al. (2004) provide a model that can organize digital forensics knowledge in a reusable way.

The proposed work adopts brain-like reasoning methods to understand the incident-related events and information. Furthermore, it combines bio-inspired approaches with graph-theoretic approaches to provide more accurate and efficient investigation results.

3.3 Threat intelligence

When an organization is threatened, or is about to be threatened, some events could be identified by checking the network or operating system. An instance of such event is called an indicator of compromise (IoC), and could be considered as the main core of cyber threat intelligence. The benefits of sharing this data are very considerable, especially when these operations, both for processing and sharing, could be done in an automated way, to speed up the entire process. Widely used standards for describing cyber threat intelligence, in a very detailed way using a common format as XML and JSON, are CybOX/STIX (OASIS 2017a, b, c, d, e), OpenIOC (2017), IODEF (Danyliw et al. 2007), CAPEC (2017) and MAEC (2016). Contrarily, a common transport mechanism for sharing this information is TAXII (OASIS 2017f), which aims to enable timely and secure sharing of threat information in cyber defender communities. It was developed by MITRE organization and considered a lot in the recent past due to its capability to suit very well with other standards, used for describing threat, developed by the same organization, such as CybOX, STIX, MAEC and CAPEC, allowing creating, processing and sharing very detailed and specific IoC, often related to Tactics, Techniques and Procedures (TTP) of an attacker. Another standard, which belongs to this category is the Traffic Light Protocol (TLP) (2017), which is more a set of designations for sharing sensitive information with the appropriate audience rather than a proper transport mechanism. However, these standards do not allow defining trust and sharing agreements, as well as access control limitations.

The proposed CyberSANE approach aims at defining advanced threat models, capable of capturing the new

communication (e.g. IoT connectivity, systems interconnectivity) and computation (e.g. cloud services, use of semi-trusted mobile devices) paradigms in IT services, and to identify hidden attack vectors, indirect and subliminal attack paths, as those utilized by Advanced Persistent Threats, ransomware and botnets.

3.4 Secure and privacy-aware information sharing

One of the most important challenges related to threat intelligence sharing regards the potential presence of sensitive data. Involved parties must be able to protect it and, for this reason, intelligence sharing must be performed in a private and confidential-enabled way, assuring secure internet sharing and preserving privacy of sensitive information. Moreover, with the new EU General Data Protection Regulation (GDPR), new challenges have to be addressed, as described in Sullivan and Burger (2017). Considering a threat intelligence sharing scenario, particular principles must be followed (Fisk et al. 2015), in order to reduce the risk of data exposure and help managing trust requirements. In this sharing environment, information exchanged must be expressed through specific formats and standards, in order to achieve interoperability, speeding up processing and analysis phases on the receiver side. For this reason, a particular emphasis should be given to privacy-preserving techniques applied to structured data, as, for example, in Ulltveit-Moe et al. (2013), where an architecture and a methodology for privacy handling in Critical Information Infrastructures is proposed, focusing upon policy-controlled authorization, anonymization and encryption of information in XML element, which is the format used in very popular threat intelligence standards such as STIX, just for versions up to 1.2, and IODEF. Another example is presented in De Fuentes et al. (2016), which proposes a protocol that provides privacy-preserving and agreeable cybersecurity information sharing, by leveraging existing format-preserving and homomorphic encryption techniques and adapting them to the particularities of standard message formats, such as the previous mentioned STIX. Finally, issues associated to the realization of an efficient and effective information sharing paradigms for actionable intelligence have been studied, focusing also on architectural solutions for ensuring privacy, considering many different standards for representing cyber threat information (Mohaisen et al. 2017).

Considering that threat incident and intelligence information could potentially contain sensitive data, the proposed CyberSANE approach, employs secure and privacy-aware sharing methods, such as data anonymization, document sanitization and implementation of cryptographic primitives and protocols, to ensure the secure distribution and storage of all forensic artifacts. In this regard, controller/processor capabilities will be deployed in the context of GDPR, allowing security experts to specify all the protection steps that have to be performed along with the required conditions to execute them and to implement incident handling business-logic decisions that will provide the establishment of a secure environment over the heterogeneous and complex interconnected CIIs.

4 CyberSANE incident handling approach

The proposed incident handling approach aims at providing a step-by-step guidance to manage incidents and breaches on CIIs occurred due to cyber attacks. On this account, CyberSANE aims to combine active approaches, that are used to detect and analyse anomaly activities and attacks in real-time, with reactive approaches that analyse the underlying infrastructure to assess an incident. Thus, the CyberSANE approach is capable of providing a holistic and integrated approach to incident handling. In this vein, CyberSANE aims to enhance the incident detection capabilities of the existing methods described in the previous section with a more efficient, elastic and scalable reasoning approach. The main characteristics of the proposed approach are the following: (1) learning from unstructured data without the need to understand the content; (2) identification of unusual activities that match the structural patterns of possible intrusions (instead of predefined rules); and (3) automatic identification and adaption to a change of the underlying infrastructure.

CyberSANE treats the handling of a cyber incident as a dynamic experimental environment that can be optimized involving all relevant CII operators and security experts. The CyberSANE approach is based on simulations to facilitate the evaluation and analysis of an identified incident and support the investigation decision making process in a rigorous manner. The pursuit of CyberSANE is to support the incident handling process with advanced correlation capabilities in terms of accuracy and efficiency strengthening the rational analysis. In this context, CyberSANE relies on existing techniques of pioneering mathematical models (Blowers and Williams 2014) (e.g. machine learning, deep learning and Global Artificial Intelligence (AI) techniques) for analysing, compiling, combining and correlating all incident-related information and data from different levels and contexts (e.g. taking into consideration information, data and opinions collected and analysed from existing risk assessment frameworks (including the CYSM, MEDUSA, MITIGATE and SAURON approaches (Papastergiou and Polemi 2017, 2018; Kalogeraki et al. 2018), knowledge and information acquired from previous incident investigations as well as evidential data extracted from the compromised cyber systems). Thus, these techniques will be able to identify, extract and analyse the most relevant parts of the information related to the initial incident, in order to find the

relationships between the compromised devices/systems and these evidences.

Moreover, efficient simulation experiments for generating multi-order evidence dependencies have been used to generate and construct secure, reliable and valid chains of evidence anticipating how the attack is progressing. In particular, CyberSANE acknowledges that the process of attacks scenarios reconstruction from both collected structured data (i.e., alarms, alerts, logs) and unstructured data (e.g. data coming from social networks such as tweets, discussions on forums) requires brain-like reasoning to understand the incident-related events and information. In this context, CyberSANE will adopt Bio-inspired approaches (Floreano and Mattiussi 2008) to self-organizing all relevant events and creating the linkage between them. Additionally, it will visualize the evidence and attack scenario on the underlying CIIs using and combining several types of graphs, including scenario graphs, logic exploitation graphs, forensics graphs, attack graphs, and evidence graphs. Additionally, taking into account the effects of a security incident, real-time insights, alerts and warnings will be produced to increase situation awareness, inform the CII stakeholders about the effects of the events and guide them how to react.

CyberSANE incorporates innovative Intelligence and Information Sharing models that disseminate and share information on useful incident-related information with relevant parties (e.g. industry cooperation groups, Computer Security Incident Response Teams - CSIRTs) about the effects and danger of incidents and the characterized diffusing threats. The adopted model includes the security and privacy-aware sharing capabilities required, considering that shared incident-related information may potentially contain sensitive data. As stated in Sect. 3.4, many works have already addressed this topic, especially focusing on privacy-preserving, such as anonymization or pseudo anonymization, and encryption techniques, applied on structured information, expressed through specific standards, such as STIX (OASIS 2017e), which is actually the most used by organizations and governments. However, most of the works done so far in this field, take into consideration simple use cases or, even worse, old versions of them and this is a high limitation, considering that, at their full potential, they do not allow to describe in a very detailed way cyber security incidents and threats, as well as software vulnerabilities and tactics, techniques and procedures (TTP) of attackers.

5 CyberSANE incident handling system

CyberSANE System is the implementation of the aforementioned approach; an innovative, knowledge based, collaborative security and response dynamic system which incorporates all phases of the Cyber incident handling lifecycle for handling incidents, incident detection, analysis activities and post-incident knowledge harvesting.

