An Information Security and Privacy Self-Assessment (ISPSA) Tool for Internet Users

Tomislav Galba[1], Kresimir Solic[2], Ivica Lukic[1]

[1] J.J. Strossmayer University, Faculty of Electrical Engineering, Kneza Trpimira 2b, Osijek, Croatia, tomislav.galba@etfos.hr, ivica.lukic@etfos.hr

[2] J.J. Strossmayer University, Faculty of Medicine, Josipa Huttlera 4, Osijek, Croatia, kresimir.solic@mefos.hr

Abstract: Privacy and overall information security are significantly affected by an Internet users' awareness, knowledge and behavior. Therefore, there is a user's awareness assessment needed before developing security solutions that will include the user of the information system. The present paper proposes a validated measurement instrument developed as a web based software security and privacy tool for self-assessment (ISPSA). This solution is based on a scientifically validated questionnaire, OWL ontology concept, evidential reasoning approach and the intelligent agent's algorithm. The main goal of this paper is to propose the solution that will raise awareness among Internet users on privacy and information security issues.

Keywords: awareness; Evidential Reasoning; Information Security; Intelligent Agent; OWL ontology; UISAQ; users

1 Introduction and Problem Statement

Many persistent and new problems regarding information security and user's privacy are causing the development of a wide range of new security concepts [1-4]. It is crucial for such new security solutions to consider the human component as well, due to the fact that users can significantly affect the overall security of an information system [5-7]. The security management should therefore integrate separate security areas in the overall security solution [8] combining risk management, infrastructure safety, hardware solutions, security protocols and users' education and control.

The most common approach to the user's low security awareness problem solution is education, as control could be considered unethical or hard to manage. The education of users on how to create better passwords [9], how to handle private data, how to be more careful towards unknown collocutors on the Internet [10],

should be based on some measurements of the users' current level of awareness and behavior. Several rare solutions for measuring users' awareness [11-13] and their potentially risky behavior [14, 15] were actually the reasons for building the present Information Security and Privacy Self-Assessment (ISPSA) tool for different users of the information systems, and more generally, for every user of the Internet.

The proposed solution is based on previously developed and validated information security questionnaire [16] combined into OWL ontology [17] with calculations relying on Evidential Reasoning approach [18] and back coupling, founded on the Intelligent Agent's algorithm [19].

This present paper is divided into following parts: Chapter two consists of a few subsections describing the complete background of the online self-assessment tool, referring to the validated UISAQ questionnaire, a brief introduction and the description of OWL ontologies, enhanced evidential reasoning algorithm and the description of the implementation of an intelligent agent. Chapter three provides a description of the complete solution with an output example and comments. Chapter four concludes the paper.

2 Background

2.1 Users' Information Security Awareness Questionnaire

The scientifically validated Users' Information Security Awareness Questionnaire (UISAQ) is a reliable measurement instrument [16]. The questionnaire has 33 items divided into two main scales: one scale measures the users' potentially risky behavior and the other measures the users' awareness. Each scale is divided into three basic subscales with 5 or 6 items presenting questions (Figure 1). For each question there are five answers proposed, graded on the Likert scale from one to five, where five means "good", as seen from the perspective of the information security and privacy engineer.

UISAQ possesses two more elements that do not belong to the validated part of the questionnaire. On the first page of the questionnaire there are demographic questions that can be changed regarding the research aims and the category of users, while on the last page there are two external questions and acknowledgments. Those two UISAQ elements were not needed and are not used as constructive elements in building the proposed self-assessment tool.

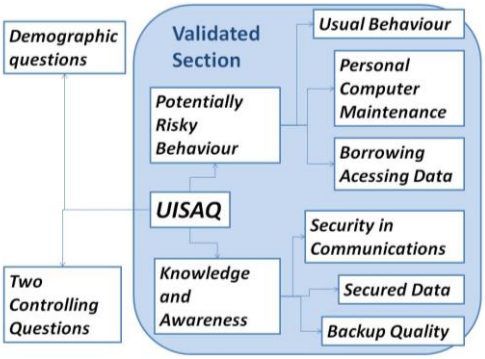

Figure 1
Segments of UISAQ

2.2 OWL Ontology

Ontology is used for formal definition of knowledge on some domain of interest. Formally, the defined knowledge should be both readable to a programmed intelligent agent and understandable to the human expert. In order to meet those requirements, there should be a language with well-defined semantics used [20]. There are many reasons for creating ontology [21]. Some of the most important reasons are:

- Reusability of domain knowledge which helps to develop a large and detailed ontology allowing integration with other new ontologies.
- Sharing common understanding for specific information domain as one of the main goals in ontology development that allows the extraction and aggregation of information from different domains.

