applied sciences

MDPI

Article

Ontology-Based System for Dynamic Risk Management in Administrative Domains

Mario Vega-Barbas [1,*], Víctor A. Villagrá [1], Fernando Monje [1], Raúl Riesco [1,2], Xavier Larriva-Novo [1] and Julio Berrocal [1]

[1] ETSI de Telecomunicación, Departamento de Ingeniería de Sistemas Telemáticos, Universidad Politécnica de Madrid, Avenida Complutense 30, 28040 Madrid, Spain; victor.villagra@upm.es (V.A.V.); f.monjer@alumnos.upm.es (F.M.); raul.riesco.granadino@alumnos.upm.es (R.R.); xlarriva@dit.upm.es (X.L.-N.); julio.berrocal@upm.es (J.B.)

[2] Spanish National Cybersecurity Institute (INCIBE), Edificio Incibe, Av. de José Aguado, 41, 24005 León, Spain

* Correspondence: mario.vega@upm.es

Received: 24 September 2019; Accepted: 23 October 2019; Published: 26 October 2019

Featured Application: The proposed system will allow for dynamic measurement and calculation of cybersecurity risk metrics in an administrative domain with just some asset estimations.

Abstract: With the increasing complexity of cyberthreats, it is necessary to have tools to understand the changing context in real-time. This document will present architecture and a prototype designed to model the risk of administrative domains, exemplifying the case of a country in real-time, specifically, Spain. In order to carry out this task, a modeling of the assets and threats detected by various sources of information has been carried out. All this information is stored as knowledge making use of ontologies, which enables the application of reasoning engines in order to infer new knowledge that can be used later in the following reasoning. This modeling and reasoning have been enriched with a dynamic system for managing the trust of the different sources of information and capabilities for increased reliability with the inclusion of additional threat intelligence information.

Keywords: dynamic risk assessment; cybersecurity metrics; ontologies

1. Introduction

Risk management systems are important cybersecurity tools for monitoring different attributes of an interactive environment to determine possible vulnerabilities the environment is under and, therefore, the security state of the environment. Usually, these kinds of supervision systems are based on a specification of risk levels for the set of attributes analyzed, determining, finally, the actions that are required to alleviate the overall risk to the environment.

Nowadays, risk management seeks to evolve the classical approaches of static risk analysis based on well-defined methodologies such as MAGERIT [1], which allow the characterization of an organization's risks. To do this, these approaches usually take a snapshot of the situation of the organization at a given moment, and then, security metrics and policies are generated. However, most of the parameters involved in a risk analysis are dynamic, constantly changing over time, so a static risk management approach can become obsolete quickly [2,3].

On the other hand, dynamic risk management systems attempt to adapt to these circumstances by including temporary variations in the elements that compose these analyses. Therefore, it is necessary to characterize these elements and their temporal variations and generate the appropriate outputs, often known as situational awareness systems. An example of the application of these dynamic systems is

the generation of metrics for calculating the level of risk in a global and complex environment, such as a country, a region, or a public sector [4–6]. In these cases, the environment's assets are unclearly defined as belonging to some administrative domain that operates with estimates of them.

In this sense, this main contribution of the article is the proposal of a formal model of an administrative domain, where assets are fuzzily defined, in order to dynamically measure the risk level of the global administrative domain or specific subdomains. The formal model is based on an ontology that formally defines all the involved components of the administrative domain (assets, threats, etc.) in order to define the adequate security metrics (which are expressed as formal rules to be executed dynamically in a reasoning engine over the instances of that ontology), which dynamically calculate the different risk levels defined for the domain or subdomains. In addition, this research will provide a methodology for defining the environment's assets, taking into account the limitation due to the blurred definition of the administrative domain and, therefore, of the amount of assets to be analyzed.

To address these objectives, this article presents a design of a system capable of handling several sources of information by means of a trustworthy, reliable management system capable of being replenished, generating new knowledge. To do this, the system will model the influence of the time elapsed since the appearance of each threat and its effect on the level of risk, while being able to represent it visually with understandable graphic display systems. Finally, in order to address the dynamism requirement, this proposed system will be able to operate in real-time.

Therefore, first of all, this article will analyze the state of the art in similar systems and approaches in order to validate its applicability to the problem. Subsequently, the proposed architecture is defined, and the design, development, and validation of a prototype carried out in a real environment is detailed.

So, with the purpose of presenting this paper's contribution, first, it will address an analysis of related works along Section 2. Then, Sections 3 and 4 detail the proposed system, presenting, first, the architecture design, and then an implementation prototype of it. In Section 5, a verification and validation process will be carried out. Finally, conclusions are pointed out in Section 6.

2. Related Works

Risk management systems based on dynamic analysis methodologies allow cybersecurity experts to supervise both with events in real-time and with new threats. For this, the use of methodologies has become a useful tool in the sense that they offer a broad vocabulary to express knowledge about the situation of an environment at various levels of detail, and thus, facilitating its modeling. In this sense, different proposals have been presented for modeling different domains related to this research, as well as the dynamic generation of security metrics.

As a starting point for the development of this research work, we find different works that have tried to generalize ontologies for the field of cybersecurity [7–9]. In [7], Herzog et al. define a generic security ontology specified in Web Ontology Language (OWL) that covers most aspects of an information security domain, providing a detailed vocabulary and reasoning capabilities by detailing the classification and relationship between all entities modeled. Fenz, S. contributed with [8] to the development of security ontologies by defining IT Security Measures. This research work pursued to be aligned with ISO 27004 standards and to be applied in real-world audit scenarios, as well as go further in the degree of automation. Obrst, L. et al. introduced in [9] an ontology for cybersecurity, as an extension of the Characterization of Attributes and Enumeration of Malware (MAEC), and using the Diamond Model of Malicious Activity as a development basis. Although these works offer a solid basis for managing to address the objective of this research, its general character prevents its application or direct use on the dynamic scope treated in this work.

