

---

## Trust and reputation management for critical infrastructure protection

---

Filipe Caldeira\*

CISUC – DEI, University of Coimbra,  
Coimbra 3030-290, Portugal  
and  
Polytechnic Institute of Viseu,  
Viseu 3504-510, Portugal  
E-mail: fmanuel@dei.uc.pt  
\*Corresponding author

Edmundo Monteiro and Paulo Simões

CISUC – DEI, University of Coimbra,  
Coimbra 3030-290, Portugal  
E-mail: edmundo@dei.uc.pt  
E-mail: psimoes@dei.uc.pt

**Abstract:** Today's critical infrastructures (CIs) depend on information and communication technologies (ICTs) to deliver their services with the required level of quality and availability. ICT security plays a major role in CI protection and risk prevention for single and also for interconnected CIs were cascading effects might occur because of the interdependencies that exist among different CIs. This work addresses the problem of ICT security in interconnected CIs. Trust and reputation management using the policy-based management paradigm is the proposed solution to be applied at the CI interconnection points for information exchange. The proposed solution is being applied to the Security Mediation Gateway being developed in the scope of the European FP7 MICIE project, to allow information exchange among interconnected CIs.

**Keywords:** CIs; critical infrastructures; ICT security; trust and reputation management; PBM; policy-based management.

**Reference** to this paper should be made as follows: Caldeira, F., Monteiro, E. and Simões, P. (2010) 'Trust and reputation management for critical infrastructure protection', *Int. J. Electronic Security and Digital Forensics*, Vol. 3, No. 3, pp.187–203.

**Biographical notes:** Filipe Caldeira is an Assistant Professor at the Department Computer Science, Polytechnic Institute of Viseu, Portugal. He graduated in Informatics Engineering from the University of Coimbra, Portugal and received his Master degree in Informatics Engineering from the same university where, currently, he is a PhD Student in Informatics Engineering. His research interests include ICT security, namely, policy-based management, trust and reputation systems and critical infrastructure protection.

Edmundo Monteiro is an Associate Professor with Habilitation at the University of Coimbra, Portugal, from where he got a PhD in Electrical Engineering,

Informatics Specialty, in 1995. His research interests are computer communications, service oriented infrastructures and security. He is the author of several publications including books, patents and over 200 papers in national and international refereed books, journals and conferences. He is a Member of the Editorial Board of several scientific journals and involved in the organisation of many international conferences and workshops and he is also a Member of IEEE Computer, IEEE Communications and ACM Communications groups. He participated in several research projects and research programmes of the European Union, various European countries, such as USA, Latin America and national organisations and companies, in the areas of information and communication technologies (ICT), critical infrastructures protection, ICT systems for energy efficiency, research and innovation, and human resources and mobility.

Paulo Simões is an Assistant Professor and Senior Researcher in the Department of Informatics Engineering at the University of Coimbra, Portugal, from where he got a PhD in Informatics Engineering, in 2002. His main research interests are network and systems management, security, computer communications and protection of critical infrastructures. He has over 70 journal and conference publications in these areas. He has been involved in several European research projects, both with technical and management activities. He also participated in several industry-funded research projects and he was co-founder of two technological spin-off companies that currently employ more than 60 persons.

*This paper is a revised and expanded version of a paper entitled [Trust and reputation management for critical infrastructure protection] presented at [ICGS3 2010 – 6th International Conference on Global Security, Safety and Sustainability, Braga, Portugal, 1–3 September 2010].*

---

## 1 Introduction

As defined by US Administration, “Critical infrastructures are those physical and cyber-based systems essential to the minimum operations of the economy and government. These systems are so vital, that their incapacity or destruction would have a debilitating impact on the defense or economic security” Clinton (1996).

From the above definition, it is clear that critical infrastructures (CIs) are one of the most information and communication technology (ICT) dependent areas of contemporary societies where we should ensure the highest security levels.

Growing interest on this matter is clear from governments initiatives, such as the critical infrastructure protection (CIP) Programme, launched in 1998 by US Administration, and the European Programme for Critical Infrastructure Protection (EPCIP) launched by the European Commission in 2006. Also citizens are becoming aware and concerned about this problem due, e.g. to a recent television series, ‘24 – season seven’, where the fiction character Jack Bauer fights a terrorist group intending to destroy some US CIs. Apart from the fiction involved, this TV series clearly demonstrates how important those infrastructures are and how weak they can be.

The EPCIP, whose main goal is to improve the protection of CIs in the European Union Commission (2006), intends to create procedures for European CIs identification; creation of expert groups; implementation of the critical infrastructure warning information network

(CIWIN); CI information sharing frameworks and the identification and analysis of CIs interdependencies. CIWIN main objective is to provide a platform for the exchange of rapid alerts, to help EC Member States and CI operators to share information on common threats and vulnerabilities.

Recent efforts have been focusing on each CI individually, launching the basis for more secure CIs with enhanced robustness, security and resiliency, introducing, e.g. fault-tolerant architectures, redundancy of components and resilient IT systems. One important aspect that still needs to be addressed relates to the interdependency existent among CIs. This interdependency can lead, in a extreme situation, to a global failure started by a single trivial incident in one CI (cascading effect).

Although large efforts have been made in modelling CIs risk analysis, the valuable information gathered from those models are still kept inside each CI and is not shared among interdependent CIs.

In this context, the lack of sharing mechanisms for risk information between interconnected CIs was identified. Those mechanisms will allow CI operators to have a real-time view on the risk level associated to services on which the modern society depends, such as power, water supply or communication lines. This shared information is also important to increase accuracy of CI risk models introduction external failures risks on those models Simões et al. (2010).