The process of building an ontology is simple and often based on defining concepts with their properties and defining relations between concepts [17]. Nowadays, the most frequently used version of the ontology is OWL, while there are many open source software solutions to choose from. OWL is an international encoding and exchanging standard with the purpose of enabling communication between computers. It is developed as an extension on two semantic web standards, the RDF (Resource Description Framework) and RDF schema. Both of those standards were endorsed by W3C. So, naturally, OWL is a valid RDF document and also a well-formed XML document which facilitates processing with the already available XML and RDF processing tools and API-s. The OWL is also characterized as a higher level of expression relative to RDF, shown in Fig. 2. This property is very important in the process of describing attributes of some enterprises, since OWL described information is considered knowledge instead of just a simple data set.

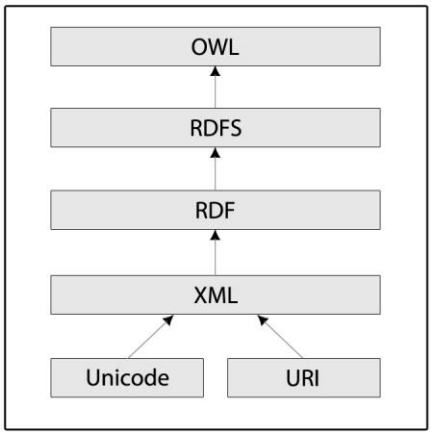

Figure 2
OWL hierarchy

There are three types of OWL expressive sub-languages [14, 22]:

- OWL Lite – is a subset of OWL DL that provides only the basics for subclass hierarchy construction which results in better performance of complete reasoners for OWL Lite.
- OWL DL – is a subset of OWL Full with less restrictions compared with the OWL Lite, designed to support description logic framework.
- OWL Full – very expressive language with no restrictions, designed as an extension to RDF. It contains all OWL language constructs with the possibility of unconstrained use of RFD constructs, resulting in full syntactic and semantic compatibility with RDF.

The ontology consists of classes, properties and individuals where [22, 23]:

- Classes are main building blocks representing sets of resources also known as individuals. Additional class description is achieved by adding properties from RDFS or OWL vocabularies. Also, a class can be related to other more general classes using subOfClassOf property which is shown in Figure 3.

```
<owl: Class rdf: ID=”Class1” />
<owl: Class rdf: ID=”Class2” />
        <rdfs: subClassOf rdf: resource=”#Class1” />
</owl: Class>
```

Figure 3
Class definition

- Properties are used to define relations between individuals. There are two main categories of properties in OWL. The object property as an instance of predefined OWL class defines a link between individuals in two different classes, and Data type property that defines the relationship between the individual and data values. As with classes, properties can be described by adding sub elements such as subPropertyOf. Also, there are lots of other things we can add to the property, like establishment of taxonomy, domain and range of a property etc., shown in Figure 4.

```
<owl: ObjectProperty rdf: ID=”hasPDescriptor” >
        <rdfs: domain rdf: resource=”#P” />
        <rdfs: range rdf: resource=”#PDescriptor” />
</owl: ObjectProperty >
```

Figure 4
Properties definition

- Instances – belong to classes and are used to express semantics of classes and properties. They are also related to other instances and data values by defining the properties. Two facts or axioms are used to define an instance:
    - Facts about properties of instances and class membership
    - Facts about identity of an instance

