Fusion of Bayesian and Ontology Approach Applied to Decision Support System for Critical Infrastructures Protection

Rafał Kozik[2], Michał Choraś[1,2], and Witold Hołubowicz[1,3]

[1] ITTI Ltd., Poznań
michal.choras@itti.com.pl
[2] Institute of Telecommunications, UT&LS Bydgoszcz
chorasm@utp.edu.pl
[3] Adam Mickiewicz University, Poznań
holubowicz@amu.edu.pl

Abstract. In this paper, a decision support system based on the ontology knowledge for Critical Infrastructure security assessment is presented. The ontology provides vulnerabilities, threats and safeguards classification and their relationships with other security aspects. Such knowledge is used to build Bayesian network, which is used to asses the severity level of the detected threats. Described approach is applied in decision support tool developed within the INSPIRE project aiming at increasing security and protection through infrastructure resilience. The major contribution of this paper is the fusion of the ontology and Bayesian approach utilized in the reasoning engine of the decision support application.

1 Introduction

Rapid success of information and communication systems had significant influence on developing new way of controlling and managing critical infrastructures. Systems monitoring and controlling critical infrastructures such as SCADA (Supervisory Control and Data Acquisition) moved from dedicated solutions for particular operator to integrated and IP-based frameworks.

However, such evolution exposed the system for cyber threats and cyber attacks and unauthorized access by hackers. This requires a new approach to critical infrastructure protection which will engage expert knowledge, decision support systems and such network elements as firewalls, intrusion and anomaly detection systems. Critical Infrastructure Protection (CIP) including cyber defense is one of the crucial security and safety aspects in EU [1].

Critical infrastructure security problems are the challenge for the EU FP7 ICT-SEC INSPIRE Project (INcreasing Security and Protection through Infrastructure REsilience). It is a two-year small or medium-scale focused research European project. More details about the project are available at: $http://www.inspire-strep.eu$.

P. Chatzimisios et al. (Eds.): MobiLight 2010, LNICST 45, pp. 451–463, 2010.

There are two main research directions in the project. The first one ("in-network") focuses on:

- analyzing and modeling dependencies between critical infrastructures and underlying communication networks,
- designing and implementing traffic engineering algorithms to provide SCADA (Supervisory Control and Data Acquisition) traffic with quantitative guarantees,
- exploiting Peer-to-Peer (P2P) overlay routing mechanisms for improving the resilience of SCADA systems,
- defining a self-reconfigurable architecture for SCADA systems,
- development of diagnosis and recovery techniques for SCADA systems.

The second research direction in the project, called "off-network" focuses on designing the INSPIRE Security Ontology and development of the decision support system to evaluate critical infrastructure security status. The role of the proposed DSS (called INSPIRE Decision Aid Tool - DAT) is to provide the SCADA operator system with all the necessary information about the threats and vulnerabilities the specific critical infrastructure is exposed to. Additionally, DAT can propose appropriate reactions and countermeasures for the particular threat.

The paper is structured as follows: in section 2 INSPIRE Decision Aid Tool (DAT) is motivated and presented. The underlying security ontology overview is given in section 2.3. Moreover, the ontology mapping mechanism and DAT knowledge organization are explained in sections 2.4 and 2.5, respectively. Then in section 3, our novel approach based on the fusion of the ontology knowledge and Bayesian network is presented in detail. Sample use case is presented and discussed in section 4. Conclusions are given afterwards.

2 INSPIRE Approach to Decision Support Systems

2.1 Overview of DSSs for Critical Infrastructures Protection

Decision Support Systems (DSS) are information systems that support human in different decision-making activities. DSS applications are successfully and widely used in industry and critical infrastructure protection (CIP). In 1987 Texas Instruments company released GADS (Gate Assignment Display System) decision support system for United Airlines. As a result, the travel delays have been reduced significantly. The system was used by the management of ground operations at various airports.

Another good example of successfully deployed decision support applications are expert systems in the banking area (expert systems for mortgages). The decision support systems are also widely used for river systems management to effectively cope with floods. For example, The German Federal Institute of Hydrology (BfG) funded the development of a Decision Support System for the Elbe river system. The great flooding in summer 2002 demonstrated the importance of such solutions.

