A Decision Support Model for Situational Awareness in National Cyber Operations Centers

Roman Graf*, Florian Skopik* and Kenny Whitebloom†
*AIT Austrian Institute of Technology GmbH, Vienna, Austria
{roman.graf, florian.skopik}@ait.ac.at
†Digital Public Library of America (DPLA)
{kenny}@dp.la

***Abstract*—Advances in situational awareness technology have led to the creation of increasingly sophisticated tools across different application domains, often involving non-textual, highly dimensional, and multimedia data. Automated tools aim to address a number of situational awareness challenges, such as complex system topology, rapidly changing technologies, high noise to signal ratio, and multi-faceted threats. These factors make real-time situational awareness of cyber operations for the National Cyber Operations Centers very difficult to achieve. Appropriate data analysis techniques, in combination with modern anomaly detection output data and user knowledge, may provide solutions in real-time that could replace human input for many situational awareness analysis tasks.**

***Keywords*—situational awareness, cyber security, data analytics**

I. Introduction

ACCORDING to the EU NIS Directive [1] - the provider of a Critical Infrastructure (CI) is obliged to report cyber incidents and information about CI's security status to the National Cyber Security Center. Modern Cyber Security tools produce a huge amount of incident and threat messages. Currently, the National Cyber Security Centers are facing problems related to the analysis and evaluation of incoming reports. The proposed Expert System can help to mitigate this problem. It is not affordable for a human analyst to handle all of them manually. The proposed expert system for situational awareness is focused on making automatic cyber defense actions in real-time when possible and where it makes sense. Such a system could replace human input for many situational awareness analysis tasks, making analysis faster and cheaper. Human operators make high quality decisions and are good for unusual system behaviour, but they cant work 24/7 and often make errors or need time for learning CI. In case of uncertainty, this system could help to categorise a problem and better visualize it.

This paper presents an approach for merging two critical sets of information data aggregated automatically from security information systems and expert knowledge related to situational awareness in cyberspace and defines and computes fuzzy models designed to provide cyber analysts with semi-automatic estimations of situational awareness for critical infrastructure metrics. Our goal is to make use of a solid knowledge base automatically aggregated from cyber systems to detect deviations and inaccuracies, thereby improving the quality of decision-making processes. Contrary to the human brain, which is well designed to derive situational awareness from the world based on a complex set of cognitive processes and schema built on experience, the artificial world of cyber operations lacks a set of well-integrated tools and models that can bridge the gap between data collection and situation comprehension. Addressing this need will provide cyber analysts with a comprehensive set of information that they can use to develop cyber security awareness and, ultimately, secure cyber operations.

In this paper, we discuss recent cyber situational awareness techniques that can support analysis and decision-making in a variety of different application scenarios. Fuzzy rules are employed for handling a level of uncertainty associated with semi-automated data aggregation from security information systems. Decision support based on the elaborated rule engine provided by cyber systems, fuzzy rules, and scored metrics are meant to support cyber analysts with suggestions in the process of analysing situational awareness. Situational awareness metrics employed in Fuzzy models are quantitative characteristics that represent the security state of a network and help measure cyber situation awareness. The modern

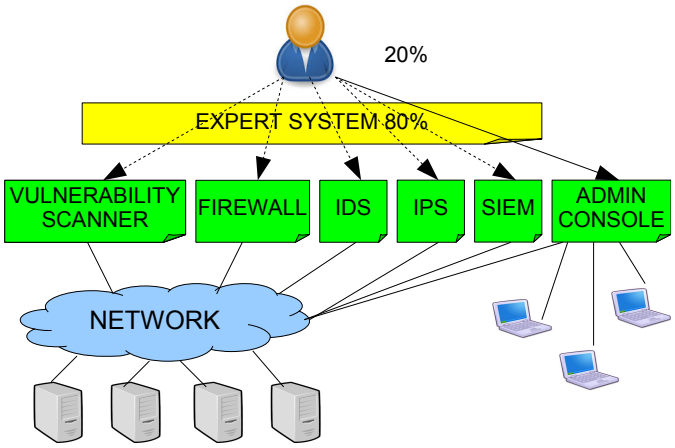

Fig. 1: The overview of the Expert System for decision support.

