

Research Article

HYBINT: A Hybrid Intelligence System for Critical Infrastructures Protection

Javier Hingant , Marcelo Zambrano, Francisco J. Pérez , Israel Pérez, and Manuel Esteve

Distributed Real Time Systems and Applications Lab, Communications Department, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain

Correspondence should be addressed to Javier Hingant; jahingme@upvnet.upv.es

Received 17 April 2018; Revised 27 July 2018; Accepted 5 August 2018; Published 27 August 2018

Academic Editor: Gregorio Martinez Perez

Copyright © 2018 Javier Hingant et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Cyberattacks, which consist of exploiting security vulnerabilities of computer networks and systems for any kind of malicious purpose (e.g., extortion, data steal, assets hijacking), have been continuously increasing worldwide in recent years. Cyberspace appears today as a new battlefield, along with physical world scenarios (land, sea, air, and space), for the organizations defence and security. Besides, by the fact that attacks from the physical world may have significant implications in the cyber world and vice versa, these dimensions cannot be understood independently. However, the most common intelligence systems offer an insufficient situational awareness exclusively focused on one of these decision spaces. This article introduces HYBINT, an enhanced intelligence system that provides the necessary decision-making support for an efficient critical infrastructures protection by combining the real-time situation of the physical and cyber domains in a single visualization space. HYBINT is a real cross-platform solution which supplies, through Big Data analytical methods and advanced representation techniques, hybrid intelligence information from significant data of both physical and cyber data sources in order to bring an adequate hybrid situational awareness (HSA) of the cyber-physical environment. The proposal will be validated in a detailed scenario in which HYBINT system will be evaluated.

1. Introduction

The accelerated technological development arose in the last decade has reached most of the society areas allowing organizations to be more efficient by expanding their operational spectrum to cyberspace and overcoming the physical limitations of markets and borders.

This recent transformation is mostly led by the Information and Communication Technologies [1], where hot research fields such as Big Data [2, 3], Artificial Intelligence [4], or Internet of Things (IoT) [5, 6] are already heading this new paradigm in which physical world (land, sea, air, and space) merges with cyberspace producing a hybrid environment where physical and cyber entities are linked together [7] (Figure 1).

This cyber-physical environment [8] brings a large set of opportunities but also new vulnerabilities and threats that must be properly managed [9]. In fact, organizations have now to face, beside the traditional threats related to physical world, the ones existing in cyberspace [10]. This task is even

harder to achieve due to the significant cross implications in both physical and cyber world caused by the attacks coming from any of them [11].

In particular, the exponential growth of cyberattacks, either limited or large scale, in recent years, has caused huge losses and damages all around the world. As reported by the Spanish National Cryptologic Centre [12] in their recent annual executive summaries [13, 14], more than 21,000 incidents were detected in Spain just in 2016, which represents an approximate growth of 98% regarding year 2010 (Figure 2). In case of the WannaCry virus, where more than 15 million of computers from more than 10,000 organizations could be affected, the economic losses reached around 200 million of euros [15]. Estonia, for its part, became a world reference due to the advanced cybersecurity technologies developed after the 2007s cyberattack that moved the country to the Stone Age by leaving blank all the government websites [16].

In order to achieve an effective defence in this hybrid scenario, where risks and threats have moved to a higher level, new strategies, security tools, and remediation plans must

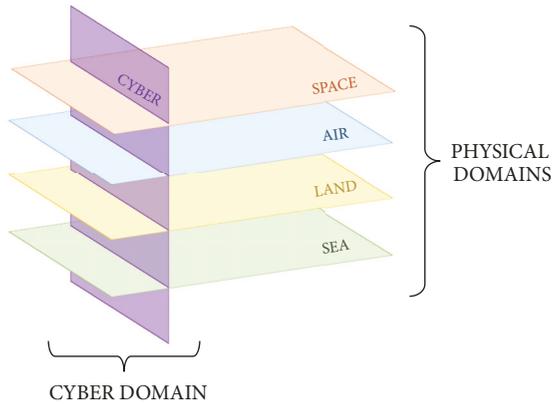


FIGURE 1: Physical and cyber domains.

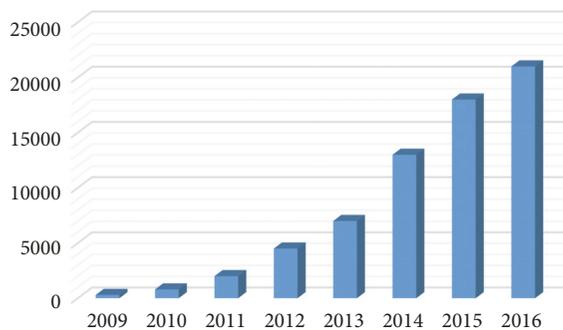


FIGURE 2: Cyber incidents detected in Spain (2009-2016) (source: <https://www.ccn-cert.cni.es/informes/informes-ccn-cert-publicos/1554-ccn-cert-ia-09-16-cyber-threats-2015-trends-2016-executive-summary/file.html>.) (source: <https://www.ccn-cert.cni.es/informes/informes-ccn-cert-publicos/2249-ccn-cert-ia-16-17-cyberthreats-trends-2017-executive-summary-1/file.html>).

be developed complying with current environment demands [17]. These requirements are even more necessary in case of critical infrastructures (CI) protection [18]. Indeed, CI refers to all of the strategic facilities and assets (e.g., power plants, medical centre, public administrations, banks, transportations) where well-functioning and maintenance are main priorities for any nation since its damage, destruction, or disruption may cause a significant negative impact on the country's security and economy [19, 20]. However, while traditional Command and Control systems-based security tools still do not efficiently integrate the situational awareness [21] of the cyber domain, the most extended cybersecurity solutions do not still provide a single and intuitive visualization space of the whole cyber-physical situation for an adequate decision-making support [22].

Intelligence systems, which seek to produce useful information for the trained security analysts from data gathered through different sources and means, play a main role in this context. An interesting approach to efficiently address the organizations security and defence in a cyber-physical (hybrid) environment could be a hybrid intelligence system which combines human [23] and cyber [24] intelligence, through collecting and processing significant data related

to both physical and cyber world, in order to achieve the adequate hybrid situational awareness [25]. For instance, in a CI like a port facility, when cyber threat intelligence information indicates that the GPS of a crane operator is locating him at the cargo area at the same time that its credentials are being used to log in on a computer placed far away, the human intelligence provided by a security personnel member inspecting these locations and reporting their situation could be the key to confirm a potential attack. In a different scenario, the intelligence information gathered from a surveillance staff member, which is notifying the suspicious behaviour of an authorized person accessing repeatedly to a strategic research centre at non-conventional working hours, could be jointly analysed and correlated with its recent activity on the organization internal network in order to determine if he/she is acting as an insider.

This paper introduces HYBINT, a real and advanced hybrid intelligence system for critical infrastructures protection. Firstly, both the current scenario and the present proposal are introduced; secondly, the oeuvres taken as main references and the motivation of this work are described; after that, HYBINT system architecture and the functionalities of each of its modules are explained in detail; then, a validation scenario and the obtained results are presented; and finally, main conclusions and future works are exposed.

2. Motivation and Related Work

Every day, current technologies provide to the organizations, CI included, new tools to improve their performance, reliability, and safety [26]. However, these same technological tools are also exploited by criminals to carry out ever more sophisticated attacks that require response actions on the physical as well as on the cyber domains.

Although multiple works focus on the study of the CI's risks and threats in the hybrid environment, it is hard to find an implemented solution that provides the necessary cyber-physical situational awareness to protect them efficiently from potential attacks in such scenario [27]. Indeed, the existing approaches do not imply real solutions able to achieve an advanced knowledge of the whole (physical and cyber) environment that enhances decision-making [28] for a complete and performing defence.

In this way, existing systems such as CoordCom [29], GEMMA [30], or GESTOP [31] are some examples of Command and Control Information Systems (C2IS) [32] which provide a situational awareness and a decision support [33] focused on the physical environment. Conversely, other relevant solutions such as the ones of Palo Alto Networks [34], IBM [35], Thales [36], or NEC [37], which provide advanced security tools for organizations protection, are specially focused on the cyber domain situation and the related threats and risks [38, 39]. This operational separation, which implies significant issues related to the systems interoperability and ontology as well as to the information standardization and representation, carry on a significant decrease in these systems performance and security deficiencies in them.