Most related to the aim of this research, we found the work developed by Singapogu, S. et al., and is detailed in [10]. They proposed an ontology for conducting an enterprise risk assessment, supporting the IT security risk analysis process. However, the environment's assets addressed by this ontology are less representative of dynamic environments since the nature of an enterprise prevents to manage undefined or uncontrolled elements.

An example of an ontology applied to a general dynamic environment can be analyzed in [11]. Erbacher, R.F. presented in [11] a packet-centered ontology called PACO, which allows for representing and capturing atomic elements of communication networks, i.e., packets and packet sequences. The proposed model represents a basis for the development of holistic approaches.

On the other hand, the area of Threat Intelligence (TI) includes all the knowledge that is possessed on the possible threats to be able to make appropriate decisions. When shared information is technical, commitment indicators (IoC) become important. However, IoC-based schemes are inefficient as they are dependent on companies, and therefore, they are only supported during the existence of the related company. In contrast to this approach, standards such as STIXTM, TAXIITM, and CybOXTM have gained strength in the context of TI exchange because they provide a framework for cybersecurity indicators, threat characterization, and different options for information exchange. Analyzing and sharing information obtained through TI in an effective manner requires common representation, standards, and exchange protocols. Again, the use of ontologies arises as an interesting approach to address this problem, and therefore, several authors have carried out studies and developments in this sense [12–16].

Ekelhart, A. et al. made a contribution to ontologies in [12] by performing a quantitative risk analysis and visualizing the damage caused by certain threats, the cost of cutting, and the recovery time. The execution of the tool with additional safeguards shows its benefits and provides objective data for decision making on what safeguards to implement, and how to avoid the installation of non-economic countermeasures.

The research works presented in [13,14] represent a clear example of the use of ontologies for knowledge sharing in the field of TI. Vergara, JEL. et al. proposed in [13] an ontology-based model for sharing alerts between different information security management systems; while Syed, Z. et al. detailed in [14] a possible integration of STIXTM and ontologies for situational consciousness, which is a very interesting approach. The authors demonstrate the benefits for different use cases (vulnerabilities associated with PDF readers, suggestions for similar software, etc.) as a very interesting contribution, for example, to the verification of the impact of supplier change. Most recently, Riesco R. et al. proposed in [15] a new dynamic risk management and threat intelligence methodology for generating inference rules to be used in different application domains. The feasibility of this research was addressed in [16], where the authors proposed blockchain and Smart contracts as a solution for fostering cyber threat and risk intelligence exchange of information. Both works were based on STIXTM and developing a new semantic version of it, and this has been selected as the basis of the present research work.

Finally, it is important to underline the use of metrics in cybersecurity decision-making. The use of these metrics allows cybersecurity systems to know the effectiveness and impact of the implemented security controls and policies. We can analyze some research works in this field, such as in [17–19].

In [17], Goluch, A. et al. integrate an ontological concept of information security into the management of risk-related business processes. The ontology developed was based on NIST, and the authors provided subontologies for describing threats, vulnerabilities, and control. Communication between the ontology security web service and the risk-dependent simulation engine was structured by XML documents. In addition, this research work proposed including the threat to human life in case of risk as a key metric in order to improve and extend the classification of threats and subsequent decision making. Mateos, V. et al. proposed in [18,19] an Automatic Intrusion Response System (AIRS) based on ontologies, which infers the optimal answers at the network level.

The research works presented so far develop different ontologies to represent knowledge related to some aspects involved in the risk management process, or propose architectures based on security ontologies previously developed. However, these researches present some shortages in terms of dynamism awareness, and with a fuzzy definition of the assets. All the risk assessments are based on a concrete set of assets, but in administrative domains, it is not possible to have a detailed definition of assets.

Therefore, for all the above, in this analysis of related works, the research presented in this article has focused on the development of a risk assessment architecture capable of handling dynamic connections in which a knowledge engine will be supported by a new security ontology. Said ontology must consider the characteristics of the administrative domains, i.e., a fuzzy definition of assets and a high degree of dynamism in relation to the assets involved (people and devices), and be capable of processing new risks not contemplated during their implementation (new income risks). In this sense, some concepts of the work developed by Villagrá in [19] have been used for the risk assessment ontology, but the definition of the global ontology for modeling administrative domains and all the security metrics that are instantiated and executed as inference engine rules are a contribution of this article. In addition, the architecture should take into account past threats to determine the final risk by means of a risk history. And finally, unlike the analyzed works, continuous feedback should be carried out in order to be able to offer an accurate risk analysis.

3. System Architecture

The architecture presented is based on the use of ontologies. However, unlike the research works analyzed along the previous section, the proposed system aims to achieve effective feedback taking into account the context of the environment and past events, all this, handling a high flow of real-time data. To create a risk calculation system that can infer new information, it is first necessary to identify the elements involved in a dynamic risk management system:

- Identification of assets, that is, determine relevant assets of the administrative domain;
- Evaluation of assets based on quantitative and qualitative values given to previously identified assets;
- Threat identification;
- Threat assessment by means of parametrization of them, both their impact on assets and the likelihood that they will materialize;
- Risk assessment, which points out the risk level of the system.

In this way, a generic risk analysis schema could be followed, such as the one shown in Figure 1.