[1] http://www.consilium.europa.eu/en/policies/cyber-security/

Situational Awareness employs multiple well-established cyber systems (see Figure 1), such as Vulnerability Scanner, Firewall, Intrusion Detection System (IDS), Intrusion Prevention System (IPS), and Security Information and Event Management (SIEM). These systems produce qualitative reports about vulnerabilities and incidents occurring in the computer network. An advantage of a Fuzzy-based system in comparison to fix-linked rule system is that a rule base can be dynamically adapted, dependent on current attack vectors. The threat level setup of aforementioned basis indicators (e.g. detected vulnerabilities and incidents) and their impact estimation can be automatically and dynamically adapted, and new rules computation can be performed. Produced reports have to be analysed, forwarded to the Admin Console, and correspondent actions should be triggered to mitigate incident impact. But it is difficult to handle having to constantly raise the number of reports manually. The proposed Expert System could facilitate a decision making process in most cases, where the decision is defined by institutional policies and can be performed automatically. Only in cases of uncertainty or new precedential cases would manual analysis be required.

The realistic use case from the notable European cyber security research projects evaluation results was used as an input for fuzzy modelling and is presented in the evaluation part of this paper, as well as an analysis of the aggregated and processed information.

Provided approach is meant for application in a state level in National Cyber Security Operations Centers. A security analyst should automatically assess, which CIs require more attention. For example, proposed system could control, which points on the map should flash to allocate more resources of the National Incident Response Team to this CI, to estimate the national SA or to enact early warning in particular sectors.

The main contribution of the paper is the employment of fuzzy rule-based technique to facilitate decision-making and costs of SA analysis in National Cyber Operations Centers.

The remainder of the paper is organized as follows: section 2 provides an overview of cyber SA concepts and related work; section 3 explains the data aggregation process; section 4 presents the fuzzy modelling and selected membership functions; section 5 presents experimental results and interpretation; and, lastly, section 6 concludes the paper and provides an outlook on planned future work.

II. Related Work

The Fuzzy Formal Concept Analysis [1] for high-level perception-oriented SA systems generates a knowledge representation lattice augmenting human perception. In contrast to the proposed system information is represented with ontological concepts, which require additional setup. The Fuzzy group decision support system [2] generates team SA in a distributed team work context with individual weights and influence degrees. A multi-agent fuzzy consensus model for SA [3] considers decision-making agents claiming their preferences on a topic of interest. These two approaches differ from proposed system, where decision support is provided for one operator. SA established though Fuzzy-based value of information [4] is calculated for complex military environments using fuzzy associative memory, which is a k-dimensional table where each dimension corresponds to the associated input rule set. This method is similar to the proposed approach with the difference that our system implements a feedback-loop that constantly automatically improves the fuzzy rules and generates new rules by means of developed software. The swarm-based semantic fuzzy reasoning for SA computing [5] is employed to focus the search for fuzzy reasoning agent, which infers relevant situations by browsing the RDF graph consisting of information about situations. Contrary to suggested approach, this method is designed for RDF graphs, needs ontologies, semantic data model and classification schema. A fuzzy cognitive SA for airport security [6] is ontology- and agent-based and employed for modelling of SA integrating ontological meta-models: situation theory ontology and saw core ontology, whereas in our method we do not employ ontologies to keep it more simple. The SIEM tool [7], [8] is very useful for automatic threat identification in an organisational network. Professional setup and maintenance is required to optimise the output and adjust it to the needs of a particular organisation. The Cyber Attack Information System [9] employs one of the modern anomaly detection algorithms and produces additional valuable input to the cluster of the common security information systems. The Vulnerability Scanner [10] is used to automatically find security vulnerabilities in web applications and report them to the security analyst. The combination of a Firewall and an Intrusion Detection System (IDS) considered in [11] is employed for the creation of an organisational security policy, and it provides an operator with useful security information. In [12] combined distributed intrusion detection and prevention systems (IDS/IPS) collects alerts and analyses them by applying data mining techniques. Fuzzy logic is one method that can be employed for modelling and reasoning based on an expert system. Prominent work in this field includes the rule-based system presented by Bernard [13]. It was designed for processing and power control in a power plant. A survey of the fuzzy logic controller (FLC) [14] evaluates the