The most advanced SCADA systems [40, 41], which constitute a subset of the traditional Industrial Control

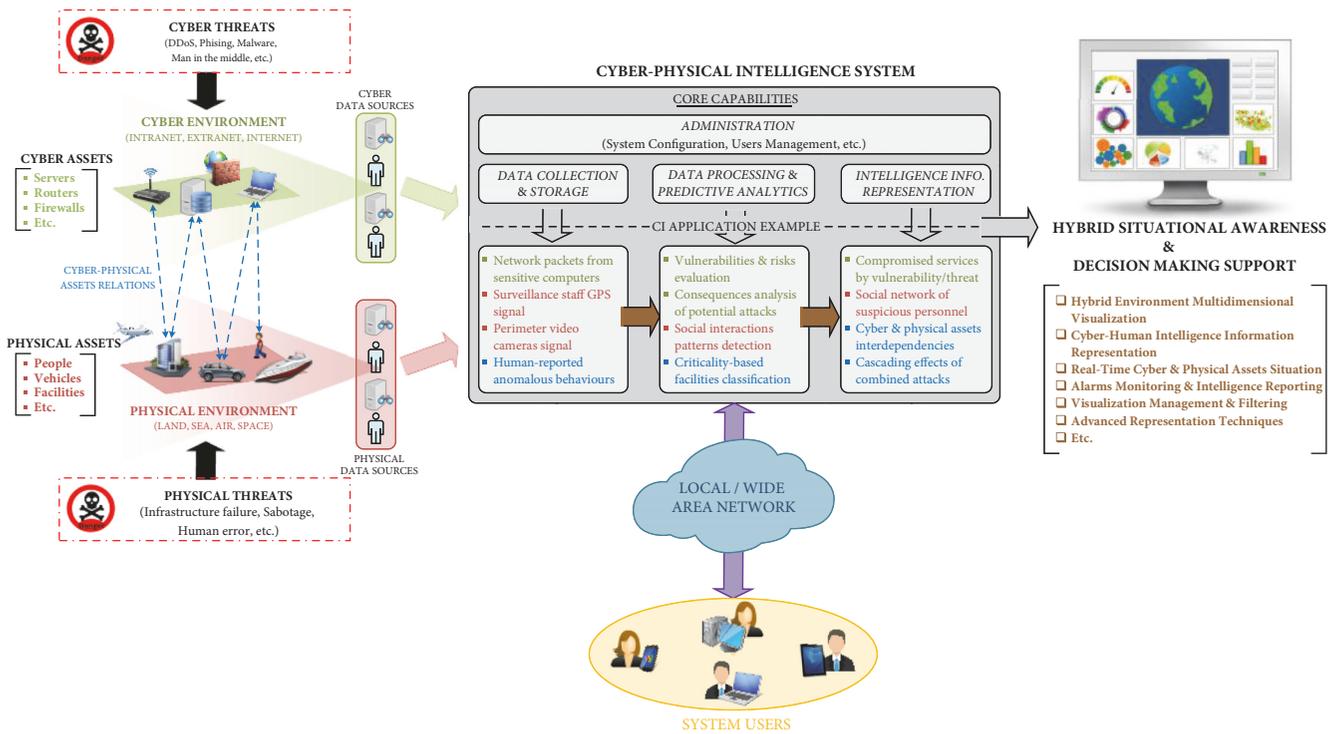


FIGURE 3: Hybrid intelligence system.

Systems (ICS) for the remote control and monitoring of sensors and high-level processes [42], integrate nowadays multiple security and defence tools such as SIEM systems, Intrusion Detection Systems (IDS), and Outage Management Systems (OMS). However, despite the fact that 4th-generation SCADA systems, based on recent technologies as cloud computing and IoT, seek to reach new application areas [43], they still remain mostly oriented to industrial sectors (oil refineries, waste management, manufacturing, etc.). Even the most recent SCADA solutions, which already integrate cyber threat intelligence capabilities, are limited to automated processes for both physical and cyber data gathering and do not still explore the benefits of human intelligence (HUMINT) techniques, which can enrich the consistently collected data with intelligence information provided by reliable human sources. Moreover, the HMI of these systems, commonly based on conventional diagrams and simple event viewer dashboards, do not often seem very intuitive and useful for their users. Thus, due to the inadequate fusion of both physical and cyber situations in a single decision-making visualization space, operators are not able to completely gain the necessary situational awareness that requires such hybrid environment.

The main motivation of this work is, therefore, beyond industrial-oriented sectors, the design and implementation of a whole and advanced cross-platform solution which provides, through combining human-based and cyber threat intelligence capabilities and taking advantage of advanced visualization techniques, the appropriate mixed (hybrid) situational awareness to achieve an agile and effective protection of any kind of CI.

3. System Architecture

This work proposes a new approach for the CI adequate protection in the current cyber-physical context through HYBINT, a hybrid intelligence solution (Figure 3) which integrates, through current analytical tools and advanced representation techniques [44, 45], the intelligence information of the hybrid environment in a single decision-making space.

HYBINT system is based on a client-server architecture in order to support scalability, availability, and security as main requirements of any kind of CI. The core of the platform, which has been designed in a flexible and uncoupled way to easily implement new functionalities and adapt to the particular specifications of a concrete infrastructure, consists of three main modules that have been independently developed following the three-tier application concept (Figure 4):

- (i) **Data Gathering Module (DGM):** it collects significant data from CI-deployed heterogeneous physical and cyber data sources (e.g., environmental sensors, network packets analysers, video cameras signal, human reports) and stores them in the system database.
- (ii) **Data Analysis Module (DAM):** it brings to the user a set of advanced analysis to produce physical, cyber, or mixed intelligence information (e.g., cyber threats evaluation, facilities classification by criticality, patterns detection from social interactions) through processing the stored raw data.
- (iii) **Data Visualization Module (DVM):** it provides a real-time hybrid situational awareness (HSA) of the