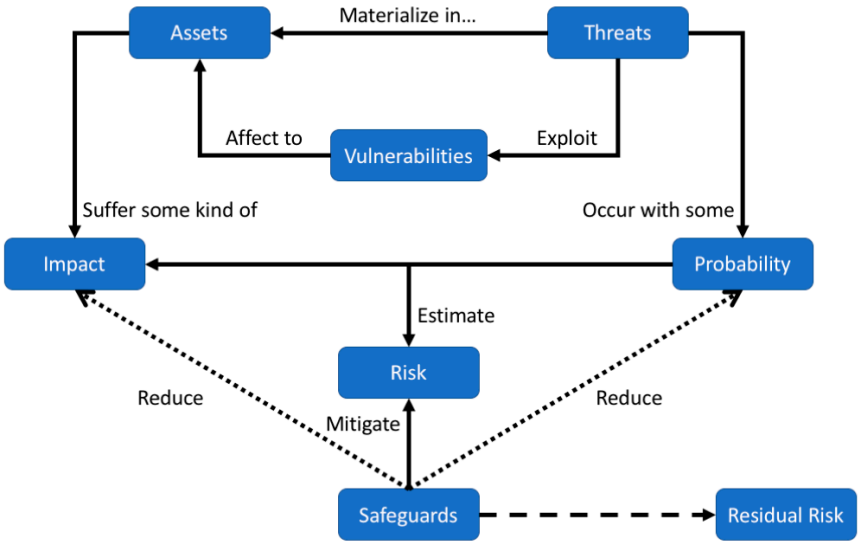

Figure 1. Generic risk analysis diagram followed for the development of the proposed architecture.

However, identifying and modeling assets in detail of an administrative domain, e.g., a country, a sector, etc., both in terms of their quantity and diversity, and the vulnerabilities of those assets that can be exploited by threats, is a challenging task. Therefore, it has been necessary to define an architecture development methodology, which is presented in Figure 2.

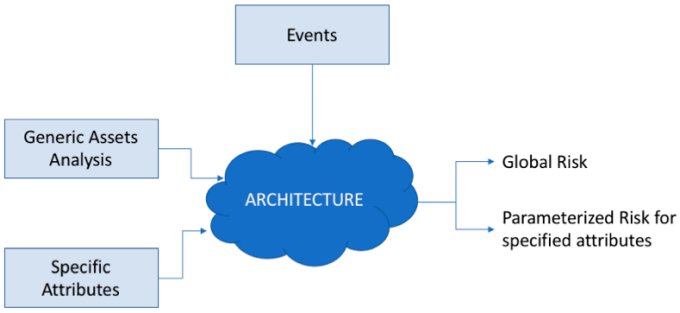

Figure 2. Design methodology followed for the architecture development.

As can be seen, the starting point is a generic analysis of the assets in the administrative domain, which are characterized by specific attributes. This information is modeled along with several events that arrive at the system representing different threats.

In addition, the architecture proposed must handle a large amount of information from several kinds of sources, so it is necessary to use appropriate techniques to deal with it while maintaining its semantics. For this reason, a system based on ontologies has been chosen capable of structuring this information and also to infer new knowledge, in this case, the level of risk. Therefore, a key task of the proposed architecture is focused on collecting relevant information in order to identify and evaluate both assets and threat events, and modeling it using an ontology language for future reasoning. For this, we selected OWL, the most widely used ontology language currently defined by W3C.

The asset modeling of an administrative domain for dynamic cybersecurity risk management has been structured in two main elements: citizens as human-being assets, and devices used by these citizens as technological assets. First, the different actors involved in the administrative domain have been defined in Table 1, with a classification according to their potential degree of exposure to cybersecurity risks. These actors (human-being assets) have been characterized by:

- Actor's technical and technological knowledge and the importance of the devices managed for the actor.
- Actor´s cybersecurity knowledge.
- Importance of the actor within the framework of the administrative domain.

Table 1. Classification of the different actors within the administrative domain.

Actor	Classification	Description
Citizen	Basic	Low technical, technological, and cybersecurity knowledge and sporadic use of technology, i.e., personal user of the technology without any incidence over the administrative framework.
Citizen	Average	Medium technical, technological, and cybersecurity knowledge and intensive use of technology, i.e., office work and automation.
Citizen	Advanced	Advanced technical knowledge with above-average technical capabilities and intensive use of technology, i.e., citizen with a high degree of impact on the administration and whose use of technology can have a direct impact on the framework.
Company	Self-Employed	General or specialized self-employed in the field of ICTs that has some kind of relationship with the administrative domain.

Table 1. *Cont.*

Actor	Classification	Description
Company	Small and Medium Enterprise (SME)	Independent firms, which employ fewer than 250 (European standards), focused on general aspects of the administration (standard SME) or very directed by ICTs (technological SME).
Company	Big Company	Companies that employ more than 250 or in which agreements or economic gains are considerably high.
Company	Industrial	Companies with industrial infrastructure and industrial control systems in an OT environment.
Company	Critical Infrastructure	Company whose niche market or contracts are related to critical infrastructure of the administration, i.e., provision of medicines, water management, defense, etc.
Research Center and University	-	Nonprofit organization focused on research or innovation.

On the other hand, we have identified another type of asset, the technological devices that human assets use to carry out their work or that they use in their daily lives. In some cases, these devices have an operating system that may be susceptible to threat. The assets corresponding to characterized devices are detailed in Table 2.

Table 2. Classification of the different devices related to each actor within the administrative domain.

Device	Penetration Operating System	Description
Mobile Phone	iOS and Android	Smart phones based on iOS or Android operating systems.
Computer	Windows, Linux, and MacOS	Any device oriented to develop work or entertainment tasks based on Windows, Linux, or MacOS.
Server	Windows Server and Linux	Workstations and computers dedicated to serving services from inside of the administration domain to external entities. Only servers based on Windows Server OS or Linux.
Router	Not applicable	Network device responsible for the interconnection of the internal network and the external network of an administration or between different LANs of the administration.
Switch	Not applicable	Network device responsible for the interconnection of different elements inside an internal network of an administration.
IoT device	Not applicable	Any kind of device with at least two features: internet access and computational capability. In the case of this research, smart watches and smart assistants (Alexa, Google Home, etc.) are included here.
Industrial Control Device	Not applicable	Devices involved within industrial control systems which are focused on controlling industrial process.