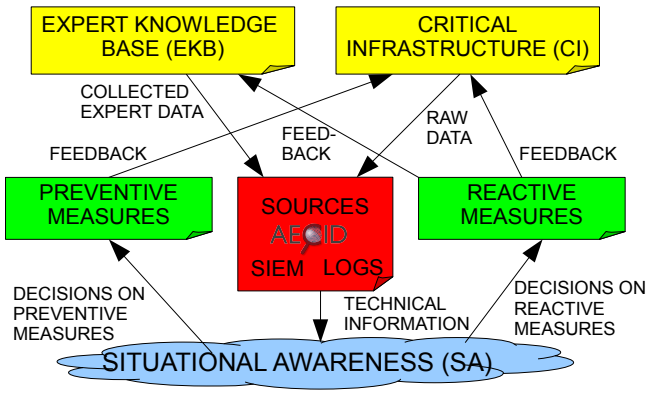

Fig. 2: The overview of establishing the Cyber Situational Awareness.

linguistic control methodologies, the derivation of the fuzzy control rules, and an analysis of fuzzy reasoning mechanisms. To evaluate a control action, the relevant parameters are measured. Actions are specified in rules. The fuzzy logic and inference engine is used to search through the knowledge base in order to identify those rules that are applicable. This approach is very similar to our Expert System organization, with the difference being that we have a different modelling by membership functions, fuzzy standards, and by fuzzy result, which can be defuzzified. The application of natural language words instead of numbers for computing and reasoning using fuzzy logic is described in [15]. The proposed Expert System makes use of this technique. This research is important since the FLC is used in our approach as a standard for fuzzy rules definition. The qualitative safety modelling in [16] is performed employing fuzzy IF - THEN rules. Compared to existing systems, the proposed system is more efficient due to the use of more complex fuzzy rules. Simple IF - THEN rule engines are not well suited for dealing with situational awareness analysis data having a level of uncertainty dependent on a particular defense strategy. A fuzzy-logic based approach may be more appropriate for situational awareness. The provided Expert System deals directly with the linguistic terms commonly used in the cyber security community for situational awareness domain. Our research focuses on the development and representation of linguistic variables and the subsequent calculation of their membership functions. These variables are then quantified using input numbers and fuzzy logic, with the goal being to decide whether or not a particular defense action should be triggered.

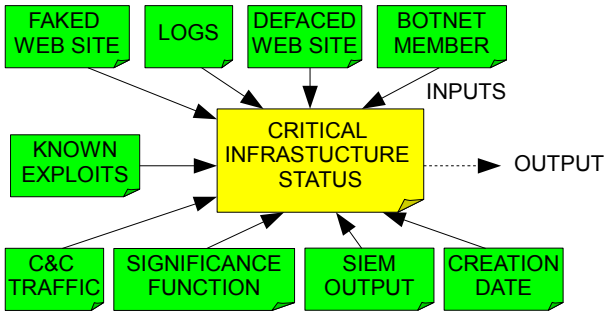

Fig. 3: An inference system model for the calculation of a defense action by employing of the associated cyber security metrics for the given critical infrastructure.