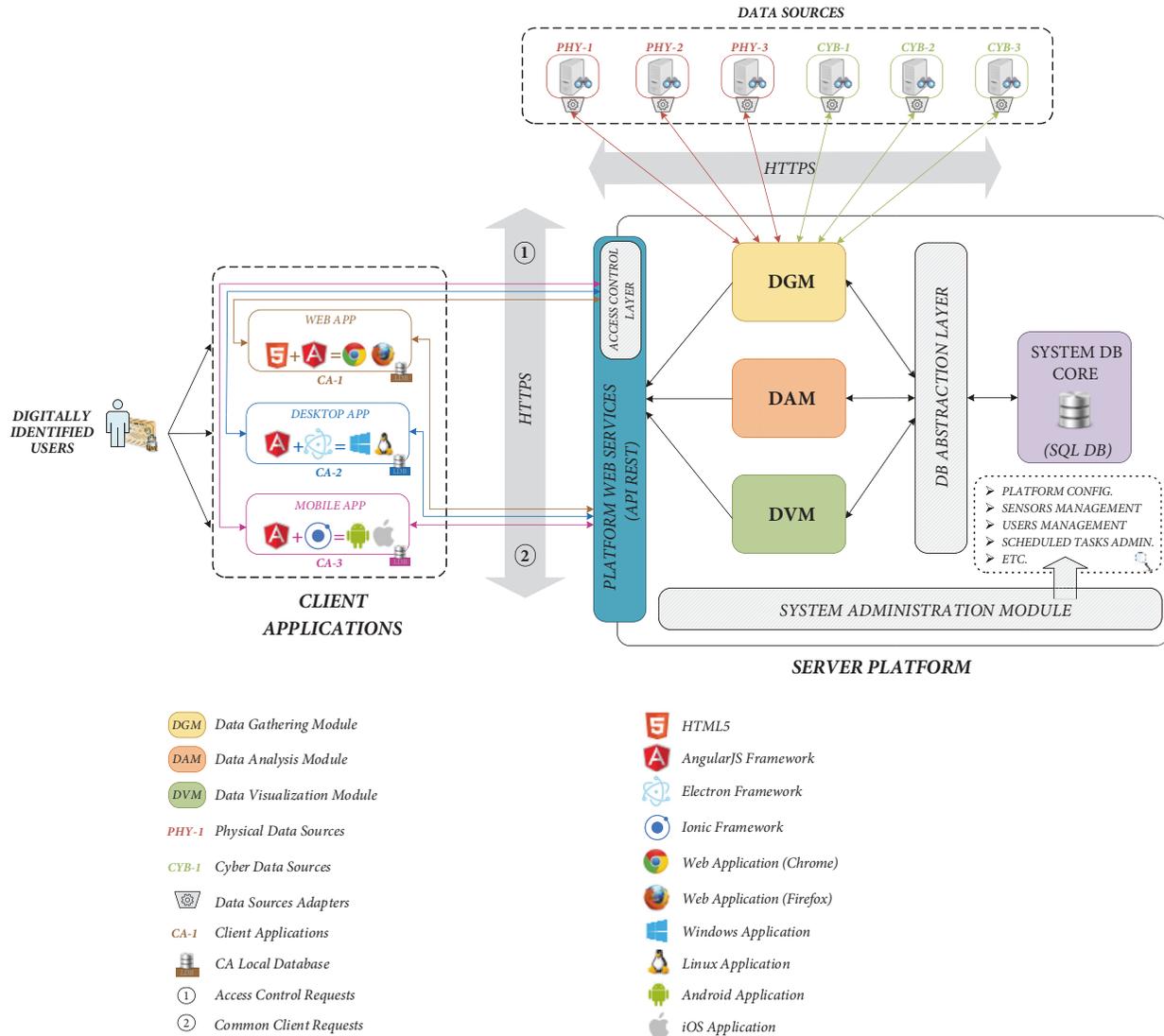


FIGURE 4: HYBINT architecture.

physical and cyber environments by a combined and georeferenced representation of the obtained intelligence information (e.g., physical and cyber assets interdependencies, social network of suspicious CI personnel, cascading effects of combined attacks).

The HYBINT platform is a Microsoft .NET Framework-based solution [46] developed in C# and deployed in a Windows Server 2016 virtual machine.

MySQL Server [47] is initially used, as SQL relational database, to store the gathered data. However, a database abstraction layer facilitates the integration and use of any other database engine by decoupling it from the system core. In any case, the database in use must be in compliance with Big Data; that is, it has to support massive data storage. From the System Administration Module (SAM), admin users are able to manage the whole HYBINT system: server platform configuration, data sources and user’s management, scheduled tasks administration, etc.

The platform’s web services, which are based on Representational State Transfer (REST) [48] architecture, are exposed through Microsoft’s Internet Information Services (IIS) [49] web server. These are accessible for any registered user through any of the three different client applications that have been implemented: web, desktop, and mobile application. The Human-Machine Interface (HMI) of these is a web-based environment developed in AngularJS [50] framework over HTML5. Desktop and mobile native applications have been, respectively, built, from the same codebase, through Electron [51] and Ionic’s [52] JavaScript frameworks. Web application can be accessed through any HTML5-supported web browser (Google Chrome, Mozilla Firefox, etc.), desktop application can be installed in Windows OS machines, and mobile application can be deployed in common Android devices.

Hypertext Transfer Protocol Secure (HTTPS) is used as secure communications protocol between client applications and the server platform in order to ensure the integrity and

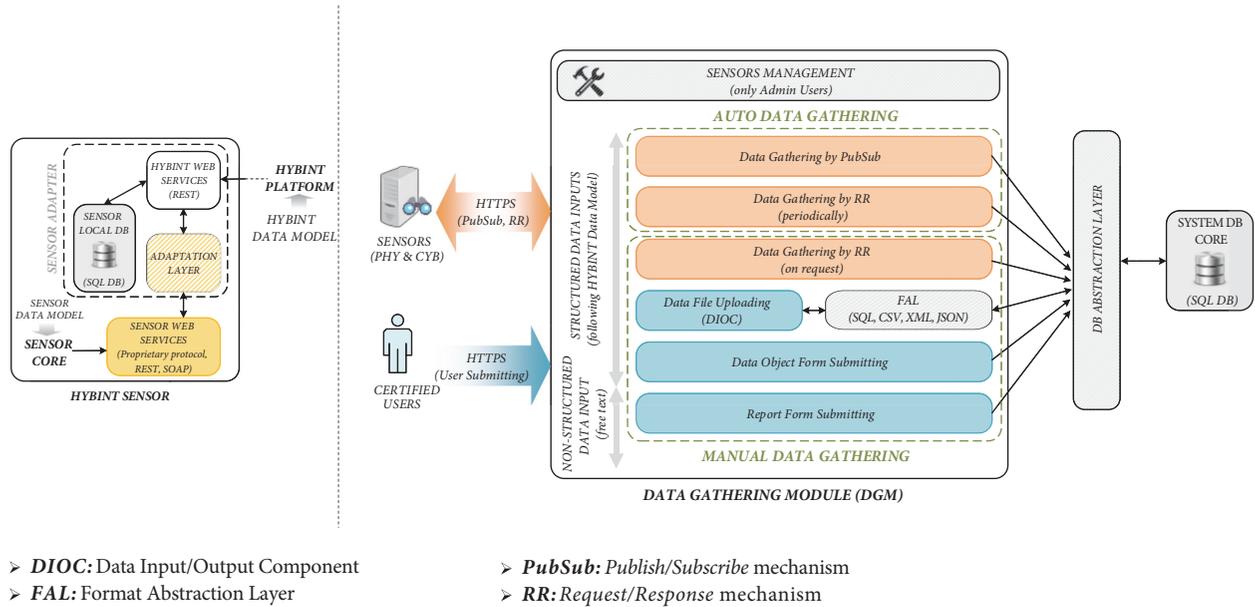


FIGURE 5: Sensor adapter (left) and Data Gathering Module (right).

privacy of the exchanged data. In addition, the inclusion of public key certificates on these clients also guarantees the digital identity of the CI personnel registered in HYBINT and their authentication in the system. An access control layer is responsible for both validating the user certificates and managing the available platform capabilities depending on the personnel category and privileges (e.g., common personnel, area supervisors, facilities surveillance staff, IT support, CI security analysts).

Moreover, an offline mode of the system has been developed as a solution to provide minimal capabilities even when no network connections, neither LAN nor WAN (Internet), are available in the client side.

3.1. Data Gathering Module. This module is responsible for collecting significant intelligence data from both physical and cyber data sources and storing it in the HYBINT database. These can be either CI members registered in the system such as technical personnel and security staff or sensors (physical or cyber) in charge of reporting any deviant behaviour or incident occurred in its facilities [53].

While CI physical sensors mainly refer to information systems (GPS, access control technologies, video surveillance systems, etc.) which provide physical data related to physical assets (people, vehicles, buildings, etc.), CI cyber sensors refer to Security Information and Event Management (SIEM) tools (OSSIM [54], MISP [55], RTIR [56], etc.) which provide cyber data related to cyber assets (servers, firewalls, routers, etc.). CI personnel with access to HYBINT can supply, for their part, human-based intelligence data referring to physical, cyber or both domains such as a non-recognized people, suspicious activities, and abnormal behaviours.

As shown in Figure 5, data gathering can be achieved both automatically and manually. Besides, the DGM is as flexible as the platform supports both structured (which follow

the HYBINT data model) and non-structured (like free text) data inputs. On the one hand, system users contribute to data collection, through CI-related human intelligence, either by submitting both data object forms and report forms or by uploading formatted data files through the Data Input/Output Component (DIOC). In this case, a format abstraction layer has been implemented to decouple the supported file types (SQL, CSV, XML, or JSON) from the database functioning. From DIOC, users can also export the current content of HYBINT database to any of these file formats.

On the other hand, the data collection from sensors deployed in the infrastructure facilities is achieved by using HTTPS as communications protocol between these and the DGM and by selecting Publish-Subscribe (PubSub) or Request-Response (RR) as data gathering mechanism. In case of RR, queries to these CI sensors can be configured to be executed periodically or just on user request. Only admin users are enabled for the sensors management either configuring the parameters (IP address, MAC address, type, geolocation, etc.) of the existing ones or registering new others.

Due to the heterogeneity of the sensors in each type of infrastructure, specific adapters have been developed to transform, through an adaptation layer, the original web services of each sensor (based on proprietary protocols, REST, SOAP, etc.) into REST-based web services that follow the HYBINT data model. Moreover, when no connectivity is available, these adapters can locally store data in an integrated database until they can be transmitted (Figure 5).