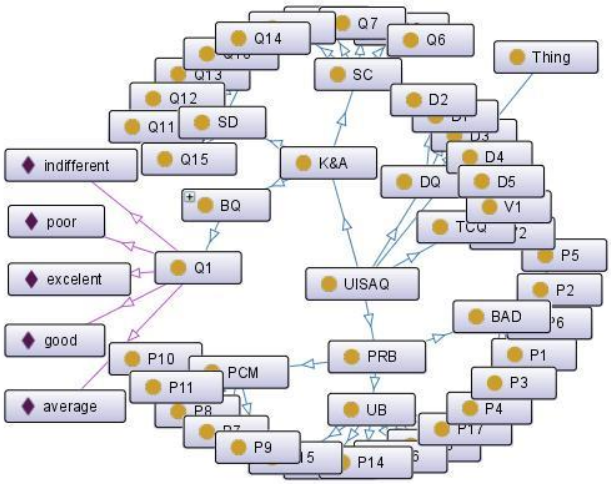

Figure 5
OWL ontology based on UISAQ built in Protégé

As stated before, classes are the main building blocks of ontology that contain instances (objects in the domain of interest). Classes are placed into superclass-subclass hierarchy. There are many benefits of using OWL ontology to formally define knowledge [17] such as re-usage among scientists and experts etc. Entities of basic subclasses are defined as grades from one to five, meaning from poor to excellent. The authors of this paper used the Protégé software solution [24] for building an OWL ontology, based on UISAQ questionnaire as shown in Figure 5.

2.3 Evidential Reasoning Algorithm

Evidential Reasoning Algorithm (ERA) is based on Dempster-Shafer theory [25, 26], decision-making theory [27] and evaluation analysis model [28]. It is also suitable for multiple-criteria decision analysis problems.

ERA is able to calculate with both quantitative and qualitative measurements considering subjective judgements with uncertainty, and objective, absolute or partial data [18]. ERA can be used to estimate the current state of technical system from many aspects. Estimated current state can be compared to the previous current state(s) of the same system, to current state(s) of another system or to previously defined referent value(s).

In this paper, the authors used the enhanced version of ERA [29]. Enhanced ERA allows aggregation of grades through a more complicated scheme with more than two levels regarding parent-child relation between the elements of the system.

The evaluation grades are defined in the same way as in the OWL ontology of UISAQ, namely as: poor, indifferent, average, good and excellent (P, I, A, G, E) while the uncertainty being calculated upon regarding of the missing answers.

The utility grade with associated utility interval is calculated from the distribution of grades with uncertainty and represents a single numerical value that is more suitable for comparing purposes. Utility interval is defined by uncertainty. Utility number can be calculated from any distribution of grades represented by any ontology subclass, single question or group of questions in the questionnaire.

An example of a calculation and aggregation through one ontology subgroup to a group of questions is shown in Table 1. More detailed explanation of ERA is available in wide scientific literature about this subject with many examples for technical systems state analysis [30-32], organizational decision making [33-35] and rarely for the evaluation of humans properties [14, 36, 37].

In order to apply ERA on ontology structure, three simple rules should be followed [14]: the hierarchical structure defined in the ontology should be strictly a cyclic graph; every direct relation should be “one-to-one” or “one-to-many” relationship; and there should be an existing crossing between classes in ontology reorganized. By additional reorganization it is possible to define mirrored classes in ontology.

Table 1
Example of calculations in the self-assessment process for the single user

UISAQ naming	**Item**	**Subscale**	**Scale**	**Total**	
Ontology naming	Class	Subclass	Superclass	Main class	
ERA naming	Basic attribute	Inter-attribute	Inter-attribute	General attribute	Utility (with uncertainty)
P1, P4	G	G(0.371)	I(0.085)	P(0.090)	U=0.772
P2, P3, P5	E	E(0.629)	A(0.045)	I(0.038)	(0.765-
P6	I	I(0.156)	G(0.222)	A(0.114)	0.779)
P7,P8	G	A(0.156)	E(0.618)	G(0.370)	
P14	E	G(0.344)	H(0.031)	E(0.375)	
P16	E	E(0.344)		H(0.013)	
P17	A				
P9	-	I(0.138)			
P10,P12-P15	E	E(0.747)			
P11	I	H(0.115)			
Q1, Q3	E	P(0.191)	P(0.194)		
Q2	G	A(0.191)	A(0.194)		
Q4	P	G(0.191)	G(0.481)		
Q5	A	E(0.429)	E(0.131)		
Q6,Q8	P	P(0.409)			
Q7,Q10	A	A(0.409)			
Q9	G	G(0.182)			
Q11-Q16	G	G(1.00)			