The main goal of the CyberSANE system is to increase the agility of the investigators and encourage continuous learning throughout the incident life cycle. In particular, the proposed system aims to improve, intensify and coordinate the overall security efforts for the effective and efficient identification; investigation, mitigation and reporting of realistic multi-dimensional attacks within the interconnected web of the cyber assets of the CIIs and security events. To realize this objective, the CyberSANE system is empowered with the newest techniques in prevention, detection and mitigation of cyber-threats, including the understanding of synthetic cyberspace through the use of Advanced Visualization Techniques (immersive interfaces, cyber 3D models, etc.). These visualization approaches will help CII operators to better comprehend the situation and to detect some traces/details that could allow them to provide incident analysis in-depth and thus to detect a potential threat/attack.

In order to capture all the CyberSANE system requirements, Gürses et al. (2005) requirements elicitation process is followed. Moreover, the elicitation, analysis and the documentation of requirements, is focused on capturing the perspectives of technical professionals, such as software developers and system engineers, who cooperate with system's end users to discover problems that have to be solved. The requirements elicitation process includes but is not limited to what the proposed system should provide, what are the expected services, the required characteristics and software constraints etc. To this intent, targeted questionnaires were developed, aiming at collecting feedback from various CIIs operators. These questionnaires serve mainly the purpose of corroborating the requirements elicited, but they are also used as a communication measure and feedback collection. In this respect, a number of questionnaire replies was gathered (approximately 30 questionnaires from organizations coming from various critical sectors, banking, maritime, transportation, healthcare, energy).

From technical perspective, the system is composed of five core components: (1) the Live Security Monitoring and Analysis (LiveNet), which is able to monitor, analyze, and visualize organizations' internal live network traffic in real time, (2) the Deep and Dark Web mining and intelligence (DarkNet), which monitors the Dark and Deep Web as a means to grasp and analyse the big picture of global malware cybersecurity activities, (3) the Data Fusion, Risk Evaluation and Event Management (HybridNet), which receives security related information on potential cyber threats from both LiveNet and Darknet accordingly with the view to analyze and evaluate the security situation inside an organization, (4) the Intelligence and Information Sharing and Dissemination (ShareNet), which disseminates and shares information of useful incident-related information to relevant parties

and (5) the Privacy and Data Protection (PrivacyNet) that undertakes the responsibility to collect, compile, process and fuse all the incident-related information in a way that their integrity and validity is well preserved. Figures 1 and 2 depict the CyberSANE Incident Handling approach and the CybeSANE system along with its main parts respectively. The following sections provide a detailed description of the five CyberSANE system core components.

5.1 Live security monitoring and analysis (LiveNet) component

The LiveNet is an advanced and scalable Live Security Monitoring and Analysis component capable of preventing and detecting threats and, in case of a declared attack, capable of mitigating the effects of an infection/intrusion. The main objective of this component is to implement the Identification, Extraction, Transformation, and Load process for collecting and preparing all the relevant information, serving as the interface between the underlying CIIs and the CyberSANE system. It includes proper cyber security monitoring sensors with network-based Intrusion Detection Systems (IDS), innovative Anomaly detection modules and endpoint protection solutions for accessing and extracting information, on a real-time basis, in order to detect complex and large-scale attacks (e.g. Advanced Persistent Threats). The incident-related information that reside in different and heterogeneous cyber systems may include various types of data, such as: active (unpatched) vulnerabilities in the technological infrastructure; misuse detection in the network or in the systems, including both host-based and network-based IDS deployment and integration; anomaly detection in the network or in the systems; system availability signals; network usage and bandwidth monitoring; industry proprietary protocol anomalies; SCADA vulnerabilities, etc.

LiveNet incorporates appropriate data management and reasoning capabilities for: (1) near real-time identification of anomalies, threats, risks and faults and the appropriate reactions; (2) proactive reaction to threats and attacks; and (3) dynamic decision making in micro, macro and global level according to the end user's needs and the identified incidents/threats. These capabilities are empowered with existing innovative algorithms based on techniques such as machine learning, deep learning and AI that identify previously unknown attacks. This component provides an abstraction of the collected information to the Data Fusion, Risk Evaluation and Event Management (HybridNet) component of the CyberSANE system. Moreover, all incidents-related

