L. Cazorla, C. Alcaraz, and J. Lopez, "Awareness and Reaction Strategies for Critical Infrastructure Protection", *Computers and Electrical Engineering*, vol. 47, pp. 299-317, 2015.
http://doi.org/10.1016/j.compeleceng.2015.08.010
NICS Lab. Publications: https://www.nics.uma.es/publications

Awareness and Reaction Strategies for Critical Infrastructure Protection

Lorena Cazorla, Cristina Alcaraz, Javier Lopez

Computer Science Department,
University of Malaga, Spain
{lorena,alcaraz,jlm}@lcc.uma.es

Dated: October, 2015

Abstract

Current Critical Infrastructures (CIs) need intelligent automatic active reaction mechanisms to protect their critical processes against cyber attacks or system anomalies, and avoid the disruptive consequences of cascading failures between interdependent and interconnected systems. In this paper we study the Intrusion Detection, Prevention and Response Systems (IDPRS) that can offer this type of protection mechanisms, their constituting elements and their applicability to critical contexts. We design a methodological framework determining the essential elements present in the IDPRS, while evaluating each of their sub-components in terms of adequacy for critical contexts. We review the different types of active and passive countermeasures available, categorizing them and assessing whether or not they are suitable for Critical Infrastructure Protection (CIP). Through our study we look at different reaction systems and learn from them how to better create IDPRS solutions for CIP.

Keywords: Critical Infrastructure Protection, Control Systems, Countermeasures, Intrusion Detection and Response Systems.

1 Introduction

Critical Infrastructures (CIs) around the globe provide the most necessary services to society, so their continuous correct operation is of paramount importance. Control systems, such as Supervisory Control and Data Acquisition (SCADA), perform the management and the regulation of behavior of the internal devices and systems of these infrastructures. They are considered a fundamental component within CIs, having an impact in the overall performance of other interconnected critical infrastructures. Thus, the protection of CIs and their control infrastructures is currently seen as an essential part of national security in numerous countries around the world.

Recent reports from the Industrial Control Systems Cyber Emergency Response Team (ICS-CERT) show that security incidents and cyber-attacks against control systems are increasing, and they are getting more aggressive and sophisticated. Known large-scale cyber attacks targeting CIs and Industrial Control Systems (ICSs), such as Stuxnet, the Nitro Attacks and the Maroochy water breach, show that CIs and ICSs are becoming increasingly targeted by different types of malicious attacks [1].

For this reason, and as dictated by governments and institutions around the globe, the integrity and availability of all these critical systems have to be protected against the numerous threats they face every day [2]. Approaches for Critical Infrastructure Protection (CIP) arise from several perspectives: preparedness and prevention, detection and response, mitigation and recovery, international cooperation, etc. [2]. As a tool to respond to this need for protection, intrusion detection has been at the center of intense research in the last decade, due to the rapid increase of cyber-attacks on computer systems.

Intrusion detection refers to a variety of techniques for detecting threats in the form of system faults (anomalies) or malicious and unauthorized activities. A technique that focalizes this detection effort is the Intrusion Detection System (IDS) [3]. IDS solutions have been proposed for multiple environments, and they could result in very valuable protection tools for ICS environments. However, their application to the protection of critical systems must comply with the strict constraints and restrictions of ICSs [2].

However, when intrusive behavior is detected by the IDS in a critical scenario such as ICSs, it is desirable to take evasive and/or corrective response actions to prevent these attacks from succeeding, and ensure the safety of the computing environment [2, 4]; such countermeasures are referred to as intrusion response. Incidentally, as threats become more abundant and sophisticated, and given the special characteristics of CIs, apart from detection mechanisms, new and more powerful solutions have to be deployed in order to safeguard them and to avoid faults and consequent cascading effects. To fight this domino effect, besides providing efficient detection mechanisms, we need to focus on the response, mitigation and recovery needs of CIs.

Solutions that can provide these functionalities are the Intrusion Prevention Systems (IPSs), also called Intrusion Response Systems (IRSs) [4] and Intrusion Detection, Prevention and Response Systems (IDPRSs) [5]. An IPS/IRS/IDPRS is "*software that has all the capabilities of an intrusion detection system and can also attempt to stop possible incidents*" [3]. In the remainder of the text, we will refer to these systems as IDPRSs, since we will use the term Response System (RS) to denominate a specific element of the whole system. The IDPRS is often integrated as an extension of the IDS, but it usually receives less attention than the IDS research due to the intrinsic complexity of developing the mechanisms to offer an automated and correct response to certain events.

Traditionally, and particularly in ICSs, the response to a threat was manually triggered by the system's human administrator, and required a high degree of expertise. However, the increasing complexity and speed of the cyber-attacks in recent years, and the intricate possible ramifications of a system's faults show the acute need for complex intelligent dynamic RSs [4]. Therefore it has become necessary to use sophisticated advanced techniques from autonomic computing, machine learning, artificial intelligence and data mining to build intelligent and smart IDPRSs. Together with the deployment

of IDS solutions in these contexts, automatic and intelligent response mechanisms have to be put in place to help protect CIs and prevent cascading failures to other interdependent infrastructures [6].

The structure of the paper is as follows. Section 2 presents a taxonomy for IDPRS solutions for CIP, where the different elements existent in these systems are analyzed in terms of applicability to CIs. Section 3 provides a review of the state of the art IDRPS solutions developed in the recent years. Section 4 presents an analysis of the reviewed solutions, studying the main countermeasures available that can be implemented in an IDPRS for CIP, and discusses the main strengths and weaknesses of the solutions. Lastly, in Section 5 the conclusions and future work are outlined.

2 Taxonomy of Intrusion Prevention and Reaction Solutions

To understand the IDPRS, it is important to also understand the nature of the event they attempt to detect, the environment where they operate, the different kinds of processes that can be triggered to protect the surveilled system, and the possible types of solutions that can be launched. In order to provide safe IDPRS solutions to protect critical systems, we need to identify those desirable features, and most importantly, the characteristics that constrain the application and deployment of response solutions in critical contexts such as CIs.

Protection mechanisms put into place to safeguard CIs must be tailored to their environment, taking into account its constraints in order to ensure the correct operation of the system as a whole. IDPRS solutions, similarly to IDSs, are designed to monitor and protect hosts (*host-based architecture*), or networks (*network-based architecture*) [3]. A host-based IDPRS must be tailored to the node where the solution is running, it must operate within the constraints imposed by the host, and therefore it should be well integrated with its environment. Network IDPRS solutions monitor the traffic of communication networks, and can be deployed in radically different contexts.

Generally, the internal networks of CIs and their ICSs can be divided into three main types: *corporate networks*, the *SCADA center* and *remote substations*. The first are the business local area networks connected to a SCADA to gain access to critical data streams on SCADA servers (e.g., historical data, alarms, etc.). Corporate networks are general-purpose complex infrastructures where the nodes of the network (e.g., servers, gateways) have moderate to high computational capabilities and the constraints of these networks are minimal. SCADA centers are in charge of constantly monitoring the controlled infrastructures, using their communication networks to reach remote substations.

The nodes connected to SCADA centers are usually powerful, e.g., SCADA servers, gateways and some powerful Remote Terminal Units (RTUs) in the main remote substations. However, the protocols they use are proprietary and restricted, thus the IDPRS solutions deployed in this environment could use powerful computational resources, but they have to be designed for the specific communications protocols. Remote substations constitute those control networks based on field devices (e.g., sensors, actu-

ators) and communication interfaces (e.g., RTUs, gateways, base stations) capable of transmitting commands from the central system to field devices deployed close to the controlled infrastructures, and sending sensorial measurements to the SCADA center.

In these remote networks, most of the nodes (sensors, base stations) have constrained computational power; moreover, the protocols used in the communication of their nodes are usually proprietary, or adapted to low-power devices. Thus the IDPRS solutions deployed in this context should have lightweight procedures, and they must be tailored to the specific protocols and restricted nodes of the networks. Summarizing, these are the requirements posed by the physical structures and components of CIs, however in our study we have to also determine the desirable features and characteristics for the IDPRS solutions that are needed in CIP.

Our research takes as its basis the methodological framework for situational awareness given in [7], which is based on two execution phases and a set of protection services, among them: detection, alerting and response. We expand the proposed methodology to describe the different methods and phases for detection and response in a critical context. The aim of our study is to analyze the features and components available for prevention and response in general-purpose networks.

To identify the characteristics that are useful for the protection of CIs, and the factors that limit the application of RSs to critical systems, we need a methodological framework for IDPRS solutions for CIP. In our study, we follow the situational awareness framework in [7], and consider the taxonomies for general-purpose IDPRS solutions proposed by N. Stakhanova et al. [4] and A. Shameli-Sendi et al. [8]. They constitute the basis we adapt to analyze the IDPRSs that better fit CIP. The framework and taxonomy we consider for our analysis is illustrated in Figure 1.

Figure 1 is divided into three main modules, corresponding to the three key components or operational characteristics provided by the IDPRS: the *detection module*, the system's *automation degree* and the *response system* (RS). These modules characterize the resulting IDPRS and the capabilities of detection, automation and response that this system will have, thus determining the degree of protection this system provides to CIs. In this section we describe the features of each element and study their advantages and disadvantages for critical systems.