For example, a basic citizen will have low knowledge of cybersecurity, with a certain importance in the framework of a national or regional administrative domain, and with an adoption of the MacOS system close to 30% and 0% in the case of servers. On the other hand, in the case of big technological companies, servers will increase their adoption considerably, and the penetration of Linux computers will be higher than the case of basic citizens.

In order to infer the risk calculation considering the previous parameters, the architecture makes use of Semantic Web Rule Language (SWRL). This reasoner tool uses a set of rules based on an action-reaction approach in such a way that if an antecedent is fulfilled, a specific consequent will be triggered. In the case of the proposed architecture, if a detected threat affects a type of device

(antecedent), the consequent is executed, i.e., the impact on the assets of the affected device is calculated. In addition, the system is designed to receive threat events from various sources. Thus, a module has also been designed to characterize the trust in information sources. That means, given a source of information, its threat events will be considered with more or less relevance depending on its trust level. Figure 3 depicts the whole system described above.

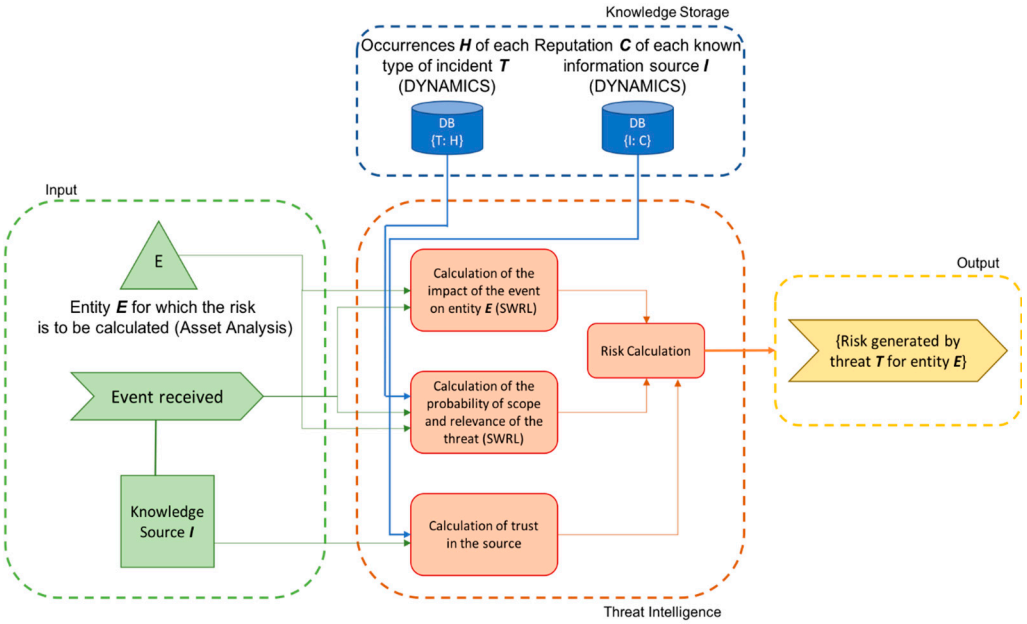

Figure 3. Overview of the architecture proposed.

This generic proposal can be adapted to the specific requirements of any administrative domain. In this case, because the project that supported the research presented in this article has been funded by the Spanish Ministry of Defense, it has been necessary to adapt its functionality to the idiosyncrasies of the administrative organization of that country. The Spanish Administration is divided into 15 autonomous communities and 2 autonomous cities (Ceuta and Melilla). In this sense, if threats are geolocalized, the system is able to calculate risks in the different geographical administrative domains. Thus, the system will calculate the following risks in different Spanish Administrative domains:

- Global risk of Spain.
- Risk of a specific asset in Spain.
- Global risk of an autonomous community.
- Risk of a specific asset in an autonomous community.

4. Formalization of the Architecture

A prototype based on the proposed architecture has been developed. This prototype has been designed to perform measurements and calculations with high regularity in order to obtain risk levels in a near real-time scenario. In terms of programming language, Java was selected because of its integration with ontologies by means of the OWLAPI [20] library.

One of the initial tasks to be performed is the formalization of the characterization of the assets defined in the previous section. For this, the prototype analyzes a configuration file, previously created, that contains the attributes that characterize each asset involved in the risk analysis process carried out. From all this information, and making use of OWLAPI, instances of assets are generated in the ontology.

From its part, Syslog [21] feed the system by means of threats, which are characterized following the CEF format. Thus, the system is able to collect the following information from the threats: date,

name of the threat (e.g., Andromeda), type of threat, source IP address, destination URL, severity, region code, and trust level of the detected threat.

In addition, the system should perform an enrichment of knowledge about the threat received by using external sources of TI, which enhances its correlation with asset analysis. Making use of INCIBE's AntiBotnet service, offered through the website of the Office of Internet Security (OIS), the system is able to collect information about types of devices affected by any threat. This search is carried out supported by a web scraper searching, which is ready to use keywords such as "Windows", "IoT", "Router", etc. In case any information is unfound, an assignment of affected devices is made based on probabilities defined within the system's configuration file. In both cases, the information is stored in separated databases, one for the threat found, and the other which registers threats for which the TI service used has no information. Figure 4 shows the flowchart that describes this functionality.