III. Data Aggregation Process For Situational Awareness

Multiple factors have an impact on establishing the Cyber Situational Awareness (SA) presented in Figure 2. The most important factor is an Expert Knowledge Base (EKB) that comprises knowledge about the network, such as external lists of malicious IPs, distribution of exploits, and known exploits (CVE databases) or known vulnerabilities. The EKB also contains information about the expected behaviour of the Critical Infrastructure (CI) modelled by the security analyst. Significant sources for analysis of Situational Awareness are raw logs from CI and output from a SIEM tool [7], [8], such as alert rates, C&C traffic, Antivirus software, Firewall output, anomaly detection by AECID system [9] and information about botnet members. A cyber security analyst may perform preventive and reactive measures to secure CIs. Preventive measures should prevent an attacker from compromising a CI. An analyst can start log analysis, employ additional logs to analysis or also employ security information system tools - such as SIEM, Vulnerability Scanner or Auditing tools for analysis - to obtain more information on a particular facet of the SA. The number of logs can also be reduced to avoid overwhelming the cyber analyst with superfluous information and help filter relevant messages. An expert is responsible for performing periodic updates and uploading patches for vulnerabilities identified by an EKB. An analyst will also execute automated or manual tests to examine a CIs status. These tests are required for regular defence measures, such as firewall and user/group access rights being circumvented by an attacker. For example, a cyber analyst can search for new hooks that an attacker may have placed to gain future access to the system or for new users with administrator rights. Furthermore, analysts may look for unusual time and port activity compared against the regular behaviour data from the EKB, or they may inspect network connection and known virus signatures. In case of test failure, a report should be created and sent to EKB to share this information with other network participants. In the event of attack, the analyst will employ reactive measures. The analyst can execute shutdown/restart or break network connections to reduce the damage to CIs and thereby prevent other network computers from being compromised.

IV. Fuzzy Modelling

In a decision support model for SA we employ cyber security metrics as depicted in Figure 3. Metrics, such as “Significance Function” or “SIEM Output” have numerical input data from correspondent tools and provide inputs to analyse the current status of CI. The output of the model is a numerical value 0 or 1 that presents an automatically computed decision about whether an alarm should be activated.

In fuzzy logic [17], variables can take not only two states (true and false) but they can also have a not fixed truth value ranging from completely false (0) to completely true (1). Fuzzy logic also makes use of linguistic variables, which in our case are defined by Cyber Security analysts that can be presented by member functions of varying complexity. The fuzzy variables describe expert rules and can be defined according to standard Fuzzy Control Logic (FCL).

The advantage of this concept is that every analyst can adjust the linguistic variables, its member functions and thresholds according to their preferences, and the requirements and policies of a particular Critical Infrastructure. To evaluate a Critical Infrastructure status, we determined a control system based on fuzzy logic that comprises nine inputs and one output presented in Figure 3.

Input variables describe an input from different security information systems. One of them is the AECID tool [9]