Use of mechanisms for sharing risk information can, along with more resilient CIs, increase the security level of multiple interdependent CIs. To achieve these service levels, a robust, resilient and inter-dependencies-aware alerting system need to be designed and implemented. This is the main goal of MICIE (tool for systemic risk analysis and secure mediation of data exchanged across linked CI information infrastructures) FP7-ICT project, aiming the design and implementation of a real-time risk level dissemination and alerting system MICIE (2008).

In this work, we briefly present the MICIE alerting system, the MICIE Secure Mediation Gateway (SMGW) and the solutions to specifically incorporate CI interdependencies in the online risk assessment framework. This alerting system is a core component of the MICIE FP7 project and is in line with the European initiative to establish a CIWIN (Commission, 2006).

The main contribution of this paper, an extended version of previous work presented in Caldeira et al. (2010c), is on the definition of the ICT security mechanisms for information exchange among CIs, namely, the trust and reputation mechanisms to be applied to the SMGWs.

The rest of this paper is organised as follows. In Section 2, we discuss related work focusing on relevant European projects in the CI area. The key components of the MICIE alerting system (including the SMGW) are presented in Section 3. Section 4 describes our approach to trust and reputation management. Section 5 discusses advantages and disadvantages of our proposal. In Section 6, we drawn some conclusions and discuss some directions for future work.

## 2 Related work

As already discussed, it is important to study and develop models for the existent interdependencies among CIs. The work of Rinaldi et al. (2001) provides an overview on the dimensions in which interdependencies can occur. Also, we can find several work related to CI modelling, namely, Sokolowski et al. (2008) use conceptual modelling to

represent an abstract, simplified view of CIs, and Panzieri et al. (2005) apply the CAS approach for CI modelling.

There are several international projects that have already addressed issues similar to the ones targeted by MICIE, but with relevant differences. European projects IRRIS (2009) and CRUTIAL (2008) are two relevant projects in this field.

CRUTIAL approach gives particular importance to CIs interdependency modelling to increase CI resilience to faults and attacks, (CRUTIAL, 2008). CRUTIAL main objectives were development of a new architecture and models applicable to heterogeneous CI; analysis of application scenarios in which a small failure could cause a high impact in the CI; research and development of distributed architectures that enable an efficient and reliable control on an electrical power distribution infrastructure. The overall CRUTIAL expected results were to gather better knowledge about CIs, enabling the development of more resilient infrastructures (Dondossola et al., 2008; Verissimo et al., 2008).

According to IRRIS (2009), IRRIS intends to develop mechanisms in order to raise the confiability, survivability and resilience of information systems related to CIs. Two main scenarios were defined, representing, respectively, a telecommunications infrastructure and a electricity production and distribution network. For each scenario, the way CI connect to the exterior by the use of convergent networks like the internet was analysed (IRRIS, 2009). Work was also developed in areas related to online risk prediction tools able to incorporate CI interdependencies, studying multiple approaches for data sharing across CIs, interdependency modelling and risk estimators (Balducelli et al., 2008).

The IRRIS project has developed a set of applications named middleware improved technology (MIT) (Balducelli et al., 2008). That made possible the communication between heterogeneous CIs. MIT main objective is to enable a simple, fast and reliable information exchange between CIs thus reducing response time to incidents that may occur in the CI by maintaining infrastructure managers well informed about the CI state (IRRIS, 2009).

Both CRUTIAL and IRRIS projects provided a strong contribution to CIP, but none of them fully addressed the problem of real-time information exchange for real-time CI risk prediction and the security issues associated with the exchange of information, among CIs adding trust and reputation. Also, for the best of our knowledge, this is the first use of policy-based management for the control of the trust and reputation on the exchanged information.

Related work in trust and reputation is mostly focused on the development of trust models with application in areas like e-commerce websites or, more generally, in situations where transactions between not identified systems or people occurs (Artz and Gil, 2007). Most of the work is focused in peer to peer (P2P) systems (Chen et al., 2009), wireless sensor networks (Zahariadis et al., 2008), online personal interactions, software agents and in evaluating generic models and formalisms for trust and reputation systems (Jøsang et al., 2007). In Spitz and Tuchelmann (2009) and Aime and Liroy (2005), authors present specific methodologies to evaluate trust and reputation, introducing factors, such as ageing of the observed values and time of the observations.

In our previous work (Caldeira et al., 2010b), we proposed a methodology for evaluating trust and reputation in the context of information exchange among CIs.

### **3 MICIE alerting system**

In line with European developments in the CIP field and with the CIWIN initiative, the MICIE project aims to contribute in three main areas: the identification and modelling of

interdependencies among CIs; development of risk models and risk prediction tools and the development of a framework enabling secure and trustfully information sharing among CIs (Capodieci et al., 2010).

The main goal of the MICIE alerting system is to provide, in real-time, each CI operator with a prediction mechanism, measuring the probability that, in the future, a service loss or degradation occurs in the interconnected CIs. The MICIE overall architecture is presented in Figure 1.

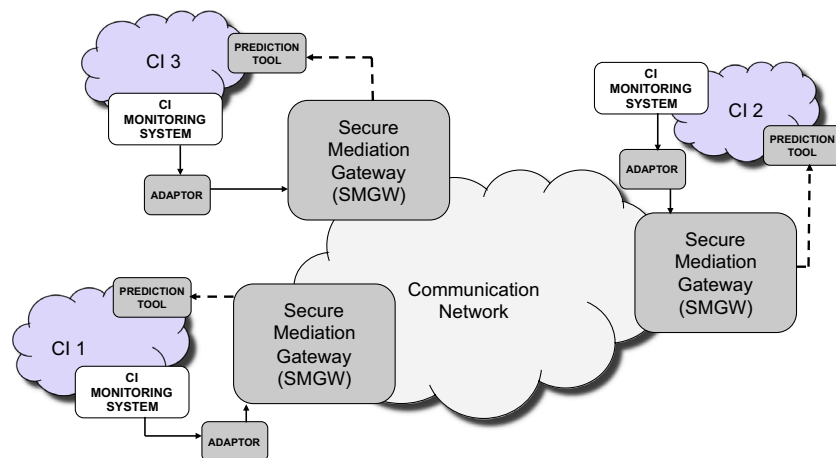
To achieve accurate risk prediction, MICIE system needs information regarding services provided by its own CI and also services provided by interdependent CIs. Following the service-oriented architecture (SoA) approach, the interdependencies among CIs can be modelled as services. Also internal relations among components or entities in one CI are, in this context, treated as services. Using this approach it is possible to introduce the notion of Quality of Service (QoS) expressing the risk of service failure (total or partial, internal or external).

