Intrusion Detection with Type-2 Fuzzy Ontologies and Similarity Measures

Robin Wikström and József Mezei

Abstract Intrusions carry a serious security risk for financial institutions. As new intrusion types appear continuously, detection systems have to be designed to be able to identify attacks that have never been experienced before. Insights provided by knowledgeable experts can contribute to a high extent to the identification of these anomalies. Based on a critical review of the relevant literature in intrusion detection and similarity measures of interval-valued fuzzy sets, we propose a framework based on fuzzy ontology and similarity measures to incorporate expert knowledge and represent and make use of imprecise information in the intrusion detection process. As an example we developed a fuzzy ontology based on the intrusion detection needs of a financial institution.

1 Introduction

Intrusion detection systems are becoming more and more important in controlling network security as the number of intrusion events is increasing rapidly due to the widespread use of internet. A general intrusion detection system operates as decision support system by making use of the (real-time) information and event reports describing previous intrusion cases to identify potential dangerous activities. There are two main approaches to intrusion detection: misuse detection (known patterns of intrusion are compared to present activities) and anomaly detection (activities that deviate from normal system behaviour but cannot be matched to any previous cases) [2, 56].

R. Wikström (✉) · J. Mezei
Department of Information Technologies, IAMSR, ÅBo Akademi University, Joukahaisenkatu
3-5B, 20520 Turku, Finland
e-mail: rowikstr@abo.fi

J. Mezei
e-mail: jmezei@abo.fi

R.R. Yager et al. (eds.), *Intelligent Methods for Cyber Warfare*,
Studies in Computational Intelligence 563, DOI 10.1007/978-3-319-08624-8_7

As sensors and other data collection methods are evolving continuously, intrusion detection systems are forced to constantly process increasing amounts of information. A fair part of this information consists of imprecise and vague knowledge [41]. A fuzzy ontology is basically an ontology that employs fuzzy logic for dealing with imprecise knowledge [6]. Using fuzzy ontologies for analysing this knowledge is important for identifying possible intrusions [20]. This is especially true regarding anomalies, as it becomes possible to identify cases in a database that are similar in a fuzzy sense. As an extension of traditional fuzzy ontologies, type-2 fuzzy ontologies are not limited to a crisp value for defining the membership function of different concepts therefore offering more possibilities to model uncertainty compared to type-1 fuzzy ontologies [36].

Similarity measures have successfully been used for several anomaly detection implementations [14]. Detecting anomalies is an important method for finding unwanted behaviour, not only in intrusion detection but also in e.g. fraud detection and military surveillance. Kernel based similarity measures (cosine and binary [40]) together with text processing techniques were applied for example for detecting host-based intrusion detection [51].

Ning and Xu [47] noticed that Intrusion Detection Systems (IDSs) are producing an increasing amount of alerts, containing a fair share of false alerts, regardless, one needs to process this data. By applying similarity measures on the collected intrusion alerts, they generated the similarity between different attacks strategies. Their basic assumption is that knowing the attack strategy, one can predict the coming moves of the attacker. The use of expert knowledge can play a crucial role in identifying anomalies and assessing the potential loss that can be caused by an intrusion. One could even state that it is necessary to include experts in intrusion detection systems, as fully automated systems seem to be impossible to achieve [13]. Some systems are able to detect malware and intrusions based on behavioural patterns, however, few come even close to automatically deciding whether the spotted abnormality is a malware or not, and therefore depend on experts for making the final decision [29].

In this chapter, we will provide an extensive literature review concerning intrusion detection systems in the financial context and similarity measures for interval-valued fuzzy sets. Based on the analysis of the literature, we will address three important issues that are not widely considered in intrusion detection systesms: (i) making use of expert knowledge in identifying anomalies; (ii) systematic representation of (imprecise) information concerning previous intrusion cases; (iii) identifying the potential causes of an intrusion in the presence of these imprecise descriptions. Our proposal is to use fuzzy ontologies to represent the available information in terms of interval-valued fuzzy sets: the combination of similarity analysis and expert opinions provides a promising tool to identify and measure the related risks of misuses and also anomalies. We use financial institutions and their operating environment as the example case to describe the model in details. The system can provide information to the users concerning two types of decisions: (i), identifying suspicious activities that can indicate intrusion, and (ii), recommendation on the countermeasures to be undertaken in a given case (which requires the estimation of the possible losses caused by the intrusion).