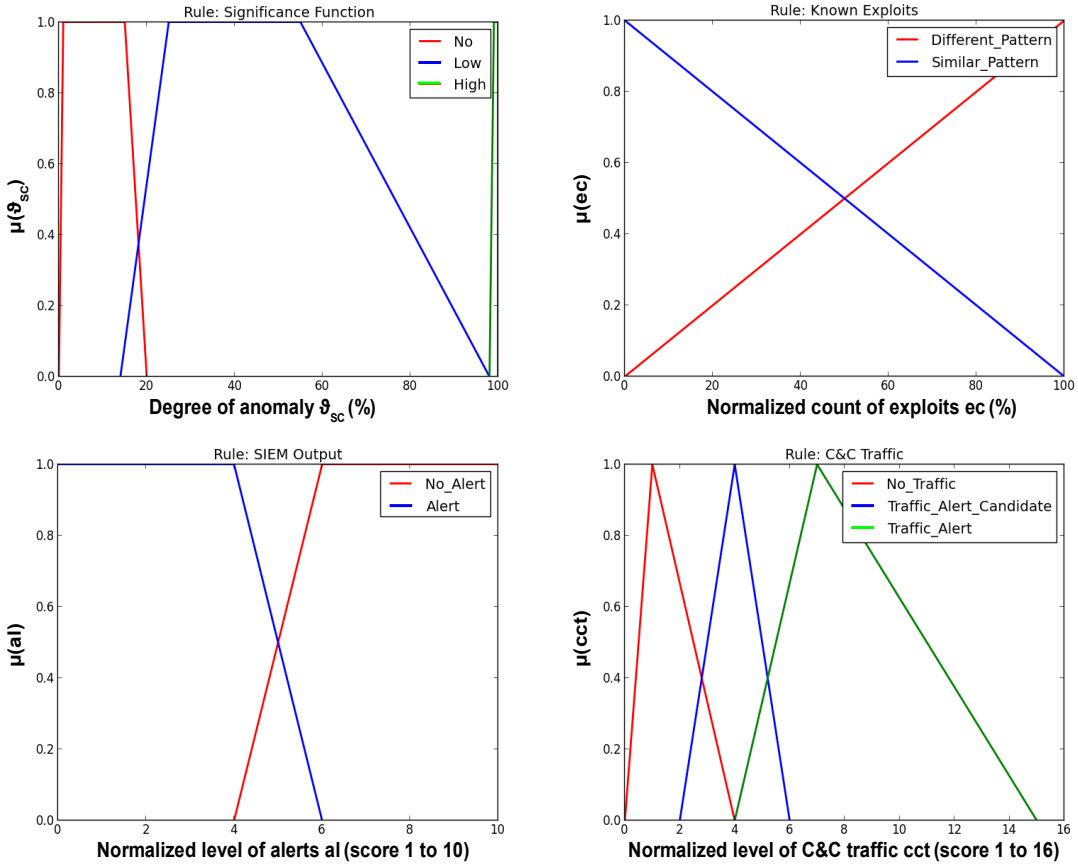

Fig. 5: Definition of selected membership functions (significance function, known exploits, SIEM output and C&C traffic) and fuzzified metrics.

providing significance function with a count of matches in "Significance Function" variable (see Figure 4). Other inputs come from "Fake Web Site", "Defaced Web Site", "Creation Date", "Botnet Member", "Known Exploits", "SIEM Output" and "C&C Traffic" variables. The output is expected in a "Critical Infrastructure Status" variable. Fuzzy inference starts with fuzzification in which numerical values are mapped to the membership functions associated with fuzzy variables. Then we apply logic defined in FCL to evaluate fuzzy rules and aggregate their output. These outputs are mapped to the output variables. Finally, these variables are mapped back to the numerical values. The example of Critical Infrastructure status evaluation is presented in the evaluation part of this paper.

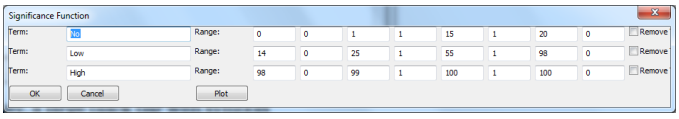

Fig. 4: Definitions of fuzzy membership functions with associated values for input variable "Significance Function".

Developed software tool employs a feedback-loop for regular automatical adjustment of membership functions. Adjustment can be performed by human or statistically. By changing of membership functions tool generates new fuzzy rules.

An input variable, "Significance Function," contains three membership functions flagged by the linguistic variables "No, Low and High", which are depicted in Figure 4. A corresponding graphical representation is shown in Figure 4, where $\mu(\vartheta_{SC})$ shows the membership function for a degree of anomaly that is dependent on the significance function sc. The values for these linguistic variables range from 0 to 1 and come from the AECID tool. For simplicity we transform these values to percentage. Therefore, significance can be defined as high if its value matches in a range between 98 and 100 percent. In contrast low significance values are between 14 and 98 percent. Finally, values between 0 and 20 percent indicate that there is no significance for analyzed Critical Infrastructure.