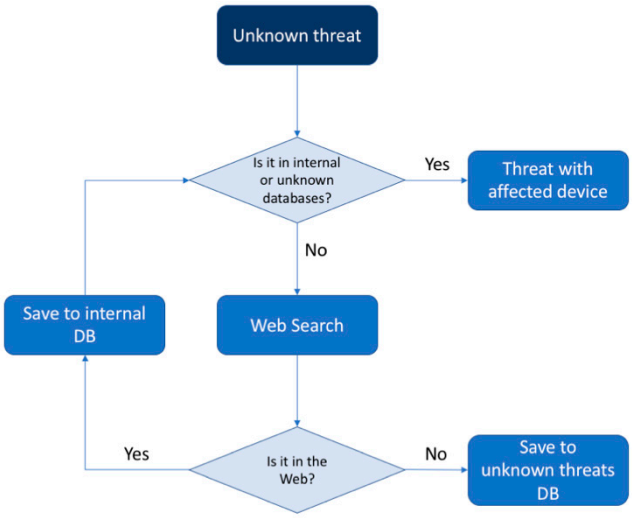

Figure 4. Graphic description of the method of enrichment of knowledge about threats followed in the development of the prototype.

With all this information, each threat event is modeled in OWL format so that it can be used by the reasoner. Figure 5 shows an example of how a threat instance would look by using Protégé.

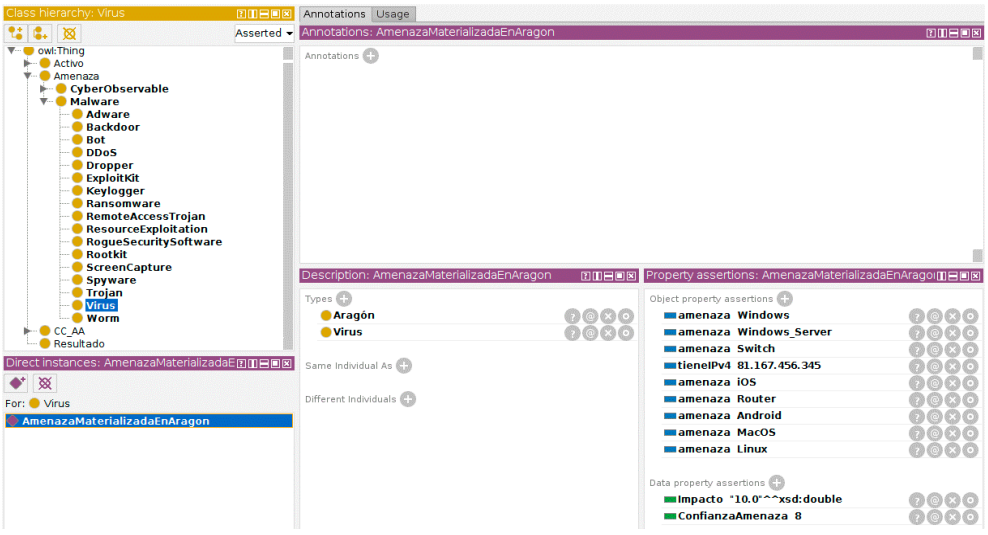

Figure 5. Information on a threat represented in an ontology.

Once both assets and threats are modeled in OWL format, the SWRL rules can be executed. These cover the whole combination between the different administrative domains, asset types, device types, and devices' operating systems, if applicable. The implemented system has a total of 2508 rules. Figure 6 shows an example of a rule used to analyze the risk of threat. In concrete, this risk was materialized on the servers of a critical company supported by a Windows Server, within the administrative domain of the autonomous community of Aragon. The analysis process described is detailed in Figure 7.

```
Modric2:Threat(?threatToAnalyze) ^
Modric2:Aragón(?threatToAnalyze) ^
Modric2:Impact(?threatToAnalyze, ?impact) ^
Modric2:Critical_Infrastucture(?kindOfIndividual) ^
Modric2:Threat(?threatToAnalyze, Windows_Server) ^
Modric2:CybersecKnowledge(?kindOfIndividual, ?knowledge) ^
Modric2:ThreatTrust(?threatToAnalyze, ?trust) ^
Modric2:WeightServer(?kindOfIndividual, ?weightDevice) ^
Modric2:Windows_Server_Server(?kindOfIndividual, ?adoptionSOpercentage) ^
swrlb:multiply(?result1, ?impact, ?knowledge) ^
swrlb:multiply(?result2, ?result1, ?adoptionSOpercentage) ^
swrlb:multiply(?result3, ?result2, ?trust) ^
swrlb:multiply(?finalRisk, ?result3, ?weightDevice) ^
Modric2:Server_Risk(Result_Aragón_Enterprise_Critical_Infrastucture, ?finalRisk)
```

Figure 6. Example of a Semantic Web Rule Language (SWRL) rule for risk calculation.

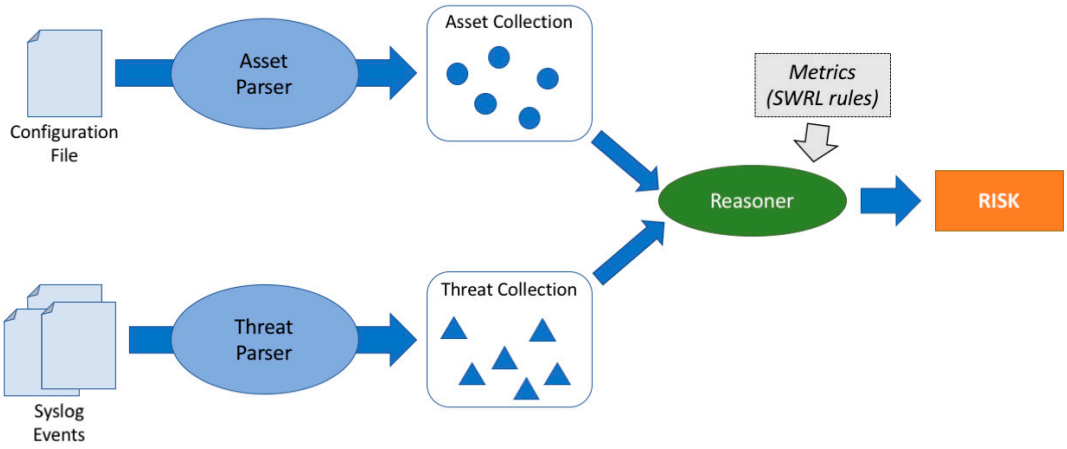

Figure 7. High level scheme of system operation.