Some examples of DSS used in the energy sector are described in [2]. DSS are also successfully deployed in nuclear power plants [3], urban water pollution control [4] or oilfield flood precaution [5].

2.2 INSPIRE DAT (Decision Aid Tool)

All the mentioned DSS examples are customized and focused on some particular branch of critical infrastructures. Decision Support Systems are usually designed for special kind of industry or application. Although they use different methodologies (Bayesian, multiagent, HMM), they rarely use ontologies description to support reasoning.

Therefore, in the INSPIRE project we proposed the security ontology, which mimics the complicated relationships between SCADA components and security aspects. Our ontology is a representation of relationships between particular classes (or instances) and as it is, cannot provide any knowledge-based reasoning or give feedback to its operator, therefore the INSPIRE Decision Aid Tool has been developed.

DAT is general and applicable to more than one critical infrastructure. It focuses on the SCADA properties to enhance protection and security of critical sectors. INSPIRE Decision Aid Tool may be also considered as a framework since it is reconfigurable by means of uploading other ontologies or various SCADA system topologies.

Two types of users (actors) are specified in DAT (Fig. 1):

- DAT User − user who wants to asses the systems security level.
- Expert − user who maintains security rules (facts about concerned system and relations between its elements), which allow to enrich the knowledge stored in ontology.

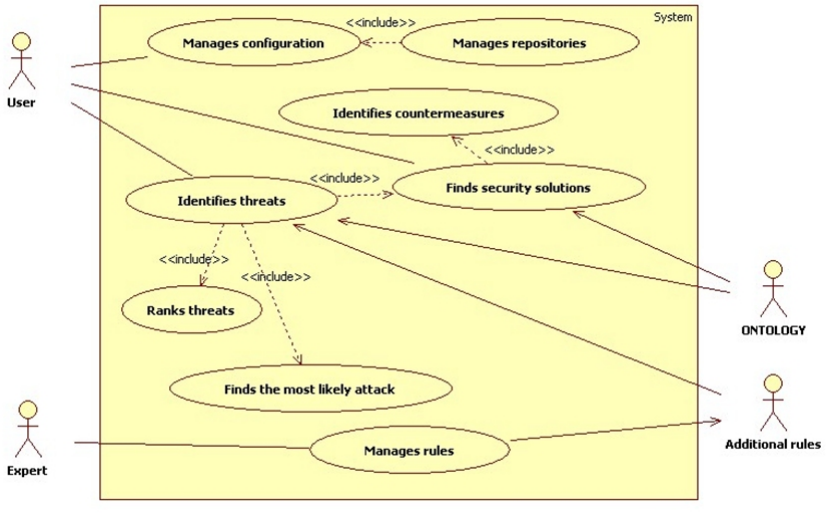

Fig. 1. DAT use cases

DAT is facilitated with user-friendly graphical interface which allows to identify and rank the threats found in critical infrastructures. Furthermore, the DAT allows user to find security solutions for particular threat and finds adequate countermeasures and strategy for minimizing the risk. What is more, DAT allows user to perform simulation scenario answering the questions such as "what may happen if particular action is taken" or "what happens (in terms of security) when particular equipment or software is added".

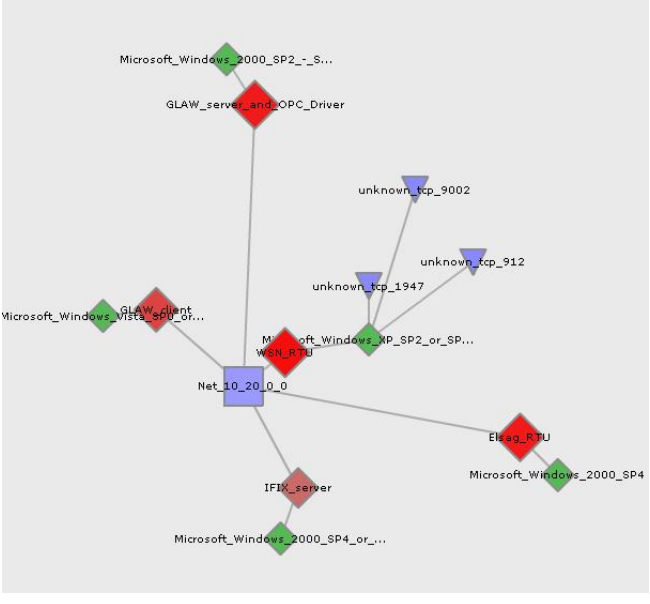

Fig. 2. Topology diagram and visualized threats (blinking red nodes)