2.1 Module 1: Detection

The IDS is the detection component of the IDPRS, and the first element of our model (see Figure 1). IDSs have several possible mechanisms to provide detection, which can influence the efficiency and applicability of the IDPRS' reaction component; thus, it is important to note the technique used for detection. According to the implementation of the IDS' engine, there are three main methods of detection: *anomaly-based detection*, *signature-based detection* and *specifications-based detection*:

- **Anomaly-based detection**: the IDS compares definitions of activity that is considered normal against observed events in order to identify significant deviations [6]. This method has the advantage of being very effective in identifying previously unknown threats. Its main drawbacks are the generation of a large rate

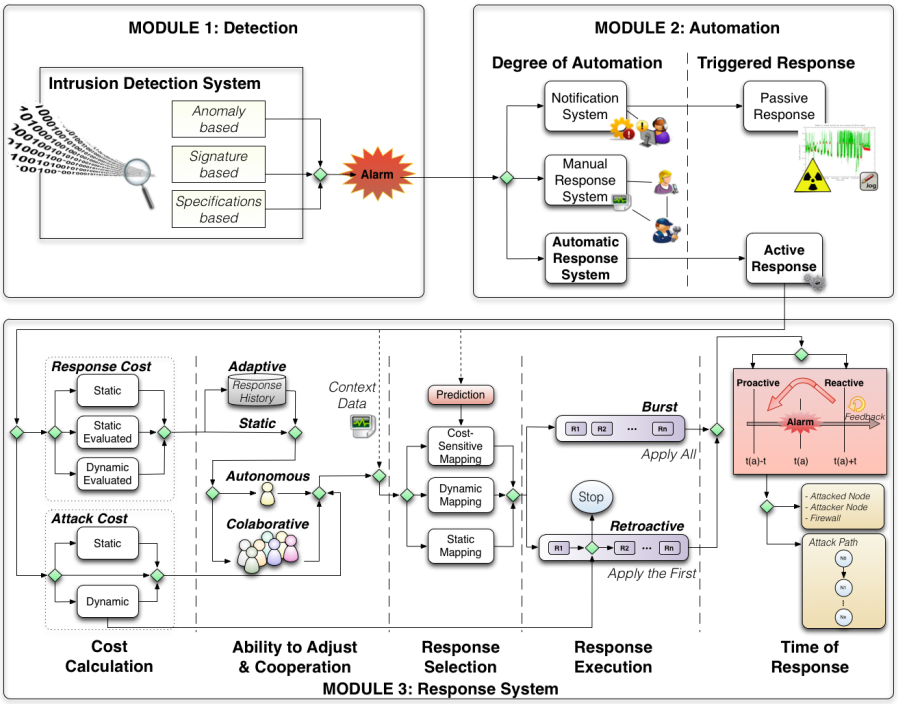

Figure 1: Methodological framework, containing a taxonomy of IDPRS, adapted from [4] and [8]

of false positives in dynamic environments, and the difficulty that arises when analyzing the causes of a given alert [3].

- **Signature-based detection**: according to the American National Institute of Standards and Technology (NIST), a signature is "*a pattern that corresponds to a known threat*" [3]. A signature-based IDS analyzes the information it gathers from the system under surveillance and compares it to signatures of known threats in order to identify possible incidents [6]. Using this method, the system looks for already identified and known attacks, thus this solution is very effective in detecting known threats, but rather ineffective in detecting previously unknown threats, threats disguised by the use of evasion techniques, and many variants of known threats.

- **Specifications-based detection**: similarly to anomaly detection, specification-based systems detect attacks as deviations from normal behavior. However, these approaches contain specifications of the system under surveillance (usually manually developed models) that capture legitimate behavior, instead of using previously seen behaviors as in the case of anomaly detection. To develop the models, legitimate systems' behaviors can be extracted from security policies and proto-

col specifications [9]. According to NIST [3], there are three main problems that arise when using stateful protocol analysis. First, the reliance on vendor-developed universal profiles that specify the use of particular protocols ties the IDS to a specific environment, since the implementation of the protocols may differ from system to system. Second, this kind of system is highly resource-intensive because of the complexity of the analysis and the overheads caused by the state tracking. And lastly, these IDSs cannot detect attacks that do not violate the characteristics of acceptable protocol behavior. Thus, the type of detection engine used in the IDS depends on the characteristics of the surveilled systems, where their selection is guided by parameters such as precision (low false alarm rate) or efficiency.

In a critical context, the parameters studied to select the adequacy of a given IDS for a constrained environment are different from the general networks' IDS solutions. Here, characteristics such as precision vary from the desire to obtain low false alarm rates (low rate of false positives) to the need to achieve a low rate of false negatives (the true attacks are not missed by the detection engine). It is also necessary to observe the techniques implemented (to avoid resource-intensive algorithms) and the level of automation achieved by the IDS for CIP [6]. Therefore, the selection of the detection engine must be aligned with the needs and constraints of the critical system under surveillance, and the type of environment where the system has to perform its tasks.

In general, taking into account their aforementioned characteristics, specifications-based detection as well as signature-based detection will work better for environments where the dynamics and behaviors are well-known [9]. Contrarily, anomaly-based detection can be recommended when the IDS is placed in an environment that constantly faces unknown behaviors and dynamics [6]. In the literature, we find different types of IDSs specifically designed for CIP [9, 10], using and adapting different types of detection engines.

2.2 Module 2: Automation

A methodology to study the levels of automation of a given IDS solution for CIP is described in [6], where the authors defend the need to develop automatic solutions to automatically and intelligently protect CIs. Here, we find that the fifth level of automation corresponds to the automation of the response mechanism implemented by the IDS. Since the IDPRS solutions provide an extension of the functionalities present in the IDS, it is necessary to determine the *degree of automation* they provide (as stated by the methodological framework for CIs given in [7]), in order to determine the automatic capabilities they implement. According to the level of automation, the RSs are capable of providing different types of *triggered responses* [4].

2.2.1 Degree of Automation

Currently, research is mainly focused on providing manual or semi-manual reaction mechanisms, due to the difficulties that arise when trying to apply correct automatic responses against determined events. Besides this difficulty, CIs have a critical need

for automatic response mechanisms [2]. On the one hand, immediate automatically triggered reactions to threats avoid undesirable effects such as cascading errors through interconnected infrastructures. On the other hand, any incorrect action performed by an automated component can potentially cause catastrophic effects on CIs. To balance these risks, IDPRSs implement automatic responses at different levels, which can be subdivided and categorized into the following types (see Figure 1) [4]:

- **Notification systems**: reaction systems based on notifications generate alarms when threatening events are detected. They provide information about the occurrence, and the system administrator is the responsible for selecting the appropriate response. The majority of the existing IDS solutions provide this kind of mechanism [6]. This approach is not designed to prevent attacks or return the system to a safe state, and its major disadvantage is the delay between the potentially harmful event and the human response [4].

- **Manual response systems**: manual response is a step ahead of the notification systems. The RS has a preconfigured set of actions that the administrator launches whenever a problem arises, and based on the characteristics of the event reported [8]. This process is not completely automated, however the countermeasures are configured in the system beforehand and the response of the operator is faster than in the previous scenario.

- **Automatic response systems**: are designed to be fully automated, thus (unlike the two methods mentioned above) human intervention is not required, and consequently there is no delay between the detection of the event and the response. However, due to the great difficulty in reaching a high level of automation in the response mechanisms, the existing systems that provide automated response are very limited [4]. The main problem of this approach is the possibility that an inappropriate response is executed when a problem arises; it is also difficult to ensure that an automatic response is able to neutralize a problem [8].

2.2.2 Triggered Response

Those IDPRSs that implement automatic reaction can be further categorized into two different classes, taking into account the type of their triggered response: the *active response* mechanisms, and the *passive response* methods (see Figure 1). This reaction determines the whole structure of the reaction system, and its intrinsic complexity.

- **Passive response**: the systems that present passive responses do not attempt to minimize the damage caused by the potentially harmful event or prevent a repeat in the future. The main objective is to notify the authority (e.g., the human operator, the SCADA center) and provide information about the occurrence [4].

- **Active response**: active systems try to locate the source of the detected event and provide active response actions to minimize the damage derived from the occurrence. The vast majority of the IDS solutions available only provide passive response, whilst the active reaction mechanisms are very limited at present [4].

Currently, most CIs are not equipped with active RSs. However, the need for dynamic incident management and response systems capable of sending alerts for anomalies caused by malfunctions or intrusive presence are widely defended by the scientific and governmental communities [2]. Most of current IDSs and IDPRSs present in critical contexts implement passive response mechanisms [10], therefore, we focus our analysis on the characteristics, functionalities and constraints of the automatic response systems.

2.3 Module 3: Response System

The third module of an IDPRS is the response system (see Figure 1), which is the component capable of selecting proper countermeasures to a given threat. This element is fed by the IDS component with insight about the threat, and has a determined degree of automation depending on the implementation of the system. The decisions taken to select the reactive measures against the threats can be made using various forms of computation. For example, they can be predefined or modified depending on the environment, they can be calculated autonomously or in a cooperative way. Also, the response can be applied in a reactive way or before the threat reaches the maximum.

Thus, there are multiple forms of implementing the IDPRSs and the varied characteristics they can add. Specifically, our methodological framework consider the following features: *response cost model*, *risk assessment method*, *ability to adjust*, *cooperation ability*, *response selection method*, *response execution method* and *time of response*. There are other different variables that could be included in the study, such as the *applying location* and the *response lifetime*. However, to determine the best solutions for CIP, we focus on the first.

2.3.1 Response Cost Model

Each threat (attack or fault) to a system entails a cost; likewise, each response (automatic or manual) to a threat has an impact on the system. Generally, the best responses are those that cost less, and it is always necessary that the cost of the selected response is less than the cost of the fault or attack. There are diverse forms of calculating the effects of the response actions, most of them belonging to the field of risk assessment [11]. There are three main ways of providing the automated IDPRS with the models for calculating the cost of a response in a particular situation: the *static cost model*, the *static evaluated cost model*, and the *dynamic evaluated cost model*.

- **Static cost model**: here the response cost is assigned statically using the opinion of an expert. This value has to be set for each response and it is usually preconfigured in the system [8].
- **Static evaluated cost model**: the cost is calculated and assigned statically to each response. Here, the evaluation mechanism usually computes the positive effects of the response (based on their consequences for the availability, confidentiality and integrity variables and performance metrics), and the negative impacts (in terms of availability and performance). The combination of the two

kinds of effects comprises the cost assigned to the response. The majority of IDPRSs use the static evaluated cost model [8].

- **Dynamic evaluated cost model**: the IDPRS dynamically calculates the cost of the response based on the dependencies between the resources and the actors in the system [8]. The resulting cost-sensitive model is usually very accurate, allowing the IDPRS to select appropriate responses that take into account the interdependencies of the system, its critical processes and the Quality of Service (QoS). However, due to its increased complexity, the majority of the IDPRS available implement static cost or static evaluated cost models.

In a critical context where the dynamics are well-known and the patterns of behavior are sufficiently static and predictable, IDPRSs that implement static cost models are a good option. Thanks to this static context, the range of possible threats and risks is limited to a restricted known set, and the countermeasure actions are equally limited. Therefore the simpler static methods perform well in these environments. However, whenever the behavior of the system has more complex dynamics, it is necessary to apply more sophisticated cost models. Especially if computationally powerful resources are available, the dynamic evaluated cost model is highly recommended, since the costs are calculated taking into account the critical processes and the multiple interdependencies existing between CIs.

2.3.2 Risk assessment (Attack Cost)

Attacks or anomalies have a negative impact on the system, and through the calculation of the cost of these events, we can help the response system determine the best course of action for protection in a particular scenario. The assessment of these costs can be done *statically* or *dynamically*:

- **Static**: consists of assigning a static value to each resource of the system [8]. This type of risk assessment has a useful basic performance, and the procedures for its application are described in many existing standards (e.g., ISO 27001 [12]). However, static risk assessment does not provide the versatility and advantages of dynamic risk assessment for a dynamic context.