A fuzzy set for "Significance Function" and its membership

functions “No”, “Low” and “High” can be described by Equations 1.

$$(U,m) = \{\frac{m(x_{NO})}{x_{NO}}, \frac{m(x_{LOW})}{x_{LOW}}, \frac{m(x_{HIGH})}{x_{HIGH}}\}, x \in U \quad (1)$$

$$m(x_{NO}) = \begin{cases} 1, & \text{if } 0 < x \leq 15, \\ -\frac{x}{5} + 4, & \text{if } 15 < x \leq 20, \end{cases} \quad (2)$$

$$m(x_{LOW}) = \begin{cases} \frac{x}{11} - 1,27, & \text{if } 14 < x \leq 25, \\ 1, & \text{if } 25 < x \leq 55, \\ -\frac{x}{43} + 2,28, & \text{if } 55 < x \leq 98, \end{cases} \quad (3)$$

$$m(x_{HIGH}) = \begin{cases} 100x - 98, & \text{if } 98 < x \leq 99, \\ 1, & \text{if } 99 < x \leq 100. \end{cases} \quad (4)$$

Where (U, m) denotes a fuzzy set whose elements x have a grade of membership - not included ($m(x) = 0$), fully included ($m(x) = 1$) or x is defined by membership function $m(x)$. The Figure 5 depicts graphical representation of selected previously defined fuzzy rules and its membership functions.

V. Experimental Results and Interpretation

The goal of evaluation is to calculate CI status and to apply experimental fuzzy rules comprised of input data from different sources on Situational Awareness analysis by making a decision about alert raising. Another goal is to adjust the accuracy of the expert rules, fuzzy input variables and its member functions for robust analysis. Additionally, a quantitative overview of the evaluated data and methods characteristics is delivered. Rules design is based on an empirical detection method. For example, by observing SIEM tool output over time and aggregating information about the regular CI behaviour, one can employ acquired data from a real system to assess and redirect the number of generated alerts via a control feedback loop back to the expert knowledge base about the regular CI behaviour. The adjustment by feedback loop improves the quality of SA analysis and reduces the number of false alerts, therefore saving the experts valuable time.

An automatic fuzzy logic based approach would significantly facilitate manual analysis and could be used by analysts for the defense of Critical Infrastructure. The suggested method would make SA analysis less cost-intensive, faster, and would perform higher throughput. However, a human-based approach performs with higher accuracy. This evaluation took place on an Intel Core i7-3520M 2.66GHz computer using Python on Windows OS. We evaluated theoretically possible input from AECID tool, additional Situational Awareness inputs, and fuzzy inference system output. Based on the model of an inference system from Figure 6, we compute a rule-based decision about CI status by employing the aforementioned cyber security metrics as input variables. Given the linguistic variables, such as “SIEM Output” and “Significance Function”, with the membership functions as defined in Figure 4, we provide this rule base to the fuzzy engine. The defuzzification operation presented in Figure 7 employs COGS (Center of Gravity) algorithm and calculates a discrete value, from the inferred fuzzy set, whereas all single results obtained by evaluating rules (e.g. RULE 387: “IF Significance Function IS High AND CandC Traffic IS No Traffic AND SIEM Output IS No Alert AND Creation Date IS Similar Date AND Known Exploits IS Different Pattern AND Botnet Member IS Not Botnet Member AND Defaced Web Site IS Defaced Web Site AND Faked Web Site IS Defaced Web Site THEN CI Status IS High;”) are combined. For given metrics Expert System has evaluated 576 rules.

The expert system starts the Situational Awareness analysis (see Figure 6) with sample numerical input values on the left side where white numbers are shown. For “Significance Function” that is the output of AECID tool the input value is 99, which means it is highly probable that an anomaly was detected. Fuzzifying this value, we map it to the associated numerical value using the FCL input variables definition. For the input variable “Significance Function,” the associated linguistic term is “High”. Other sample numerical input variables mean that there is a report about 1 defaced Web site, given CI belongs to one botnet, the CI has one known