The analysis performed by DAT is divided into three steps. Firstly, the topology diagram is rendered (Fig.2) to provide the operator with the information e.g. about CI elements interconnections, used applications etc. In the next step, the threats are visualized by the red blinking nodes (Fig. 2). Afterwards, the detailed security report is created to provide user with detailed information (Fig. 10). The used inference engine and ranking methodology remain transparent for the user. However, these can be configured and customized via the configuration panel.

Moreover, the mechanism for adding additional rules (expert security rules) has been created for experts and security operators. Such rules about security aspects can enrich the ontology knowledge and improve reasoning. The GUI for adding expert rules is presented in Fig. 3.

2.3 INSPIRE Security Ontology

Our approach to security ontology is based on ISO/IEC 133351:2004 standard [6]. According to the standard, vulnerabilities are considered as properties of a

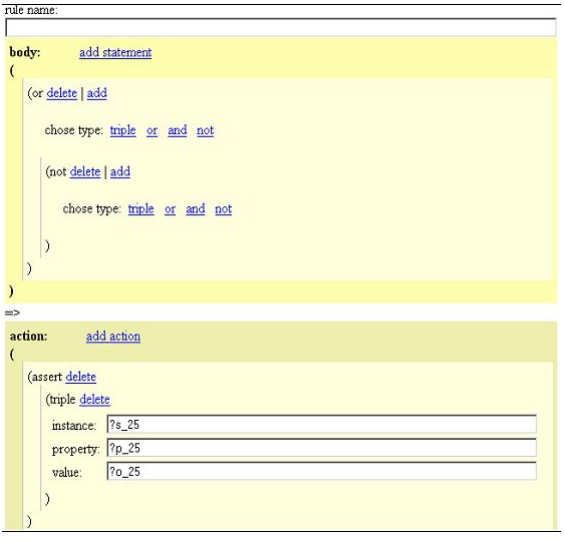

Fig. 3. GUI for expert rules generation

network security system. In such approach assets and components have weak points named vulnerabilities. These vulnerabilities can be exploited by threats, leading to attacks. This security system is depicted into a form of classification with properties and relationships between various security aspects [7][8].

2.4 Ontology Mapping

The proposed ontology is used by the Decision Aid Tool (DAT). However, the ontology format is not directly accepted by the inference engine and requires mapping.

Therefore instances (and also relation between instances) stored in ontology are mapped into facts and SWRL rules are mapped into production rules [9]. Afterwards the reasoning can be performed by the inference engine (in this case JESS inference engine has been adopted). DAT uses ontology classes and instances to acquire the knowledge as RDF triples and processes them in the rule engine [10]. Each RDF triple consists of:

- Subject,
- Predicate,
- Object.

Each triple is able to fully describe one property of the instance. The interpretation of a triple is that "subject" has property "predicate" whose value is "object". Such strategy allows DAT to be more flexible to ontology schema changes, because adding new properties to the particular instance has no impact on the mapping mechanism and no impact on DAT source code.

In example the relation "Asset x hasVulnerability y" is mapped into (triple (subject x) (predicate 'hasVulnerability') (object y)). The SWRL rules are mapped into production rules as follows:

$$hasParent(?x1, ?x2) \wedge hasBrother(?x2, ?x3) \Rightarrow hasUncle(?x1, ?x3)$$

is mapped into:

```
(defrule rule-1
(triple (predicate 'hasParent') (subject ?x1) (object ?x2) )
(triple (predicate 'hasBrother') (subject ?x2) (object ?x3) )
⇒
(assert (tiple (predicate 'hasUncle') (subject ?x1) (object ?x3) ) ). )
```

2.5 DAT - Knowledge Organization

The knowledge about the critical infrastructure maintained in the ontology is large and requires classification and additional organization in order to be efficiently used by the inferencing engine. The most reasonable solution is to organize it in a hierarchical manner (from low level facts to high level ones as in Fig.4).