- **Dynamic**: the assessment of risk dynamically provides a real time process for the evaluation of risk indices in the system [8]. The model is dynamically created by propagating the impact of the security variables through service dependency models [13], attack graphs [14], or general models based on metrics [8, 15]. Online risk assessment, although computationally complex, minimizes the costs of the threats and response events in the system, and allows the IDPRS to work, taking into account the context (state) of the system [16].

According to our study of the literature, a fully-fledged dynamic risk assessment component is present in few IDPRS existing solutions (described further on in Section 3). However, given its importance in determining the best possible countermeasures for an RS, this component should be addressed for the IDPRS, if only in its

simpler forms. Static risk assessment mechanisms like the one proposed in [11], provide risk metrics such as the *average loss* or the *variance of the losses*. In [11], an heuristic is proposed to minimize risk by estimating the potential losses, to identify the high priority equipment (sensors) and to invest resources in protecting them. This kind of component is especially well suited to constrained, low-computational power sub-systems, such as the field networks in CIs.

More sophisticated approaches include dynamic risk assessment components which perform their evaluation of the risk level using different techniques. Y. Haimes et al. [17] study risk from different perspectives, and at different levels (e.g., physical risk, logical and information risk) to create models of risk. The Network Security Risk Model (NSRM) is capable of assessing the risk of cyber attacks on process control networks using models. They show the different attack scenarios, and enable studying the progressions and consequences of the different attacks modeled and the risk levels introduced by the selected response strategies.

Also based on modeling, the risk-aware framework proposed in [18] contains an online component which measures the likelihood of success of an ongoing threat or attack, and the cumulative impacts (cost) of the threat and the response. These measures help the RS determine the need for activation or deactivation of the system's policies as countermeasures. Specifically, this framework is proposed for complex infrastructures and systems, with multiple interdependencies and constraints present in their normal operation. Another risk-aware RS is presented in [19], where the authors propose a framework for risk assessment in the smart grid.

R. Habash et al. identify metrics and factors evaluating the level of risk in this environment, and the level of risk remaining after applying countermeasures to tackle the threat. In their paper, A. Shameli-Sendi et al. [16] present ARITO, an RS using accurate risk impact tolerance. This RS contains a risk assessment component that measures the risk impact in real time. This system is also capable of providing feedback mechanisms for the countermeasures applied by the RS, provided by the calculations of the response goodness, which helps indicate the new risk level after applying the selected responses. After the risk evaluation, the system can decide whether to activate a response or not, and the strength of the response, according to the network risk level and the risk impact of each countermeasure.

Thus, given the different types of solutions at hand, we firmly believe that IDPRS solutions for CIs should have risk assessment components. The RS must have mechanisms to evaluate the costs of both the threat and the response actions occurring within the system in order to determine the best response to anomalous situations. This feature is essential to apply accurate countermeasures to the threat, while avoiding causing havoc and widespread operative disruption due to a mistakenly applied response. Ideally, risk assessment should implement dynamic procedures, however, in constrained areas of the network, it is possible to implement static well-adjusted procedures.

2.3.3 Ability to Adjust

The capability of a response system to adjust to new situations and scenarios is a desirable characteristic that supports its robustness and continuity. However, its implementation is costly in terms of computational complexity and difficulty of design, and

in some cases deemed unnecessary. Thus we can find in the literature two positions regarding the adjustability of the RS: *adaptive solutions* and *static solutions*.

- **Adaptive**: adaptability is "*an ability of the system to dynamically adjust the response selection to the changing environment*" when facing harmful events [4]. Adaptive models usually adjust their actions based on response history [8], in the form of: (i) the adjustment of the system's resources devoted to intrusion response; e.g., activation of additional IDSs; or (ii) the consideration of the success or failure of the previously used responses.

- **Static (non-adaptive)**: models provide a static response selection procedure, which remains the same throughout the lifetime of the IDPRS software. There is no mechanism for tracing the behavior of the applied responses, and the support to the system is manual (periodic upgrades). Although static, this kind of response model is simple and easy to maintain [4], so in some cases it is preferred by the administrators.

There are different types of networks within CIs, each of them having different environmental characteristics, thus the adjustability of an IDPRS is critical in these sections of the network where continuous changes happen, with frequent updates of the networks' nodes and continuously changing dynamics (e.g., in the corporate networks of CIs). However, other sections of the networks usually have very static patterns of behavior, and the need for an adaptive IDPRS there is not critical. This is the case, for example, in the networks that connect the SCADA with the RTU (or destination gateway), and in the communication networks between RTUs (or gateways) and sensors/actuators.

2.3.4 Cooperation Ability

RSs can be designed in such a way that they perform their tasks *autonomously*, being capable of monitoring localized areas of the surveilled system or the holistic behaviors of the system. Conversely, they can be implemented *cooperatively*, which allows multiple IDPRSs to communicate with each other and collaborate at different levels.

- **Autonomous**: the autonomous IDPRS solutions handle events independently, without communication with other components, at the level the threats are detected [4]. These solutions are usually aware of just a restricted part of the context of the system, thus the responses they can provide are generally localized.

- **Cooperative**: it refers to the set of RSs that combine efforts to respond to an intrusion. They consist of several autonomous systems, capable of detecting and responding to intrusions locally, but with a final response strategy determined and applied globally. Sometimes, IDPRSs are directly built to operate cooperatively, which makes them perform better in terms of response speed and contained damage volume. Nonetheless, they are also more complex and require strong coordination and communication between their components [4].

Complex networks, with complicated dynamics and interconnected dependent systems, benefit from cooperative IDPRS solutions. It is possible to deploy multiple autonomous IDPRS modules for protection at different levels in the varied networks of CIs. However, cooperative detection systems could be able to correlate different events (occurring at different levels or locations of the infrastructure) to provide behavior forecasting and improve the overall performance of the IDPRS (proactive protection). Nevertheless, it is not always cost-effective to increase the complexity of the security mechanisms, since they can diminish the efficiency of the operation of the infrastructure. Thus in constrained environments (e.g., remote substations), autonomous lightweight IDPRSs are the recommended option.

2.3.5 Response Selection Method

Once the IDS has provided information about the threat that is placing the system at risk, and the costs have been calculated, the RS must determine the best actions to carry out to counteract the threat. The IDPRS makes such decisions by associating the threat alerts with a determined set of actions. This association can be done in three different ways, with increasing degrees of complexity: *static mapping*, *dynamic mapping* and *cost-sensitive mapping*.

- **Static mapping**: here, each alert is mapped to a predefined response [8]. The resulting model is easy to build and maintain, however, it also makes the system predictable and thus vulnerable to intrusions (in particular Denial of Service (DoS) attacks). Moreover, static mapping systems have an inherent inability to consider the current state of the whole system, therefore the actions triggered represent an isolated effort to mitigate a problem, without considering the current condition and impact of the response on the system. The application of this technique for large systems is infeasible and prone to errors, since the number of threat scenarios needed to be analyzed and the constant changes in system policies make this process extremely complex [4].

- **Dynamic mapping**: these RSs are more sophisticated than static mapping systems, since their response selection considers certain attack metrics (e.g., confidence, criticality) and network policies [20]. Each alert is associated with a set of response actions and when a potentially harmful event occurs, the system chooses, in real time, the best countermeasures from the corresponding set, taking into account the characteristics of the particular threat [4]. This approach provides a fine-grained control over the automatic response of the system through adjustments to the metrics. However, its main drawback is that the RS does not learn lessons from one situation to the next; thus its intelligence level remains constant until the next system upgrade [8].

- **Cost-sensitive mapping**: these RSs attempt to balance the cost of the damage caused by the harmful event and the cost of the response [4]. The optimal countermeasure is determined through the cost-sensitive model of the IDPRS, which includes cost and risk factors related to the event, and to each response [21]. Traditionally, cost-sensitive approaches use an offline risk assessment procedure,

where the cost and risk factors are calculated in advance and the values are static. In some cases, this mechanism is completely manual, and it is the system administrator's task to update these values over time [4].

To improve the static procedures and lower the burden on the system's operators, online risk assessment components have been proposed to measure the cost of attacks, faults and automated responses [8]. Therefore, whenever the equipment capabilities allow it, it would be preferable to use the most sophisticated response selection techniques for the RS. Cost-sensitive or dynamic mapping would assist the operators in making the best decisions for incident response. However, this advantage comes with the price of complexity and computational cost: it is necessary to analyze a vast number of factors (intrusion cause and effect, identification of optimal response, state of the system, maintainability) and to completely understand the problems addressed to provide optimal responses [4].

Additionally, it is necessary to adjust the measures of accuracy and adequacy of a response to the selection method implemented by the IDPRS. The maintainability of the system also increases whenever automatic sophisticated techniques are used for response selection. Furthermore, in ICSs, the application of incorrect countermeasure actions can be devastating to the operation of the surveilled system, and can cause fatal errors to cascade throughout the infrastructure, and into other dependent interconnected CIs. Therefore, the response selection process must be supervised by the system's operator, and the RS must adjust its functionality to the constraints of its environment.

In a critical context it is thus desirable to have an IDPRS well integrated with its environment. Within ICSs, there are sectors with heavy-duty equipment and powerful computational capabilities, such as the SCADA center or the main remote substations (based on gateways as main interfaces), able to run IDPRS solutions that provide the most sophisticated response selection methods (i.e., cost-sensitive mapping). Also, the ICS contains sectors with constrained resources, where the equipment is not capable of performing high-complexity computations, e.g., in the field networks, where there are mainly lightweight sensors with low computational power.

In this case, it is also interesting to protect the system through IDSs and IDPRSs, but these solutions have to be tailored to a constrained scenario. Thus, lightweight IDSs and IDPRSs can be deployed, where the methods used to perform their tasks consume fewer resources, e.g., providing an IDPRS with static mapping techniques for response selection. Then, by deploying the IDPRS that best suits each part of the system, it is possible to protect the infrastructure without the need to modify or add equipment to execute these tasks.

2.3.6 Response Execution

When the response system has selected the most suitable actions to counteract the threat, there are two ways of executing them: in a *burst*, or using *retroactive feedback*.

- **Burst**: this mode of execution does not take into account any mechanism to measure the risk once the selected response (or set of responses) is applied. This means that all the countermeasures are always applied, disregarding the possibility that a subset of the actions could be enough to mitigate the threat. This is

the response execution mode usually present in the literature, its main weakness being the performance cost [8].