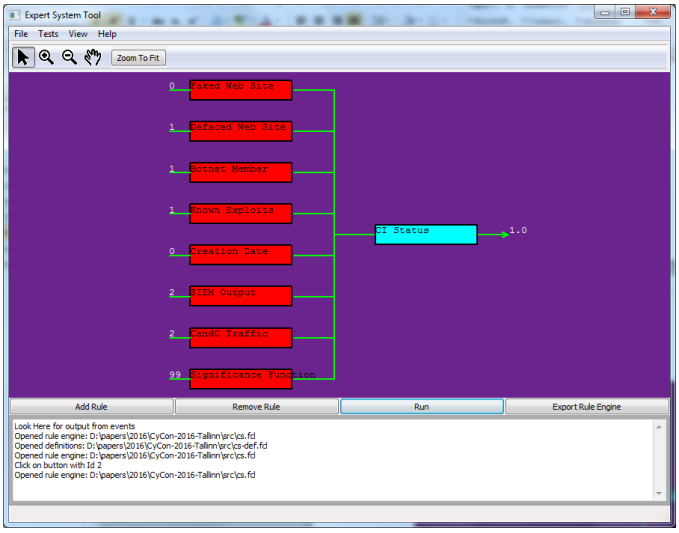

Fig. 6: A fuzzy inference system, developed by authors, for calculation of the truth degree for raising an alert.

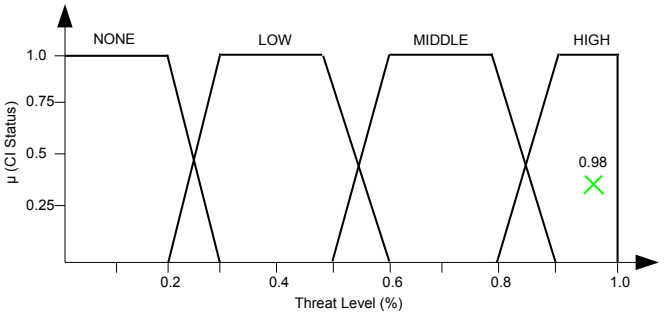

Fig. 7: Defuzzification by determining the center of gravity.

exploit and no C&C traffic was detected for this CI (value 2 could be a noise value).

$$\mu(x_{HIGH}) \quad = \quad \begin{cases} \frac{x}{10} - 8, & \text{if } 80 < x \leq 90, \\ 1, & \text{if } 90 < x \leq 100. \end{cases} \tag{5}$$

The Equation 5 shows the defuzzification function $\mu(x_{HIGH})$ that returns value on a range 0.0 to 1.0 that corresponds to one of the defined labels "None", "Low", "Middle" or "High" standing for a threat level. Aggregating all rule outputs we defuzzify the output threat level value of the total "CI Status" for "High" and obtain the related value of 0.98. This value is a result of correlation by basis indicators, such us "Defaced Websites", "Botnet Detected", "Detected Vulnerabilities", "Discovered Malware" and so on. Therefore, for the given CI should be raised an alert with degree of truth 0.98.

VI. Conclusion

In this paper we presented an approach for providing solutions in real-time which could replace human analyst inputs for some basic but frequent decisions that can be automated. This method can make SA analysis faster and cheaper. The expert system brings together information automatically aggregated from security information systems and expert knowledge from the field of Cyber Security. Large input data from Critical Infrastructure requires automatic decision support. The main contribution of this work is the definition of fuzzy input and output variables and the application of a fuzzy inference system to support Situational Awareness analysts in the decision-making process of raising alerts. The novelty of proposed approach is that our system implements a feedback-loop that constantly automatically improves the fuzzy rules and generates new rules by means of developed software. During evaluation we adjusted thresholds for member functions of fuzzy input variables in order to make the expert system more robust for new evaluations. Therefore, we propose a system that supports Situational Awareness handling for Critical Infrastructures in Cyber Security domain. The evaluation demonstrates that the given approach enables the integration of complex rules, provides automated decision support, and helps solve practical Situational Awareness issues such as the decision whether or not to raise an alert. The calculated fuzzy results and their linguistic interpretation provided by developed Expert System are about the reduction of uncertainty by Situational Awareness analysis. In future work we plan to increase the amount of fuzzy input variables, to extend an Expert System with additional fuzzy input variables, to improve the accuracy of member functions and to increase the quality of the outputs. We also plan to evaluate dynamic adaptation of a rule base along with threat level and severity, dependent on current attack vectors.