2.4 Intelligent Reflex Agent

When talking about intelligent agents and environments in which they act, we mainly perceive them as software or hardware implementations. Both of these implementations are based on some input from sensors which can also be software (user input form, data from database etc.) or hardware (temperature sensor, moisture sensor etc.), taking actions through actuators as a result of an analysis. An actuator can be a robotic arm, or in other cases actions can be taken on the software level either in terms of showing some data to the user, by changing the data in the database, or by an automatic creation of documents, reports etc. Mathematically speaking, the behavior of an agent is described with a function which transforms any input to adequate action [19]. There are four types of intelligent agents which satisfy the above-mentioned:

- Simple reflex agent – the simplest type of an intelligent agent where every action is based only on current input regardless of everything else.

- Model-based reflex agent – a more advanced type of an intelligent agent in relation to a simple reflex agent where action depends on current input from sensors and the history of previous actions for different inputs.
- Goal-based agent – an intelligent agent with predefined goals, similar to the model-based agent, only with the difference in checking the impact of a certain action on a defined goal.
- Utility based agent – operates through a utility function which is used to map a state. The result of that function is some kind of measure which defines how desirable a particular state of an agent is.

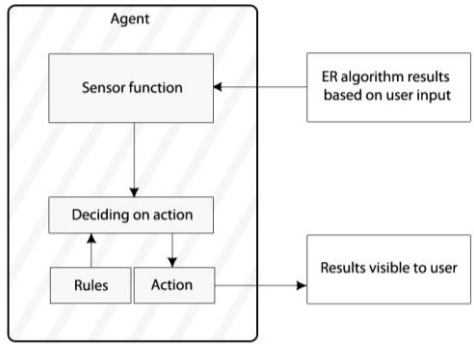

Figure 6
Simple reflex agent

This paper presents an ongoing work on the intelligent agent for online information security and privacy self-assessment tool. Since the main task of our intelligent agent in the self-assessment tool is to make decisions based on the final grade for the level of security, there is the model of a simple-reflex agent used. The use of this model is satisfying because there is no need to look at past, but only on current states. The expected output from the agent is information about whether to increase the level of security and the need to emphasize critical elements or sub-elements of the user behavior. The structure of our intelligent agent is shown in Figure 6.

Input variables for our intelligent agent are as follows:

- Rp – Referent value for average safe security level referring to the desired level of security which is predefined, based on previous testing and comparison with previous information system security assessment.
- U – Utility value given by information security assessment based on enhanced evidential reasoning algorithm.

Referent values are defined in previous testing conducted on the sample of 701 Internet users with different age, gender, technical knowledge, level of education, working position, coming from different institutions and business subjects [38]. The referent values are shown in Figure 8.

As stated in [19], a simple reflex agent brings simple decisions based on the current environment state, and since our intelligent agent is based on the simple reflex agent the decisions that had to be made are as follows:

- The proposed corrections of certain security segments if overall utility value is below the desired level of security.
- If the overall utility value is greater than the desired utility value including the correction value, then the intelligent agent searches for worst-rated basic attributes or items.

As an addition to the above stated, the intelligent agent has a predefined set of critical questions, to which, attention has to be paid, under all circumstances.

3 Software Web Solution

The present section introduces a new solution which implements all the elements stated earlier in this paper. One of the main goals of this work is to consider a person as negative influence on the information security system, in order to improve the current solutions and make future implementations better. Figure 7 shows a self-assessment tool structure with OWL Ontology structure based on UISAQ described in section 2.2 and the intelligent agent described in section 2.4.