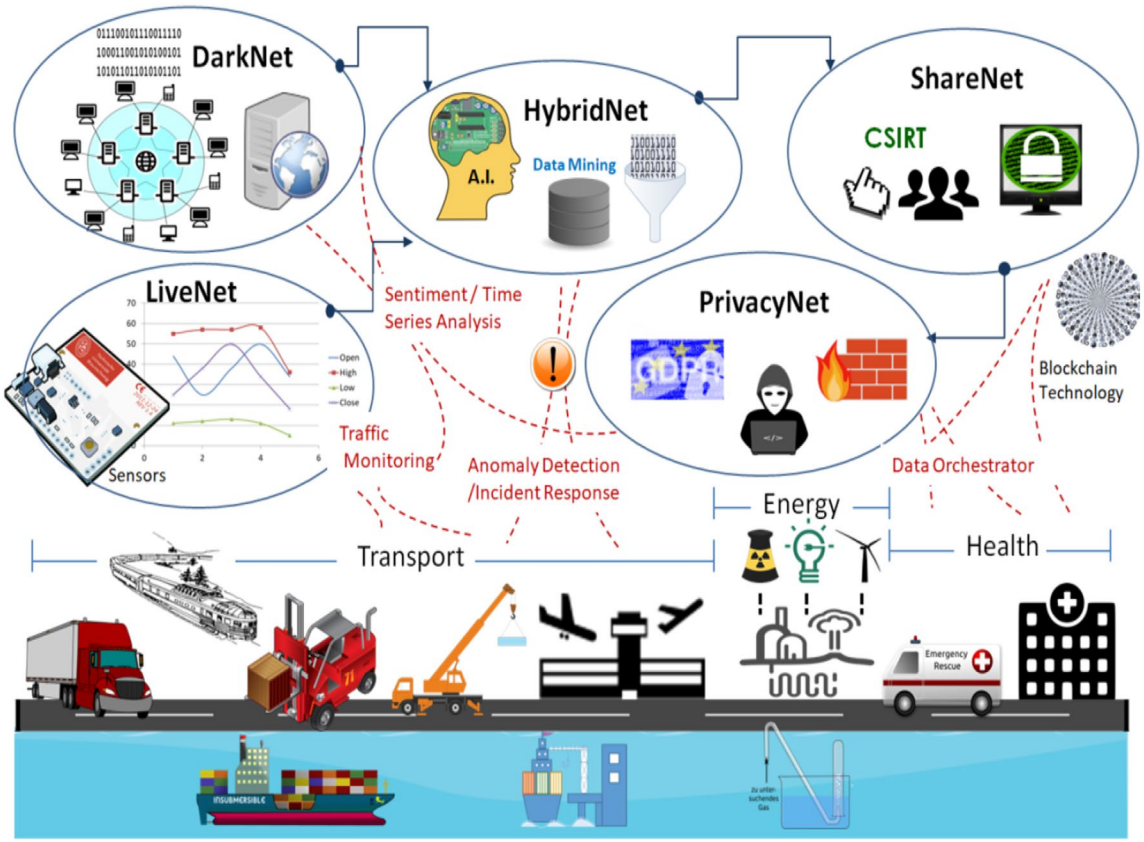

Fig. 1 CyberSANE incident handling approach

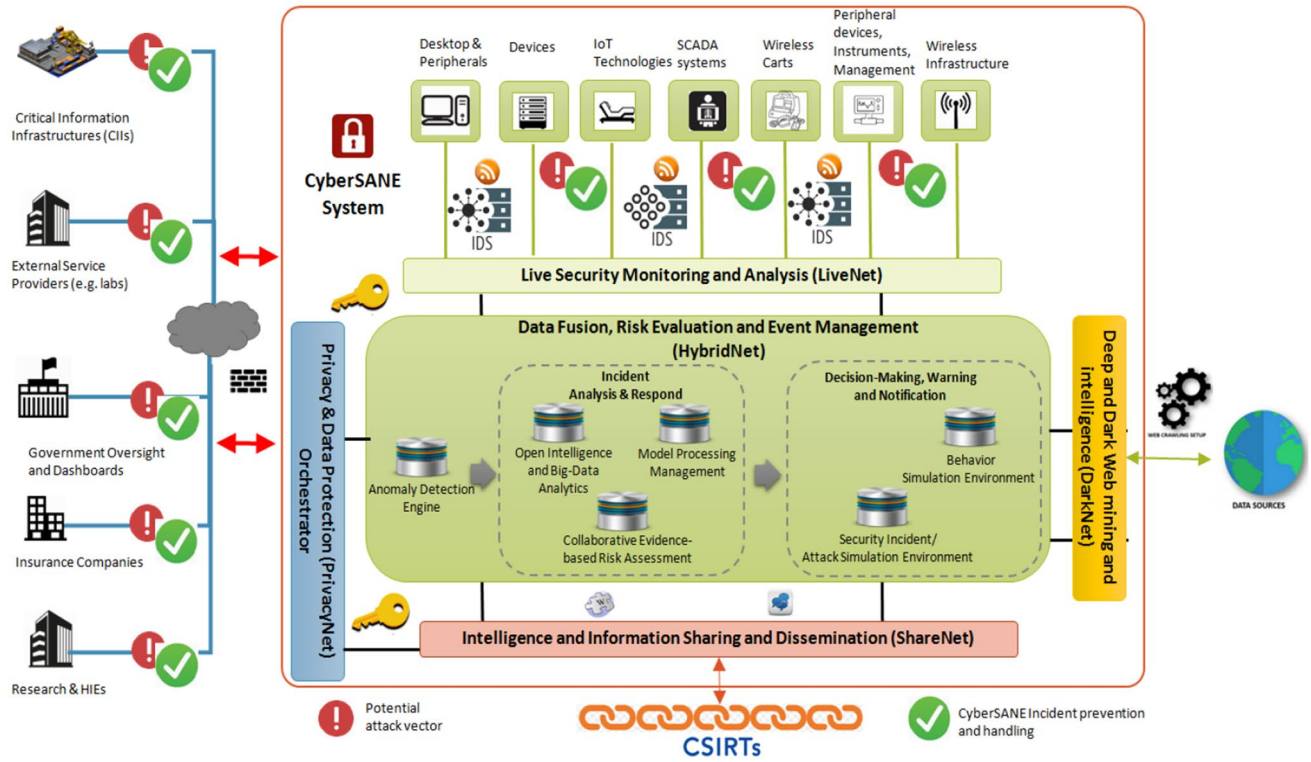

Fig. 2 CyberSANE system and components

information captured from LiveNet will be parsed, filtered, harmonized and enriched to ensure that only the data necessary for the multivariate and multidimensional analysis are available to the other components (e.g. HybridNet). Thus, LiveNet contributes as follows: (1) preventing a flood of irrelevant or repeated information from cluttering the HybridNet processing component; and (2) consolidating the different data contents and formats towards a uniform perspective in order to provide the upper components a unified and convenient way to handle the information.

5.2 Deep and dark web mining and intelligence (DarkNet) component

The Deep and Dark Web mining and intelligence (DarkNet) component provides the appropriate Social Information Mining capabilities that will allow the exploitation and analysis of security, risks and threats related information embedded in user-generated content (UGC). This is achieved via the analysis of both the textual and meta-data content available from such streams. Textual information is processed to extract data from otherwise disparate and distributed sources that may offer unique insights on possible cyber threats. Examples include the identification of situations that can become a threat for the CIIs with significant legal, regulatory and technical considerations. Such situations are: organization of hacktivist activities in underground forums or IRC channels; external situations that can become a potential threat to the CIIs (e.g. relevant geopolitical changes); disclosure of zero day vulnerabilities; sockpuppets impersonating real profiles in social networks etc. Entities (e.g., events, places) and security-realated information will be uniquely extracted from textual content using advanced Natural Language Processing (NLP) techniques, such as sentiment analysis.

5.3 Data fusion, risk evaluation and event management (HybridNet) component

The Data Fusion, Risk Evaluation and Event Management (HybridNet) component provides the intelligence needed to perform effective and efficient analysis of a security event based on: (1) information derived and acquired by the LiveNet and DarkNet components; and (2) information and data produced and extracted from this component. In particular, HybridNet component retrieves incidents-related data via the LiveNet component from the underlying CIIs and data from unstructured and structured sources (e.g. from Deep and Dark Web) consolidated in a unified longitudinal view which are linked, analysed and correlated, in order to achieve semantic meaning and provide a more comprehensive and detailed view of the incident.

In CyberSANE, a formal and uniform representation of digital evidence along with their relationships has been used to encapsulate all concepts of the forensic field and provide a common understanding of the structure of all information linking to evidence among the CIIs' operators and the forensics investigators. The main goal of the analysis process is to continuously carry out the assessment (e.g. identification of on-going attacks and related information, such as what is the stage of the attack and where is the attacker) and prediction (i.e. identification of possible scenarios of future attacks through forecasting models). HybridNet incorporates fusion models based on existing mathematical models (e.g. data mining, AI, deep learning, machine learning and visualization techniques). These models will support and provide reasoning capabilities for the near real-time identification of anomalies, threats and attacks, assessing any possible malicious actions in the cyber assets such as abnormal behaviors or malicious connections to identify unusual activities that match the structural patterns of possible intrusions.

In order to meet its objectives, the HybridNet will consist of three elements *Anomaly Detection Engine*, *Incident Analysis & Respond* and *Decision-Making, Warning and Notification* which are further described below.

The *Anomaly Detection Engine* will undertake the responsibility to process a large amount of data delivered from the abovementioned components. The objective of this engine is to analyse the received data in order to further evaluate and correlate attack-related patterns associated with specific malicious or anomalous activities in the CIIs. Thus, when the engine identifies unusual activities that match the structural patterns of possible intrusions, generates alerts to show that these activities require a more intensified analysis.

Once an event has been considered by the *Anomaly Detection Engine* as a real security incident, the *Incident Analysis & Respond* element is responsible to further investigate it. The analysis is performed based on data and information produced from the following subsystems: (1) the *Collaborative Evidence-based Risk Assessment* subsystem implements the main steps required for the identification, evaluation and mitigation of all vulnerabilities, threats and risks associated with the CIIs, in a graphical way using visualization tools, simulation processes, automated routines and structured content. (2) The *Collective Intelligence and Big-Data Analytics* subsystem implements a wide array of reasoning, data mining and big data analytics techniques which will incorporate and leverage a variety of data sources and data types in order to enhance and optimize the investigation, analysis and response of a security incident. (3) The *Model Processing Management* subsystem includes a variety of modeling tools and methods in order to easily visualize CIIs and identify all type of interdependencies (at physical, system, technology and business levels).

Finally, the *Decision-Making, Warning and Notification* element is responsible to orchestrate and facilitate the analysis, which includes the scrutiny of the attacker's actions and identification of the means that were employed by the attacker, and in overall and understanding of how the attack originated and evolved. This subsystem takes into account the incident-related information processed by the *Anomaly Detection Engine* and the *Incident Analysis & Respond* in order to design and execute the necessary simulation experiments through the *Security Incident/Attack Simulation Environment* and the *Behavior Simulation Environment*.

The *Security Incident/Attack Simulation Environment* comprises a set of novel mathematical instruments, including mathematical models for simulating, analysing, optimizing, validating, monitoring simulation data and optimizing security incident handling process. Specifically, these instruments include: (1) a bundle of novel process/attack analysis and simulation techniques for designing, executing, analyzing and optimizing threat and attack simulation experiments that will produce appropriate evidence and information that facilitate the identification, assessment and mitigation of the CII-related risks; (2) graph theory to implement attack graph generation, to perform security incident analysis and to strengthen the prognosis of future malefactor steps; (3) pioneering mathematical techniques for analyzing, compiling and combining information and evidences about security incidents and attacks/threats patterns and paths in order to find relationships between the recovered forensic artefacts and piecing the evidential data together to develop a set of useful chain of evidence (linked evidence) associated with a specific incident; (4) innovative simulation techniques which will optimize the automatic analysis of diverse data; (5) innovative techniques in order to link optimization and simulation. In this context, this simulation environment is fed with information about an incident and proceeds to calculate and generate a number of possible attack graphs (routes of possible attacks) and graphs of linked evidence (chains of evidence) and also compute probabilities for a sequence of events on top of these graphs. The resulting probabilistic estimate for the compromised CIIs' assets will be used to identify, model and represent the course of an attack as it propagates across the CIIs. It should be noted the HybridNet component continuously updates the simulation engine with information collected and piece of information, thereby enabling both understanding which assets might have been compromised, as well as gain more accurate estimates on the likelihood that other assets might be compromised in the future.