Acknowledgment

This work was supported by the EU FP7 Project ECOSSIAN (GA:607577) http://ecossian.eu/

References

[1] G. Benincasa and et al, "Intelligent systems'2014: Proceedings of the 7th ieee international conference intelligent systems is'2014, september 24?26, 2014, warsaw, poland, volume 1: Mathematical foundations, theory, analyses." Cham: Springer International Publishing, 2015, pp. 813–824. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-11313-5_71

[2] J. Lu, G. Zhang, and W. F., "Team situation awareness using web-based fuzzy group decision support system," *International Journal of Computational Intelligence Systems*, vol. 1, no. 1, pp. 50–59, 2008.

[3] G. D'Aniello, V. Loia, and F. Orciuoli, "A multi-agent fuzzy consensus model in a situation awareness framework," *Applied Soft Computing*, vol. 30, pp. 430 – 440, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1568494615000824

[4] T. Hanratty and et. al., "Enhancing battlefield situational awareness through fuzzy-based value of information," *46th Hawaii International Conference on System Sciences*, pp. 1402–1411, 2013.

[5] C. D. Maio, G. Fenza, D. Furno, and V. Loia, "Swarm-based semantic fuzzy reasoning for situation awareness computing," in *Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on*, June 2012, pp. 1–7.

[6] L. V. Furno, D. and M. Veniero, "A fuzzy cognitive situation awareness for airport security," *Control and cybernetics 39(4)*, vol. 39, no. 4, pp. 959–982, 2010.

[7] D. Swift, "A practical application of sim/sem/siem, automating threat identification," *SANS Institute*, p. 3, 2006.

[8] I. Kotenko and A. Chechulin, "Common framework for attack modeling and security evaluation in siem systems," *in Green Computing and Communications (GreenCom), 2012 IEEE International Conference on*, pp. 94–101, 2012.

[9] H. Leopold, T. Bleier, and F. Skopik, "Cyber attack information system," *Springer Vieweg*, 2015.

[10] A. Doupe and et. al., "Enemy of the state: A state-aware black-box web vulnerability scanner," *21st USENIX Security Symposium (USENIX Security 12)*, 2012.

[11] U. Roedig, R. Ackermann, M. Tresse, L. Wolf, and R. Steinmetz, "Verbesserte systemsicherheit durch kombination von ids und firewall," *Vieweg+Teubner Verlag*, pp. 117–128, 2000.

[12] H. Nen-Fu, C.-N. Kao, H.-W. Hun, G.-Y. Jai, and C.-L. Lin, "Apply data mining to defense-in-depth network security system," *in Advanced Information Networking and Applications, AINA 2005*, pp. 159–162, 2005.

[13] J. A. Bernard, "Use of rule-based system for process control," *IEEE Contr. Sys. Mag*, pp. 3–13, 1988.

[14] H. S. Sii, T. Ruxton, and J. Wang, "A fuzzy-logic-based approach to qualitative safety modelling for marine systems," *Reliability Engineering & System Safety*, vol. 73, no. 1, pp. 19 – 34, 2001. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0951832001000230

[15] C.-C. Lee, "Fuzzy logic in control systems: fuzzy logic controller. i," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 20, no. 2, pp. 404–418, 1990.

[16] L. A. Zadeh, "Fuzzy logic = computing with words," *Fuzzy Systems, IEEE Transactions on*, vol. 4, no. 2, pp. 103–111, 1996.

[17] K. J. Zadeh, Lotfi A., "Fuzzy logic for the management of uncertainty," *John Wiley and Sons, Inc., NY, USA*, 1992.