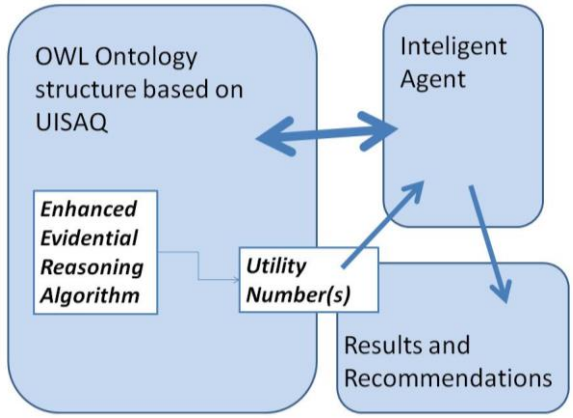

Figure 7
Self-Assessment Software Tool Elements

For a successful self-assessment, a user needs to pass over 33 predefined points divided into two major segments, which are then divided into six sub-segments, from which, every point has a different meaning (frequency, degree of security, degree of belief, degree of importance).

After passing through 33 points (it is not necessary to answer all questions), the algorithm for enhanced evidential reasoning is applied. The resultant values are shown in the form of a graph in Figure 8.

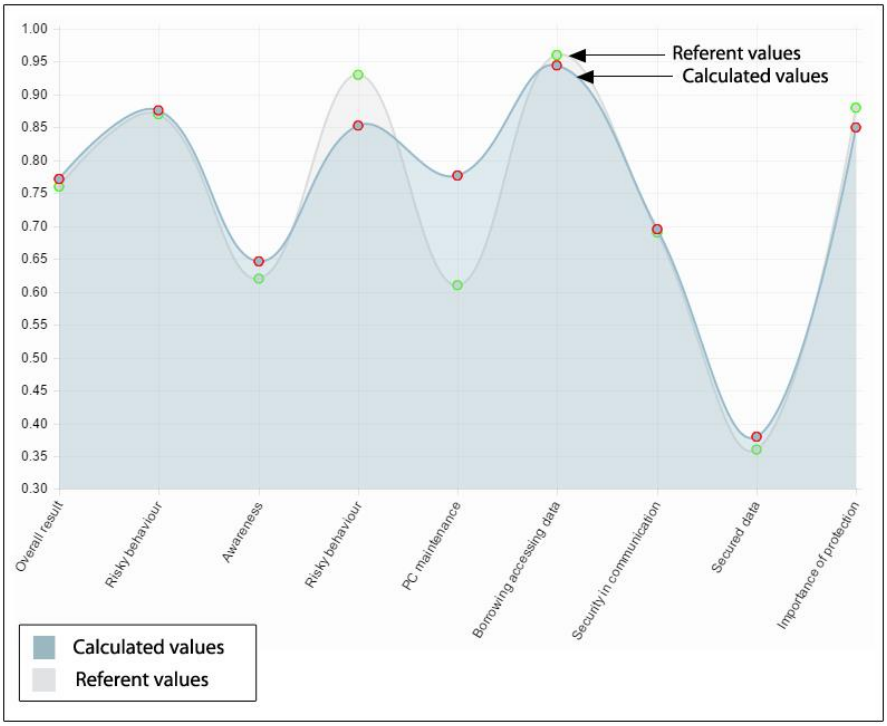

Figure 8
Self-assessment tool resulting graph

Figure 8 shows graph results for user input. From the calculated values we can conclude that the users have very good habits in most of the categories except few concerning the awareness, borrowing access data, protecting the data and poor backup habits. Except graph results, at the end, the user will the results from the intelligent agent, which emphasize the worst utility values and give recommendations for improvement. The working version of the questionnaire can be found in the referenced work[1].

[1] Towards Information Security and Privacy Self-Assessment (ISPSA) Tool for Internet Users link available at: http://vns.etfos.hr/Samoprocjena/

4 Discussion

The proposed self-assessment web solution tool has proven to be a valid measurement instrument that can be used to raise awareness among Internet users concerning privacy and information security issues. Once the user has passed through the 33 points in the self-assessment tool, the overall result is calculated. Also, the results for each of the two areas regarding behavior and knowledge, and the results for six subareas: usual behavior, PC maintenance, borrowing access data, security in communications, security of data and quality of backup are calculated. Moreover, there is an intelligent agent function which compares and analyses the overall user result with referent values representing the general user's behavior, knowledge and awareness called. Depending on comparison results, the user is pointed to critical security issues and provided with relevant recommendations for security improvement.