The second simulation environment is the *Behavior Simulation Environment* that aims to stimulate the behavior of CIIs' operators and Treat Actors taking into consideration their cyber interactions and interdependencies to measure the cascading effects of various cyber-attack patterns and

security incidents within the digital ecosystem. Based on the estimated effects, the environment is able to formulate extensive plans to mitigate the effects of such incidents.

5.4 Intelligence and information sharing and dissemination (ShareNet) component

The ShareNet component provides the necessary threat intelligence and information sharing capabilities within the CIIs and with relevant parties (e.g. industry cooperation groups, CSIRTs). It is responsible for the instantiation of the adopted intelligence model; in particular, ShareNet undertakes the identification and dissemination of, the right and sanitized information that have to be shared in a usable format and in a timely manner. This environment produces and circulates notifications containing critical information, enhancing the perception of the current situation and improving the projection into the future. It should be noted that all potential evidence from the systems that are suspected to be part of the infrastructure being investigated are forensically captured, stored and exchanged in a way that their integrity is maintained using the security and data protection methods of the PrivacyNet Orchestrator.

To this end, ShareNet follows a trusted and distributed intelligence and incident sharing approach to facilitate and promote the collaboration and secure and privacy-aware information sharing of the CIIs' operators with relevant parties (e.g. industry cooperation groups, CSIRTs), in order to exchange risk incident-related information, through specific standards and/or formats (STIX) (OASIS 2017a, b, c, d, e), improving overall cyber risk understanding and reduction. Privacy preserving is another important issue considered at every phase of sharing, applying methods such as anonymization or pseudo anonymization and encryption techniques incorporated in and made available from PrivacyNet Orchestrator. This brings forward a mixture of several cryptographic techniques that holds certain security guarantees.

5.5 Privacy and data protection (PrivacyNet) Orchestrator

Through the specific "Privacy and Data Protection Orchestrator" (PrivacyNet), it is possible to coordinate the abovementioned components of the CyberSANE system in order to ensure desired-levels of data protection for sensitive incident-related information, enabling the possibility to apply such protection in all phases of cyber security incident handling flow. The main purpose the PrivacyNet is to manage and orchestrate the application of the innovative privacy mechanisms and maximize achievable levels of confidentiality and data protection towards compliance with the highly-demanding provisions in the GDPR in the context of protecting sensitive incident-related information within and outside CIIs. To this end, PrivacyNet sets up the security and data "protection configurations" allowing security experts and members of the incident response team to specify all the protection steps that have to be performed and the required conditions to execute them, which can be referred to GDPR-based rules (and to other guidance for its application by the European Data Protection Board, formerly Art. 29 Data Protection Working Party).

In addition, the orchestration approach of the CyberSANE allows applying the most appropriate security and data protection methods depending on the user's privacy requirements, which cover a wide range of techniques including anonymization, location privacy, obfuscation, pseudonymization, searchable encryption, multi-party computation and verifiable computation, in order to meet the highly demanding regulatory compliance obligations, for example in relation to accountability towards data protection supervisory authorities, for adequate management of informed consent etc. For this reason, novel techniques and processes for enhancing the secure distribution and storage of all forensic artifacts in order to protect them from unauthorized deletion, tampering revision and sharing (e.g. Attribute-based Encryption (ABE) and block-chain technologies) have been combined.

6 Data flow and system operation of the CyberSANE components

The CyberSANE approach aims to assist all the phases of cybersecurity assessment and reduction of cyber risks on relevant operational scenarios, such as those that are described in the next section. Figure 3 illustrates the data flows and the inherent connections between the CyberSANE core components, which are enumerated in a subsequent step format and presented in the following. In the perception phase (step 1), members of the incident response team of the CIIs use the CyberSANE's LiveNet Component for extracting all source data (e.g. their servers, systems, logs) and discovered the required evidence, such as changes in the backend servers, traces of suspicious traffic as well as malicious software or back door Trojan installation. Thus, they can conclude whether their organization has fallen victim to a cyberattack. However, at this stage with current technologies they are unable to identify exactly how the intrusion happened, or who the attackers were. The CyberSANE system helps with that through the LiveNet components, which undergoes a transformation process following a common semantic format, allowing collected incidents-related information (operational/static/network traffic data, existing vulnerabilities) to be parsed, filtered, harmonized and enriched, fed and fused to CyberSANE's HybridNet Component (step 2). HybridNet is responsible to correlate real-time information fed by the

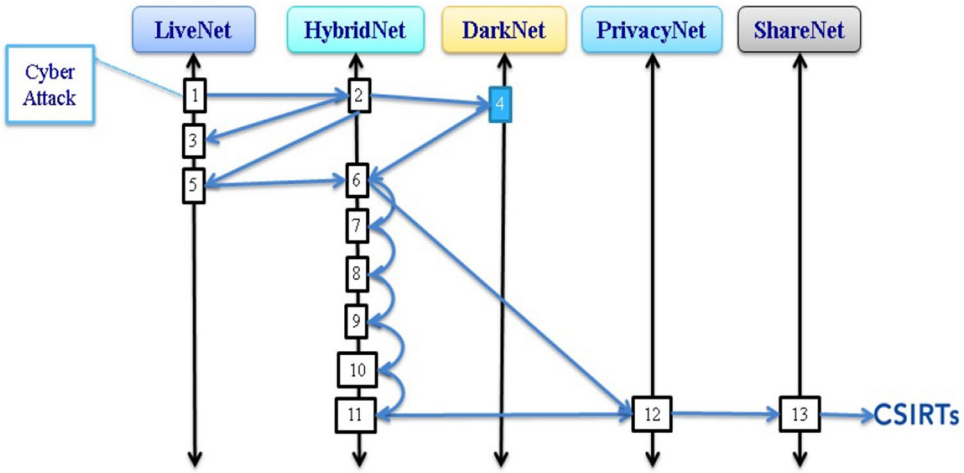

Fig. 3 Data flows and operation of the CyberSANE system's core components

LiveNet (step 3). with the most updated information about latest mechanisms of cyber-attacks collected by DarkNet (step 4) in order to carry out the investigation of the identified incident. In particular, HybridNet uses the collected operational data that describes the configuration of systems and software (e.g., network topologies and existing vulnerabilities), static data that describe general risk (e.g., if an identified vulnerability has an exploit that is publicly available), operational network traffic logs corresponding to the studied systems and software as well as information about malicious activities published in underground forums to calculate an attack graph and compute a Bayesian Network on top of this attack graph (Reazul et al. 2017). Using the produced graphs combined with non-technical factors (e.g. adversaries' intensions), the services begin to map out all systems/devices (e.g. the container management system or the energy management sensors) needed to be analysed and prompt to capture as much information as possible related to these systems and network. After this planning stage, HybridNet utilizes LiveNet components start collecting relevant information (e.g. any traffic identified as malicious) which could be useful for the investigation (e.g. Keywords, IP/MAC addresses) and will support the evaluation process (step 5). The resulting data are used to re-generate and re-calculate the graphs and the Bayesian Networks corresponding to this new piece of information, thereby enabling both understanding which assets might have been compromised and which vulnerabilities were exploited (step 6). In this way, these services enable the victims to identify all evidence related to the initial incident, to discover and depict the relationships between devices and the evidence and to produce a timeline of the incident, including a map of affected devices and chains of evidence (step 7). In addition, HybridNet assesses threats in terms of estimating the probability that some actor will initiate a cyber-attack by taking non-technical factors, such as intent, into account (step 8). The probability estimates is used compare mitigation strategies and de-risking steps in order to find the most cost-efficient actions to implement (step 9). HybridNet proposes mitigation strategies and re-risking steps at multiple levels (step 10). In this case, relevant de-risking actions include: patching existing vulnerabilities, investing in a better intrusion detection system, implementing new policy's/procedures and getting a new tailor made insurance (step 11). Now, having understood the cause of the incident, CyberSANE system uses the PrivacyNet to apply the appropriate security and data protection mechanisms in order to guarantee the secure communication, maintenance and storage of the incident-related information (step 12). Finally, the anonymized incident-related information will be shared with and distributed to accredited CSIRTs in order to enhance the cyber security aware (step 13).

7 Cyber-attack showcase scenarios

As aforementioned, CyberSANE will be applicable to various security scenarios of CIIs'. In order to evaluate the proposed solution, CyberSANE system will be validated in the scope of three realistic cyber-attacks and incident management showcases scenarios described in the following sections, covering three (3) industry sectors: energy, transportation and healthcare, which are critical according to the Council Directive 2008/114/EC (2008) and to the Homeland Security Presidential Directive 7 (Homeland Security 2003).