3.2. Data Analysis Module. An understanding of the stored data is needed in order to obtain an advanced HSA: that task is the responsibility of the DAM. To this end, the system interface binds to the CI security administrators a

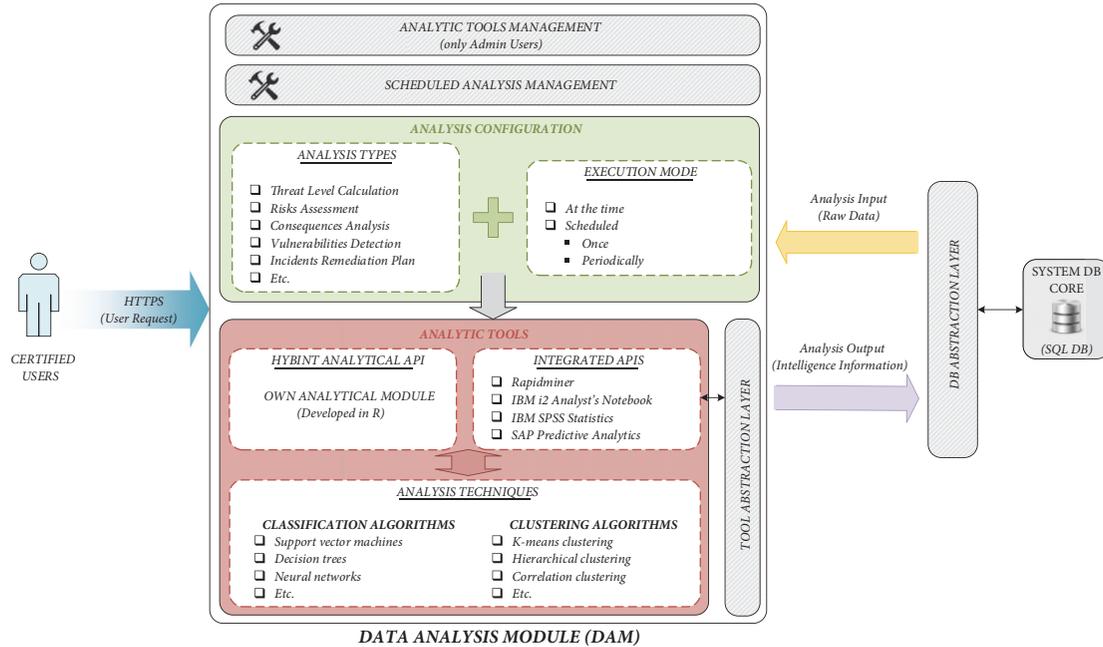


FIGURE 6: Data Analysis Module.

set of preconfigured analysis (infrastructure risks assessment, security vulnerabilities detection, criticality-based facilities classification, patterns detection from social interactions, etc.) to generate the desired physical, cyber, or hybrid intelligence information as a result of processing the appropriate raw data. For any new analytical task, the platform also allows to set either at the time or scheduled (running once or periodically) as its execution mode (Figure 6).

All the proposed intelligence analyses are linked to an analytic tool in which they will be executed. These implement data mining methods [57], through statistical analysis techniques, to find out meaningful patterns and discover unknown knowledge from wide datasets (that is, Big Data) that helps to better decision-making [58]. Besides, non-structured data (as the ones provided by CI personnel report forms) are also supported: in that particular case, text mining methods are used.

The analysis tools mainly refer to third-party analytical tools (Rapidminer [59], IBM i2 Analyst's Notebook [60], IBM SPSS Statistics [61], and SAP Predictive Analytics [62]) that have been integrated in HYBINT. Nevertheless, an own analytical module with several statistical functions developed in R [63] has been additionally implemented in the DAM, as a complement to the integrated tools capabilities, in order to offer more flexible and specific data analysis. All of the third-party integrated tools have been physically deployed in the HYBINT core machine. Depending on each case, the access to these is carried out either through their programmable API or through REST calls to their exposed web services. Only admin users are responsible for assigning, as far as possible, each preconfigured analysis to the most appropriate tool for it or discarding the use of any of them.

The computing techniques performed by the analytic tools consist of algorithms based on machine learning [64].

In other words, they carry out predictive analytics to estimate future behaviours by learning from trained data. These algorithms can be either supervised, as classification algorithms (support vector machines, decision trees, neural networks, etc.) [65] or unsupervised, as clustering algorithms (k-means clustering, hierarchical clustering, correlation clustering, etc.).

When an analysis task must be executed, the DAM queries HYBINT database to get the appropriate stored data depending on the analysis type and brings them as input data to the corresponding analytic tool through their respective Application Programming Interfaces (API). The intelligence information provided once data processing finishes is registered in the database in order to be available for its representation in the DVM. A tool abstraction layer is used for decoupling the analytic tool's logic of the database access. Every CI member with access to DAM is able to modify at any moment the configuration of his own scheduled analytical tasks.

3.3. Data Visualization Module. The aim of the DVM is to provide the necessary HSA to enhance decision-making through a mixed and georeferenced representation of both CI's physical and cyber world current situation in a unique visualization space. As shown in Figure 7, his HMI is therefore an interactive visualization environment, mostly composed of a web-based Geographic Information System (GIS) and several visualization panels, where physical, cyber, or hybrid intelligence information is shown in real time.

The key feature of this module is the flexibility to represent any kind of system-registered data following the most adequate type of visualization. To this end, through a representation management interface, HYBINT users can define any query related to analysis results (e.g., cascading

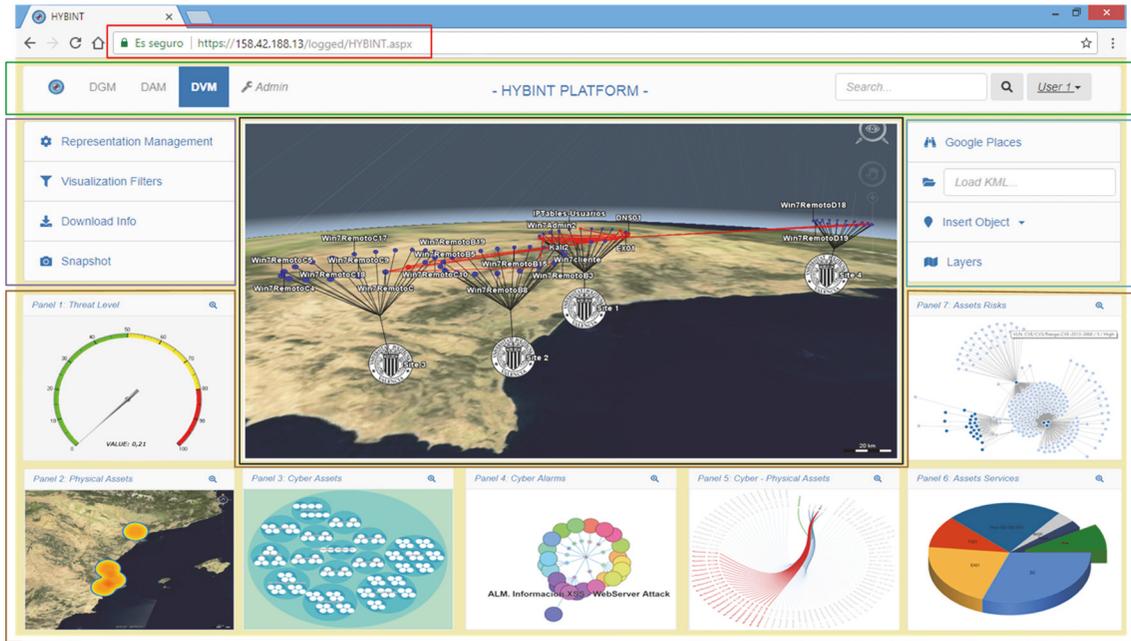


FIGURE 7: HMI of Data Visualization Module.

effects of combined attacks, social network of suspicious personnel, cyber and physical assets interdependencies) but also to any other stored raw data. Either the main GIS map or any of the visualization panels has to be selected as destination for the generated request's representation. Moreover, previous requests can be reloaded at any time if they were saved in the system database when they were created.