The chapter is structured as follows: Sect. 2 presents relevant concepts and definitions of intrusion detection systems focusing on financial institutions. Section 3 discusses the role of expert (linguistic) knowledge in information systems and a possible representation in the form of a fuzzy ontology. A discussion on the most important similarity measures for interval-valued fuzzy sets is provided in Sect. 4. In Sect. 5, an *OWA* operator-based distance is presented for interval-valued fuzzy numbers that can be used to calculate similarities. Section 6 presents a fuzzy ontology based on a financial institution taxonomy. Finally, some conclusions are given in Sect. 7.

2 Intrusion Detection Systems and Financial Institutions

The following section introduces relevant concepts and definitions of Intrusion Detection Systems in financial institutions.

2.1 Intrusion Detection

Today, as roughly 1/3 of the earth's population[1] have access to the internet and the penetration rate is rapidly increasing, naturally not only the private users are active online, but also an increasing amount of businesses. With increased amounts of users and business connected to the internet, the risk of intrusions and other complications is amplified. As a consequence, in this context, intrusion detection systems are constantly becoming more important.

Applications and software that help users to protect their devices from viruses and malware constitute an important research topic. Ontologies have turned out to be an useful method to be used for intrusion detection tasks, as they offer possibilities to analyse, for instance, patterns generated by intruders; this way also previously unknown attack methods can be detected [37].

Dai et al. [18] observe that it is a well known fact that hackers tend to be one step ahead of systems created for protection. This results in an endless circle of data losses and a constant demand for new software to fix previous errors. As hackers and their methods are adaptive, behaviour-based approaches have gained an increased interest in developing systems to protect data, as they are effective when dealing with previously unknown attacks [29].

Malware is the common name used for the software's that perform the attacks on computers. They often employ anti-reverse engineering techniques to avoid being detected by analytically-based software. Wagener et al. [59] propose a possible solution for this problem, by applying similarity and distance measures on malware behaviour to implement a better classification of malware types. Different comparisons of similarity and distance measures in the context of malware have also been

[1] http://www.internetworldstats.com/stats.htm

realized, e.g. by [3]. Due to the complexity of malwares, ontologies and especially fuzzy ontologies have attracted significant interest, regarding their possibilities to aid in these kinds of tasks, notably when dealing with imprecise knowledge.

2.2 Financial Institutions

A financial institution can be defined as an institution which offers financial services, working as an intermediary by providing, for example: loans, deposits, currency exchanges and investments. Banks and insurance companies are examples of financial institutions. We chose to use financial institutions as the example case as they tend to be an important object for cyber-attacks. These institutions own sensitive data and also significant amount of monetary funds. It has to be noted that in this context not every attack is conducted for personal monetary gain, e.g. stealing funds, but more as a challenge for achieving credibility in online communities or getting noted by the global media. Financial institutions are attractive targets also for this purpose, as people tend to react when their savings are "in danger". As a result, security systems protecting the institutions are designed in a way that is difficult to break, i.e. the one managing to break it deserves some credit. Recently, there has been a global increase in attacks directed towards financial institutions. As these institutions can be considered to be one of the prime targets even on an nationwide scale, the treat of cyber-terrorism can not be overlooked [26, 48].

The risk of intrusions taking place in the financial sector is consequently increasing steadily. Reports constantly indicate that, for instance, bank website outage hours are increasing every month and that more online banking frauds occur. An old but still relevant financial malware is called Zeus. It was recognized already in 2006, but since then it has been remodelled and re-customized several times, each version requiring more preventive work by the institutions. This malware originates from Russian cybercrime organisations. However, the leakage of the source code in 2011 opened the door for basically anyone to modify Zeus and use it for intrusion purposes. Before the leakage of the code, the software was available only for those willing to pay for it, somewhat limiting the usage. Nowadays, there are communities devoted to sharing and trading "plug-ins" for the Zeus malware. As Zeus is not the only malware available, the risk of intrusions happening to a financial institution is far from unlikely [50]. Recently, there have been several publications about how one could prevent different types of attacks specifically aimed towards the financial sector [35, 49].