- **Retroactive**: in the retroactive execution, there is a feedback mechanism that measures the response effect taking into account the results of applying the most recent set of countermeasure actions. This measurement helps the system to make decisions before applying the next set of actions [8]. This kind of system was first presented by C. Mu et al. in [22], where the authors indicate several ways to implement the retroactive approach:

  - Selection window: each response has a static risk threshold associated with it, and to run the countermeasure it is necessary to consider the current risk index of the system. If its value is higher than the static threshold of the action, the next response can be activated. With a selection window, the most effective countermeasures are selected to repel intrusions.
  - Independent responses: this method involves measuring the risks associated with the countermeasure applied in order to make a decision about the application of the next one. Since responses are evaluated independently taking into account their cost impact, this step-by-step execution mechanism is more conservative than the previous approach, and more suitable to be applied in critical contexts.
  - Grouped responses: when calculating the risk of a single response does not provide enough information to make the decision about running the next one, it is interesting to build groups of countermeasures. The decision to run the next round of responses is based on the general risk of the system. Once a group of countermeasures have been applied, the risk needs to be re-calculated. The challenge of this method is to determine how many responses in a round are considered enough to neutralize an attack.

ICSs are complex and delicate systems, thus when an undesired event happens, it is necessary to palliate its effects as soon as possible and in the least harmful way possible. Two approaches have been proposed: the burst response execution, which applies the countermeasures in bulk, as a whole, and the retroactive response execution. Although it is possible to find both types of RSs in the literature, the latter provides the advantage of stopping the process of responding to the event to evaluate the effects of the countermeasures already applied. Since, in a given situation, a subset of the countermeasures selected to palliate the anomalous state could be enough to restore the system to a normal operation, retroactivity capabilities are desirable.

Moreover, in a critical environment, executing countermeasures in a burst without evaluating their impact on the system could bring it to a critical state. Thus, automatic non-evaluated responses must not be executed in CIs without the supervision of a human operator. The reason behind this is the criticality of any action performed within CIs, since any mistaken activity (automatic or manual) can disrupt the operation of the system with potentially devastating consequences. Therefore, in this context it is better to execute countermeasures retroactively. However, these feedback mechanisms, as is the case with all adaptive approaches, face some challenges that make their use

difficult, e.g., measuring the success of the most recently applied response or handling multiple threatening events.

There are several ways to provide feedback to the RS, ranging from simple static system's metrics, to a more dynamic approach such as including a risk assessment component. As suggested in Section 2.3.2, risk assessment can help determine the costs of the responses, to provide the desired feedback to the IDPRS. Here, in order to make the IDPRS more precise, risk assessment should be conducted dynamically (online). There are systems in the literature capable of tackling risk assessment in different scenarios, varying from general purpose environments [11, 16, 17] to complex cyber-physical infrastructures such as the telecommunications industry or the smart grid [18, 19].

However, since these online dynamic mechanisms make the IDPRS costlier in terms of the system's resources [8], they can be included in areas of the network with sufficient computation capabilities. In constrained areas of the network, it is possible to use simpler methods than [11] to determine the suitability of the countermeasures for a determined situation. This can help computationally constrained devices to include RSs with retroactive, although rudimentary, response execution capabilities.

Therefore, in a critical context, the unsupervised automated response execution of countermeasures in a burst should not be applied. Instead, supervised response execution, or retroactive execution mechanisms should be put in place to prevent the application of countermeasures that exceed the risk cost of the threat. It is, therefore, necessary to delicately execute the responses, maintaining a control of the risk associated with the automatic procedures of the IDPRS. In this case, running and assessing the responses independently, or in small related groups can deliver adequate automatic responses, while minimizing the risks and costs of the automatic reaction.

2.3.7 Time of Response

It is possible to classify IDPRS solutions into *proactive* and *reactive* systems, taking into account the time instant when the IDPRS launches the response actions, with reference to whether the threat (e.g., an attack) has been already been confirmed or not.

- **Proactive (preemptive)**: proactive RSs foresee the incoming (potentially harmful) event and launch the response actions to help control the threat before it has affected the resource. This prediction is complicated to make and usually relies on probability measures and analysis of the system's behavior. Proactive solutions require the detection and response mechanisms to be tightly coupled, so the countermeasures can be triggered as soon as the event has been identified [4]. However, and although early response is highly desirable, it is difficult to guarantee the correctness of the triggered response action; thus the proactivity of the system has to be balanced with the correctness of the responses provided.

- **Reactive (delayed)**: in these RSs, the reaction is delayed until the threat has been confirmed [4]. The threat can be confirmed using confidence metrics in the IDS or by the matching of the event's trace with an existing signature in the IDS. Clear distinction exists between the proactive mechanisms (calling them incident prevention systems) and the delayed IDPRSs (calling them intrusion

handling/response systems). The proactive response usually includes actions to restore the system state to its normal operation [4].

Reactive systems, because they are less complex compared to proactive mechanisms, are widely used in IDPRS solutions [8]. These RSs do not trigger any countermeasure actions until the threat has been detected. The problem with reactive solutions compared to preemptive systems is that, generally, the delayed response leaves the event "unattended" for a longer period, consequently allowing more damage to occur, and setting a greater burden on the recovery mechanisms and the system administrators. While this might not cause too much trouble in general networks, a delayed response is not suitable for critical systems [4].

N. Anuar et al. [5] detail the disadvantages of a reactive RS, defending the difficulty to return an affected system to its normal operation, while having to consider that the system remains in an unsafe state in the time window until the response actions are applied. However, deploying a proactive RS in a critical environment is challenging, since false positively detected threats might trigger countermeasure actions from the IDPRS and bring the system to a unstable state. To prevent this problem and benefit from the advantages of a proactive IDPRS, it is necessary to finely tune the detection engine to make sure the rate of false positives is as low as possible.

Additionally, critical systems should not implement entirely automatic response systems, without human supervision, to avoid executing mistaken actions. Therefore, semi-supervised or supervised proactive IDPRS solutions are recommended in this environment. Of course, it is not always possible to deploy sophisticated and computationally costly solutions in several areas of CIs, due to their constrained nature. However, there are simple techniques, such as statistics or rule-based detection, which allow implementing lightweight preemptive systems. They are capable of detecting certain patterns and behaviors that deviate from the standard operation and which precede a threat, and launching adequate responses, even in a constrained environment such as the field sensor networks.

Thus, sophisticated proactive IDPRSs can be applied to ICSs in a context where the nodes have sufficient computational power, e.g., to the SCADA center, or to the networks that connect the SCADA center with the main remote substations. There, network dynamics are simple and the nodes are powerful. Another area of ICSs that could make use of proactive solutions are the corporate networks, where there are complex dynamics and behaviors, but the resources are sufficiently powerful to apply behavior-based forecasting. Finally, as mentioned, sectors of ICSs where the equipment has constrained capabilities (e.g., field sensor networks) can benefit from simpler, lightweight proactive RS solutions.

3 Review of the State of the Art: Approaches, Techniques and Tools

A review of the literature shows the different approaches taken to build IDPRS solutions through the recent years. This analysis is based on the methodological framework and protection methods described in Section 2 and focuses on non-commercial

general-purpose academic IDPRSs. There is a line of commercial tools that provide several interesting solutions regarding IDPRS mechanisms, however, the features implemented by these systems have a proprietary nature which restricts their study to the characteristics publicized by the provider of these tools. N. Anuar et al. present a study based on Gartners report on network IDPRSs in their paper [5]. They analyze the level of response applied in commercial and research products, looking at IDS and IDRPS technologies, as well as Security Information and Event Management (SIEM) products (tools for real-time analysis and management of security alerts within a system and its network). Therefore commercial solutions are not included in our study, as we cannot identify the essential features needed for the protection of the CIs.

One of the earliest IDPRSs found in the literature is the SoSMART system, by S. Mnsman and P. Flesher [23], an agent-based IDPRS with a statically mapped response selection procedure. The incident cases are designed by the user and mapped to the appropriate countermeasures. Additionally, the system uses case-base reasoning as an adaptation mechanism in charge of determining if the current solution corresponds to intrusive behavior.

Another autonomous IDPRS is the PH system, designed by A. Somayaji and S. Forrest [24]. PH is based on a behavioral profile of the system composed of sequences of system calls. The calls that deviate from the normal behavior are considered anomalous and can therefore be marked to be counteracted. The system only implements two simple kinds of response actions: suspending the suspicious processes or aborting them permanently. As do the majority of IDPRSs, PH implements a delayed response mechanism, i.e., it waits until the intrusion has been confirmed.

W. Lee et al. propose a cost-sensitive IDPRS [25] based on three cost factors: (i) *operational cost*, the cost of processing and analyzing data for detecting intrusions; (ii) *damage cost*, the amount of damage that can be caused by an attack when the IDS is ineffective; and (iii) *response cost*, the cost of applying the response when the attack has been detected. These factors combined present the total cost of the intrusion, and they help the system select the appropriate countermeasure in every case.

One of the most complex dynamic mapping approaches of the 2000's decade is the Adaptive, Agent-based Intrusion Response System (AAIRS) based on an agent architecture [20]. AAIRS is a an agent-based complex system, where multiple IDSs monitor a host and generate alarms. These agents operate at the different layers of the response process. Firstly, the intrusion alarms are processed by the master analysis agent, which calculates the confidence level based on pre-set decision tables and classifies the attack as new or ongoing. The system then passes this information to the analysis agent, which provides an action plan based on a seven-dimensions response taxonomy: degree of suspicion, attack time, attacker type, attack type, attack implications, response goal, and policy constraints.

Lastly, the tactics agent decomposes the response plan into particular actions and activates the appropriate components of the response toolkit. It is capable of adapting its countermeasures to each situation using the IDS' confidence metrics, which indicate the number of false positive alarms against the correct number of intrusions generated by each IDS. Similarly, the success metrics indicate the response actions that were successful in the past. A drawback to AAIRS is that it requires the intervention of the system administrator after each incident.

In 2001, S. M. Lewandowski et al. presented another cooperative RS, Survivable Autonomic Response Architecture (SARA) in [26]. It is composed of several components that function as: sensors (gathering of information), detectors (analysis of sensor data), arbitrators (selection of adequate response actions), and responders (implementation of response). SARA's components are arranged to provide the highest possible levels of detection and prevention. For example, each host can have an arbitrator to provide intrusion response while the selection response comes from a global (cooperative) strategy.

The Cooperative Intrusion Traceback and Response Architecture (CITRA), presented by D. Schnackenberg et al. in [27], is another cooperative agent-based system. CITRA uses a neighborhood structure to propagate the intrusion information until it reaches a centralized authority called the discovery coordinator, which determines the optimal response to the intrusion. The discovery coordinator centralizes the global response, however, the local CITRA agents are in charge of delivering the local response actions.

CITRA's framework is composed of network-based IDS, security management systems and network components (e.g., routers). Their aim is to detect the intrusion, trace it back to the source and coordinate the suitable reactions. Two factors guide the response mechanism: the certainty (likelihood of the event being an actual intrusion) and the severity of the intrusion (potential damage to the system). Once these two parameters, which define the characteristics of the event, have been determined, the response action is chosen from a pre-determined set.