The 3D graph representation, where the current state of the infrastructure's physical and cyber entities is shown in a hybrid and multidimensional view, is the most relevant for the main GIS map in order to get a real-time geolocated visualization of the whole CI's current situation, that is, HSA (Figure 7). However, heatmap and KML-based representations are also other representation types supported by the main map. All of these capabilities have been developed in two different web-based geospatial tools that have been built-in in HYBINT: Luciad [66] (through implemented functions following its proprietary API) and Cesium [67] (through REST calls to its exposed web services). Moreover, a GIS abstraction layer allows easy switching between these maps providers and makes easier the future integration of new ones in the system.

Regarding the visualization panels, in addition to the geolocated representations, the representation manager offers a wide set of data chart visualizations summarized in two main categories: 2D/3D static charts (bar chart, area chart, pie

chart, etc.) and 3D interactive diagrams (Hebbian dynamics, bubble chart, force directed graph, etc.), which are both provided through JavaScript open source libraries as Chart.js [68] and Data-Driven Documents D3.js [69].

Besides, additional features have been developed, related to both GIS map and visualization panels, in order to enrich the DVM's capabilities. From the system interface, CI staff with access to DVM can query to Google Places database through his proprietary API, load local KML files into the map, insert new georeferenced objects on it, or manage the existing map layers. Moreover, visualization filters can be set, current view's snapshot are downloadable as image file, and any kind of intelligence information can also be locally downloaded. In this case, depending on the desired type of information, this can be either an intelligence report (text file), a data chart (image file), or a geolocated information (KML file).

3.4. System Availability and Secure Access. HYBINT is a cross-platform solution that ensures the scalability requirement since it is accessible from a wide range of network-connected devices (laptops, smartphones, tablets, etc.) through the system web services. However, access to the platform may not be guaranteed depending on the connectivity in the client side. To enhance the availability requirement, when network connections are unavailable on the HYBINT's client applications, these automatically switch

TABLE 1: System access levels and associated functionalities.

Level	Accessible modules	Main capabilities
1	DGM	Forms submitting and data files uploading
2	DGM	Level 1 + Own data management
3	DGM	Level 1 + All data management and sensors data gathering
4	DGM + DAM	Level 2 + Own data analysis
5	DGM + DAM	Level 3 + All data analysis and analysis scheduling
6	DGM + DAM + DVM	Level 4 + Representation of own analysis results
7	DGM + DAM + DVM	Level 5 + Representation of all analysis results
8	DGM + DAM + DVM	Level 7 + Admin capabilities

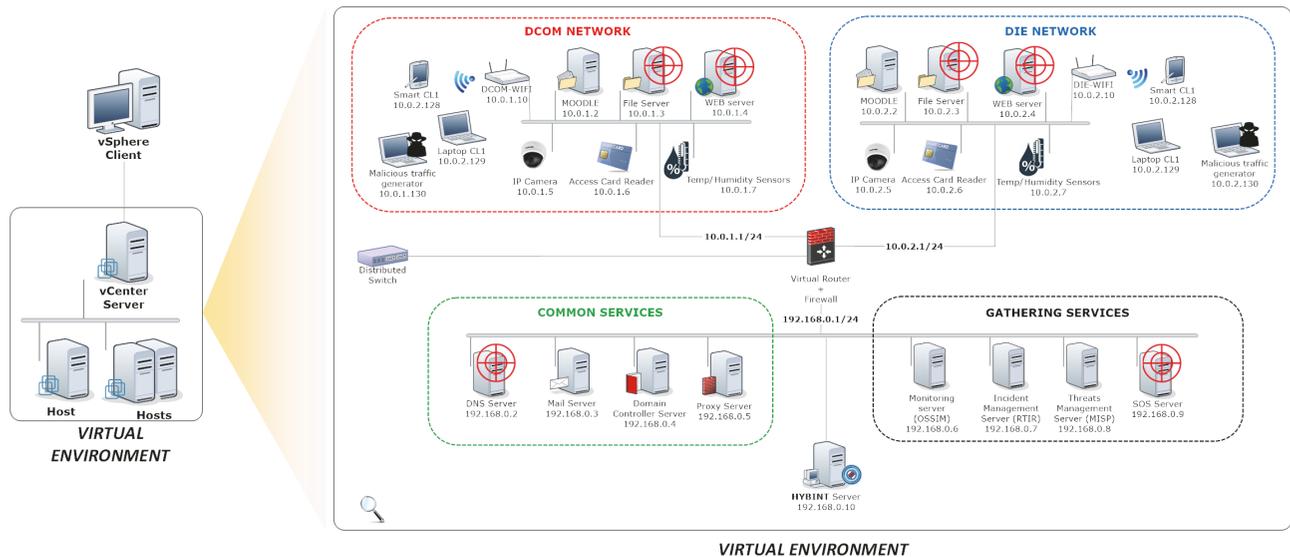


FIGURE 8: HYBINT validation scenario.

to an offline mode where only capabilities related to data local loading are enabled. In a transparent manner for user, all the data collected from loaded files and filled forms (data object forms or report forms) are locally stored in minimal database integrated on the client applications (Figure 4) until a local service automatically reestablishes HYBINT's direct mode when connectivity is back.

Regarding the security requirement, beyond the use of secure communications protocols and public key certificates, system access levels have been defined to restrict the available functionalities according to the position of the personnel in the CI staff. Table 1 summarizes the default-defined HYBINT's access profiles and their main capabilities.

By default, the lowest access level is initially assigned to any new registered user. Only admin users are allowed to provide privileges to the users of their CI as well as managing the platform access profiles by modifying them or creating new ones.

4. Validation and Results

HYBINT system has been validated through functionality tests to an implemented prototype based on the architecture

described in this paper. These tests were carried out within a virtualized platform [70] where were installed, on the one hand, the ESXi software as Hypervisor in a cluster of servers HP Proliant ML110 Gen9 and, on the other hand, a vCenter server in a main server Fujitsu PRIMERGY TX1310 M1 which allows the automatic configuration and control of the whole machines and network elements required to cover the different use cases of the current scenario.

As shown in Figure 8, the network configuration of two different Universitat Politècnica de València's departments has been simulated in the deployed virtual environment. On the upper left part, the red box corresponds with the subnetwork 10.0.1/24, where both physical and cyber assets related to the Communications Department's (DCOM) users are connected. On the upper right part, the blue box represents the subnetwork 10.0.2/24 where both physical and cyber assets related to the Electronic Engineering Department's (DIE) users are connected. Additionally a third subnetwork 192.168.0.1/24 has been configured and split into 2 different boxes: at the bottom left part, the green box contains all the common and management services (DNS, E-Mail, Domain Controller, and Proxy) for both departments; at the bottom right part, the black box contains all the virtual machines where several SIEMs have been deployed (OSSIM, MISP,

```

root@kali: ~
File Edit View Search Terminal Help
VULNERABLE:
Remote Code Execution vulnerability in Microsoft SMBv1 servers (ms17-010)
State: VULNERABLE
IDs: CVE:CVE-2017-0143
Risk factor: HIGH
A critical remote code execution vulnerability exists in Microsoft
SMBv1
servers (ms17-010).
Disclosure date: 2017-03-14
References:
https://technet.microsoft.com/en-us/library/security/ms17-010.aspx
https://blogs.technet.microsoft.com/msrc/2017/05/12/customer-guidance-for-wannacrypt-attacks/
root@kali: ~
File Edit View Search Terminal Help
root@kali:~# ./ms17-010-m4ss-sc4nn3r.py 10.0.1.3
[+] 10.0.1.3 is likely VULNERABLE to MS17-010 (Windows 7 Enterprise 7601 Service Pack 1)
root@kali:~#

root@kali: ~
File Edit View Search Terminal Help
Tired of typing 'set RHOSTS'? Click & pwn with Metasploit Pro
Learn more on http://rapid7.com/metasploit
[ metasploit v4.14.10-dev ]
+ -- --[ 1641 exploits - 945 auxiliary - 289 post ]
+ -- --[ 472 payloads - 40 encoders - 9 nops ]
+ -- --[ Free Metasploit Pro trial: http://r-7.co/trypsp ]

msf > use exploit/windows/smb/eternalblue_doublepulsar
msf exploit(eternalblue_doublepulsar) > show targets

Exploit targets:

Id  Name
--  ---
0   Automatic
1   Windows XP (all services pack) (x86) (x64)
2   Windows Server 2003 SP0 (x86)
3   Windows Server 2003 SP1/SP2 (x86)
4   Windows Server 2003 (x64)
5   Windows Vista (x86)
6   Windows Vista (x64)
7   Windows Server 2008 (x86)
8   Windows Server 2008 R2 (x86) (x64)
9   Windows 7 (all services pack) (x86) (x64)

msf exploit(eternalblue_doublepulsar) > set TARGET 9
TARGET => 9
msf exploit(eternalblue_doublepulsar) > set PAYLOAD windows/meterpreter/reverse_tcp
PAYLOAD => windows/meterpreter/reverse_tcp
msf exploit(eternalblue_doublepulsar) > set PROCESSINJECT explorer.exe
PROCESSINJECT => explorer.exe
msf exploit(eternalblue_doublepulsar) > set LHOST 10.0.1.3
LHOST => 10.0.1.3
msf exploit(eternalblue_doublepulsar) > set LPORT 4444
LPORT => 4444
msf exploit(eternalblue_doublepulsar) > set RHOST 10.0.1.3
RHOST => 10.0.1.3
msf exploit(eternalblue_doublepulsar) > exploit