Lowest layer (Fig. 4) of the knowledge stack simply represents the knowledge obtained from OWL thanks to mapping described in section 2.4. The RDF triples based description allows to extract information about basic relations between elements and particularly identify the root classes and instances belonging to that class (Asset, Vulnerability, Threat, Safeguard). These concepts allow to extract, so called "root facts" about critical infrastructure, thus such knowledge, describing relations "asset-vulnerability-threat", gives the user information, which is similar to this which can be found in typical vulnerability databases.

Therefore, we provided the second layer which additionally extracts the given critical infrastructure topology. Hereby, information about particular node, its connection to the network, running applications etc. can be found. This allows to identify additional facts about analyzed environment such as faults in elements connectivity, configuration faults, etc.

On top of this knowledge, DAT allows the user to provide security rules. Particularly, operator has ability to asses what may happen if particular action is taken.

Let us consider the scenario where the operator plans to take down the router during the maintenance of the CI network. DAT using the information about the connectivity and information about business importance about detached nodes (detached by shutting down the router) will alert that this may cause serious malfunction of CI (or alternatively, that it has no impact on it). Also operation of adding new machine to CI infrastructure may also be validated by DAT prior the physical manipulations, saving the time and eventually the money.

However, the knowledge described above seems condition-action based (if-then structure). Therefore, one more layer of knowledge in the stack is introduced to

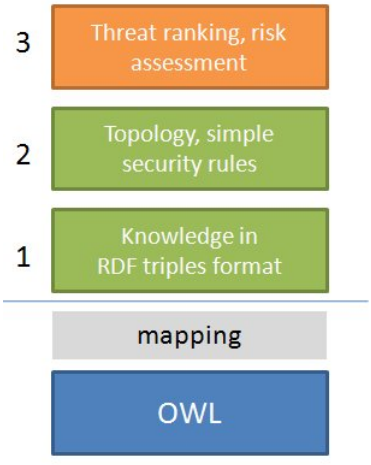

Fig. 4. DAT knowledge logical layers. Thanks to mapping the ontology (OWL file) is used to build two layers of DAT knowledge used to perform reasoning.

asses each threat (found in CI) severity level. It is done via the Bayesian network combined with the facts and rules maintained in bottom layers. More details follow in section 3.

3 Fusion of Bayesian Network and Ontology Knowledge

In this section the proposed approach to asses threats severity levels by means of Bayesian Network is presented. The general idea of fusion is presented in Fig. 5. The BN output computes the particular threat severity using observation about CI gained from ontology. The input is obtained from the first and the second layer of the knowledge stack. The first knowledge layer provides the network with the basic information about the threaten asset, threat itself, vulnerability and safeguard. The second layer, besides introducing the new threats using the security rules, allows to update the BN prior probabilities. For example added rules can increment the assets counters when new element is added to CI, or shape the particular node business value using some predefined conditions.

3.1 Structure of Proposed Bayesian Network

The structure of the proposed BN is presented in Fig. 6. The right arrows represent the input (observations), while the left arrows represent the posterior probability of fact the node is threaten by attack given the AT (Asset Type), VR (Vulnerability Risk) and SA (Safeguard Applied) observation.

Those observations are extracted by DAT using the knowledge about the CI. The AT observations represent the asset type. The information is adapted to emphasize the fact that some assets (elements in the CI) are more valuable than others.

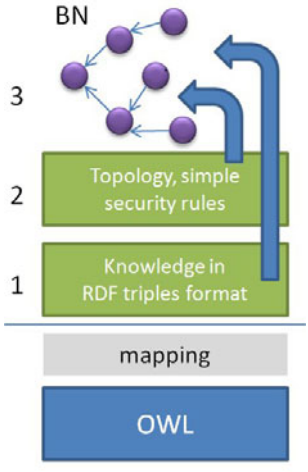

Fig. 5. Bayesian network is fed by the facts about CI from RDF triples, topology and simple security rules

Furthermore, the number of valuable assets also influences the total risk value. Particularly the AV (Asset Value) is used to increase the importance of SCADA servers, routers, RTUs and other critical elements. According to ISO standard [6], each network element may have vulnerabilities, which eventually put system in danger, therefore BN uses the VR (Vulnerability Risk) to evaluate its severity. The VR depends on the observation of fact that asset is applied and the VSL (Vulnerability Severity Level).