The first step to make risk prediction is to discover all the distributed information that is relevant for the alerting system. Each CI collects, in real time, information regarding the status of its components. This information is filtered and adapted by an specific adaptor (see Figure 1). The adaptor selects proper information and handles format and semantic conversions. The next step is performed by the prediction tool. This tool makes use of risk models to assess the risk level of the monitored services. In this entity, CI own risk level is associated with the risk levels for services received from partner (interdependent) CIs. Each CI using MICIE has at least one prediction tool.

Status information as results from the prediction tools can be exchanged across partner CIs using a SMGW allowing CIs to work in a fully cooperative distributed environment for risk prediction.

SMGW main functions can be summarised as: provision of a secure and trustfully cross-CI communication infrastructure; collection of information about the local CI; retrieval of information about the other interdependent CIs in the system; exchanging of information related to local CI with remote CIs; composition of CIs critical events and semantic inference and delivering of all the collected information to the prediction tool (Caldeira et al., 2010a).

**Figure 1** MICIE overall system architecture (see online version for colours)



Source: Caldeira et al. (2010a).

Each SMGW has a discovery mechanism that supports dynamic discovery information available on the local CI and on all interconnected CIs. Besides being used to evaluate risk prediction, this information also provides CI operator a powerful real-time view about identified risks and alerts.

The sensitive nature of the exchanged information poses special attention on the security requirements. The SMGW design has to guarantee strict security requirements, such as confidentiality, integrity, availability, non-repudiation and auditability/traceability. Trust and reputation management are also essential requirements for the information exchange among SMGWs. The proposal and evaluation of the trust and reputation management mechanisms to be implemented in SMGWs are the main contribution of the present work.

## 4 Trust and reputation management in SMGWs

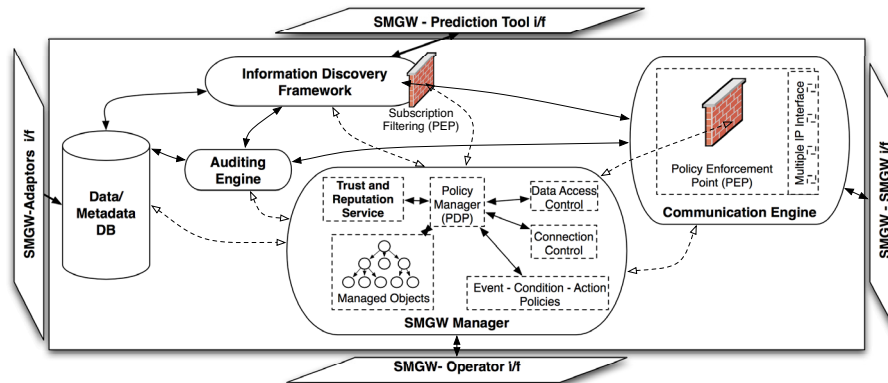
In this section, the SMGW manager, the SMGW entity responsible for the SMGW management is described, detailing the application of the trust and reputation indicators in the management process.

### 4.1 SMGW management

Management of the SMGW is the role of the SMGW manager. Developed according to the policy-based management (PBM) paradigm. The SMGW manager intends to manage all SMGW aspects. SMGW manager also performs monitoring with the help of the auditing engine and can also act as intrusion prevention and detection engine by configuring firewalls that enforce security in the communication process. SMGW architecture is presented in Figure 2.

The proposed approach provides the CI operator with a tool where he/she can define, in a high abstraction level, the behaviour he/she pretends for the system. Traditional approaches are mainly oriented to the management of individual components, not considering the system structure as a whole. In this proposal, we use the PBM (Poylisher and Chadha, 2008) that includes the policy decision point – the SMGW manager – and policy enforcement point – the entities that enforce policies, for instance the communication engine and the subscription filter.

**Figure 2** SMGW architecture (see online version for colours)



The SMGW manager handles issues regarding authorisation, authentication and accounting controlling both interaction with peer SMGWs and all the internal operation of the SMGW, including also monitoring, alarming, management of intrusion prevention and detection functions and the management of the trust and reputation service (TRS).

The CI operator can define policies that will address the relations among local SMGW and foreign SMGWs, including defining how each particular CI can connect and data access policies. The SMGW manager graphical user interface (GUI) will allow the operator to browse through existent information and define actions that remote SMGWs can perform (e.g. write and/or read risk information). All data access controls are implemented with a high level of granularity thus maintaining simplicity and abstraction from equipment configuration details.

The use of policies supports definition, verification and deployment of security policies related to the information gathered by the MICIE system. For instance, those policies will permit:

- the definition of how and to whom each particular piece of information can be sent
- the definition of trust relations among different CI
- the enforcement of different communications protocols/technologies in each particular context
- the enforcement of service level agreements or service level specifications between CIs
- the decision on how received events will be managed by the SMGW.

The CI operator defines policies using a policy specification language and/or using the provided GUI. This GUI has the representation of all managed entities allowing the operator to easily define relations among them (policies). The management policies are stored in a policy repository.