3 Expert Based Knowledge in IDS

The role of the human experts in information systems has taken a slightly ambiguous role. Wang et al. [61] and Huang et al. [29], amongst others, state that current systems are unable to completely provide full protection against attacks and intrusions. One of the critical issues is the inclusion or exclusion of the human experts. The goal seems to be to exclude the expert as much as possible, however, currently all systems require human input at some stages of the process [13].

Experts usually express themselves using linguistic terms, i.e. "the activity is quite low" and "one should block some of the intrusions". These imprecise linguistic terms, fully understandable for other experts, are hard to interpret for a computer as it is designed for computing with precise data. Linguistic modelling [34], fuzzy logic [66] and other methodologies have been proposed as possible solutions for making human-computer communications feasible.

3.1 Fuzzy Ontologies For Intrusion/Malware Detection

Ontologies, in the context of the Semantic Web, provide a structure and blueprint of the tacit data inherent in different domains. An ontology reveals the relations and connections between the different instances, facilitating the reasoning and decision making. Furthermore, ontologies can be combined and reused by/with different internet-based techniques, encouraging interoperability [5, 24]. This, in turn, gives computers the possibility to reason in a more human-like way, as they can grasp some of our tacit knowledge of the world [33]. There has been an increase in using ontologies for the purpose of intrusion and malware detection.

Undercoffer et al. [58] constructed one of the first ontologies for intrusion detection in the context of computer attacks. They used the DAML+OIL ontology modelling language (a precursor to OWL). Before this approach, mainly taxonomies were used for this purpose. Introducing ontologies enabled better meta-data modelling and ontologies can naturally subsume taxonomies. Simmonds et al. [52] developed an ontology for defending against attacks aimed at networks, emphasising that one should also prepare for what happens if the attack is successful and how the designed system reacts in that scenario.

The rapid development of mobile devices created a completely new field vulnerable to intrusions and malwares. Chiang and Tsaur [17] therefore took the first steps towards extending ontologies also towards the protection of mobile devices. They modelled an ontology based on the behaviours of known mobile malware. Hung et al. [30] created an extensive ID ontology, which also included a feature allowing users to model the ontology application from a conceptual level. This broadens the possible range of users, meaning that even non-expert users could contribute to intrusion detection processes.

However, it has several times been stated that traditional, non-fuzzy, ontologies are not suitable to deal with imprecise and vague knowledge [27, 42]. Avoiding imprecise data in the online world is close to impossible, hence, the combination of fuzzy logic and ontologies has recently gained an increased interest from the research community.

In recent years, the combination of fuzzy logic and ontologies has been an emerging topic in intrusion detection [9, 20, 56]. Huang et al. [27–29] developed a

IT2FS-based ontology, as a novel approach for malware behavior analysis (MiT). Using the Fuzzy Markup Language (FML) [1] and the Web Ontology Language (OWL), they managed to create a fully operational system, able to analyse collected data and extract behavioural information.

Tafazzoli et al. [55] created a fuzzy malware ontology for the Semantic Web. The ontology represents relevant concepts inherent in the malware field. The relationships between the different malwares are modelled with the help of fuzzy linguistic terms, such as: **weak relation** and **very good relation**. Considering that it was created with the Semantic Web in mind, it can also be used for sharing information online.

As it can be noticed, there is a fair amount of positive results that have emerged with the fuzzy ontology approach. We believe that it is reasonable to state that more research in how fuzzy ontologies can benefit the task of intrusion detection is needed.

The Web Ontology Language (OWL) is the main language used for creating ontologies on the Semantic Web [25]. By settling on one standard and using it for the Semantic Web and its ontology modelling needs, it makes the co-operation between different domain ontologies (for instance) more straightforward and smoother, facilitating the expansion of OWL to the non-expert users. The fuzzy ontology created in this chapter (and presented in more detail later on) is modelled in OWL.

4 Similarity Measures for Interval-Valued Fuzzy Sets

Similarity measures have become an important technique for handling imprecise information in the context of information systems [60]. The easily embraceable notion behind these measures, comparing how similar two instances are, has made them widely used in various topics. Numerous applications and implementations based on similarity measures exist also for intrusion detection issues.