Also in 2001, X. Wang et al. [28] presented TBAIR, the Tracing Based Active Intrusion Response system. TBAIR is a dynamic-mapping non-adaptive IDPRS capable of tracing the intrusion back to the source host to dynamically select a proper response to mitigate its effects; e.g., by blocking the intruder remotely or isolating the affected hosts. The model proposed by T. Toth and C. Kruegel [21] considers the cost and benefits of the countermeasures. It implements a network RS capable of modeling dependencies between the services and resources of the system, in the form of a tree. This model can reveal priorities in targets and helps evaluate the impact of the response strategy on the system. Likewise, the algorithm for response selection can take into account the assigned static penalty cost of having a resource unavailable, so it can indicate the impact that the response strategy has on the system. Considering this model, the algorithm applies the set of actions that has the least negative impact.

S. Tanachaiwiwat et al. [29] present a cost-sensitive, static IDPRS. The system is non-adaptive given the difficulty of calculating the effectiveness of a given countermeasure. The IDPRS is based on the efficiency of the IDS, the alarm frequency per week (indicating the number of alarms triggered per attack) and the potential damage cost. These variables serve to identify the best reaction strategy from a predefined list of responses. Similar to [21], the system proposed in [13] presents a cost-sensitive, dynamic RS capable of modeling dependencies between the services of the system to identify the impact of the different countermeasures. This system implements a delayed model, which suspends any action until after the threat has been confirmed.

The IDPRS uses host IDS and provides two ways of classifying the resources of the system: a resource hierarchy or a system map. The resource hierarchy is a directed graph, where its nodes are specific system resources and the graph edges represent

dependencies between them. Each node is associated with a set of reactions to restore its working state when attacked. The response is selected using: (i) the reaction cost, corresponding to the sum of the resources affected by the response; (ii) the reaction benefit, through the sum of nodes previously affected and restored to a working state; and (iii) the cost of the threatened resource. When there are alerts about nodes that have not previously considered, the system adds them dynamically to the map/hierarchy.

In 2005, B. Foo et al. presented ADEPTS, an adaptive and proactive RS, in [30]. ADEPTS uses Intrusion Graphs (I-Graph) to model intrusions, which identify attack targets, the possible spread of the intrusion, and those nodes where it is possible to apply successful responses. The RS maps the alarms provided by the IDS against the I-Graph and selects the countermeasures taking into account their calculated effectiveness, their potential to cause disruption and the level of confidence of the system being intruded. ADEPTS uses feedback mechanisms coming from the affected nodes, where parameters such as the confidence level of the attack and previous similar experiences help in estimating the success or failure of the applied response. As opposed to AAIRS [20], this system is capable of automatically updating the response effectiveness metric.

FLIPS, Feedback Learning Intrusion Prevention System [31], is another proactive IDPRS, which emulates the applications in a restricted environment before their execution. Thanks to this previous emulation, the system can recognize code injection attacks with only a few bytes of data, and prevent the system from executing their malicious code. M. Papadaki and S. Furnell present a cost-sensitive RS capable of evaluating the static and dynamic context of an attack in [32] using a database of their characteristics (e.g., target, applications, vulnerabilities). It also takes into account the characteristics of the responses available (e.g., counter-effects, stopping power, transparency, efficiency, and confidence level) to propose different kinds of countermeasures according to the attack, and it is capable of adapting the response to changes in the environment.

Similar to FLIPS [31] and ADEPTS [30], N. Stakhanova et al. propose in [33] another proactive, cost-sensitive IDPRS designed to detect anomalous behaviors in terms of system calls. The IDS tries to match the sequences of system calls with sets of normal and abnormal patterns to determine whether there is an attack or not. If the IDS finds no signature (pattern) matches, a machine learning engine is used to discern whether the behavior is normal or anomalous. Since this system is proactive, the reactions are triggered before the attack is completed. To operate in advance, the IDPRS has to have a predetermined mapping between system resources, countermeasures and intrusion patterns. When a sequence of system calls matches an abnormal pattern, the RS chooses the proper reactions available that have the least negative effect. The effectiveness of the response is measured and considered for future events.

K. Haslum et al. presented the Distributed Intrusion Prediction and Prevention System (DIPS) in [15], a cost-sensitive, real time IDPRS with prediction and risk assessment modules based on fuzzy models. Fuzzy logic is used to automatically estimate and infer risk, taking over this task from the security and risk experts. DIPS implements a hidden Markov model to represent the interaction between the attacker and the system's network. Also in 2007, M. Jahnke et al. [14] proposed a cost-sensitive IDPRS that uses graph-based mechanisms for risk assessment. Graphs model the effect of attacks in the resources, and the effects of the countermeasures in terms of availability. This system expands on the idea of T. Toth and C. Kruegel [21], using directed graphs

to model dependencies between the resources, and to calculate differences between system states.

One of the few early adaptive solutions is presented by C. Strasburg et al. in [34]. The authors propose a structured methodology to evaluate the cost of a countermeasure based on three parameters: (i) operational cost: the cost of preparing and developing responses; (ii) impact of the reaction on the system: which measures the negative effect of the response action on the system; and (iii) response goodness: based on the number of possible intrusions that the countermeasure can cope with, and also the number of resources that can be protected by the countermeasure. The total response cost is a combination of these parameters.

C. Mu and Y. Li propose IDAM&IRS (Intrusion Detection Alert Management and Intrusion Response System) in [22], which includes an RS based on hierarchical task networks. Each countermeasure of the RS has an associated static risk threshold calculated using its ratio of positive and negative effects. The reaction is chosen using a response selection window that shows the most effective countermeasures. IDAM&IRS triggers a response when its value is higher than its static impact risk index. The action is selected according to the goal of the RS: analyzing, capturing or masking the attack, maximizing the system's confidentiality or integrity, etc. Each goal has its own sequence of responses according to the risks they imply, e.g., weak responses earlier in time, and strong responses later.

W. Kanoun et al. presented in [18] a risk-aware framework composed of an online model and its architecture, which allows activating or deactivating response policies. They focus on the need of having deactivation mechanisms which allow the RS to stop applying responses when it calculates that the impact of the reaction surpasses an sufficient threshold. This proactive model bases its decisions on parameters such as the likelihood of the success of an on-going threat and the accumulative impact of the threat and the response.

Also in 2010, N. Kheir et al. [35] proposed a proactive solution based on dependency graphs. It extends the propagation process developed in [14], by integrating the evaluation of the impact of an attack on the security variables Confidentiality, Integrity and Availability (CIA). Each resource in the dependency graph has an associated CIA vector, where the variables are updated by active monitoring estimation mechanisms or by extrapolation. The dependencies in the graph can be structural or functional.

In 2013, S. Wang et al. presented in [36] a middleware RS, with the aim of providing cost-benefit security hardening. The authors' approach is the use of attack-graph models together with Hidden Markov Models to explore the probabilistic relation between system observations and states. With this probabilistic insight, the IDPRS runs heuristic searching algorithms for cost-benefit analysis, to determine the best security hardening measures available for the defense of the system.

Also in 2013, A. Shameli-Sendi and M. Dagenais published in [16] a cyber-attack RS using Accurate Risk Impact Tolerance (ARITO) (see Section 2.3.2). The main component of ARITO is the online risk assessment module, which evaluates in real-time the risk impact. This model also provides a feedback mechanism for retroactive response execution, it measures the goodness of the applied countermeasure, and indicates the new risk level after the application of the selected action(s).

M. Zaghdoud and M. Al-Kahtani describe in [37] an RS based on contextual fuzzy

cognitive maps and ontology-based knowledge representation. This IDRPS has three layers, the first layer uses ontologies to recognize the intrusions in the system. The second layer uses fuzzy cognitive maps to determine the context of the effect of the intrusion on the target system and to diagnose it. In the third layer, there are response agents that select the suitable remedies available and react in a passive (alerting the system administrator) or active (applying determined countermeasures) way.

In 2014, S.A. Zonouz et al. presented in [38] a game theory-based IDPRS, where the RS and the adversaries are modeled as opponents in a two-player stochastic game. Its cost-sensitive response selection method uses attack-response tree structures, which, solving partially observable competitive Markov decision processes, derive the optimal reactions in each case. The RS tries always to minimize the mathematical costs, while maximizing the benefit of the reactions applied. B. Fessi et al. specify in [39] a genetic algorithm-based IDPRS, fed by a double-IDS cooperative schema (network and host IDSs). This RS uses a weighted linear combination model to standardize the multiple attribute alternatives prior to the decision analysis. The genetic algorithms determine the appropriate countermeasures using a decision model, which takes into account the financial cost, reputation loss, and processing resources.

After our thorough review of the literature, we find that all of the aforementioned solutions, which implement active reaction mechanisms, have been developed for general purpose networks. It is possible to identify automatic IDS solutions with notification systems capable of sophisticated alerting processes [10]; but, they rely heavily on the presence of human operators to confirm or perform the countermeasure actions needed. Therefore, we can state that (to the best of our knowledge) in the public domain there are still no specific IDPRS solutions for the field of CIP which are capable of implementing active and automated response solutions.

Additionally, there are other solutions (manual RSs) for CIP that provide passive countermeasures in the form of notifications to the operators and they are capable of automatically performing harmless actions such as logging of data records for forensic analysis. However, and as we have stated, active IDPRS solutions are widely needed for the critical infrastructure protection [2]. Thus, it is necessary to employ more effort in developing balanced solutions for automated response in critical contexts.

4 Analysis of Solutions and Countermeasures

In the previous section, we described some of the different IDPRSs existing in the literature and their evolution from their first appearance in the 2000's to the present day. This section is dedicated to analyzing and assessing these solutions, evaluating the different methods used for intrusion response and prevention against potentially harmful events occurring in CIs. Taking into account the methodological framework and the taxonomy provided in Section 2, the IDPRSs are categorized according to the characteristics of the current solutions, and they are presented in Table 1.

The systems are chronologically ordered, and the fields of the taxonomy that do not apply to a specific solution are left blank. In Table 1, it is possible to appreciate that the majority of the existing solutions implement reactive mechanisms, executing the response actions without stopping (in a burst). We can also observe the trend over the

years to move from static solutions with respect to the calculation of the costs of the intrusion and the response, to more dynamic, even cost-sensitive ones. In recent years, we have seen several attempts to provide adaptive solutions, which help adjust to the specific situation of the system.