These critical sectors have been selected as a suitable test bed for the CyberSANE platform. In the context of cybersecurity, there are not major differences between different domains (e.g. transportation, airport, energy, telco, healthcare and maritime) and all of them are facing similar security challenges. Therefore, the generic nature of the CyberSANE approach is envisaged to render it applicable to a variety of Critical Information Infrastructures of different sizes and different business activities. Furthermore, the CyberSANE

system can be used by various organizations coming from multiple sectors to handle cyberscurity incidents and manage the security risks of their ICT environments.

Within this framework, a number of use cases will be run to ensure the system is functioning properly works as intended and to validate the system's capability for detecting and preventing any types of malicious activities that may allow the adversaries to launch and conduct attacks on the CIIs. It should be noted that these scenarios were identified as critical to security and economics taking into consideration knowledge gained from various research activities dealing with CIIs security and protection (i.e. CORE, MEDUSA, MITIGATE and SAURON). In addition, the targeted scenarios play a central role in the assessment and validation of CyberSANE's proposed innovations.

The showcases presented in the following, engage security operations in heterogeneous, large-scale, cross-border CIIs, that are characterized by the following features: (1) complex, highly distributed, and large-scale cyber systems (including IoT and cyberphysical) regarding the number of entities involved; (2) heterogeneity of the underlying networks interconnecting the physical-cyber systems; and (3) different levels of exposure to attacks.

7.1 Scenario 1: "solar energy production, storage and distribution service"

In the era of Smart Energy, various energy providers have developed and offered advanced and innovative energy production, storage and distribution services. To this end, they operate disruptive, transformational and intelligent energy management solutions and offer a number of digital services on top, helping energy "procumers", utilities and grid operators to optimize power flows, secure the electricity grid and finally reduce the cost of electricity. Such solutions incorporate a bundle of components such as: a range of web apps for the end user that enable users to see in real time the power flow between the solar system, the battery and the grid of their household; intermediate devices between sensors, smart meters, inverters, the battery and appliances to collect and monitor the data; smart grid networks of a population of distributed and interconnected assets; back-office applications which automate the entire process of the business; electric panels designed to accelerate the installation process of the system and eliminate connectivity errors etc. Figure 4 depicts the Solar Energy Production, Storage and Distribution Service.

Attacks on "Solar Energy Production, Storage and Distribution Service": various combined cyber-attacks may affect the examined solar energy service. From the cyber part, attacks against back-end components such as gaining unauthenticated remote access to IoT components and other entities to disrupt services and change their data set points or state. Other cyber-attacks may target against the IT and communication systems that are used to process the sensed data and transmit them to the corresponding IT systems.

7.2 Scenario 2: "container cargo transportation service"

Nowadays, commercial ports are considered as the primary "hubs" within the larger network of flows and interactions and are at the core of the national, regional and European economic recovery. The automation of port terminal and intermodal handling operations is very important and requires the use of highly complex and involve numerous different systems. In order to meet their objectives, the commercial ports operate over three main family of systems: Information Technology (IT) systems which includes databases with operational and business

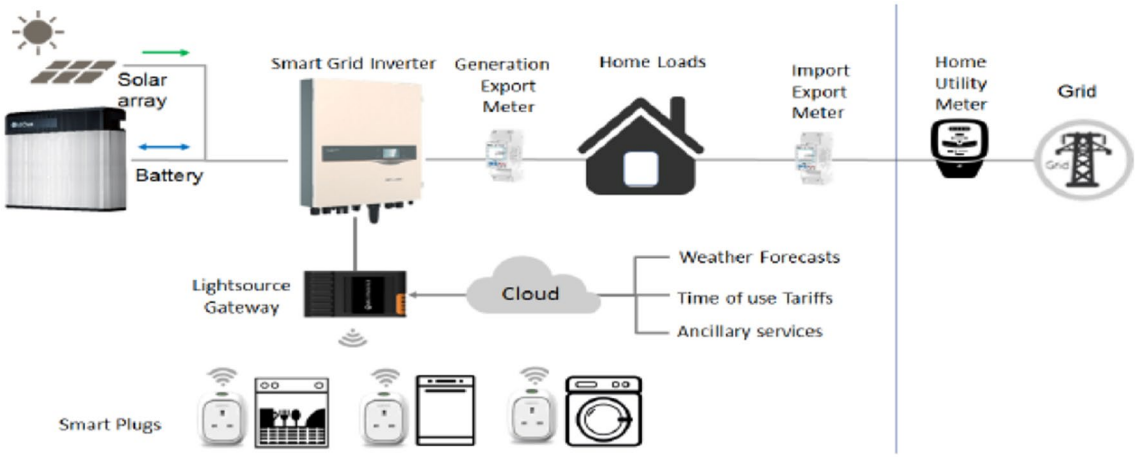

Fig. 4 Solar energy production, storage & distribution service

information, Operational Technology (OT) systems like Supervisory Control and Data Acquisition (SCADA) systems, which control physical processes e.g. unloading from the vessel is performed by yard tractors and forklifts with their auxiliary equipment, which are monitored by a SCADA system, and the Port Community System (PCS) which generally supports port's services such as management of docking and shipping of vessels. Any attack to the above mentioned systems may result in a serious threat to the correct functioning of the port and in consequence for the national security. Figure 5 presents the Container Cargo Transportation Service.

Attacks on "Container Cargo Transportation Service": the Container Cargo Transportation Service can be subject of a number of possible threat scenarios that can be realized by conducting a combination/series of specific cyber attacks in port's IT component. In this context, malicious users/adversaries are able to realize complex threat scenarios for the purpose of disrupting port's operations or facilitating illegal activities (such as smuggle illegal material of any kinds (such as drugs, weapons etc.) or illegal immigrants, or event to destroy a major/critical Infrastructure) aimed at obtaining financial, political/military or even ideological gain and benefits. To this end, malicious attackers can launch targeted attacks in order to gain unauthorized access to the victims' systems and use them as stepping stone to launch further, more sophisticated, attacks and go deeper into their cyber infrastructure. For example, they can infiltrate the port's wireless by obtaining the network identifier or through other network vulnerabilities in order to sniff, modify or inject falsified data to achieve the expected results. In addition, they can continue to exploit other vulnerabilities in various systems, either in the corporate network or the SCADA network, to gain unauthorized access to places such as corporate networks, SCADA systems, interconnections, and access links. Therefore, the adversary can target the Port Community System of the port and takes advantage of specific software bugs and flaws that may have. In this way, the attacker can interfere with the authorization process, allowing a vessel carrying illegal or hazardous materials to enter and dock at the port or even to bypass the inspection procedure.

7.3 Scenario 3: "real-time patient monitoring and treatment service"

Over the past decade, the medical field has experienced a massive digitization. Electronic Health Record (EHR)/Electronic Medical Record (EMR) have appeared and clinical processes have been automated. Modern digital health care organizations rely more and more on the aid of technological advancement including IoT, Cloud Computing, Big Data, and advances on medical equipment to increase the degree of flexibility, scalability, and efficiency in the communication and coordination of health care services. In particular, in order to automatically collect and process the medical data, various medical devices and instruments are connected, through wired or wireless communications, with the EHR/EMR systems (Stellios et al. 2018). For example, smart insertable cardiac monitoring devices beds may be connected to automatically inform the doctors with patient data, or they may be used by nurses to note the daily treatment/medicines received by a patient. Also, medical instruments such as medical radiation devices can be connected to EHR/EMR IT systems to assist doctors during medical treatment. Other IT equipment may involve secondary services, such as web presence. Outside the hospital, various medical Internet of Things (IoT) technologies can also be used to extend the medical services provided. For example, Implantable and Wearable Medical Devices (IMD/WMD) can be used to monitor (sense) patient data and also to remotely treat a patient in emergency situations, such as inject insulin when the sensed data indicate that this is urgent. The IMD/WMD devices may be controlled by home monitoring or programming devices, which communicate with the IMD/WMD using short-range wireless communication protocols, while they communicate with the in-hospital EHR/EMR IT systems using Internet access (Stellios et al. 2018).