The SMGW manager uses Ponder2 toolkit Twidle et al. (2009) where each SMGW entity is represented using Ponder2 concept of managed object. The complete set of SMGW entities form a Ponder2 self-managed cell (SMC). This SMC contains the representation of all system-managed objects. In a simplified approach, we can identify two main types of managed objects: connections to the SMGW, represented with remote SMGW managed object and alert data represented by the defined alert parameters. Policy enforcement is based on Ponder2 authorisation policies and event condition action concepts.

Apart from the existing Ponder2 communication modes, a dedicated API to manage communication aspects between SMGW components and the SMGW manager was developed. This API allows the full configuration of the SMGW, e.g.

- change state (attributes) of managed objects (e.g. change of connection type of a remote SMGW). This can be treated as event condition action in Ponder context;
- send authorisation requests. This can be treated as authorisation policies in Ponder context.

By exploring the features provided with this management approach, the management aspects were improved introducing the concept of trust and reputation in the context of communication among CIs.

### 4.2 Trust and reputation management

Although the MICIE context can be seen as a closed system where it is supposed that partners trust each other, it is possible for a partner CI to provide inaccurate information, either maliciously (e.g. if their system is somehow compromised) or due to a faulty components in its monitoring framework.

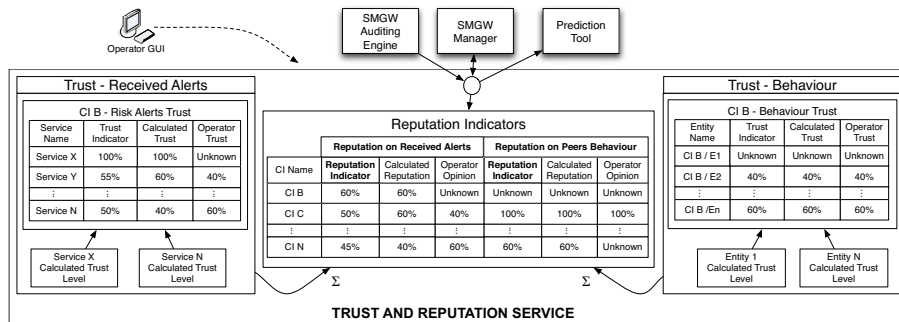
To deal with the above situation, the need of a TRS on each SMGW was identified, to maintain real-time trust and reputation information about peering SMGWs. This service will monitor information exchanged among peer SMGWs and partner behaviour in order to manage reputation and to infer a trust level for each entity.

There are two main areas where trust and reputation are applied in the MICIE project (see Figure 3). Firstly, a trust indicator about the information received from partner CIs. This indicator is evaluated at two levels: service level, evaluating each service subscribed to remote CI, reflecting the trust on information received on one specific service; and at CI level, evaluating a trust indicator for each interconnected CI, representing the trust in that particular CI (reputation indicator). Secondly, the TRS is also capable of understand the interdependent CIs behaviour in terms of ICT security. The interactions between peers are monitored to gather intelligence about the partnership.

The information needed to evaluate trust and reputation is gathered and evaluated from multiple sources, namely:

- 1 *Analysis of past data provided by partners/services*: the TRS will compare the risk estimates provided overtime, for each service, against the current service levels in order to infer the trustiness of future estimates.
- 2 *Analysis of partner behaviour*: based on the knowledge gathered by the security entities existent in the CI, the TRS analyses the partner behaviour in terms of ICT security. For instance, if the partner CI behaves abnormally (e.g. trying to access non-authorized data or using non-authorized credentials) the TRS should downgrade the level of trustiness associated with that partner CI as this could indicate that the partner is faulty or does not have good intentions.
- 3 *Human factors*: operator perception about each partner/service. The operator may have information about each partner/service that he/she wants to incorporate in the TRS.

Figure 3 Trust and reputation service



The TRS gathers the information needed using two agents: one agent that detects and calculates the risk alert event accuracy and another used to receive behaviour events. The TRS computes, in real time, the trust and reputation indicators. Computed indicators are provided to external entities using a web service. A graphical interface provides the CI operator with an overall view about trust and reputation indicators.

#### 4.2.1 Trust and reputation indicators on risk alerts

As said before, the trust indicator is evaluated for each service using information available on the SMGW. Monitoring information received from peer CIs is kept to estimate trust for each service simply by evaluating provided risk estimation (risk alert level –  $RI_t$ ) against actual service behaviour (measured service level –  $MI_t$ ). For example, if a partner CI keeps informing that one service is going to stop with a probability of 90% and the service never fails we can infer that this particular alert is not credible.

The trust indicator, based on past experience, is evaluated using a statistical approach where the operator can define parameters like the penalisation he/she wants to give on detected anomalies on observed services and risk levels or the fast or slow ageing of past events. This last parameter has special importance to avoid situations where for a long period, received informations are accurate and then starts to be faulty. A simple average will maintain a high trust level while the ageing factor (weight moving average) will give more attention to the new information.

To be able to evaluate trust aspects related to the received risk information, the first step is to define an accuracy value for each received risk. For this purpose, the concept of risk alert event is introduced. A risk alert event is triggered when:

- 1 a service decreases its QoS (this event ends when the QoS exceeds the threshold or, if an alert is received, the event ends when the alert is removed)
- 2 after the reception of a risk alert message.

The SMGW is monitoring, in real time, the risk alert levels ( $RI_t$ ) and the current measured service levels ( $MI_t$ ) in order to detect events. The accuracy of each event  $A$ (event) is defined as the average of the comparisons made during the event (value  $T$ ), between measured service level and risk alert level (Equation (1)). Function  $f(MI_t, RI_t)$  is a discrete function so a sample rate for the time factor is needed. This sample rate can be different for each service and will depend on the information available on the system. A high sample rate will allow more realistic observations.

$$A(\text{Event}_n) = \frac{\sum_{t=1}^T (f(MI_t, RI_t))}{T} \quad (1)$$

where  $f(MI_t, RI_t) = |MI_t - RI_t|^\kappa$ ,  $\kappa \in R^+$ . The value  $k$  allows to penalise the larger differences or the small differences and should be assigned considering the degree of importance of each service. Basically, in some services a small difference can be more relevant than in others.