The review of the state of the art in the previous section covers solutions from diverse fields of knowledge and application. In this section, we also evaluate these solutions according to their possible application in the field of CIP. In order to do so, it is important to analyze the most common intrusion responses, and to provide a simple taxonomy to study which kinds of solutions better fit in with a critical scenario. Table 2, originally based on the study available in [4] and later extended for our purposes, provides an overview of the main types of responses that are found in the literature (note that this table is not exhaustive, since it leaves out most of the proprietary solutions).

In Table 2, and according to our IDPRS taxonomy (see Figure 1), we divide the possible countermeasures that an IDPRS can provide into: **passive** and **active** responses. In the first place, it is important to consider passive reactions, which are the most abundant solutions in the literature and they are usually included in the normal operation of some IDSs [4]. We classify **passive solutions** into three categories, namely: *administrator notification*, *prevention measures* and *others*. The first category corresponds to those systems whose mission is to log the system's information and state, and also alert the system's administrator or human operators to control the situation. As we have mentioned, notifications to administrators are the most common operations implemented in deployed IDS/IDPRS systems for CIP. Prevention measures are mechanisms that are sometimes present in protected systems. In our study, we have distinguished 6 main types of *preventive mechanisms*:

Table 1: Classification of the existing IDPRS solutions according to our taxonomy

IDPRS	Year	Response Cost	Attack Cost	Risk Assessment	Ability to Adjust	Cooperation Ability	Response Selection	Response Execution	Response Time
[23]	2000	Static			Static	Cooperative	Static Mapping	Burst	Reactive
[24]	2000	Static			Static	Autonomous	Static Mapping	Burst	Reactive
[25]	2000	Static	Static		Static	Autonomous	Cost Sensitive	Burst	Reactive
[20]	2000	Static Evaluated			Adaptive	Autonomous	Dynamic Mapping	Burst	Reactive
[26]	2001	Static			Static	Cooperative	Dynamic Mapping	Burst	Reactive
[27]	2001	Static			Static	Cooperative	Dynamic Mapping	Burst	Reactive
[28]	2001	Static			Static	Cooperative	Dynamic Mapping	Burst	Reactive
[21]	2002	Dynamic Evaluated	Static		Static	Cooperative	Cost Sensitive	Burst	Reactive
[29]	2002	Static	Static		Static	Autonomous	Cost Sensitive	Burst	Reactive
[13]	2003	Dynamic Evaluated	Dynamic	Dependencies	Static	Autonomous	Cost Sensitive	Burst	Reactive
[30]	2005	Static	Static		Adaptive	Autonomous	Cost Sensitive	Burst	Proactive
[31]	2005	Static Evaluated	Static		Static	Autonomous	Static Mapping	Burst	Proactive
[32]	2006	Static Evaluated	Static		Static	Autonomous	Cost Sensitive	Burst	Reactive
[33]	2007	Static Evaluated	Static		Adaptive	Autonomous	Cost Sensitive	Burst	Proactive
[15]	2007	Static	Dynamic	Attack Metrics	Static	Cooperative	Cost Sensitive	Burst	Proactive
[14]	2007	Dynamic Evaluated	Dynamic	Attack Graph	Static	Autonomous	Cost Sensitive	Burst	Reactive
[34]	2009	Static Evaluated	Static		Adaptive	Autonomous	Cost Sensitive	Burst	Reactive
[22]	2010	Static Evaluated	Dynamic	Attack Metrics	Static	Cooperative	Cost Sensitive	Retroactive	Reactive
[18]	2010	Static Evaluated	Dynamic	Dependencies	Adaptive	Autonomous	Cost Sensitive	Burst	Proactive
[35]	2010	Dynamic Evaluated	Dynamic	Dependencies	Static	Autonomous	Cost Sensitive	Burst	Proactive
[36]	2013	Dynamic Evaluated	Dynamic	Security Metrics	Static	Autonomous	Cost Sensitive	Burst	Reactive
[16]	2013	Dynamic Evaluated	Dynamic	Metrics	Adaptive	Autonomous	Cost Sensitive	Retroactive	Reactive
[37]	2013	Dynamic Evaluated	Dynamic	Graphs	Adaptive	Autonomous	Dynamic Evaluated	Burst	Reactive
[39]	2014	Dynamic Evaluated	Dynamic	Graphs	Adaptive	Cooperative	Cost Sensitive	Retroactive	Reactive
[38]	2014	Dynamic Evaluated	Dynamic	Graphs	Adaptive	Cooperative	Cost Sensitive	Retroactive	Reactive

Table 2: Taxonomy of intrusion response actions

Triggered Response	Type / Field of Response	Response Action	Details and Examples	
Passive	*Administrator Notification*	Generate Alarm	e-Mail	Online / Pager Notification
		Generate Report	Attack Target	Criticality
			Time	Source IP / User Account
			Description of Suspicious Packets	Intrusion Statistics
	Prevention Measures	Cryptography	Authentication Mechanisms	Masking: Dummy Operations
		Security Policies	Access Control	Information Protection
			User Account	
		Monitoring	IDS / IDPRS	Host-Based Vulnerability Scans
			Security Self-Awareness	
		Protective / Defensive Infrastructure	Demilitarized Zones	Firewall
			Proxy	
		Low-Level Prevention	Synchronized Docks	Directional Antennas
		Session Measures	Cookies	Packet-Leashes
	Others		Enable Local / Remote / Network Activity Logging	Enable Additional IDS
			Enable Intrusion Analysis Tools	Backup Tampered Files
			Trace Connection for Information Gathering	
Active	*Host - Based Response*	Operations on Files	Deny Full / Selective Access to File	Delete Tampered File
			Restore Tampered File from Backup	Allow to Operate on Fake File
		Operations on User Accounts	Restrict User Activity	Disable User Account
		Operations on Processes and Services	Shutdown Compromised Service/Host	Restart Suspicious Process
			Terminate Suspicious Process	Disable Compromised Services
			Abort Suspicious System Calls	Delay Suspicious System Calls
		Trust Mechanisms	Reputation	Credit and Token-Based Trust
	Network - Based Response	Disable / Block Operations	Block Suspicious Incoming / Outgoing Connections	Enable / Disable Firewall Rules
			Block Connection to Port / IP Addresses	
		Isolation Actions	Isolate / Quarantine Node	
		Routing	Multi-Hop Routing	Trust-Based Routing
			Secure Routing	Routing Discovery
			Evolutionary Algorithms for Routing Discovery + Trust	
		Deceiver Devices	Honeypots and Honeypets	

- Cryptography: using cryptography for data encryption is an effective approach to prevent attackers from understanding captured data. Bus and memory encryption increases the difficulty of successfully attacking a device. It is also possible to support the Operating System (OS) security by providing secure execution of cryptographic primitives as OS services. There are other mechanisms such as link-layer encryption and authentication, multi-path routing, identity verification, bidirectional link verification, and authenticated broadcast that can help protect networks against intrusions and attacks.

- Security Policies: define the measures taken by the organizations to provide security to their entity. Usually they address constraints, restrictions and rules imposed on the systems and on their operators/users. Since each organization is free to define or adopt the security policies to be implemented, there are many variants of them. Security policies for CIs are called *CIP policies*, and they follow governmental guidelines; IDPRS solutions, following these guidelines, can identify violations in the security policies of the surveilled system [3].

- Monitoring: the monitoring system par excellence is the IDS, which supervises the local (host or network) system's operations and state. IDS solutions for CIP implement different degrees of automation using varied types of detection engines, e.g., A. Carcano et al. [10] propose a state-based IDS using rules to detect complex attack scenarios; H. Lin et al. [9] have designed a specification-based IDS, relying on the formal specification of the system under surveillance to verify the correct use of the network packets.

- Protective/defensive infrastructure: consists of devices or system configurations designed as prevention mechanisms, capable of performing protection tasks. Following the National Infrastructure Security Co-ordination Centre (NISCC) good practices guidelines, this line of defense can include *firewalls*, *Demilitarized Zones* (DMZ) and *proxies* [3]. Firewalls are software or hardware security systems that control the incoming and outgoing network traffic based on specific rule sets. The DMZs are physical or logical subnetworks that provide services to an external untrusted network, allowing only the access from the outside to the DMZ nodes, but not to other nodes of the internal system. Proxies represent intermediary interfaces capable of helping their clients to make indirect connections to other network services, and filtering incoming requests. These measures are usually present in current CIs at several different levels of the SCADA infrastructures.

- Low-level preventive mechanisms: are physical measures or procedures implemented at the lower layers of the communication systems to prevent intrusions. Examples of such measures are directional antennas in wireless devices or synchronized clocks [40], which limit the possibilities of traffic interference and disruptions by restricting the access to the communication channels. These measures do not impose any overheads to the protected infrastructures, thus they are highly recommendable mechanisms to add to CIs.

- Session/communication measures: are techniques that add security at the session or communications level, e.g., packet leashes or cookies. A *leash* is the information added inside a packet to restrict its transmission distance [40]. *Cookies* are small pieces of data sent from a website and stored in the user's web browser to remember stateful information. These measures are frequent in general-purpose networks, however, they might not be present in CIs' industrial protocols. Nevertheless, they can be implemented in sections of the CIs' networks that use general-purpose communication protocols, e.g., the corporate networks.

Lastly, in Table 2 we provide another category to contain a broader range of passive response methods that are not directly related to the categories mentioned above. Examples are the activation of additional IDS, logging mechanisms, intrusion analysis tools, etc. It is important to note that we have reflected in Table 2 only a sample of the many possible countermeasures applicable to critical infrastructures, always taking into account the need to observe the constraints of CIs and the specific context where the response actions are applied. Concerning the **active reaction mechanisms**, we can divide them into two groups [4]: *host-based* and *network-based* response actions.

Host-based responses refer to those local procedures which perform operations to modify parameters or processes that run within the affected nodes, e.g., operations on files (restore, delete files), operations on user accounts (restrict activities, disable accounts), operations on processes and services (shutdown, restart, disable, abort, delay actions), and also trust mechanisms. Trust-based mechanisms such as the use of reputation, credit-based trust or token-based trust have gained relevance in the protection of information in communication networks [40]. Network-based response, conversely, corresponds to those activities performed in the communications network and that affect communication's services and parameters. Here we can differentiate responses such as disabling or blocking network operations, isolating segments of the network, modifying routing parameters or setting up deceiver devices [40].

Currently, IDS/IDPRS solutions designed specifically for CIP implement some of the passive methods present in Table 2, such as the generation of reports and logs for later forensic analysis, the presence of defensive infrastructure, security policies [9], etc. However, some CIs still lack important passive prevention mechanisms; for example, systems based on the Modbus TCP protocol, which is commonly used in SCADA and Distributed Control System (DCS) networks for process control, lack authentication of the source of a request. This gives a chance for an adversary to attempt to gather information on the industrial system and the network in general.