In this section we provide an extensive overview of the existing similarity measures for interval-valued fuzzy sets. We also included the similarity measures proposed for intuitionistic fuzzy sets by reformulating the definitions using the traditional transformation between interval-valued and intuitionistic fuzzy sets. We will use the following definitions for interval-valued fuzzy sets:

Definition 1 ([22]) An interval-valued fuzzy set A defined on X is given by

$$A = \left\{(x, [\mu_A^L(x), \mu_A^U(x)])\right\}, x \in X,$$

where $\mu_A^L(x), \mu_A^U(x) : X \to [0, 1]; \forall x \in X, \mu_A^L(x) \leq \mu_A^U(x)$, and the ordinary fuzzy sets $\mu_A^L(x)$ and $\mu_A^U(x)$ are called lower fuzzy set and upper fuzzy set of A, respectively.

A starting point for evaluating different similarity measures can be a set of predefined properties that are expected to be satisfied by a measure to be called as similarity. As we will see, in many cases these properties only ensure a basic reliability of the measures: there are many examples for similarity measures that satisfy all the properties but provide non-intuitive values for specific fuzzy sets. In this article, we will adopt the four properties specified in [62]:

1. Reflexivity: $s(A, B) = 1 \Longrightarrow A = B$
2. Symmetry: $s(A, B) = s(B, A)$
3. Transitivity: If $A \leq B \leq C$ than $s(A, B) \geq s(A, C)$
4. Overlapping: If $A \cap B \neq \emptyset$ than $s(A, B) > 0$

Additionally, we can require the similarity measure to be normalized: $0 \leq s(A, B) \leq 1$.

The first group of similarity measures is based on the extension of traditional distance measures (Hamming, Euclidean) to fuzzy sets. This family of similarities has been developed for type-1 fuzzy sets before the 1990s (see [57]), and later on to interval-valued and intuitionistic fuzzy sets. The traditional way to obtain the similarity from a normalized distance measure d is to calculate

$$s(A, B) = 1 - d(A, B).$$

As Zeng and Guo described in a systematic analysis, there are numerous other ways to generate similarity measures from distances.

Theorem 1 ([67]) *Given a real function $f : [0, 1] \to [0, 1]$, if f is a strictly monotone decreasing function, and d is a normalized distance on interval-valued fuzzy sets, then*

$$s(A, B) = \frac{f(d(A, B)) - f(1)}{f(0) - f(1)}$$

is a similarity measure.

For example if $f(x) = \frac{1}{1+x}$, then $s(A, B) = \frac{1-d(A, B)}{1+d(A, B)}$ and for $f(x) = 1 - x^2$ one can obtain the similarity $s(A, B) = 1 - d^2(A, B)$.

Burillo and Bustince [10] were the first ones to extend the Hamming and Euclidean distances to (discrete) interval-valued (and intuitionistic) fuzzy sets in the following way (the normalized distances are presented):

- Hamming distance:

$$d_H(A, B) = \frac{1}{n} \sum_{i=1}^{n} \frac{| A^L(x_i) - B^L(x_i) | + | A^U(x_i) - B^U(x_i) |}{2};$$

- Euclidean distance

$$d_2(A, B) = \left[\frac{1}{2n} \sum_{i=1}^{n} (A^L(x_i) - B^L(x_i))^2 + (A^U(x_i) - B^U(x_i))^2 \right]^{0.5}.$$

In [4], Atannasov presented similar definitions for intuitionistic fuzzy sets. Szmidt and Kacprzyk [53, 54] further improved the definitions for intuitionistic fuzzy sets by incorporating the intuitionistic fuzzy index (and their proposal was extended by using geometric distance and weights in [64]). As Grzegorzewski [23] pointed out, these modifications are not properly motivated and result only in marginal differences and improvements; for this reason, he used the Hausdorff metric to modify the definitions in a natural way that is easy to use for applications. The proposed (normalized) distances can be formulated as:

- Hamming distance:

$$d_H(A,B) = \frac{1}{n}\sum_{i=1}^{n} \max(\mid A^L(x_i) - B^L(x_i) \mid, \mid A^U(x_i) - B^U(x_i) \mid);$$

- Euclidean distance

$$d_2(A,B) = \left[\frac{1}{n}\sum_{i=1}^{n} \max(A^L(x_i) - B^L(x_i))^2, (A^U(x_i) - B^U(x_i))^2)\right]^{0.5}.$$

Zeng and Li [68] proposed the same definition of the Hamming-distance independently as a result of a transformation procedure to connect entropy and similarity measures for interval-valued fuzzy sets. They additionally defined the continuous version of the definition as

$$s(A,B) = \frac{1}{2(b-a)}\int_a^b \mid A^L(x_i) - B^L(x_i) \mid + \mid A^U(x_i) - B^U(x_i) \mid \mathrm{d}x,$$

with the support of the interval-valued fuzzy sets is in the $[a,b]$ interval. Zhang and Fu [69] introduced weight values (functions) in the two cases of the Hamming distance and showed that it is a similarity measure in a more general sense as it can be used for any L-fuzzy sets.