However, each SWRL rule only makes an instantaneous contribution to the risk of an asset since it only calculates the risk on a specific device or operating system for a specific asset in a specific administrative domain. The system implemented has been applied to the geographical context of Spain, based on two administrative domains, a national level (Spain) and a community level composed of a set of autonomous communities (ACs) with their own identity. So, the system has to analyze the interrelations between all elements.

The sum of the values given by the rules concerning an asset in a specific AC will constitute the risk of that asset, without taking into account the threats materialized in the rest of the communities. Therefore, to obtain the risk of such an asset at the national level, it will be sufficient to apply Equation (1), where j indexes the type of asset, and i indexes each AC.

$$RiskCountry_j = \frac{\sum_i RiskAC_{j,i} * Asset_i}{\sum_i Asset_i} \quad (1)$$

In an AC, the weighted sum of assets, as specified in the initial configuration, will result in the risk of AC. Thus, the risk throughout the national territory will be obtained according to Equation (2).

$$Risk_{Spain} = \frac{\sum_i RiskCountry_i * Asset_i}{\sum_i Asset_i} \tag{2}$$

Although these operations determine the risk associated with Spain, they are useless for calculating the risk related to an AC or to an asset within an AC. This is due to the fact that (1) and (2) do not model that threats materialized in a given AC could be materialized within other ACs. In order to characterize this risk transfer, the correlation parameter ρ between AC is used, which indicates how much of the risk of the rest of the ACs affects the current AC analyzed. With this new appreciation, the risk of an asset j in an autonomous community k would be modeled by Equation (3).

$$RiskAC_{j,k} = \frac{Risk_{j,k} * Asset_k + \rho * \sum_{i \neq k} Risk_{j,i} * Asset_i}{\sum_i Asset_i} \tag{3}$$

In addition, the risk associated with an AC, k, could consider the risk of the rest of the ACs according to Equation (4).

$$RiskAC_k = \frac{Risk_k * Asset_k + \rho * \sum_{i \neq k} Risk_i * Asset_i}{\sum_i Asset_i} \tag{4}$$

Nevertheless, previous equations only offer instantaneous risk calculations as they operate on a measurement made at a specific time. In order to take into account the results of previous measurements, the system has been enhanced with memory capability. When a measurement is made, the risk levels are calculated following the process described above, but also the measurements made in the past are taken into account, and through the use of a time function, a new risk level is computed, which is no longer instantaneous. This function must also take into account the age of each past measure to avoid the collapse of the calculation due to the continuous growth of the measurement history. For this, the configuration file defines a forgetfulness variable ʒ that represents the number of temporary instants (seconds, milliseconds, etc.) that will determine when a measurement should be discarded, that is, the system memory. Equation (5) shows the function designed for this task. The behavior of (5) can be analyzed in Figure 8.

$$mem\_funct(t) = e^{-\frac{4t^3}{zh^3}} \tag{5}$$

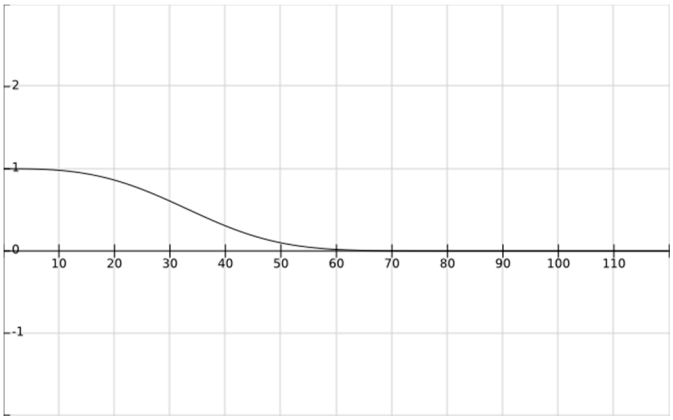

Figure 8. Representation of the memory management function proposed for a time of 3600 s (ʒ = 3600).

As can be seen, as the measures analyzed get old, they are penalized by this function. It is also necessary to weigh the number of threats analyzed in each measurement since the greater the number

of threats analyzed, the greater their representativeness in the histogram will have to be. The use of the function of time *f(t)* and the number of threats can be seen in Equation (6),

$$Current\ Risk = \frac{\sum_i Risk_{t_i} * Number\_threats_{t_i} * mem\_funct(t_i)}{\sum_i Number\_threats_{t_i} * mem\_funct(t_i)} \quad (6)$$

where t_0 is the current instant of time and therefore $f(t_0) = 1$.

Finally, once a measurement has been made and all the risk calculations have been performed, the result must be stored. In this sense, all results generated by the system are specified following a specific JSON format. In addition, a graphical interface has also been developed for a better visualization of the information that will be seen in more detail in the following section.