As detailed in our previous work Caldeira et al. (2010b), the trust that  $CI_A$  has in risk alerts received for service  $X$  provided by  $CI_B$  is represented by  $T_{(A,B,X)}$  and is calculated by the average of the accuracy of each past event between those two CIs for that particular service (Equation (2)). The concept of ageing is used, applying a weight factor  $D$ , to give more weight to recent events. The ageing factor should always depend on the context. In our model, the ageing factor needs to be defined on a per peer/service basis. In this context,

$T'_{(A,B,X)}$  can be computed for the  $N$ th event as:

$$T'_{(A,B,X)} = \frac{(D(N-1)T_{(A,B,X)} + A(\text{Event}_N))}{D(N-1) + 1} \quad (2)$$

$D$  will be a value in the  $[0 \dots 1]$  interval and a small value of  $D$  will raise the importance of the past events while a value of  $D$  near 1 will provide less ageing to oldest events.

During trust evaluation, CI operator can have an active role introducing his own subjective trust regarding specific services or globally about one CI (Equation (3)). This indicator introduces aspects not detectable by automatic trust evaluation. For instance, CI operator can know that some CI was having faulty equipment during a time period. During this period is likely that our trust indicator decreases. In this case, operator can act, raising his own confidence parameter and consequently do not letting the global trust value decrease in the future.

$$T(\text{Final})_{(A,B,X,t)} = (1 - \alpha)(T_{(A,B,X)}) + \alpha (TO_{(A,B,X)}) \quad (3)$$

The factor  $\alpha$  is in the range  $[0 \dots 1]$  and is assigned by the CI operator depending on the confidence he/she has in  $(TO_{(A,B,X)})$ .  $T(\text{Final})_{(A,B,X,t)}$  represents the TRS confidence in alerts taking into account also the CI operator perspective. In order to understand how trust indicators evolve overtime, and to define a relation among them, a time value is associated with each  $T(\text{Final})$ .

The reputation of each CI is evaluated using Equation (4) where  $GT'_{(A,B,t)}$  represents the reputation that CI  $A$  as about CI  $B$  on time  $t$ .  $GT_{(A,B)}$  represents the last evaluated indicator.  $W_i$  is the weight associated to service  $i$  provided by CI  $B$  (this value is gathered from the MICIE risk models).  $N$  is the number of evaluations.  $S$  represents the services that  $A$  receives from  $B$  and  $D$  is the ageing factor.  $T(\text{Final})_{(A,B,i)}$  represents the last indicator calculated for service  $i$ . Equation (4) should be evaluated every time a service indicator changes.

$$GT'_{(A,B,t)} = \frac{(D(N-1)GT_{(A,B)}) + \left( \left( \sum_{i=1}^S (T(\text{Final})_{(A,B,i)} W_i) \right) / \left( \sum_{i=1}^S W_i \right) \right)}{D(N-1) + 1} \quad (4)$$

As described in our previous work, the reputation indicator also includes the operator contribution.

#### 4.2.2 *Trust and reputation indicators on peer behaviour*

The proposed security model allows the TRS to collect and analyse data related to security aspects used to infer a trust indicator for each peer behaviour. The behaviour trust agent receives security alerts from the entities responsible for maintaining security in the SMGW and sends those events to the TRS in order to be incorporated in the behaviour trust indicators.

The events used to evaluate trust on CI behaviour are all the interactions among peer CIs in terms of security (internal or external). For instance, the events can be intrusion detection system (IDS) alerts, failed connection attempts, attempts to read/write information without permission. As the event values are gathered from heterogeneous sources, a normalisation methodology, presented in our previous work must be applied.

In the security behaviour context, we expect to receive alerts only when a misbehaviour is detected, leading to a situation where almost only negative events are received and used

in the evaluation. This situation would generate low behaviour trust overtime. To evaluate a precise indicator, the factor time and the concept of inactivity were introduced. Time is divided into a set of time slots. Inactivity in one slot means that the peer behaviour indicators have the maximum value. If information is received during one slot, the indicators value for that slot value become the average of all values received during that slot Caldeira et al. (2010b).

For the time slot  $s$ , the trust on entity  $E$  regarding  $CI_B$  ( $T'_{(E,B,s)}$ ) is calculated using Equation (5) where  $D$  is the ageing factor,  $T_{(E,B)}$  is the indicator evaluated for the slot  $(s - 1)$  and  $Event_{(Slot\ s)}$  is the event value of the slot  $s$ .

$$T'_{(E,B,s)} = \frac{(D(s - 1)T_{(E,B)}) + Event_{(Slot\ s)}}{D(s - 1) + 1} \quad (5)$$

Using Equation (6), the operator trust is included. The  $\theta$  factor is assigned by the CI operator representing the confidence on the subjective trust ( $TO_{(E,B)}$ ) that he or she has on the behaviour of  $CI_B$  concerning security entity  $E$ .

$$T(\text{Final})_{(E,B)} = \theta(TO_{(E,B)}) + (1 - \theta)(T_{(E,B)}), \quad (0 < \theta < 1) \quad (6)$$

The TRS also evaluates indicators encompassing all monitored security entities. Using a weight factor for each entity, the behaviour reputation for each CI (or group of entities) can be computed, considering also the operator information. This indicator,  $TBehaviour'_{(B,t)}$ , represents the reputation of the behaviour of  $CI_B$  at time  $t$  Caldeira et al. (2010b).

Presented trust and reputation indicators can then be used to enhance risk prediction models, make decision based on defined policies and also help the CI operator to evaluate the partnership between his CI and their peers.