After 2002, as a different direction to extend distance measures, numerous articles were published in the journal Patter Recognition Letters mainly originating from the approach introduced in the article of Dengfeng and Chuntian [19]. They used the middle points of the membership interval of the interval-valued fuzzy sets to obtain a type-1 fuzzy set and then they calculate the distance of the resulting type-1 fuzzy sets. Using the notation

$$f_A(x) = \frac{A^L(x) + A^U(x)}{2},$$

the following formulas can be obtained:

- in the discrete case

$$s_d^p(A, B) = 1 - \frac{1}{n^{1/p}} \left[\sum_{i=1}^{n} (f_A(x_i) - f_B(x_i))^p \right]^{1/p} ;$$

- in the continuous case

$$s_c^p(A, B) = 1 - \frac{1}{(b-a)^{1/p}} \left[\int_a^b (f_A(x) - f_B(x))^p \mathrm{d}x \right]^{1/p} .$$

Both definitions can be used with weight functions, for example in the continuous case

$$s_{cw}^p(A, B) = 1 - \left[\int_a^b w(x)(f_A(x) - f_B(x))^p \mathrm{d}x \right]^{1/p} .$$

One important disadvantage of this method is that equality of the interval-valued fuzzy sets is only a sufficient but not necessary condition for the similarity measure to take its maximal value. After illustrating this with a simple example, Mitchell [45] proposed an improvement to the method: instead of calculating the similarity of type-1 fuzzy sets obtained as the average of the lower and upper membership functions, the similarity of the two upper fuzzy sets and the two lower fuzzy sets are computed and the overall similarity is the average of these two values.

Noticing the same problem with the Li-Cheng approach, Liang and Shi [39] proposed new families of similarity measures by using a more complex combination of the upper and lower membership functions:

$$f_{l_{AB}}(x) = \frac{\mid A^L(x) + B^L(x) \mid}{2}; \; f_{u_{AB}}(x) = \frac{\mid A^U(x) + B^U(x) \mid}{2}.$$

Interestingly, this approach in a special case provides a formula similar to the Euclidean-distance based approach originally proposed in [10]. Liang and Shi proposed further modifications by incorporating the median of the interval membership into the f values and showed that the new definitions provide more intuitive results. As a summary of the proposals utilizing distance measures to calculate similarity, Li et al. [38] provided a detailed description of the different proposals published before 2007 (focusing on intuitionistic and vague sets) and created a selection process to identify the best method: they concluded that the measure proposed in [39] is the only one that does not result in counter intuitive values in any case.

The second main group of similarity measures consists of definitions based on set-theoretic measures and arithmetic operations on fuzzy sets. As discussed in [57], for type-1 fuzzy sets, similarity measures belonging to this family are as popular as the distance based similarities, but we can find significantly less measures when investigating interval-valued fuzzy sets. Probably the most general formula for similarity measures was given by Bustince [11]. He proved that the combination of an inclusion grade indicator and any t-norms will result in an interval-valued similarity measure. Using a similar approach, Zhang et al. [70] defined a new inclusion measure and combined it with a t-norm to obtain a similarity value.