Eventually, the VR and AV are combined to asses the final value of the risk probability.

$$p(A = T|VRL, SA, AT) = \frac{p(A, VRL, SA, AT)}{p(VRL, SA, AT)} \tag{1}$$

3.2 Prior Probabilities Estimation Problem

The nominator in the equation 1 (based on the BN structure shown in Fig. 6) can be rewritten as in eq. 2:

$$\begin{aligned} p(A, VRL, SA, AT) = {} & p(VSL)p(AT)p(AS) \\ & p(AV|AT)p(VR|AS, VSL) \\ & p(A|VR, AV) \end{aligned} \tag{2}$$

In our approach, it is proposed to obtain probabilities distributions via the inferencing engine as is it shown in Fig. 5. Particularly the marginal probabilities distributions $(p(VSL), p(AT), p(AS))$ can be easily obtained via the histogram-based estimation method. In the example, the AT (asset type) variable can be

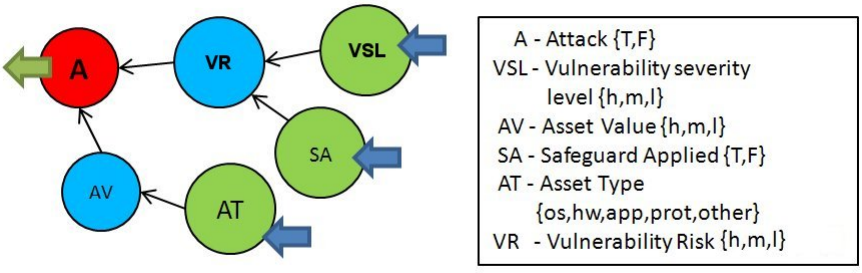

Fig. 6. Bayesian network details (h=high, m=medium, l=low, t=true, f=false, os=operating system, app=application, prot=protocol, hw=hardware)

assigned one of the Asset classes names (particularly these classes in ontology which have instantiated individuals). The probability of particular AT is computed via single rule (JESS's engine production rule) which counts the number of instances belonging to particular class and divides this value by total number of all instances building the CI. The same approach is used for other marginal probabilities. The advantage is the fact when new elements are introduced (not only hardware but also software) these distributions are updated and eventually the estimated risk value is different.

In our approach the conditional probabilities represent user defined preferences. Particularly $p(AV|AT)$ asses the given asset business value given the knowledge about its type, allows to stress the fact that some assets are more valuable than the others. In example user may define rule: "If AT is RTU then $P(AV = high|AT = RTU)$ is 0.99". The same approach is applied to $p(VR|AS, VSL)$ distribution.

The $p(A|VR, AV)$ (probability of attack given the VR an AV observation) is also strictly user depended (different users have different sense of balance between "asset business level" and "vulnerability risk"), and is computed using the same approach as for $p(AV|AT)$.

4 DAT Use Case Example

In this section the Decision Aid Tool demo is presented. The goal is to show "simulation mode" of the DAT, which allows to evaluate the risk of particular action prior to the physical manipulation. The demo uses ontology provided by "Topology Discovering Tool" and "Expert knowledge" provided by expert via DAT GUI interface and concerns following steps, where user:

- uses DAT to visualize topology graph
- turns off firewall and ant-virus applications on one of the routers
- identifies the risks and visualizes topology again
- simulates installation of firewall and anti-virus application on affected OS

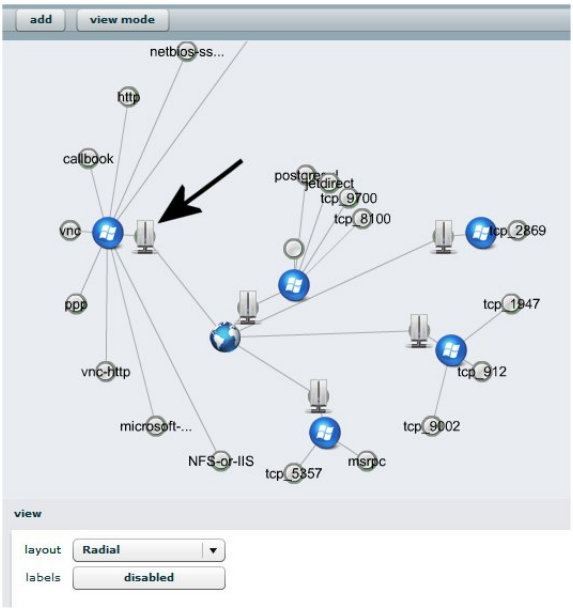

Fig. 7. Topology diagram (before security level assesment). The arrow indicates the router where the firewall and anti-viurs software will be switched off.