## 5 Evaluation

The proposed framework was already validated using simulation techniques Caldeira et al. (2010b). This section presents some results concerning the experimental validation of the trust and reputation indicators and discusses the advantages and disadvantages of our proposal.

Several simulation scenarios were developed and tested. One subset of those scenarios is described in Table 1, representing the following situations: (S1) the system behaves as expected with only small errors with the event accuracy always above 60% and mainly between 90% and 100%; (S2) system is not accurate but can still be trustworthy, as evaluated event accuracy is always above 40%; (S3) received alerts are not as expected with above 40% of inaccurate indications but never rising above 60%; (S4) the system is not trustworthy as 90% of the events has an accuracy lower than 20%.

**Table 1** Simulation scenario (% of events for each range of event accuracy values)

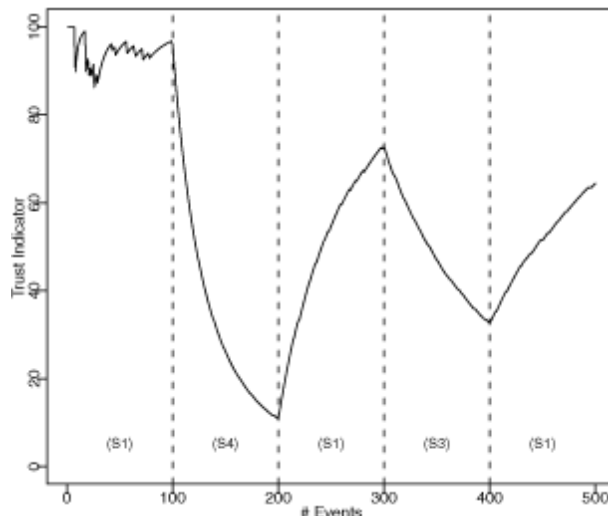
Scenarios	Event % of occurrence									
	[0–10]	[10–20]	[20–30]	[30–40]	[40–50]	[50–60]	[60–70]	[70–80]	[80–90]	[90–100]
S1	0	0	0	0	0	0	5	5	10	80
S2	0	0	0	0	10	10	10	10	20	40
S3	40	20	10	10	10	10	0	0	0	0
S4	80	10	5	5	0	0	0	0	0	0

For the presented simulation, the risk alert and behaviour events were generated using a normal distribution and the following parameters are used: penalisation factor  $k = 2$ ; ageing factor  $D = 0.3$ . A threshold of 10%, meaning that event accuracy values above 90% are rated as 100%. The simulation as been performed using the R environment for statistical computing and graphics R (2005).

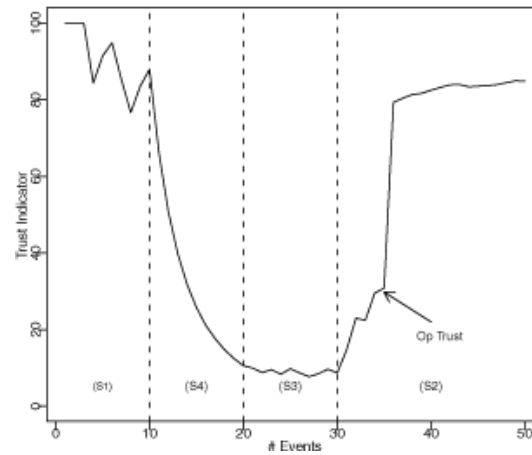
Figure 4(a) represents an attack or faulty component situation were the scenario change on every 100 events. The better scenario (S1) abruptly changes to the worst (S4) after the first 100 events. This change is almost immediately incorporated in the trust indicator. After 100 events from scenario S4, the trust indicator clearly means that the CI should not trust on the received alerts for this specific service. After 200 events the received risk alerts become, again, reliable. As we use the ageing factor, the trust indicator rapidly incorporates this new scenario. From the 300th to the 400th event, the scenario changes to S3, decreasing the trust indicator (in this case the indicator decreases more slowly than in the last case). This simulation shows that the trust indicator can rapidly react when the scenario changes. It is also clear that even with an abrupt change of scenario, the indicator changes gradually due to the ageing factor.

In Figure 4(b), we use less events than in the last simulation, showing that the TRS is still accurate even with a small number of received events. In this simulation, the received alerts are not reliable between the 20th and the 30th event. This leads to a very low-trust value that gradually starts to grow after the 30th event (grows gradually as the received events are based on scenario S2). Supposing that the CI operator knows that the situation that lead to this scenario is already solved, he can rapidly update the trust indicator with his opinion about the service. In this case, the operator assigned a value of trust as being 90% and defined a contribution of 0.8 to the trust indicator. Although the indicator continues to incorporate changes from the events, his value will be higher. In this case, it is important that the operator know the consequence of his action as his contribution to the indicator will continue to be applied until he changes it.

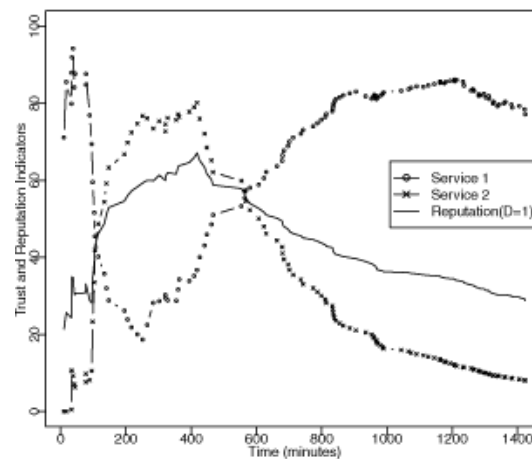
**Figure 4** Trust on received risk alerts



(a)

**Figure 4** Trust on received risk alerts (continued)

(b)