It should be noted that the evolving digital interconnectivity has also changed the threat landscape, producing a

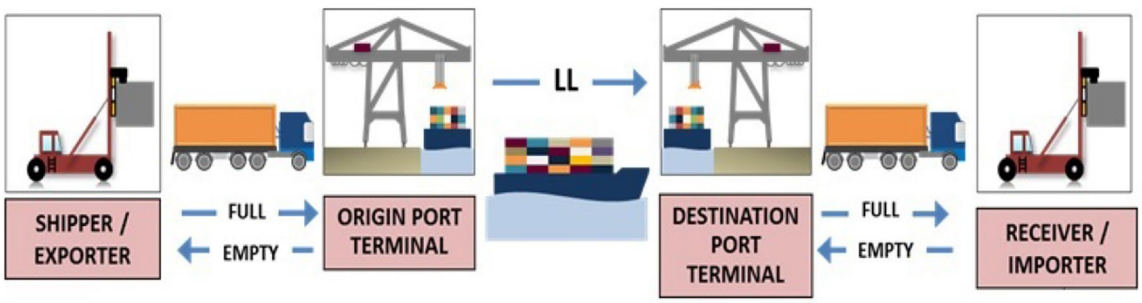

Fig. 5 Container cargo transportation service

wide range of security and privacy challenges and increasing the danger of potential cybersecurity attacks in the digital healthcare ecosystem. In fact, the situation regarding cybersecurity in healthcare facilities is quite alarming, with the number of attacks and breaches related to health care infrastructures growing. Figure 6 illustrates the Real-Time Patient Monitoring and Treatment Service.

Attacks on "Real-time patient monitoring and treatment service": the Patient Monitoring Service is subject to various cyber-attacks on both sides. At the hospital side, EHR/EMR IT and file systems are very attractive targets for ransomware attacks, due to their importance for all medical data and services (Stellios et al. 2018). Also, privacy loss is highly important, due to privacy regulations. However, various Internet connected medical devices with IoT capabilities are easy entry points for hackers, due to their low security level (Stellios et al. 2018). In real security events, such medical devices have been used by hackers as an entry point in order to escalate their access and attack their actual target IT systems or to exfiltrate sensitive data. On the user side, the use of vulnerable wireless communications can (and in many real cases has) been used in order to attack the Service and even cause physical damage to a patient. For example, according to Stellios et al. (2018) by replaying or by manipulating commands at the API used by the IMD/WMD devices, it is possible to inject commands that may change the dosage of an insulin pump, thus directly affecting the health of the patient.

7.4 Benefits and added value of employing CyberSANE on the showcases scenarios

Although the presence of such large number of interconnected ICT components, complex ICT infrastructures and emerging technologies (e.g. IoT) as described in the three above scenarios, have significant benefits for CIIs' operators, there are also significant risks and potential vulnerabilities, which CyberSANE will explore through the application of the CyberSANE system and accompanying components to the above use cases:

- Various Internet connected systems (e.g. energy management sensors or medical devices with IoT capabilities) are also easy entry points for attackers, due to their low security level. In real security incidents such systems (e.g. IoT devices have been used by attackers as part of APT attacks and as an entry point towards escalating their access.
- Most CIIs' cyber assets are not designed with security in mind. Devices are mostly built based on "intended use" cases, and what a reasonable person might do. Hacking and other network-borne accidents are "unintended use" or "abuse" cases. This posture leads to a number of systemic vulnerabilities and risks throughout the CIIs.
- Security decisions made locally for a specific device can have global impacts. In many cases devices were designed without the specific intent to be connected to a network (sometimes specifically intended to remain isolated)—that requirement came later and was bolted on. The communication between smart devices and legacy systems can also create gaps and give space for malicious attackers to gain illegal access to systems and data.

So far, existing HealthCare, Transportation and Energy CIIs have made a huge effort in isolating all its systems against cyber-attacks, malware strains or many other malicious techniques. In this line, in the aim to fight against cyber-attacks and to protect critical national infrastructures, they have installed a set of security probes in several points of port's technological ecosystem formed by the IT, OT and PCS. These probes are devices or programs that monitor and collect network data as well as activity data from systems and provide mechanisms for early detection of attacks and vulnerabilities of systems in order to quickly proceed with counter actions to overcome these threats. Despite the enormous work done in terms of cyber security, there is still a

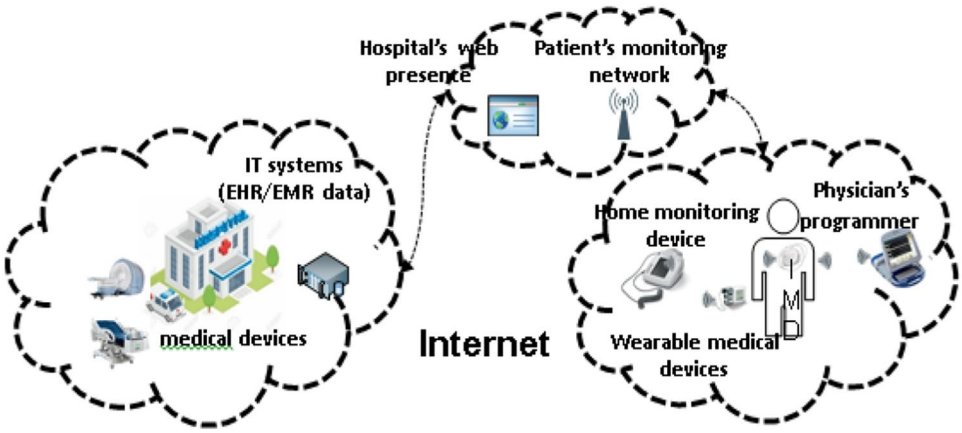

Fig. 6 Real-time patient monitoring and treatment service

need for keeping cyber security tools updated in near real time (ideally in the range of minutes) as new ways of cyber-intrusion appear. An over-the-top solution to this problem would be to use the CyberSANE's system in order to correlate, through HybridNet, real-time information coming from the of-the-self CIIs' security framework (raw data, processed data, local security warnings, security test results, etc.) via LiveNet with the most updated information about the latest mechanisms of cyber-attacks collected by DarkNet. With the CyberSANE system, the CIIs will definitely be able to boost the detection process of near future threats and ongoing attacks, as well as to minimize potential damages when an attack occurs. At the same time, ShareNet would be able to inform about incidents to interested key external players, which are connected to CIIs, in order to react on time and avoid potential cyber-attack propagations.

The selected showcases engage APT attack scenarios, target at illustrating how the proposed system can detect and respond to advanced, sophisticated attacks in various industry CIIs. Additionally they aim to show how the CybeSANE system can be used by industry stakeholders as a tool to get a comprehensible insight view of the cybersecurity processes in cross-sector partnerships, learn on the existing vulnerabilities of their operating CIIs and get advised to improve the management of cyber security threats.

The validation and evaluation of the CyberSANE system will rely on the Validation and Verification best practices described by the International Council on Systems Engineering (INCOSE) (2019) and the Guide to the Systems Engineering Body of Knowledge (SEBoK) (2019) created by the Body of Knowledge and Curriculum to Advance Systems Engineering (BKCASE) project.

8 Innovative aspects of CyberSANE system

CyberSANE will develop a system that addresses technical related to identification, prevention and protection against attacks. In particular, the CyberSANE system collects, compiles, processed and fuses attack related data from multiple perspective, through its main three components: The Live Security Monitoring and Analysis (LiveNet) component, the Deep and Dark Web mining and Intelligence (DarkNet) component and the Data Fusion, Risk Evaluation and Event Management (HybridNet) component. In order for these components to optimize the identification and analysis of security incidents and privacy breaches, they combine existing machine learning techniques, such as clustering and hidden Markov models (Jain and Abouzakhar 2012), with deep learning and Global Artificial Intelligence (AI) to develop an innovative way that optimizes the automatic analysis of huge amounts of events, information and evidence. To identify malicious actions in the cyber assets, such as abnormal behaviours, they combine both structured data (e.g. logs and network traffic) and unstructured data (e.g. data coming from social networks and dark web) in a privacy-aware manner. Furthermore, they adopt deep semantic analysis techniques together with Natural Language Processing (NLP) methods (e.g. Named Entity Recognition and Word Sense Disambiguation) to extract important information from multilingual security-related contexts, facilitating multilingual data generation and exploitation within the networked ecosystem.

In addition, the ShareNet component optimizes the collaboration and the interaction of the CIIs' stakeholders with relevant parties (e.g. industry cooperation groups, Computer Security Incident Response Teams—CSIRTs), in order to exchange risk incident-related information, through specific standards and/or formats, improving overall cyber risk understanding and reduction.