This Information Security and Privacy Self-Assessment Tool (ISPSA tool) for internet users is modular, based on scientifically validated UISAQ questionnaire, OWL ontology concept, enhanced Evidential Reasoning approach and intelligent agent's algorithm. ISPSA tool therefore benefits from each element's properties such as: measurement quality of the questionnaire, human machine utilization of the ontology, calculations with subjective assessment of evidential reasoning, and the agent's automation of analysis, as well as the presentation of results.

Future work will include some additional testing of the English version and making the self-assessment tool available freely to all Internet users. Also, with international collaboration, it should be possible to develop a better questionnaire, one which is more applicable to the world-wide Internet users' knowledge and habits and also more suitable to newly emerging information security issues. The modularity of the proposed solution also allows for the improvement of the different segments.

References

[1] Sallai, G.: Future Internet Vision and Research Clusters, Acta Polytechnica Hungarica, 2014, Vol. 11, No. 7, pp. 5-24

[2] Tot, L., Grubor, G., Takacs, M.: Introducing the Information Security Management System in Cloud Computing Environment, Acta Polytechnica Hungarica, 2015, Vol. 12, No. 3, pp. 147-166

[3] Vokorokos, L., Baláž, A., Madoš, B.: Application Security through Sandbox Virtualization, Acta Polytechnica Hungarica, 2015, Vol. 12, No. 1, pp. 83-101

[4] Vokorokos, L., Pekár, A., Norbert, A.: Yet Another Attempt in User Authentication, Acta Polytechnica Hungarica, 2013, Vol. 10, No. 3, pp. 37-50

[5] H, Thompson: The Human Element of Information Security, IEEE Security&Privacy 2013, Vol. 11, pp. 32-35

[6] Lukasik, S. J.: Protecting Users of the Cyber Commons, Communications of the ACM. 54, 9(2011) pp. 54-61

[7] Solic, K., Sebo, D., Jovic, F., Ilakovac, V.: Possible Decrease of Spam in the Email Communication, IEEE MIPRO 2011, pp. 170-173

[8] Michelberger Jr, P., Labodi, C.: After Information Security - Before a Paradigm Change (A Complex Enterprise Security Model), Acta polytechnica Hungarica, 2012, Vol. 9, No.4, pp. 101-116

[9] Keszthelyi, A.: About Passwords, Acta Polytechnica Hungarica, 2013, Vol. 10, No. 6, pp. 99-118

[10] I, Kirlappos., M. A, Sasse.: Security Education against Phishing: A Modest Proposal for a Major Rethink, IEEE Security&Privacy, 2012, Vol. 10, pp. 24-32

[11] Vuković, M., Katušić, D., Skočir, P., Jevtić, D., Delonga, L., Trutin, D.: User Privacy Risk Calculator, Proceedings of the 22nd International Conference on Software, Telecommunications and Computer Networks, Split, 2014, pp. 1-6

[12] Aggeliki, T., Spyros, K., Maria, K., Evangelos, K.: Process-Variance Models in Information Security Awareness Research, Information Management & Computer Security, 2008, Vol. 16, pp. 271 - 287

[13] Petri, P., Mikko, S.: Improving Employees' Compliance through Information Systems Security Training: an Action Research Study, Puhakainen & Siponen Appendices, 2010, Vol. 34, pp. 757-778

[14] Solic, K., Jovic, F., Blazevic, D.: An Approach to the Assessment of Potentially Risky Behavior of ICT systems' users, Tehnički vjesnik - Technical Gazette, 2013, pp. 335-342

[15] Novakovic, L., McGill, T., Dixon, M.: Understanding User Behavior towards Passwords through Acceptance and Use Modelling, International Journal of Information Security and Privacy (IJISP) 2009, Vol. 3, pp. 9-27

[16] Velki, T., Solic, K., Ocevcic, H.: Development of Users' Information Security Awareness Questionaire (UISAQ) - Ongoing Work, MIPRO 2014

[17] Horridge, M.: A Practical Guide To Building OWL Ontologies Using Protege 4 and CO-ODE Tools,The University of Manchester, 2011, URL: http://owl.cs.manchester.ac.uk/tutorials/protegeowltutorial/resources/ProtegeOWLTutorialP4_v1_3.pdf