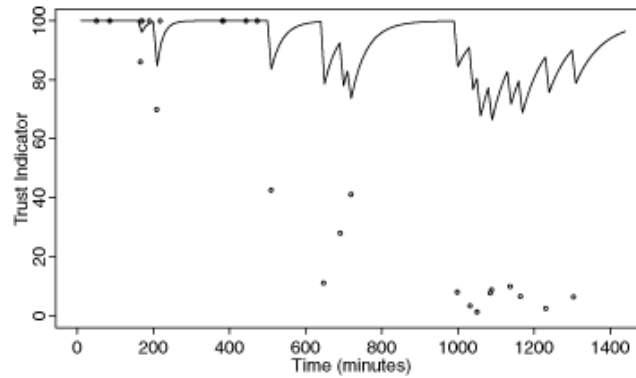


(c)

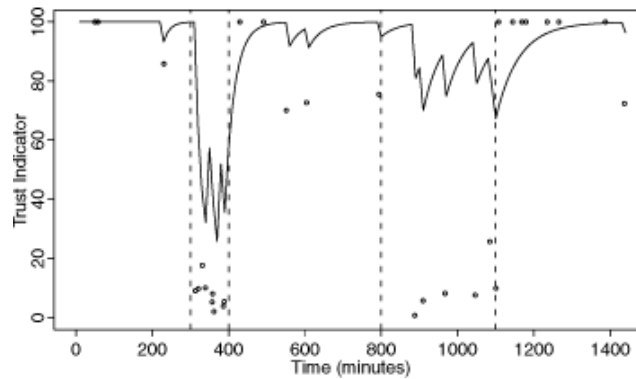
In Figure 4(c), two separated services received from the same CI were used in order to evaluate the reputation indicator for that CI. Each service has an average of five events per 60 min from a mixture of scenarios one to four. The operator assigned a weight of 0.3 to service 1 and 0.7 to service 2. A value of  $D = 1$  (simple average) was used to calculate the reputation indicator. In this simulation, when the service, more important (due to the weight assigned by the operator), is becoming unreliable, then the CI reputation is decaying even when the other service is trustworthy.

In order to simulate the security events (behaviour), the arrival time of each event is generated using an exponential distribution, with an average of  $x$  events per minute. The values for each event is generated based on the scenarios defined in Table 1.

Figure 5 shows the results of two different simulations with common parameters time slot = 10, ageing  $D = 0.05$  and simulation period = 1,140 min. The defined ageing factor allows the incorporation of new situations in the behaviour indicator. The dots on Figure 5(a) and 5(b) represent the received events.

**Figure 5** Trust on peer behaviour

(a)



(b)

The first simulation (Figure 5(a)) has a rate of one event per each 60 min. from multiple scenarios, namely, S1, S3 and S4. In this simulation, with few events, the trust indicator does not drop below 60% due to influence of the slots where the system is behaving well. This simulation demonstrates how important is the value defined for the time slot. In this scenario, a larger time slot would lead to a lower trust indicator. It is also important for the CI operator to know how to interpret the received indicators, in order to properly configure the system.

The results obtained when simulating a situation with two possible attacks or misbehaviour is presented in Figure 5(b). In this case, during the first 300 min, events from scenario 1 arrive with a rate of 1/60 min. During a period of 100 min, the scenario changes to S1 (the worst scenario) with an event rate of 5/60 min. In that moment, the behaviour trust indicator rapidly decays below 50% clearly indicating that something is wrong. Next, the peer behaviour is simulated for S2 with a lower event rate leading the indicator to raise. Between the 800th and 1,100th min, the scenario changes to S4, now with a event rate of 1/60 minutes. It is visible that even with only few events the CI operator can infer that the peer behaviour is not normal. With this indicator, the SMGW manager can act, for instance blocking the access from that CI. The last simulated minutes represent the scenario S1 at a rate of 1/60 min. On this scenario and with a lower event rate, the trust indicator clearly indicates the resolution of the past situation.

From the presented results, it is clear that the trust and reputation indicators allow to enhance contributions expected from MICIE project, increasing the quality of the MICIE risk alerts. MICIE system provides, in real time, risk levels measuring the probability that a CI loses the capacity to provide services or receive services. This information is based on internal data and on data received from peer CIs.

Trust and reputation indicators are incorporated in the MICIE prediction tool as a way to improve its accuracy and its resilience to inconsistent information provided by peer CIs making possible, for instance, to give more weight to highly trusted data or ignore data provided by low-trust partner.

Also, system management can become more dynamical with the use of trust and reputation indicators, reacting autonomously when those indicators change. For instance, if our trust regarding the behaviour of one peer decreases below a defined threshold a new policy is triggered and the SMGW can stop accepting connections from that peer.

Prototypes for the SMGW manager and for the TRS are already developed and tested. The validation of the presented framework is fundamental as the trust and reputation indicators will influence our risk models. Also, we expect to test this proposal along with MICIE project starting with a simple reference scenario that encompasses a small portion of an electricity distribution network and an interdependent telecommunications network Capodieci et al. (2010). Planned validation work for the MICIE project will also include more complex scenarios, provided by Israel Electric Corporation and including multiple CIs.

## **6 Conclusions**

This work reports some research achievements in the scope of the FP7 ICT-SEC MICIE project on the development of a real-time risk level prediction, dissemination and alerting system. In order to reach MICIE objectives, one of the main key challenge to be addressed is the design and the implementation of a SMGW, namely a new innovative network element able to:

- 1 discover CI status information
- 2 overcome information heterogeneity
- 3 provide a secure communication of such information among peer CIs.

Author's contribution to this enhancement is described in this paper, namely the development of a PBM tool for the SMGW and the incorporation of the concept of TRS in the SMGW.

Trust and reputation indicators and the use of policies can enhance risk indicators accuracy, help incorporating trust in system management and also help CI operator to evaluate the relation between his CI and their peers.