Finally, the PrivacyNet component implements the necessary security and privacy-related algorithms, techniques and approaches that ensures that all potential evidence from the systems that are suspected to be part of the infrastructure being investigated are forensically captured, stored and exchanged in a way that their integrity is maintained. Along with encryption methods (e.g. Attribute-based Encryption (ABE)) that have the potential to achieve the privacy and security goals of the project, PrivacyNet uses innovative methods (Cresitello-Dittmar 2016) (e.g. blockchain technologies, anonymization techniques) for secure distribution and storage processes, to protect all forensic information from deletion, tampering, and revision in order the data to be used as evidence in the court.

The abovementioned components are combined with innovative visualization techniques based on virtual reality technologies in order to create a multidimensional environment that allows the operations managers to have a better comprehension of the cyber environment. This new approach includes the use of novel immersive HMI for the operators in order to give them a new perspective for visualizing some parameters or graphs that impede the sensorial overload inherent to the large amount of data management and analysis process. As "situation awareness" is a decision makers' mental perception, the use of immersive multidimensional HMI allows them to reach the more precise situation status in order to go beyond and provide the "situation understanding" mental perception to the decision makers. Thus, the decision makers will be able to take proactive as well as reactive decisions for reducing the response time and improving the efficiency.

9 Conclusions

Current work presented an innovative, knowledge-based, collaborative security and response dynamic approach, which engages an overall view of the cyber incident handling

lifecycle, in order to detect and handle security incidents, scrutinize activities and provide post-incident knowledge harvesting. Furthermore, the approach targets at leveraging collected security information, for the purpose of finding new ways of protection for ICT assets, enabling the entity at risk to evaluate the risk and invest to limit that risk in an optimal way. The presented work provides a holistic and integrated approach of incident handling, which aims to enhance the incident detection capabilities of existing methods with efficient, elastic and scalable reasoning insights. On this account, it aims to facilitate the detection and analysis of Advanced Persistent Threats (APTs) and anomalous activities due to complex cyber attacks on Critical Information Infrastructures in the Transportation, Healthcare and Energy industries and thus expand the sector awareness on cyber threats and risks. Concerning this, the proposed work provides a way to securely collect both structured data (e.g. logs and network traffic) and unstructured data (e.g. data coming from social networks and dark web) making them available for analysis, fostering new innovations that will only unravel after having access to such data, harnessing its full potential. In this vein, CyberSANE's has a twofold aim; to minimize the exposure to security risks/threats and help CIIs' operators to respond successfully to relevant incidents.

The ground-breaking nature of the proposed incident handling approach is based on: (1) the identification of attacks and incidents using innovative approaches and algorithms of existing unobserved components techniques and linear state-space models producing meaningful information from cyber systems, (2) the combination of active incident handling approaches with reactive approaches producing real-time insights, alerts and warnings about cyber events, (3) innovative normalization process that unifies all relevant incident-related information gathered from heterogeneous CIIs, (4) novel attacks scenarios and evidence representation with simulation techniques and visualization tools that increase the efficiency of investigation results, (5) hybridization forms of existing mathematical models and combinations of data mining, Global artificial intelligence, machine learning that optimize evidential data from different sources.

The presented showcases are real-life demonstration scenarios upon which the proposed solution will be evaluated. They have a twofold objective: to illustrate the security requirements and challenges in the scope of various and different types and architectures of CIIs and to provide guidelines for the successful deployment of the CyberSANE system. However, as challenges in the context of cybersecurity have not major differences across industries, the CyberSANE solution could be utilized by organizations coming from multiple industry sectors.

The CyberSANE approach achieves to combine active approaches that are used to detect and analyze anomaly activities and attacks in real-time with reactive approaches that deals with the analysis of the underlying infrastructure to assess an incident in order to provide a more holistic and integrated approach to incident handling.

The proposed CyberSANE solution meets its objectives embedding core security features allowing faster and better operation of advanced cyber security functionalities. These aspects comprise an innovative, knowledge based, collaborative security and response dynamic system which increases the agility of the investigators and encourages continuous learning throughout the incident life cycle.

Acknowledgements Open access funding provided by Stockholm University. This work has been supported by the European Union's Horizon 2020 Project "CyberSANE" under Grant Agreement No. 833683, the European Union's Horizon 2020 Project "CyberSec4Europe" under Grant Agreement No. 830929 and the European Union's Horizon 2020 Project "SAURON" under Grant Agreement No. 740477 addressing the topic CIP-01-2016-2017. The authors would like to thank all projects' members for their valuable insights. Finally, special thanks to the University of Piraeus, Research Centre for its continuous support.

References

Ab Rahman NH, Choo KKR (2015) A survey of information security incident handling in the cloud. Comput Secur 49:45–69

Ahmad A, Hadgkiss J, Ruighaver AB (2012) Incident response teams-challenges in supporting the organisational security function. Comput Secur 31(5):643–652

Blowers M, Williams J (2014) Machine learning applied to cyber operations. In: Pino RE (ed) Network science and cybersecurity. Springer, New York, pp 155–175

British Standards Institution (2011) BS ISO/IEC 27035:2011—information technology. Security Techniques. Information Security Incident Management

Bruschi D, Monga M, Martignoni L (2004) How to reuse knowledge about forensic investigations. In: Digital forensics research workshop, Linthicum, Maryland

CAPEC (2017) CAPEC common attack pattern enumeration and classification. https://capec.mitre.org/. Accessed 09 Oct 2019

Casey E (2006) Investigating sophisticated security breaches. Commun ACM 49(2):48–55

Cichonski P, Scarfone K (2012) Computer security incident handling guide recommendations of the National Institute of Standards and Technology (NIST). NIST, Gaithersburg

Connell A, Palko T, Yasar H (2013) Celebro: a platform for collaborative incident response and investigation. In: 2013 international