Active reaction solutions, consequently, are rarely applied to critical systems. There are two main reasons; firstly, CIs need stable well-behaved environments to perform their functions and it is critical that this restriction is always observed, since any change in this scenario could have a big impact on the correct operation of the infrastructures and cause disruptions in the critical services they provide. Thus any active reaction mechanism must not (directly or indirectly) cause disruptions in the normal operation of CIs. Secondly, the legacy systems and proprietary protocols and components traditionally present in CIs make the development of new security mechanisms very difficult for the academic community, therefore there is little research into this topic.

However, and despite these difficulties, it is necessary to develop and deploy ac-

tive response solutions in CIs, thus we have selected some of the possible countermeasure mechanisms that come from IDPRS solutions for general-purpose networks, which can be adapted to protect critical systems (see Table 2). We divide the responses into: host-based response, where the response system applies the security measures to the host system, such as modifications to the user accounts and permissions and trust mechanisms; and network-based response, focused on network operations, e.g., the installation of deceiver devices (honeypots, honeypets), active routing techniques, etc.

4.1 Discussion

In the previous sections, based on a thorough study of the literature, we have designed a framework to integrate IDPRS solutions in critical contexts such as the ICS networks of CIs. As seen in Figure 1, our methodological framework is composed of three main modules, each of them corresponding to the main requirements needed to be addressed in this type of networks: *detection*, *automation* and *response*. The detection module provides the system with insight about the behavior of the system and sends alerts in the case of occurring threats or anomalies. These processes can be performed by an IDS and, optionally, an alert management component.

The automation module (see Section 2.2) compiles the different levels of automation that can be present in an IDPRS solution [6]. The level of automation selected for the infrastructure will determine the components integrated in the third module, the RS. And, in the event that the ICS implements active response mechanisms, the methodological framework in its RS module categorizes the different characteristics that it can implement.

This third module of our framework is adapted from the work of N. Stakhanova et al. [4] and A. Shameli-Sendi et al. [8], which present taxonomies of RS designed for general-purpose networks. It has been our task to carefully analyze these subcomponents and properties in order to establish whether or not they can be used for CIs. From Section 2.3.1 to Section 2.3.7 we have presented our assessment and made recommendations on their use according to the characteristics of the networks where the different RSs can be deployed. We have therefore studied the general-purpose RS components present in these taxonomies, contrasted them with the literature on ICS protection systems, examined their constraints and scopes, and reflected in our framework the different possibilities for integrating IDPRS solutions in critical environments.

The methodological framework, together with our assessment and recommendations, integrates the most important components that should be included when designing an IDPRS solution for ICSs. The main objective of this framework is to create a framework focused on preparedness and response, to guide the design and development of IDPRSs for critical infrastructures. Some of the components are not present in current ICS-dedicated IDPRSs, however they can help address the challenges of prevention and protection set by the institutions around the globe [2].

Once the different components that can be present in the RS have been analyzed, we examine the possible countermeasures that the system can implement. In order to do this, we have reviewed the academic literature in search of different types of protective mechanisms. We did not restrict our search to ICS-related protection, but rather to

different types of networks, and we have tried to assess whether these measures can help protect critical environments. We discuss some of the academic non-proprietary methods of protection, providing examples of response actions in Table 2. We have divided them into passive and active solutions.

Passive methods, such as security policies, cryptographic mechanisms, etc., are put in place to prevent malicious attacks and anomalies within the system, addressing the need to enhance the *preparedness and prevention* in CIs, one of the five pillars of CIP identified by the European Network and Information Security Agency (ENISA) [2]. Active methods, such as dynamic routing techniques, are those deployed to address the need for more efficient early warning and response capabilities in the CIs, corresponding to the second pillar of CIP identified by ENISA, *detection and response* [2]. Therefore, and given these analyses on IDPRS constraints and countermeasure actions, it is possible to reach several conclusions and recommendations about the inclusion of IDPRS solutions in critical contexts:

- An IDPRS solution deployed in a critical context should implement both passive and active defense mechanisms, to be able to protect the surveilled infrastructure by providing early detection and reaction measures, but avoid any interference with its correct and normal operation. Most CIs are equipped with some type of IDPRS that mainly provides passive protection, however, both institutions [2] and researchers [7] report the need to have active response solutions in order to provide early detection and prevention mechanisms.

- Our review of the existing systems in the literature shows that, to the best of our knowledge, in the public domain there are currently no specific IDPRSs that implement automatic active response mechanisms for the protection of CIs. However, through the revision of more recent literature on IDPRS solutions for general-purpose networks, and through the analysis of the countermeasures taxonomy, it is possible to evaluate the advantages and disadvantages of the different implementations of their components, and determine the capabilities that they offer, and constraints they pose for CIs.

- Whenever the resources of the infrastructure are sufficiently powerful to allow the necessary computations, the decisions taken by the RS should be balanced, taking into account the costs of the anomalies and the costs of the response actions applied in each case. According to N. Stakhanova et al. [4], the ideal cost model should be evaluated dynamically. Due to a better accuracy achieved using dynamic evaluated cost models, the performance of the IDPRS increases and is able to select the optimal responses to threats, taking into account the interdependencies and criticality of the affected components of the system.

- It is generally better to provide proactive responses instead of delayed mechanisms, since the latter need to wait until the intrusion or the anomaly is finally confirmed to trigger the reaction. In the case of a cyber attack, this postponement of the response gives the adversary time to start and maybe complete malicious actions, and the system is susceptible to deteriorating fast and unleashing harmful cascading effects. In the case of an anomaly, the continuation of the normal

operation of the system in the presence of faults may make the errors cascade through the interdependent systems before the countermeasures take place. Contrarily, proactive mechanisms could detect changes in the dynamics of the system earlier and send an alert or try to correct them from an early stage. However, proactive responses for CIP must always be evaluated (taking into account costs and risks of the countermeasures) and preferably supervised by human operators.

- The response selection method implemented by the IDPRS is also important, as we consider that dynamic mapping and cost-sensitive mapping methods provide higher benefits to the IDPRS than the static mapping-based solutions. Dynamic mapping systems take into account metrics and policies to select adequate countermeasures, while cost-sensitive mapping systems consider the costs of the attack and the response to make better choices in the selection of the reaction triggered. Thus, these two mechanisms are more likely to perform better than a static selection procedure, and make the IDPRS more effective and safe for CIP.

- Risk assessment methods can help assess the costs and implications of the attacks and anomalies occurring within the system. An IDPRS including a risk assessment component can determine the impact of a given threat, but also the impact of the countermeasure actions that are applied to palliate the problem. This module can introduce feedback mechanisms into the RS, in order to execute the countermeasures gradually, preventing the responses applied from generating undesired effects. According to our analysis, and if the system has sufficient capabilities, it is better to provide the IDPRS with a dynamic online risk assessment component, since it will provide real time evaluation of the system's risk status.

- Our taxonomy of response actions (see Table 2), provides a general vision of the main passive and active protection mechanisms for critical contexts that can be implemented in an IDPRS. However, it is difficult to provide insight about which of the techniques perform better on a generic critical scenario, particularly as the active reaction mechanisms are usually not present in CIs. Nevertheless, we categorize the different techniques according to their nature and where are they applied, which will allow a further study of their suitability for CIs.

In light of these facts, it is possible to identify the needs that current IDPRSs have in the context of critical systems. As we have said, according to the institutions around the globe [2], it is necessary to tackle the issues of preparedness and prevention, and detection and response for the protection of the critical infrastructures. One of the technologies that help protect ICSs are the IDPRS solutions, since these tools serve as elements of detection and preparedness, and when implementing effective RSs, they act as prevention and response components.

Thus, we believe that further research is needed in order to develop efficient and cost-effective IDPRSs capable of automatic active reactions for CIP. Through the literature on general-purpose solutions, it is possible to examine the components and features usually present in IDPRSs and determine the ones that have the best fit for critical contexts. Through the analysis of their internal features we gain insight about their

capabilities and requirements (computational complexity, QoS demands, etc.), thereby identifying the ones that would behave better in a constrained environment.

Firstly, we find that there is a great need to include dynamic adaptive online evaluation of costs in the IDPRS for ICSs. Approaches that address this topic are mainly designed for general-purpose networks [16, 38, 39], instead of CIP. It would be especially useful to channel this evaluation of costs through a risk assessment component in charge of determining the impact of threats and responses in ICSs [19]. Additionally, these feedback mechanisms would positively affect the capability of IDPRSs to implement retroactive response execution procedures.

Few of the general-purpose RSs in the literature implement retroactive execution, something which would be very useful, particularly in critical contexts. Since CIs are very sensitive to changes in dynamics and configurations, the countermeasures must be applied most carefully. Retroactive feedback mechanisms would therefore be especially useful in this context, because they would allow the system to always measure the risk levels and evaluate the system's state before launching the responses. Therefore, research efforts should be put in place in order to develop effective risk and cost evaluation mechanisms capable of performing correctly in a critical environment.

In the same vein, it is important to consider improving proactive RSs. These systems should be capable of detecting subtle changes in the behavior patterns of the system, and detect and identify threatening dynamics for the ICS early on. Once these early detection mechanisms have been deployed and tuned to the system, they can trigger actions from the RS capable of stopping attacks and palliate system anomalies in order to avoid errors and threats to cascade through the infrastructure to other interdependent systems. The analysis and prevention of cascading (domino) effects is currently a very active area of research in the field of CIP and much effort is being made to tackle this problem.

It is our belief that IDPRSs capable of providing the characteristics mentioned above would be highly useful for CIP. Capabilities such as dynamic online risk and cost evaluation would make the control systems sufficiently trustworthy for ICS operators to transfer some of their functions and supervision tasks to the automated management of the infrastructure. Note that most critical countermeasures should be always supervised by human operators, however, in the event of a semi-supervised ICS implementing automatic countermeasure responses, the IDPRS could always report its cost and risk analysis to the operator to determine the best actions to take. It is our aim that this study will be useful as a guide to future design and development of IDPRS solutions specifically created to protect CIs.