```

FIGURE 9: Cyberattacks execution example.

and RTIR) responsible to gather significant data related to the cyber environment, a Sensor Observation Service (SOS) server which exposes the necessary services to feed and retrieve information from all the physical assets simulated in the scenario [71], and a common computer in order to allow, through accessing to HYBINT platform, the system administrator to perform manual insertions (intelligence reports, data object submissions, etc.). These three subnetworks are interconnected with each other through a virtual router based on a Linux distribution (VyOS developed by Vyatta) able to provide software-based network routing, firewall, and VPN functionalities [72].

This virtual router has been configured with the necessary rules to routing and firewalling the network connections in order that the different machines could be reachable from any other in case it is needed.

Once the deployment of all the virtual machines, network configurations, and virtual sensors had been performed, HYBINT executed a scheduled vulnerability scan using OSSIM over the proposed scenario identifying potential open doors to both physical and cyber assets. According to the vulnerabilities found, several scripts were scheduled to perform different attack types [73] over the course of a week, during the 24 hours of day, using the Kali Linux distribution [74], which contains a set of preinstalled tools to execute diverse types of penetration tests and security analysis.

These simulated cyberattacks mainly consisted of a set of intrusion actions performed, through multiple scripts and exploits, against different virtualized assets in the form of denial-of-service attacks (DoS), brute-force blocking attacks, and spoofing attacks, among others. These performed intrusions caused multiple incidents and failures in DNS servers,

file servers, and web servers due to open ports, wrong firewall configurations, and others. As an example, Figure 9 shows some attacks performed against the DCOM Network's File Server (IP: 10.0.1.3), where several exploits have been executed to access without authorization to a targeted asset in which a vulnerability has been previously identified by Nmap tool of Kali Linux distribution.

Regarding the physical environment, attacks to physical assets were difficult to be performed due to the virtualization of the corresponding sensors. Therefore, a set of scripts, which was designed to simulate multiple attacks that could be performed against physical assets in the context of CI's protection (as intrusive access, sensors sabotage, surveillance cameras signal disruption, or sensors disconnection among others), was also executed.

As a consequence of these simulations, different anomalies and failures like unauthorized access, anomalous sensors data, or video surveillance camera signal lost were either detected by deployed physical and cyber sensors or reported via Human Intelligence Reports (HIR). HYBINT system was able to collect and combine data from all its deployed data sources through the DGM, to use the DAM in order to extract relevant patterns related to physical and cyber worlds as well as classifying incidents through analytic techniques such as Support Vector Machine, Decision Tree, or Random Forest, and finally to represent in the DVM the generated physical, cyber, and hybrid intelligence information in a useful and enhanced decision-making space.

Figure 10, which summarizes the results of the system evaluation, shows both the cyber and physical vulnerabilities found as a result of an initial HYBINT analysis in this virtualized hybrid environment, the amount and type of

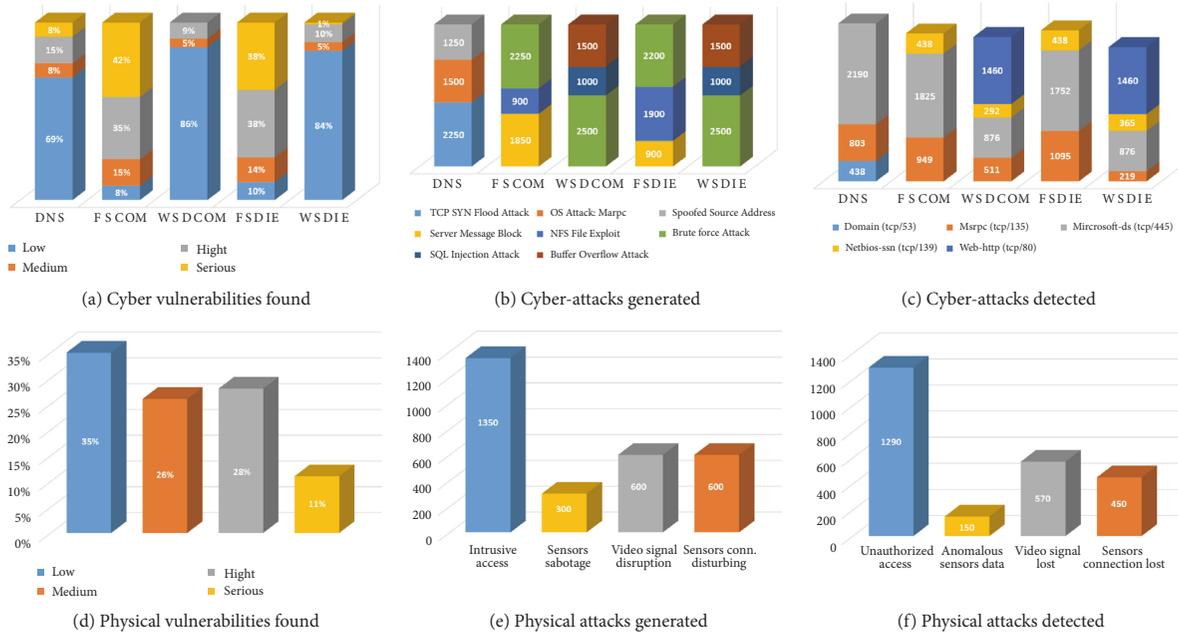


FIGURE 10: HYBINT evaluation results.

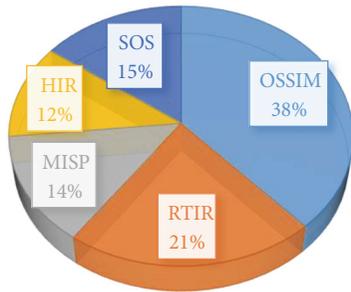


FIGURE 11: Individual ratio of reported attacks.

attacks performed against the simulated cyber and physical assets, and finally the amount and type of attacks identified thanks to the HYBINT analysis tools. The obtained results from the functionality tests performed in the current scenario were highly satisfactory since HYBINT was able to identify, to a greater or lesser extent, the majority of the attacks performed in both cyber and physical domains.

Conceived as a holistic solution, HYBINT advantages arise from a set of modules and tools that cannot be assessed in an isolated way. In this sense, unlike current CI protection approaches, by integrating multiple cyber-physical data sources with HUMINT capabilities and advanced Big Data analytic methods, this cyber-human intelligence system is able to provide a more complete knowledge of the CI's hybrid environment that leads to an enhancement of the decision-making process. Indeed, as shown in Figure 11, while each particular data source is able to report a limited ratio of the performed events and attacks, HYBINT provides a more accurate operational picture that overtakes the individual performance of the current CI protection methods

by combining, in a single common view, the intelligence information delivered by the different SIEMs connected to the platform, the physical sensors information, and also the Human Intelligence Reports (HIR).

Moreover, the hybrid situational awareness concept and therefore the current proposal are also validated since trained security analysts, through visualizing in a more useful and intuitively way the up-to-date cyber-physical situation of the whole scenario thanks to the use of advanced representation techniques, are now able to achieve a more effective, reactive, and faster protection of their own CI in the cyber-physical context.