conference on technologies for homeland security (HST). Waltham, MA, 2013. IEEE, pp 241–245
Cresitello-Dittmar B (2016) Application of the blockchain for authentication and verification of identity
Cusick JJ, Ma G (2010) Creating an ITIL inspired incident management approach: roots, response, and results. In: Network operations and management symposium workshops (NOMS Wksps), 2010 IEEE/IFIP. IEEE, pp 142–148
Danyliw R, Meijer J, Demchenko Y (2007) The incident object description exchange format, 5070
De Fuentes JM, González-Manzano L, Tapiador J, Peris-Lopez P (2016) PRACIS: privacy-preserving and aggregatable cybersecurity information sharing. Comput Secur 69:127–141
ENISA (2019) CSIRTs by country-interactive map. https://www.enisa.europa.eu/topics/csirts-in-europe/csirt-inventory/certs-by-country-interactive-map. Accessed 09 Oct 2019
Filkins B (2016) IT security spending trends. SANS Institute, Fredericksburg
FireEye (2013) The need for speed: 2013 incident response survey
Fisk G, Ardi C, Pickett N, Heidemann J, Fisk M, Papadopoulos C (2015) Privacy principles for sharing cyber security data. In: Security and privacy workshops (SPW), 2015 IEEE. IEEE, pp 193–197
Floreano D, Mattiussi C (2008) Bio-inspired artificial intelligence: theories, methods, and technologies. MIT Press, Cambridge
Gladyshev P (2004) Formalising event reconstruction in digital investigations. Doctoral dissertation, University College Dublin
Grimes J (2007) National information assurance approach to incident management. Committee for National Security Systems. CNSS-048-07
Grispos G, Glisson WB, Storer T (2014) Rethinking security incident response: the integration of agile principles. arXiv:1408.2431
Grobauer B, Schreck T (2010) Towards incident handling in the cloud. In: Proceedings of the 2010 ACM workshop on cloud computing security workshop (CCSW 10), pp 77–85
Guide to the Systems Engineering Body of Knowledge (SEBoK) (2019) SEBoK v.2.1. https://www.sebokwiki.org/wiki/Guide_to_the_Systems_Engineering_Body_of_Knowledge_(SEBoK). Accessed 09 Oct 2019
Gürses S, Jahnke JH, Obry C, Onabajo A, Santen T, Price M (2005) Eliciting confidentiality requirements in practice. In: Proceedings of the 2005 conference of the centre for advanced studies on collaborative research, pp 101–116
Homeland Security (2003) Homeland security presidential directive 7: critical infrastructure identification, prioritization, and protection division of emergency management. https://www.cisa.gov/homeland-security-presidential-directive-7. Accessed 09 Oct 2019
International Council on Systems Engineering (INCOSE) (2019). https://www.incose.org/. Accessed 09 Oct 2019
Jain R, Abouzakhar NS (2012) Hidden Markov model based anomaly intrusion detection. In: 2012 international conference for internet technology and secured transactions, London, 2012, pp 528–533
Kalogeraki E-M, Papastergiou S, Polemi N, Douligeris C (2018) SAURON real-life scenario: a terrorist coordinated attack in a EU port. Marit Interdiction Oper J 16(1):22–27
Khurana H, Basney J, Bakht M, Freemon M, Welch V, Butler R (2009) Palantir: a framework for collaborative incident response and investigation. In: Proceedings of the 8th symposium on identity and trust on the internet, p 38e51
Leucari V (2012) Analysis of complex patterns of evidence in legal cases: Wigmore charts vs. Bayesian networks. Artif Intell Law 4:173–182
Line MB (2013) A case study: preparing for the smart grids-identifying current practice for information security incident management in the power industry. In: IT security incident management and IT forensics (IMF). In: 2013 7 international conference on IT security incident management and IT forensics. IEEE, pp 26–32
Liu C, Singhal A, Wijesekera D (2013) Merging sub evidence graphs to an integrated evidence graph for network forensics analysis. Adv Digit Forensics IX:227–241
MAEC (2016) Malware attribute enumeration and characterization. http://maec.mitre.org/. Accessed 9 Oct 2019
Marsh report (2018) Could energy industry dynamics be creating an impending cyber storm? https://www.marsh.com/uk/insights/research/energy-industry-dynamics-be-creating-an-impending-cyber-storm.html?utm_source=publicrelations&utm_medium=referral-link&utm_campaign=eic-2018. Accessed 09 Oct 2019
Mohaisen A, Al-Ibrahim O, Kamhoua C, Kwiat K, Njilla L (2017) Rethinking information sharing for actionable threat intelligence. arXiv:1702.00548
Monfared A, Jaatun MG (2012) Handling compromised components in an IaaS cloud installation. J Cloud Comput Adv Syst Appl 1:16
MTI Network (2015) An MTI network special report: maritime cyber security. http://www.mtinetwork.com/mti-network-special-report-maritime-cyber-security/. Accessed 09 Oct 2019
Mukherjee S (2017) Why health care is especially vulnerable to ransomware attacks. http://fortune.com/2017/05/15/ransomware-attack-healthcare/. Accessed 09 Oct 2019
Neralla S, Bhaskari DL, Avadhani PS (2013) A novel graph model for e-mail forensics: evidence activity analysis graph. Int J Eng Sci Technol 5(10):1750
Nnoli H, Lindskog D, Zavarsky P, Aghili S, Ruhl R (2012) The governance of corporate forensics using COBIT, NIST and increased automated forensic approaches. In: 2012 international conference on privacy, security, risk and trust. IEEE
Northcutt S (2003) Computer security incident handling version 2.3.1
OASIS (2017a) STIX™ version 2.0. Part 1: STIX core concepts. https://docs.oasis-open.org/cti/stix/v2.0/stix-v2.0-part1-stix-core.pdf. Accessed 09 Oct 2019
OASIS (2017b) STIX™ version 2.0. Part 2: STIX objects. https://docs.oasis-open.org/cti/stix/v2.0/stix-v2.0-part2-stix-objects.pdf. Accessed 09 Oct 2019
OASIS (2017c) STIX™ version 2.0. Part 3: cyber observable core concepts. https://docs.oasis-open.org/cti/stix/v2.0/stix-v2.0-part3-cyber-observable-core.pdf. Accessed 09 Oct 2019
OASIS (2017d) STIX™ version 2.0. Part 4: cyber observable objects. https://docs.oasis-open.org/cti/stix/v2.0/stix-v2.0-part4-cyber-observable-objects.pdf. Accessed 09 Oct 2019
OASIS (2017e) STIX™ version 2.0. Part 5: STIX patterning. https://docs.oasis-open.org/cti/stix/v2.0/stix-v2.0-part5-stix-patterning.pdf. Accessed 09 Oct 2019
OASIS (2017f) TAXII™ version 2.0. https://docs.oasis-open.org/cti/taxii/v2.0/taxii-v2.0.pdf. Accessed 09 Oct 2019
OpenIOC (2017) OpenIOC: an open framework for sharing threat intelligence. http://www.openioc.org/. Accessed 09 Oct 2019
Papastergiou S, Polemi D (2017) Securing maritime logistics and supply chain: the medusa and mitigate approaches in proceedings of 2nd nmiotic conference on cyber security. Marit Interdiction Oper J 14(1):42–48
Papastergiou S, Polemi N (2018) MITIGATE: a dynamic supply chain cyber risk assessment methodology. In: Yang XS, Nagar A, Joshi A (eds) Smart trends in systems, security and sustainability. Lecture notes in networks and systems, vol 18. Springer, pp 1–9
Phillips C, Swiler LP (1998) A graph-based system for network-vulnerability analysis. In: Proceedings of the 1998 workshop on new security paradigms. ACM, pp 71–79
Reazul MK, Onik AR, Samad T (2017) A network intrusion detection framework based on Bayesian network using wrapper approach. Int J Comput Appl 166(4):13–17

Scott M (2018) Energy firms are worried about cyber attacks, but don't really know what to do, 2018 (by Mike Scott). https://www.forbes.com/sites/mikescott/2018/03/07/energy-industry-worried-about-cyber-attacks-but-doesnt-really-know-what-to-do/#621beac768bb. Accessed 09 Oct 2019

Shedden P, Ahmad A, Ruighaver AB (2011) Informal learning in security incident response teams. In: 2011 Australasian conference on information systems

Sheyner O, Haines J, Jha S, Lippmann R, Wing JM (2002) Automated generation and analysis of attack graphs. In: Proceedings. 2002 IEEE symposium on security and privacy, 2002. IEEE, pp 273–284

Stellios I, Kotzanikolaou P, Psarakis M, Alcaraz C, Lopez J (2018) A survey of iot-enabled cyberattacks: assessing attack paths to critical infrastructures and services. IEEE Commun Surv Tutor 20(4):3453–3495

Sullivan C, Burger E (2017) In the public interest: the privacy implications of international business-to-business sharing of cyber-threat intelligence. Comput Law Secur Rev 33(1):14–29

Swiler LP, Phillips C, Ellis D, Chakerian S (2001) Computer-attack graph generation tool. In: DARPA information survivability conference & exposition II, 2001. DISCEX'01. Proceedings, vol 2, pp 307–332

Tan T, Ruighaver T, Ahmad A (2003) Incident handling: where the need for planning is often not recognised. In: 1st Australian computer, network & information forensics conference

The Council of the European Union (2008) Council directive 2008/114/EC of 8 December 2008 on the identification and designation of European critical infrastructures and the assessment of the need to improve their protection. Off J Eur Union

Traffic Light Protocol (2017) Traffic Light Protocol (TLP) definition and usage. https://www.us-cert.gov/tlp. Accessed 09 Oct 2019

Ulltveit-Moe N, Gjøsæter T, Assev SM, Køien GM, Oleshchuk V (2013) Privacy handling for critical information infrastructures. In: 11th IEEE international conference on industrial informatics (INDIN), 2013. IEEE, pp 688–694

Vangelos M (2011) Incident response: managing. Encyclopedia of information assurance. Taylor & Francis, Milton Park, pp 1442–1449

Wang W, Daniels TE (2005) Building evidence graphs for network forensics analysis. In: Computer security applications conference, 21st annual. IEEE, p 11

Wang W, Daniels TE (2006) Diffusion and graph spectral methods for network forensic analysis. In: Proceedings of the 2006 workshop on new security paradigms. ACM, pp 99–106

Werlinger R, Muldner K, Hawkey K, Beznosov K (2010) Preparation, detection, and analysis: the diagnostic work of it security incident response. Inf Manag Comput Secur 18(1):26–42

West-Brown MJ, Stikvoort D, Kossakowski KP, Killcrece G, Ruefle R (2003a) Handbook for computer security incident response teams (csirts) (No. CMU/SEI-2003-HB-002). Carnegie-Mellon University, Pittsburgh, PA, Software Engineering Institute

Widup S (2018) Introducing the 2018 protected health information data breach report. https://www.verizon.com/about/news/new-report-puts-healthcare-cybersecurity-back-under-microscope. Accessed 09 Oct 2019

Wiik J, Kossakowski KP (2005) Dynamics of incident response. In: 17th annual FIRST conference on computer security incident handling, Singapore