5. Proof of Concept and Validation

In order to test the capabilities of the system, it has been subjected to several tests. First, the system has been tested with a large number of incoming threats, being capable of analyzing each one in an average time of 0.01 s. Then, the system was also tested with threats from more than one source of information, some with a high degree of confidence and others with a very low level. At the end of each analysis, a JSON document with the results is generated. An example of this kind of document is shown in Figure 9.


```
CalculationDate:2019/03/25 17:02:39
riskSpain:3.990046153846283
riskSpainWithMemory:3.296827160476196
TotalNumberOfThreats:126
▾ numberOfThreatsOfInformationSources{2}
    TRUSTED:126
    INTERNET:0
▸ individuals{12}
▾ individualsWithMemory{12}
    Citizen_Average:4.0020403207625215
    Citizen_Advanced:3.66715404929009
    Citizen_Basic:4.557418226082124
     ⋮
▾ communities       [19]
   communityName:Andalucía
   communityRisk:2.1675076923077388
   communityRiskWithMemory:2.129086239298577
   numberOfThreatsInCommunity:48
   ▸ communityIndividuals{12}
   ▾ communityIndividualsWithMemory{12}
       Citizen_Average:2.2994025102746045
       Citizen_Advanced:2.094598559546031
       Citizen_Basic:2.593361070135563
            ⋮
```


Figure 9. Example of JSON generated at the end of a measurement.

Here, 126 threats were analyzed, which came from sources of information with a high level of confidence, and therefore, all of them were influenced by the confidence assigned to that source. As can be seen, there is a considerable difference between the general risk for Spain (riskSpain) and the risk when taking into account the history of attacks, that is, the memory of the system related to how old the attacks are by applying equation 5 (riskSpainWithMemory). If this situation were maintained over time, the risk with memory would end up converging with the instantaneous risk. In addition, the risk within an AC, i.e., Andalusia, is calculated based on 48 threats, both with and without memory.

In order to simplify the visualization of the data produced by the system, a graphical interface was implemented by using Java's library JFreeChart. Figure 10 shows the history of the general risk in Spain, both with memory and the instantaneous one.

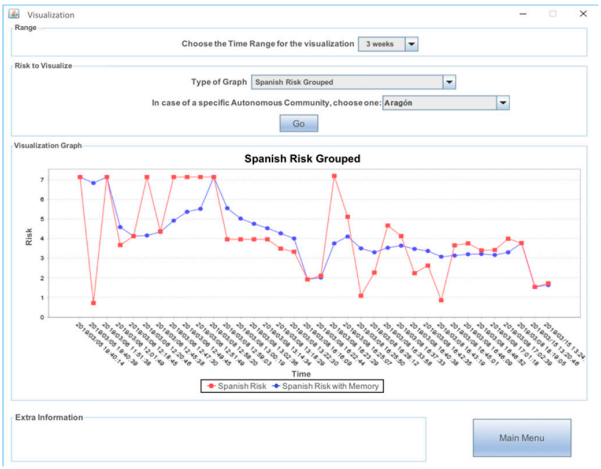

Figure 10. Visualization of the risk history of Spain. The first section allows the analyst to select a temporary viewing window, in this case, 3 weeks of risk history. The second section offers the option of selecting the graph to be displayed; in this case, the grouped risk of Spain. Finally, the lower section shows the visualization of the selected data in the time window defined at the top, differentiating the instantaneous risk (square red), and the risk with memory (blue circle line).

As can be observed, the risk with memory tends to follow the trend of the instantaneous, but if the variations are very fast, it tends to represent the average of these. The strongest peaks of memory risk are due to the fact that certain measurements are composed of a large number of threats and, therefore, great representativeness. The history described in Figure 10 is composed of 36 risk measurements, each generated with a range of threat events from 100 to 500. Each threat event has been compounded by its geographic location, IP, a timestamp, severity index, the event confidence, name of the detected threat, etc. The threats simulated were balanced in terms of impact level and confidence in the source of information. This is why the instantaneous trend, represented by the red line, has such pronounced peaks.

The system also provides the analyst with a detailed asset histogram, which can be grouped by a specific AC, as Figure 11 shows. Finally, the risk in real-time can also be observed. In this sense, Figure 12 shows the risk in real-time as measurements are produced.

Figure 11. Visualization of the histogram of the risks of the different autonomous communities (ACs). The upper section offers a selection menu to determine the temporal analysis window, in this case, 3 weeks. The second section allows you to select the administration you want to view, being able to select one or all ACs. The lower section shows the selected display. This example shows a histogram that includes all the ACs of Spain.

Figure 12. Real-time visualization of the global risk of Spain. The current risk is shown in red over a maximum possible in blue.

6. Conclusions

The research work presented within this document details a solution to the problem of obtaining the level of risk in an administrative domain in real-time. As the main result, this research provides a system architecture ready to infer in real-time security risk metrics for environments with a diffuse definition of assets. To solve this, the operation domains have been modeled by using an organic definition of assets, dividing them into citizens, companies, and university and research organizations, and characterizing them with attributes, e.g., knowledge about cybersecurity, impact over the administrative domain, and the adoption and impact of each type of device. In order to correlate the threats detected with assets, groups of devices have been defined, e.g., mobiles, personal computers, and IoT devices.

By using ontologies and SWRL rules, the system can process the relationship between assets, threats, and both. That means, for example, if the system processes a threat event that affects Android, the semantic reasoner can infer all assets related to a mobile phone with this operating system will be affected. In addition, the system architecture has been designed to be aware of several information sources about threat events, setting its utility dynamically in terms of the trust of the source.

To test the viability of the system architecture proposed, a proof of concept based on the development of a prototype has been carried out. This prototype has been validated against tests performed in a real operation environment, the Spanish administrative domain. Although this is a specific case study, the level of abstraction of the proposed metrics allows its application to be applied in any administrative domain, as long as an adequate adaptation of the domain characteristics and asset models is made during the initial configuration process.

Author Contributions: Conceptualization, V.A.V., R.R., and J.B.; methodology, M.V.-B., V.A.V., and F.M.; software, F.M., R.R., and X.L.-N.; validation, M.V.-B., and V.A.V.; investigation, V.A.V., M.V.-B., and F.M.; data curation, F.M., R.R. and X.L.-N.; writing—original draft preparation, F.M., M.V.-B. and V.A.V.; writing—review and editing, M.V.-B. and V.A.V.; supervision, V.A.V., J.B. and M.V.-B.; project administration, V.A.V.; funding acquisition, V.A.V., and J.B.