5. Final Notes and Conclusions

Nowadays, organizations must address in conjunction the threats from both the physical and the cyber world for their efficient defence. To this end, a hybrid intelligence system that provides a real-time enhanced situational awareness compliant with the requirements of this cyber-physical environment is the key to provide the adequate decision-making support for the effective protection of critical infrastructures in such scenario. This work introduces HYBINT, a real and complete cross-platform solution of an advanced intelligence system; describes its main capabilities in detail; and exposes the results obtained of its evaluation in a simulated scenario.

The core of HYBINT system resides in its three main capabilities: data gathering, data analysis, and data visualization. In this way, significant data collected from heterogeneous cyber and physical data sources are processed through latest analytical procedures in order to produce real-time intelligence information which is shown, through advanced representation techniques, in a single visualization space. HYBINT concept has been validated through an

implemented prototype based on the proposal described in this article, which was tested in a simulated scenario, deployed in a virtual environment, through both physical and cyberattacks. The evaluation results demonstrated that HYBINT was able to offer an advanced situational awareness of the whole hybrid environment by detecting the existing vulnerabilities in two subnetworks which corresponds to different departments of the Universitat Politècnica de València as well as identifying a high percentage of performed attacks against different cyber and physical assets deployed in the scenario.

Currently an implemented prototype is deployed as a pilot test in a small subnetwork of the Communications Department analysing more than fifty cyber and physical assets including servers, personal computers, arduinos, sensors, and others. Since its deployment, the majority of the discovered vulnerabilities were solved thanks to the mitigation actions suggested by the system. Moreover, several people intrusions, either real or generated via supervised attacks, were also detected and remediated thanks to the cyber-physical situational awareness provided by HYBINT.

Regarding future works, a new alternative is being explored regarding data storage thanks to the increasing number of cloud and Big Data solutions [75]. To this end, a review of Big Data solutions is required to identify the best way to make HYBINT consistent with both cloud and NoSQL's alternatives [76, 77]. Moreover, the use of distributed solutions is a must in order to achieve a useful solution able to work in more complex and heterogeneous environments. Current technologies as Cassandra [78] or MongoDB [79] seem compliant with HYBINT system requirements, providing distributed data sharing and capacity to store huge amount of data, as it will be expected in real environments. Additionally, by deploying the platform in multiple nodes as a distributed infrastructure, it would be possible to replicate the HYBINT's critical services in order to provide, in conjunction with the integration of NoSQL technologies [80], server side's high availability and fault tolerance in case of network problems or high number of requests. To address possible scalability problems when accessing to DAM's 3rd party systems services with private license, the proposal could be improved by including messages queue service to manage the number of requests by each proprietary license.

Finally, integrating some of the new data analysis and intelligence tools that are continuously appearing, an even more complete and updated version of HYBINT solution can be obtained which will be able to face future requirements.

Data Availability

The underlying data related to this submission have been generated and collected by our own means through the computing equipment of our research group's laboratory (at the Universitat Politècnica de València) and could be available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] B. Genge, C. Siaterlis, and M. Hohenadel, "Impact of network infrastructure parameters to the effectiveness of cyber attacks against Industrial Control Systems," *International Journal of Computers, Communications & Control*, vol. 7, no. 4, pp. 674–687, 2012.
- [2] A. Colombo, T. Bangemann, S. Karnouskos, J. Delsing, and P. Stluka, *Industrial Cloud-Based Cyber-Physical Systems: The IMC-AESOP Approach*, Springer, 2014.
- [3] A. Ferreira, L. Weigang, J. A. Fregnani, and I. Romani, "Big data management and processing in the context of the system wide information management," in *Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1066–1073, October 2017.
- [4] G. Kumar and K. Kumar, "The use of artificial-intelligence-based ensembles for intrusion detection: a review," *Applied Computational Intelligence and Soft Computing*, vol. 2012, Article ID 850160, 20 pages, 2012.
- [5] M. A. Ferrag, L. A. Maglaras, H. Janicke, J. Jiang, and L. Shu, "Authentication protocols for internet of things: a comprehensive survey," *Security and Communication Networks*, vol. 2017, Article ID 6562953, 41 pages, 2017.
- [6] S. Zhao, Y. Zhang, B. Cheng, and J.-L. Chen, "A Feedback-corrected Collaborative Filtering for Personalized Real-world Service Recommendation," *International Journal of Computers Communications & Control (IJCCC)*, vol. 9, no. 3, pp. 356–369, 2014.
- [7] M. Conti, S. K. Das, C. Bisdikian et al., "Looking ahead in pervasive computing: challenges and opportunities in the era of cyberphysical convergence," *Pervasive and Mobile Computing*, vol. 8, no. 1, pp. 2–21, 2012.
- [8] H. Zhuge, "Semantic linking through spaces for cyber-physical-socio intelligence: A methodology," *Artificial Intelligence*, vol. 175, no. 5–6, pp. 988–1019, 2011.
- [9] R. J. Robles, M.-K. Choi, E.-S. Cho, S.-S. Kim, and G.-C. Park, "Common Threats and Vulnerabilities of Critical Infrastructures," *International Journal of Control and Automation (IJCA)*, vol. 1, no. 1, pp. 17–22, 2008.
- [10] K. Coffey, R. Smith, L. Maglaras, and H. Janicke, "Vulnerability Analysis of Network Scanning on SCADA Systems," *Security and Communication Networks*, vol. 2018, Article ID 3794603, 21 pages, 2018.
- [11] S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly, "Identifying, understanding, and analyzing critical infrastructure interdependencies," *IEEE Control Systems Magazine*, vol. 21, no. 6, pp. 11–25, 2001.
- [12] Spanish National Cryptologic Centre (CCN-CERT), <https://www.ccn-cert.cni.es/en>.
- [13] Spanish National Cryptologic Centre (CCN-CERT), *Cyber Threats and Trends*, 2016.
- [14] Spanish National Cryptologic Centre (CCN-CERT), *Cyber Threats and Trends*, 2017.
- [15] S. Mohurle and M. Patil, "A brief study of Wannacry Threat: Ransomware Attack 2017," *International Journal of Advanced Research in Computer Science (IJARCS)*, vol. 8, no. 5, pp. 1938–1940, 2017.
- [16] J. Healey, "Winning and losing in cyberspace," in *Proceedings of the 8th International Conference on Cyber Conflict, CyCon 2016*, pp. 37–49, June 2016.
- [17] D. DiMase, Z. A. Collier, K. Heffner, and I. Linkov, "Systems engineering framework for cyber physical security and