Besides validation by simulation discussed in this paper, Authors have developed prototypes of the present solutions and are evaluating them using the demonstrator that MICIE project is developing and testing in the field.

Improving MICIE project beyond his initial objectives, described work, represents a step forward in CIs interoperation.

## Acknowledgements

Work partially financed by FP7 ICT-SEC MICIE project (MICIE, 2008) Grant Agreement No. 225353, and by the Portuguese Foundation for Science and Technology (SFRH/BD/35772/2007). The authors want to thank all the involved partners for their valuable support to this work.

## References

- Aime, M. and Liou, A. (2005) 'Incremental trust: building trust from past experience', *Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks, 2005. WoWMoM 2005*, pp.603–608.
- Artz, D. and Gil, Y. (2007) 'A survey of trust in computer science and the semantic web', *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 5, No. 2, pp.58–71.
- Balducelli, C., Di Pietro, A., Lavalle, L. and Vicoli, G. (2008) 'A middleware improved technology (mit) to mitigate interdependencies between critical infrastructures', in R. de Lemos, F. Di Giandomenico, C. Gacek, H. Muccini and M. Vieira (Eds.), *Architecting Dependable Systems V*, Vol. 5135 of *Lecture Notes in Computer Science*, Berlin, Heidelberg: Springer, pp.28–51.
- Caldeira, F., et al. (2010a) 'Secure mediation gateway architecture enabling the communication among critical infrastructures', *Future Network and MobileSummit 2010 Conference*.
- Caldeira, F., Monteiro, E. and Simões, P. (2010b) 'Trust and reputation for information exchange in critical infrastructures', *5th International Conference on Critical Infrastructures Information Security (CRITIS 2010)*, Athens, Greece.
- Caldeira, F., Monteiro, E. and Simões, P. (2010c) 'Trust and reputation management for critical infrastructure protection', in S. Tenreiro de Magalhães, H. Jahankhani and A.G. Hessami (Eds.), *Global Security, Safety, and Sustainability*, Vol 92 of *Communications in Computer and Information Science*, Berlin, Heidelberg: Springer, pp.39–47.
- Capodiecici, P., et al. (2010) 'Improving resilience of interdependent critical infrastructures via an on-line alerting system', *Complexity in Engineering, 2010. COMPENG '10*, pp.88–90.
- Chen, S., Zhang, Y. and Yang, G. (2009) 'Trust and reputation algorithms for unstructured P2P networks', *International Symposium on Computer Network and Multimedia Technology, 2009. CNMT 2009*, pp.1–4.
- Clinton, W.J. (1996) 'Executive order 13010 – critical infrastructure protection', *Federal Register*, Vol. 6, No. 138, p.37347.
- Commission, E. (2006) 'Communication from the commission on a european programme for critical infrastructure protection', COM/2006/0786 final.
- CRUTIAL (2008) 'Crutial project web site', Available at: <http://crutial.cesiricerca.it>.
- Dondossola, G., Garrone, F., Szanto, J. and Gennaro, F. (2008) 'A laboratory testbed for the evaluation of cyber attacks to interacting ict infrastructures of power grid operators', *SmartGrids for Distribution, 2008. IET-CIRED. CIRED Seminar*, pp.1–4.
- IRRIIS (2009) 'Irriis project web site', Available at: <http://www.irriis.org>.
- Josang, A., Ismail, R. and Boyd, C. (2007) 'A survey of trust and reputation systems for online service provision', *Decision Support System*, Vol. 43, No. 2, pp.618–644.
- MICIE (2008) 'MICIE – tool for systemic risk analysis and secure mediation of data ex-changed across linked ci information infrastructures', *FP7-ICT-SEC-2007.1.7 – 225353 – Annex I – "Description of Work"*.

- Panzieri, S., Setola, R. and Ulivi, G. (2005) 'An approach to model complex interdependent infrastructures', *16th IFAC World Congress*. CISIA, Critical Infrastructures.
- Poylisher, A. and Chadha, R. (2008) 'PBNM technology evaluation: practical criteria', *IEEE Workshop on Policies for Distributed Systems and Networks, 2008. POLICY 2008*, pp.105–108.
- R (2005) Development core team, R: a language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. ISBN 3-900051-07-0. Available at: <http://www.r-project.org>.
- Rinaldi, S.M., Peerenboom, J.P. and Kelly, T.K. (2001) 'Identifying, understanding, and analyzing critical infrastructure interdependencies', *IEEE Control Systems Magazine*, Vol. 21, pp.11–25.
- Simões, P., et al. (2010) 'An alerting system for interdependent critical infrastructures', *ECIW 2010 – 9th European Conference on Information Warfare and Security*.
- Sokolowski, J., Turnitsa, C. and Diallo, S. (2008) 'A conceptual modeling method for critical infrastructure modeling', *41st Annual Simulation Symposium (ANSS 2008)*, pp.203–211.
- Spitz, S. and Tuchelmann, Y. (2009) 'A trust model considering the aspects of time', *Second International Conference on Computer and Electrical Engineering, 2009. ICCEE '09*, Vol. 1, pp.550–554.
- Twidle, K., Dulay, N., Lupu, E. and Sloman, M. (2009) 'Ponder2: a policy system for autonomous pervasive environments', *Fifth International Conference on Autonomic and Autonomous Systems, 2009. ICAS '09*, pp.330–335.
- Veríssimo, P., et al. (2008) 'The crucial architecture for critical information infrastructures', R. de Lemos, F. Di Giandomenico, C. Gacek, H. Muccini and M. Vieira (Eds.), *Architecting Dependable Systems V*, Vol. 5135 of *Lecture Notes in Computer Science*, Berlin, Heidelberg: Springer, pp.1–27.
- Zahariadis, T., et al. (2008) 'Trust models for sensor networks', *ELMAR, 2008. 50th International Symposium*, Vol. 2, pp.511–514.