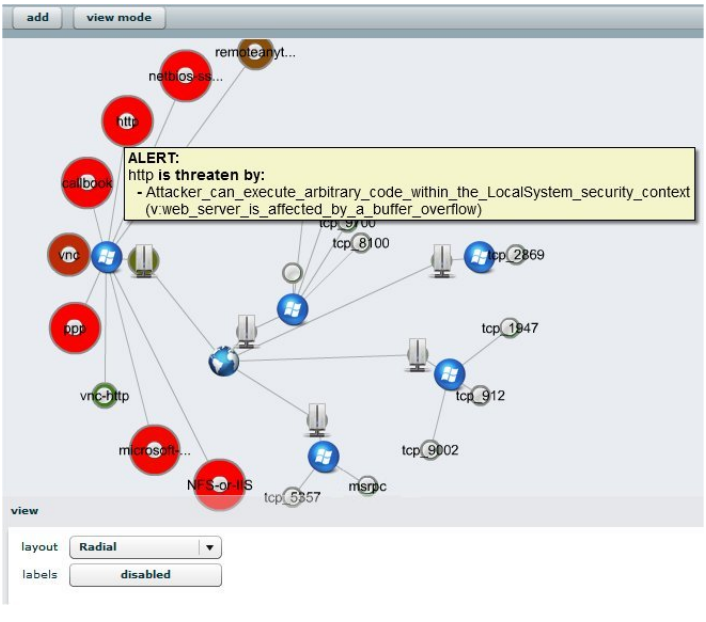

Fig. 8. Topology diagram and visualized threatened nodes

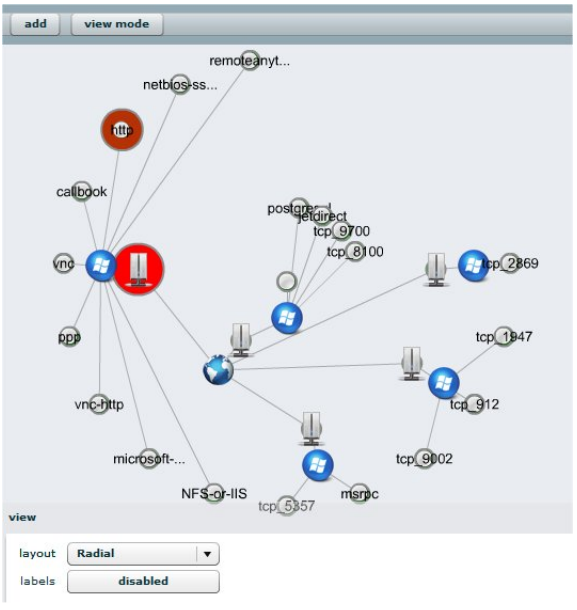

Fig. 9. Topology diagram and visualized threatened nodes after installing the firewall and anti-virus applications

Threats:

Asset	GLAW server and OPC Driver
Severity level	89,136
Solution	software fw
Threat	Lack of firewall between asset and WAN
Exploits	No firewall between asset and WAN
Asset	GLAW client
Severity level	89,136
Solution	software fw
Threat	Lack of firewall between asset and WAN
Exploits	No firewall between asset and WAN
Asset	IFIX server
Severity level	89,136
Solution	software fw
Threat	Lack of firewall between asset and WAN
Exploits	No firewall between asset and WAN
Asset	WSN RTU
Severity level	89,136
Solution	software fw
Threat	Lack of firewall between asset and WAN
Exploits	No firewall between asset and WAN
Asset	Elsag RTU
Severity level	89,136
Solution	software fw
Threat	Lack of firewall between asset and WAN
Exploits	No firewall between asset and WAN