5 Conclusions

In this paper we have studied the available tools and mechanisms capable of fulfilling the existing need of providing ICSs around the globe with automatic intelligent reaction mechanisms for defense and protection. To this end, we have reviewed the main characteristics that these defense systems need to implement, and the main constraints present in CIs, in order to analyze the application of IDPRS solutions for CIP. Initially, through an analysis of recent literature, we have identified the main compo-

nents that are normally present in general-purpose IDPRSs. The study of the modules and features provided by the currently available RSs have helped us gain insight into the applicability of these solutions to ICSs and the constraints imposed by the critical environments on the automatic response systems. Taking this analysis as our basis, we have composed a taxonomy model for IDPRS solutions in CIs, which contains the main elements and characteristics desirable for the protection of critical systems. We have also reviewed, analyzed and categorized the different active and passive countermeasure actions that IDPRS scan implement for protection, discussing their strengths and weaknesses in the context of CIP. From this study we have extracted general recommendations to guide the adaptation or development of new IDPRS solutions for critical contexts. As we have discussed in this paper, critical systems are in great need of automated intelligent solutions capable of providing early detection and protection mechanisms. Since the IDPRS solutions found in the literature that were specifically designed for CIP do not implement automatic active reaction mechanisms to the best of our knowledge, future work should be devoted to developing IDPRSs that are capable of automatically and actively protecting CIs, without interfering in their performance and normal operation.

Acknowledgments

The first author has been funded by a FPI fellowship from the Junta de Andalucía through the project FISICCO (P11-TIC-07223). The second author has received funding from the Marie Curie COFUND programme "U-Mobility" co-financed by Universidad de Málaga, the EC FP7 under GA No. 246550 and the Ministerio de Economía y Competitividad (COFUND2013-40259). The third author has been partially funded by the research projects PISCIS (P10-TIC-06334) and PERSIST (TIN2013-41739-R).

References

[1] C. Wueest, Targeted Attacks against the Energy Sector, Symantec Security Response, Mountain View, CA (2014).

[2] European Commission, COM(2009) 149 - Protecting Europe from Large Scale Cyber-Attacks and Disruptions: Enhancing Preparedness, Security and Resilience, Publications Office, 2009.

[3] K. Scarfone, P. Mell, Guide to Intrusion Detection and Prevention Systems (IDPS), NIST Special Publication 800 (2007) 94.

[4] N. Stakhanova, S. Basu, J. Wong, A Taxonomy of Intrusion Response Systems, International Journal of Information and Computer Security 1 (2007) 169–184.

[5] N. B. Anuar, M. Papadaki, S. Furnell, N. Clarke, An Investigation and Survey of Response Options for Intrusion Response Systems (IRSs), in: Information Security for South Africa, IEEE, 2010, pp. 1–8.

[6] L. Cazorla, C. Alcaraz, J. Lopez, Towards Automatic Critical Infrastructure Protection through Machine Learning, in: Critical Information Infrastructures Security, Springer, 2013, pp. 197–203.

[7] C. Alcaraz, J. Lopez, Wide-Area Situational Awareness for Critical Infrastructure Protection, IEEE Computer 46 (4) (2013) 30–37.

[8] A. Shameli-Sendi, N. Ezzati-Jivan, M. Jabbarifar, M. Dagenais, Intrusion Response Systems: Survey and Taxonomy, Int. J. Comput. Sci. Network Security 12 (1) (2012) 1–14.

[9] H. Lin, A. Slagell, C. D. Martino, Z. Kalbarczyk, R. Iyer, Adapting Bro into SCADA: Building a Specification-based Intrusion Detection System for the DNP3 Protocol, in: Proceedings of Cyber Security and Information Intelligence Research Workshop, ACM, 2013, p. 5.

[10] A. Carcano, I.Fovino, M. Masera, A. Trombetta, State-Based Network Intrusion Detection Systems for SCADA Protocols: a Proof of Concept, Critical Information Infrastructures Security (2010) 138–150.

[11] A. A. Cárdenas, S. Amin, Z. S. Lin, Y. L. Huang, C. Y. Huang, S. Sastry, Attacks Against Process Control Systems: Risk Assessment, Detection, and Response, in: Proceedings of Symposium on Information, Computer and Communications Security, ACM, 2011, pp. 355–366.

[12] International Standard Organization, ISO/IEC 27005 - Information Security Risk Management, ISO/IEC (2008).

[13] I. Balepin, S. Maltsev, J. Rowe, K. Levitt, Using Specification-Based Intrusion Detection for Automated Responses, in: Recent Advances in Intrusion Detection, Springer, 2003, pp. 136–154.

[14] M. Jahnke, C. Thul, P. Martini, Graph Based Metrics for Intrusion Response Measures in Computer Networks, in: Proceedings of Conference on Local Computer Networks, IEEE, 2007, pp. 1035–1042.

[15] K. Haslum, A. Abraham, S. Knapskog, DIPS: A Framework for Distributed Intrusion Prediction and Prevention Using Hidden Markov Models and Online Fuzzy Risk Assessment, in: International Symposium on Information Assurance and Security, IEEE, 2007, pp. 183–190.

[16] A. Shameli-Sendi, M. Dagenais, ARITO: Cyber-Attack Response System Using Accurate Risk Impact Tolerance, International Journal of Information Security (2013) 1–24.

[17] Y. Haimes, J. Santos, K. Crowther, M. Henry, C. Lian, Z. Yan, Risk Analysis in Interdependent Infrastructures, in: Critical Infrastructure Protection, Springer, 2007, pp. 297–310.

[18] W. Kanoun, N. Cuppens-Boulahia, F. Cuppens, S. Dubus, Risk-Aware Framework for Activating and Deactivating Policy-Based Response, in: Proceedings of Conference on Network and System Security, IEEE, 2010, pp. 207–215.

[19] R. W. Habash, V. Groza, D. Krewski, G. Paoli, A risk assessment framework for the smart grid, in: Electrical Power & Energy Conference, IEEE, 2013, pp. 1–6.

[20] J. Carver, A. Curtis, Adaptive Agent-Based Intrusion Response, Tech. rep., DTIC Document (2001).

[21] T. Toth, C. Kruegel, Evaluating the Impact of Automated Intrusion Response Mechanisms, in: Proceedings of Computer Security Applications Conference, IEEE, 2002, pp. 301–310.

[22] C. Mu, Y. Li, An Intrusion Response Decision-Making Model Based on Hierarchical Task Network Planning, Expert systems with applications 37 (3) (2010) 2465–2472.

[23] S. Mnsman, P. Flesher, System or Security Managers Adaptive Response Tool, in: Proceedings of DARPA Information Survivability Conference and Exposition, Vol. 2, IEEE, 2000, pp. 56–68.

[24] A. Somayaji, S. Forrest, Automated Response Using System-Call Delays, in: Proceedings of the 9th USENIX Security Symposium, Vol. 70, 2000.

[25] W. Lee, W. Fan, M. Miller, S. Stolfo, E. Zadok, Toward Cost-Sensitive Modeling for Intrusion Detection and Response, Journal of Computer Security 10 (1) (2002) 5–22.

[26] S. Lewandowski, D. V. Hook, G. O'Leary, J. Haines, L. Rossey, SARA: Survivable Autonomic Response Architecture, in: Proceedings of DARPA Information Survivability Conference, Vol. 1, IEEE, 2001, pp. 77–88.

[27] D. Schnackengerg, H. Holliday, R. Smith, K. Djahandari, D. Sterne, Cooperative Intrusion Traceback and Response Architecture (CITRA), in: Proceedings of DARPA Information Survivability Conference, Vol. 1, IEEE, 2001, pp. 56–68.

[28] X. Wang, D. Reeves, S. F. Wu, Tracing Based Active Intrusion Response, Journal of Information Warfare 1 (1) (2001) 50–61.

[29] S. Tanachaiwiwat, K. Hwang, Y. Chen, Adaptive Intrusion Response to Minimize Risk Over Multiple Network Attacks, ACM Trans on Information and System Security 19 (2002) 1–30.

[30] B. Foo, Y. Wu, Y. Mao, S. Bagchi, E. Spafford, ADEPTS: Adaptive Intrusion Response Using Attack Graphs in an e-Commerce Environment, in: Proceedings of Conference on Dependable Systems and Networks, IEEE, 2005, pp. 508–517.

[31] M. E. Locasto, K. Wang, A. D. Keromytis, S. J. Stolfo, Flips: Hybrid Adaptive Intrusion Prevention, in: Recent Advances in Intrusion Detection, Springer, 2006, pp. 82–101.

[32] M. Papadaki, S. Furnell, Achieving Automated Intrusion Response: A Prototype Implementation, Information Management & Computer Security 14 (3) (2006) 235–251.

[33] N. Stakhanova, S. Basu, J. Wong, A Cost-Sensitive Model for Preemptive Intrusion Response Systems, in: AINA, Vol. 7, 2007, pp. 428–435.

[34] C. Strasburg, N. Stakhanova, S. Basu, J. S. Wong, A Framework for Cost Sensitive Assessment of Intrusion Response Selection, in: Conference of Computer Software and Applications, Vol. 1, IEEE, 2009, pp. 355–360.

[35] N. Kheir, N. Cuppens-Boulahia, F. Cuppens, H. Debar, A Service Dependency Model for Cost-Sensitive Intrusion Response, in: ESORICS, Springer, 2010, pp. 626–642.

[36] S. Wang, Z. Zhang, Y. Kadobayashi, Exploring Attack Graph for Cost-Benefit Security Hardening: A Probabilistic Approach, Computers & Security 32 (2013) 158–169.

[37] M. Zaghdoud, M. S. Al-Kahtani, Contextual Fuzzy Cognitive Map for Intrusion Response System, Computer 2 (3) (2013) 471–478.

[38] S. A. Zonouz, H. Khurana, W. H. Sanders, T. M. Yardley, RRE: A Game-Theoretic Intrusion Response and Recovery Engine, IEEE Transactions on Parallel and Distributed Systems 25 (2) (2014) 395–406.

[39] B. A. Fessi, S. Benabdallah, N. Boudriga, M. Hamdi, A Multi-Attribute Decision Model for Intrusion Response System, Information Sciences 270 (2014) 237–254.

[40] M. Meghdadi, S. Ozdemir, I. Güler, A Survey of Wormhole-based Attacks and their Countermeasures in Wireless Sensor Networks, IETE Technical Review (2011).

Lorena Cazorla is a Ph.D. student at NICS Lab. She received a Master in Computer Science from University of Malaga in 2012, and currently she is working on her Ph.D. in Information Security with a focus on Critical Information Infrastructure Protection. Her research interests include Critical Infrastructure Protection, response systems and machine learning.

Cristina Alcaraz is a Postdoctoral Researcher at NICS Lab. She received her Ph.D. in Computer Science in 2011 from the University of Malaga. Her research activities are focused on Critical Information Infrastructure Protection, on secure monitoring of critical infrastructures, security of SCADA systems and Smart Grids, as well as the use of Wireless Sensor Networks for protection of critical systems.

Javier Lopez received his MSc and Ph.D. degrees in Computer Science in 1992 and 2000, respectively, from University of Malaga. He is Head of NICS Research Lab, and his research activities are mainly focused on network security and Critical Infrastructure Protection, leading a number of national and international research projects in those areas.