- resilience,” *Environment Systems and Decisions*, vol. 35, no. 2, pp. 291–300, 2015.
- [18] Y. Deng, L. Song, Z. Zhou, and P. Liu, “Complexity and Vulnerability Analysis of Critical Infrastructures: A Methodological Approach,” *Mathematical Problems in Engineering*, vol. 2017, Article ID 8673143, 12 pages, 2017.
- [19] European Union, “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,” *Official Journal of the European Union*, 2008.
- [20] C. Alcaraz and S. Zeadally, “Critical infrastructure protection: Requirements and challenges for the 21st century,” *International Journal of Critical Infrastructure Protection*, vol. 8, pp. 53–66, 2015.
- [21] P. Salmon, N. Stanton, G. Walker, and D. Green, “Situation awareness measurement: A review of applicability for C4i environments,” *Applied Ergonomics*, vol. 37, no. 2, pp. 225–238, 2006.
- [22] J. P. Shim, M. Warkentin, J. F. Courtney, D. J. Power, R. Sharda, and C. Carlsson, “Past, present, and future of decision support technology,” *Decision Support Systems*, vol. 33, no. 2, pp. 111–126, 2002.
- [23] NATO Standardization Office, *AAP-06 NATO Glossary of Terms and Definitions*, 2017.
- [24] V. Mavroeidis and S. Bromander, “Cyber Threat Intelligence Model: An Evaluation of Taxonomies, Sharing Standards, and Ontologies within Cyber Threat Intelligence,” in *Proceedings of the 2017 European Intelligence and Security Informatics Conference (EISIC)*, pp. 91–98, Athens, September 2017.
- [25] C. Alcaraz and J. Lopez, “Wide-area situational awareness for critical infrastructure protection,” *The Computer Journal*, vol. 46, no. 4, pp. 30–37, 2013.
- [26] I. Lendak, S. Vukmirovic, E. Varga, A. Erdeljan, K. Nenadic, and N. Ivancevic, “Client side internet technologies in critical infrastructure systems,” *International Journal of Computers, Communications & Control*, vol. 7, no. 5, pp. 879–891, 2012.
- [27] B. Cheng, J. Zhang, G. P. Hancke, S. Karnouskos, and A. W. Colombo, “Industrial Cyberphysical Systems: Realizing Cloud-Based Big Data Infrastructures,” *IEEE Industrial Electronics Magazine*, vol. 12, no. 1, pp. 25–35, 2018.
- [28] P. Ribino, A. Augello, G. Lo Re, and S. Gaglio, “A knowledge management and decision support model for enterprises,” *Advances in Decision Sciences*, vol. 2011, Article ID 425820, 16 pages, 2011.
- [29] Carmenta: Superior Situational Awareness: Carmenta Coord-Com., <https://www.carmenta.com/en/products/carmenta-coord-com>.
- [30] Atos, *Atos Global Emergency Management (GEMMA)*, <https://atos.net/en/products/defense-mission-critical/homeland-security/emergency-management>.
- [31] TR Sistemas, GESTOP, <http://www.trsisistemas.com/productos.php>.
- [32] M. Athans, “Command and Control (C2) Theory: A Challenge to Control Science,” *IEEE Transactions on Automatic Control*, vol. 32, no. 4, pp. 286–293, 1987.
- [33] F. Antunes and J. P. Costa, “Integrating decision support and social networks,” *Advances in Human Computer Interaction*, vol. 2012, Article ID 574276, 10 pages, 2012.
- [34] Palo Alto Networks, <https://www.paloaltonetworks.com/products>.
- [35] IBM Security, <https://www.ibm.com/security/solutions>.
- [36] Thales Group, <https://www.thalesgroup.com/en/global/activities/security/critical-information-systems-and-cybersecurity>.
- [37] NEC Cyber Security Solutions, <https://www.nec.com/en/global/solutions/cybersecurity/solutions/index.html>.
- [38] U. Franke and J. Brynielsson, “Cyber situational awareness - A systematic review of the literature,” *Computers & Security*, vol. 46, pp. 18–31, 2014.
- [39] G. P. Tadda and J. S. Salerno, “Overview of cyber situation awareness,” *Advances in Information Security*, vol. 46, pp. 15–35, 2010.
- [40] Assac Networks, Graphene, <https://assacnetworks.com/scada-cyber-security>.
- [41] Tibbo Systems: AggreGate SCADA/HMI, <http://aggregate.tibbo.com/solutions/scada-hmi.html>.
- [42] R. Johnson, “Survey of SCADA security challenges and potential attack vectors,” in *Proceedings of the 2010 International Conference for Internet Technology and Secured Transactions, ICITST 2010*, November 2010.
- [43] S. Karnouskos and A. W. Colombo, “Architecting the next generation of service-based SCADA/DCS system of systems,” in *Proceedings of the IECON 2011 - 37th Annual Conference of IEEE Industrial Electronics*, pp. 1–6, November 2011.
- [44] G.-D. Sun, Y.-C. Wu, R.-H. Liang, and S.-X. Liu, “A survey of visual analytics techniques and applications: State-of-the-art research and future challenges,” *Journal of Computer Science and Technology*, vol. 28, no. 5, pp. 852–867, 2013.
- [45] E. Bertini and D. Lalanne, “Investigating and reflecting on the integration of automatic data analysis and visualization in knowledge discovery,” *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 2, p. 9, 2009.
- [46] Microsoft, Microsoft .NET Framework, <https://www.microsoft.com/.NET>.
- [47] Oracle MySQL, MySQL Server Database, <https://www.mysql.com/products/enterprise/database>.
- [48] B. Costa, P. F. Pires, F. C. Delicato, and P. Merson, “Evaluating a Representational State Transfer (REST) architecture: What is the impact of REST in my architecture?” in *Proceedings of the 11th Working IEEE/IFIP Conference on Software Architecture, WICSA 2014*, pp. 105–114, April 2014.
- [49] Microsoft, Microsoft Internet Information Services (IIS), <https://www.iis.net>.
- [50] Angular.js., <https://angularjs.org>.
- [51] Electron.js, <https://electronjs.org>.
- [52] Ionic, <https://ionicframework.com>.
- [53] M. Iturbe, I. Garitano, U. Zurutuza, and R. Uribeetxeberria, “Towards Large-Scale, Heterogeneous Anomaly Detection Systems in Industrial Networks: A Survey of Current Trends,” *Security and Communication Networks*, vol. 2017, Article ID 9150965, 17 pages, 2017.
- [54] AlienVault: Open Source Security Information Management (OSSIM), <https://www.alienvault.com/products/ossim>.
- [55] Malware Information Sharing Platform and Threat Sharing (MISP), <http://www.misp-project.org>.
- [56] Best Practical Solutions LLC; Request Tracker for Incident Response (RTIR), <https://bestpractical.com/rtir>.
- [57] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*, Elsevier, 3rd edition, 2011.
- [58] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, John Wiley & Sons, 2nd edition, 2012.
- [59] Rapidminer: Rapidminer, <https://rapidminer.com>.

- [60] IBM, IBM i2 Analyst's Notebook, <https://www.ibm.com/uk-en/marketplace/analysts-notebook>.
- [61] IBM, IBM SPSS Statistics, <https://www.ibm.com/uk-en/marketplace/spss-statistics>.
- [62] SAP: SAP Predictive Analytics, <https://www.sap.com/uk/products/predictive-analytics.html>.
- [63] R Language., <https://www.r-project.org>.
- [64] C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- [65] C. Bishop, *Neural Networks for Pattern Recognition*, Clarendon Press, 1995.
- [66] Luciad Company, Luciad Solutions, <http://www.luciad.com>.
- [67] Cesium Consortium: Cesium, <https://cesiumjs.org>.
- [68] Chart.js., <https://www.chartjs.org>.
- [69] Data-Driven Documents (D3.js), <https://d3js.org>.
- [70] Q. Ali, H. Zheng, T. Mann, and R. Srinivasan, "Power aware NUMA scheduler in VMware's ESXi hypervisor," in *Proceedings of the IEEE International Symposium on Workload Characterization, IISWC 2015*, pp. 193–202, October 2015.
- [71] L. Zhou, N. Chen, and C. Hu, "A method of coupling with multi-source heterogeneous remote sensing data system based on SOS web service," *International Conference on Geoinformatics*, pp. 1–52, 2013.
- [72] Y. Rebahi, S. Hohberg, L. Shi et al., "Virtual security appliances: The next generation security," in *Proceedings of the International Conference on Computing, Management and Telecommunications, ComManTel 2015*, pp. 103–110, December 2015.
- [73] M. Denis, C. Zena, and T. Hayajneh, "Penetration testing: Concepts, attack methods, and defense strategies," in *Proceedings of the IEEE Long Island Systems, Applications and Technology Conference, LISAT 2016*, pp. 1–6, 2016.
- [74] Kali.org, Kali Linux, <https://www.kali.org>.
- [75] N. Phaphoom, X. Wang, and P. Abrahamsson, "Foundations and technological landscape of cloud computing," *ISRN Software Engineering*, vol. 2013, Article ID 782174, 31 pages, 2013.
- [76] F. Gessert and N. Ritter, "Scalable data management: NoSQL data stores in research and practice," in *Proceedings of the 32nd IEEE International Conference on Data Engineering, ICDE 2016*, pp. 1420–1423, May 2016.
- [77] S. Sakr, A. Liu, D. M. Batista, and M. Alomari, "A survey of large scale data management approaches in cloud environments," *IEEE Communications Surveys & Tutorials*, vol. 13, no. 3, pp. 311–336, 2011.
- [78] Apache Software Foundation, Cassandra, <http://cassandra.apache.org>.
- [79] Mongo DB., <https://www.mongodb.com>.
- [80] J. M. Clarence, S. Aravindh, and A. B. Shreeharsha, "Comparative study of the new generation, agile, scalable, high performance NOSQL databases," *International Journal of Computer Applications*, vol. 48, no. 20, pp. 1–4, 2012.



Hindawi

Submit your manuscripts at
www.hindawi.com