Funding: This work has been partially funded with the support of the Spanish MINECO (DHARMA project, Dynamic Heterogeneous Threats Risk Management and Assessment, with code TIN2014-59023-C2-2-R) and by the European Commission (FEDER/ERDF).

Acknowledgments: The authors are grateful for the support provided by the National Institute of Cybersecurity (INCIBE) in the undertaking of this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lucero Gómez, A.J.; Valverde Padilla, J.O. Análisis y Gestión de Riesgos de los Sistemas de la Cooperativa de Ahorro y Crédito Jardín Azuayo, Utilizando la Metodología Magerit. Bachelor's Thesis, Universidad de Cuenca, Cuenca, Ecuador, 2012.
2. Li, J.; Ou, X.; Rajagopalan, R. Uncertainty and risk management in cyber situational awareness. In *Cyber Situational Awareness*; Springer: Heidelberg, Germany, 2010; pp. 51–68.
3. Williams, P.A.H.; Woodward, A.J. Cybersecurity vulnerabilities in medical devices: a complex environment and multifaceted problem. *Med. Devices (Auckland, NZ)* **2015**, *8*, 305. [CrossRef] [PubMed]
4. Wirtz, B.W.; Weyerer, J.C. Cyberterrorism and Cyber Attacks in the Public Sector: How Public Administration Copes with Digital Threats. *Int. J. Public Adm.* **2017**, *40*, 1085–1100. [CrossRef]
5. Coppolino, L.; D'Antonio, S.; Mazzeo, G.; Romano, L.; Sgaglione, L. How to Protect Public Administration from Cybersecurity Threats: The COMPACT Project. In Proceedings of the 2018 32nd IEEE International Conference on Advanced Information Networking and Applications Workshops (WAINA), Cracow, Poland, 16–18 May 2018; pp. 573–578.
6. Crovini, C.; Ossola, G.; Marchini, P.L. Cyber Risk. The New Enemy for Risk Management in the Age of Globalisation. *Manag. Control* **2018**, *2*, 135–155. [CrossRef]
7. Herzog, A.; Shahmehri, N.; Duma, C. An ontology of information security. *Int. J. Inf. Secur. Priv.* **2007**, *1*, 1–23. [CrossRef]
8. Fenz, S. Ontology-based generation of IT-security metrics. In Proceedings of the 2010 ACM Symposium on Applied Computing, Sierre, Switzerland, 22–26 March 2010; pp. 1833–1839.
9. Obrst, L.; Chase, P.; Markeloff, R. Developing an Ontology of the Cyber Security Domain. In Proceedings of the STIDS, Fairfax, VA, USA, 23–26 October 2012; pp. 49–56.
10. Singhal, A.; Singapogu, S. Security ontologies for modeling enterprise level risk assessment. In Proceedings of the 2012 Annual Computer Security Applications Conference, Orlando, FL, USA, 3–7 December 2012.
11. Ben-Asher, N.; Oltramari, A.; Erbacher, R.F.; Gonzalez, C. Ontology-based Adaptive Systems of Cyber Defense. In Proceedings of the STIDS, Fairfax, VA, USA, 18–20 November 2015; pp. 34–41.
12. Ekelhart, A.; Fenz, S.; Klemen, M.; Weippl, E. Security ontologies: Improving quantitative risk analysis. In Proceedings of the 2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07), Waikoloa, HI, USA, 3–6 January 2007; p. 156a.
13. De Vergara, J.E.L.; Villagra, V.A.; Holgado, P.; De Frutos, E.; Sanz, I. A semantic web approach to share alerts among Security Information Management Systems. In Proceedings of the Iberic Web Application Security Conference, Madrid, Spain, 10–11 December 2009; pp. 27–38.
14. Syed, Z.; Padia, A.; Finin, T.; Mathews, L.; Joshi, A. UCO: A unified cybersecurity ontology. In Proceedings of the Workshops at the Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016.
15. Riesco, R.; Villagrá, V.A. Leveraging cyber threat intelligence for a dynamic risk framework. *Int. J. Inf. Secur.* **2019**, *18*, 715–739. [CrossRef]
16. Riesco, R.; Larriva-Novo, X.; Villagra, V.A. Cybersecurity threat intelligence knowledge exchange based on blockchain. *Telecommun. Syst.* **2019**, *72*, 1–30. [CrossRef]
17. Goluch, G.; Ekelhart, A.; Fenz, S.; Jakoubi, S.; Tjoa, S.; Muck, T. Integration of an ontological information security concept in risk aware business process management. In Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008), Waikoloa, HI, USA, 7–10 January 2008; p. 377.
18. Lanchas, V.M.; González, V.A.V.; Bueno, F.R. Ontologies-based automated intrusion response system. In *Computational Intelligence in Security for Information Systems 2010*; Springer: Heidelberg, Germany, 2010; pp. 99–106.
19. Mateos, V.; Villagrá, V.A.; Berrocal, J. Application of ontologies and formal behaviour definitions for automated intrusion response systems. *J. Res. Pract. Inf. Technol.* **2014**, *46*, 197.

20. Horridge, M.; Bechhofer, S. The owl api: A java api for owl ontologies. *Semant. Web* **2011**, *2*, 11–21.
21. Abe, H.; Shima, K.; Miyamoto, D.; Sekiya, Y.; Ishihara, T.; Okada, K.; Nakamura, R.; Matsuura, S. Distributed Hayabusa: Scalable Syslog Search Engine Optimized for Time-Dimensional Search. In Proceedings of the Asian Internet Engineering Conference, ACM, Phuket, Thailand, 7–9 August 2019; pp. 9–16.