[18] Jian-Bo, Y., Dong-Ling, X.: On the Evidential Reasoning Algorithm for Multiple Attribute Decision Analysis under Uncertainty, IEEE Transactions on Systems, Man and Cybernetics, 2002, Vol. 32, pp. 289-304

[19] Russell, S., Norvig, P.: Artificial Intelligence, A Modern Approach. Third Edition, PEARSON, 2010, pp. 46-50

[20] Gali, A., Chen, C. X., Claypool, K. T., Uceda-Sosa, R.: From Ontology to Relational Databases. Shan Wang et al (Eds.): Conceptual Modeling for Advanced Application Domains, LNCS, 2005, Vol. 3289, pp. 278-289

[21] Noy, N. F., McGuiness, D. L.: Ontology Development 101: A Guide to Creating Your First Ontology, Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SML-2001-0880, 2001

[22] OWL Web Ontology Language Overview available at: http://www.w3.org/TR/owl-features/

[23] Korda, N., Astrova, I., Kalja, A.: Storing Owl Ontologies in SQL Relational Databases, Engineering and Technology, 2007, Vol. 29

[24] Protégé software tool, Stanford Center for Biomedical Informatics Research. URL: http://protege.stanford.edu/

[25] Dempster, A. P.: Upper and Lower Probabilities Induced by a Multivalued Mapping, Ann Math Stat, 1967, pp. 325-339

[26] Shafer, G.: A Mathematical Theory of Evidence, Princeton University Press, 1976, New Jersey

[27] Zhou, M., Liu, X. B., Yang, J. B.: Evidential Reasoning-Based Nonlinear Programming Model for MCDA under Fuzzy Weights and Utilities, International Journal of Intelligent Systems, 2010, pp. 31-58

[28] Zhang, Z. J., Yang, J. B., Xu, D. L.: A Hierarchical Analysis Model for Multiobjective Decision making, Analysis, Design and Evaluation of Man–Machine Systems, Pergamon, Oxford, U.K, 1990, pp. 13-18

[29] Yang, J. B., Singh, M. G.: An Evidential Reasoning Approach for Multiple Attribute Decision Making with Uncertainty, IEEE Trans Syst Man Cybern, 1994), pp. 1-18

[30] Jagnjic, Z., Slavek, N., Blazevic, D.: Condition Based Maintenance of Power Distribution System, EUROSIM, 2004, pp. 13-14

[31] Xin-Bao, L., Mi, Z., Jian-Bo, Y., Shan-Lin, Y.: Assessment of Strategic R&D Projects for Car Manufacturers Based on the Evidential Reasoning Approach, Int J Comput Intell Syst, 2008, Vol. 1, pp. 24-49

[32] Zhang, X. D., Zhao, H., Wei, S. Z.: Research on Subjective and Objective Evidence Fusion Method in Oil Reserve Forecast, J Syst Simul, 2005, pp. 2537-2540

[33] Beynon, M., Cosker, D., Marshall, D.: An Expert System for Multi-Criteria Decision Making Using Dempster–Shafer Theory, Expert Syst Appl, 2001, pp. 357-367

[34] Wu, W. Z., Zhang, M., Li, H. Z., Mi, J. S.: Knowledge Reduction in Random Information Systems via Dempster–Shafer Theory of Evidence, Inf Sci. 174, 2005, pp. 143-164

[35] Srivastava, R. P., Liu, L.: Applications of Belief Functions in Business Decisions: a Review, Inf Syst Frontiers. 5, 2003, pp. 359-378

[36] Kong, G., Dong-Ling, X., Body, R., Jian-Bo, Y., Mackway-Jones, K., Carley, S.: A Belief Rule-based Decision Support System for Clinical Risk Assessment of Cardiac Chest Pain, European Journal of Operational Research 2012, pp. 564-573

[37] Solic, K., Kramaric-Ratkovic, K., Antolovic, Z.: Applying Evidential Reasoning Approach in Biomedicine - Example on the APGAR Score, Simpozij HDMI Medicinska Informatika, Dubrovnik, 2013, pp. 7-9

[38] Solic, K., Velki, T., Galba, T.: Empirical Study on ICT System's Users' Risky Behavior and Security Awareness, IEEE MIPRO / ISS, 2015, pp. 1623-1627