Fig. 10. Security report with the ranked threats

At the beginning of the demo the system topology is analyzed. The example topology of our test bed can be shown in Fig. 7. The arrow indicated the router where firewall and anti-virus applications will be turned off. As is it shown in Fig. 8 single action has impact on many nodes being in relation with affected router. The cascading effect can be noticed. Turned off firewall and anti-virus protection exposes the W2K operating systems to different network attacks, which eventually has impact on the provisioning of the applications hosted by that OS. What is more it is also assumed that the IIS WWW server with WebDav service is enabled on W2K OS (default configuration is assumed due to the lack of detailed information about ran applications and services). The majority of discovered problems can be solved by installing anti-virus and firewall applications on W2K OS. Therefore the user inserts into DAT the information that such an action has been taken and as result new security report is obtained (the visualized topology can be shown in Fig. 9).

5 Conclusion

In this paper, the fusion of the ontology-based approach and the Bayesian network is proposed. Such innovative solution is applied in Decision Aid Tool for critical infrastructures security status assessment.

The sample result of CI system security evaluation by DAT is presented in Figure 10. The presented security report contains the ranked threats for discovered assets with their threat severity value calculated by the Bayesian network. Moreover, in the security report, details about the detected threats and the proposed solutions are given.

Acknowledgment

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 225553 (INSPIRE Project).

References

1. European Parliament legislative resolution of 10 July 2007, on the proposal for a Council directive on the identification and designation of European Critical Infrastructure and the assessment of the need to improve their protection (COM(2006)0787 C6-0053/2007 2006/0276(CNS)) (July 2007)
2. XiaoFeng, D., YuJiong, G., Kun, Y.: Study on Intelligent Maintenance Decision Support System Using for Power Plant Equipment. In: Proc. of the IEEE International Conference on Automation and Logistics Qingdao, China, 96100 (September 2008)
3. Lee, S.J., Mo, K., Seong, P.H.: Development of an Integrated Decision Support System to Aid the Cognitive Activities of Operators in Main Control Rooms of Nuclear Power Plants. In: Proc. of IEEE Symposium on Computational Intelligence in Multicriteria Decision Making (MCDM), pp. 146–152 (2007)

4. Zhang, B., Wu, G., Shang, S.: Research on Decision Support System of Water Pollution Control Based On Immune Agent. In: Proc. of International Symposium on Computer Science and Computational Technology, ISCSCT, vol. 1, pp. 114–117 (2008)
5. Xie, L., Wang, Z., Bian, L.: The Research of Oileld Flood Precaution Decision Support System. In: Proc. of International Seminar on Business and Information Management, ISBIM 2008, vol. 2, pp. 236–239 (December 2008)
6. ISO/IEC 13335-1:2004, Information Technology Security Techniques Management of information and communications technology security Part 1: Concepts and models for information and communications technology security management (2004)
7. Choras, M., Stachowicz, A., Kozik, R., Flizikowski, A., Renk, R.: Ontology-based approach to SCADA systems vulnerabilities representation for CIP. Electronics 11, 35–38 (2009)
8. Choras, M., Flizikowski, A., Kozik, R., Renk, R., Holubowicz, W.: Ontology-Based Reasoning Combined with Inference Engine for SCADA-ICT Interdependencies, Vulnerabilities and Threats Analysis. In: Pre-Proc. of 4th International Workshop on Critical Information Infrastructures Security, CRITIS 2009, Bonn, Germany, pp. 203–214. Fraunhofer IAIS (2009)
9. SWRL: A Semantic Web Rule Language Combning OWL and RuleML, W3C Member Submission, `http://www.w3.org/Submission/SWRL/`
10. Deliverable D2.3, Ontological approach and inference engine, INSPIRE Project (2009)
11. Macaulay, T.: Critical infrastructure: Understanding Its Component Parts, Vulnerabilities, Operating Risks, and Interdependencies (August 2008)
12. Lewis, T.G.: Critical Infrastructure Protection in Homeland Security: Defending a Networked Nation. Wiley-Interscience, Hoboken (2006)
13. McClanahan, R.H.: The benefits of networked SCADA systems utilizing IP-enabled networks. IEEE, Los Alamitos (2002)