An important measure of similarity in set theory is the Jaccard index. In fuzzy set theory, one can find many different generalizations also for interval-valued fuzzy sets. The first approach is to calculate the Jaccard index of the upper membership values and the lower membership values separately and combine them to obtain an overall similarity: Zheng et al. [71] calculated the average as

$$s(A,B)=\frac{1}{2}\left(\frac{\int_X \min(A^U(x),B^U(x))\mathrm{d}x}{\int_X \max(A^U(x),B^U(x))\mathrm{d}x}+\frac{\int_X \min(A^L(x),B^L(x))\mathrm{d}x}{\int_X \max(A^L(x),B^L(x))\mathrm{d}x}\right),$$

while Hwang and Yang [31] used the minimum of the Jaccard index of the lower membership function and the Jaccard index of the complement of the upper memberships as the similarity:

$$s(A,B)=\min\left(\frac{\sum_{x\in X}\min(A^{U_C}(x),B^{U_C}(x))}{\sum_{x\in X}\max(A^{U_C}(x),B^{U_C}(x))},\frac{\sum_{x\in X}\min(A^L(x),B^L(x))}{\sum_{x\in X}\max(A^L(x),B^L(x))}\right).$$

A different approach making use of the Jaccard index is of calculating the similarity directly from the interval-valued memberships and not as a combination of upper and lower values. This method results in the following formula defined by Wu and Mendel [62] (this is an improved version of the previously defined vector similarity measure [63] as the authors noticed that it does not satisfy the overlapping property:)

$$s(A,B)=\frac{\int_X \min(A^U(x),B^U(x))+\min(A^L(x),B^L(x))\mathrm{d}x}{\int_X \max(A^U(x),B^U(x))+\max(A^L(x),B^L(x))\mathrm{d}x}.$$

The theory of similarity for general type-2 fuzzy sets is still in the early stages, we only mention two approaches that can naturally be applied to interval-valued fuzzy sets (as special cases of general type-2 fuzzy sets): Zheng et al. [72] defined a similarity measure employing the footprint of uncertainty and secondary membership function of type-2 fuzzy sets motivated by a clustering application. McCulloh et al. [43] created a framework to extend any similarities of interval-valued fuzzy sets to general type-2 fuzzy sets.

Another commonly used approach is to calculate the similarity of specific type-1 fuzzy sets and aggregate them into the overall similarity of interval-valued fuzzy sets. Mitchell [46] proposes to choose N embedded fuzzy sets randomly and calculate the average similarity value (any type-1 similarity measure can be used). Motivated by risk analysis problems, Chen et al. [15, 16] proposed several similarity measures based on arithmetic operations by comparing the upper and lower similarities. Feng and Liu [21] used the similarity for the upper and lower membership functions and the kernel function of interval-valued fuzzy sets to obtain a new similarity measure.

5 Similarity of Interval-Valued Fuzzy Sets Based on the OWAD Operator

In this section we will define a similarity measure for interval-valued fuzzy numbers (IVFN's) based on the concept of the ordered weighted averaging distance operator (*OWAD*) introduced by Xu and Chen [65]:

Definition 2 An OWAD operator of dimension n is a mapping $OWAD : \mathbb{R}^n \times \mathbb{R}^n \to [0, 1]$ that has an associated weighting vector W with $\sum_{j=1}^{n} W_j = 1$ and $W_j \in [0, 1]$ such that:

$$OWAD\left(\langle \mu_1^{(1)}, \mu_1^{(2)} \rangle, \ldots, \langle \mu_n^{(1)}, \mu_n^{(2)} \rangle\right) = \sum\nolimits_{j=1}^{n} w_j D_j,$$

where D_j represents the jth largest of the $|\mu_i^{(1)} - \mu_i^{(2)}|$.

In order to extend the definition to the family of IVFN's, we use the mean value of an interval-valued fuzzy number to measure the distance of IVFN's.

Definition 3 ([12]) The mean (or expected) value of $A \in$ IVFN is defined as

$$E(A) = \int_0^1 \alpha(M(U_\alpha) + M(L_\alpha))\mathrm{d}\alpha, \tag{1}$$

where U_α and L_α are uniform probability distributions defined on $[A^U]^\alpha$ and $[A^L]^\alpha$, respectively, and M stands for the probabilistic mean operator.

The distance of two IVFN's, $d : \text{IVFN} \times \text{IVFN} \to \mathbb{R}$, is defined as

$$d(A, B) = |E(A) - E(B)|. \tag{2}$$

The distance (2) satisfies the four properties of a distance measure:

1. Non-negativity: $|E(A) - E(B)| \geq 0$
2. Commutativity: $|E(A) - E(B)| = |E(B) - E(A)|$
3. Reflexivity: $|E(A) - E(A)| = 0$
4. Triangle inequality: $|E(A) - E(B)| + |E(B) - E(C)| \geq |E(A) - E(C)|$.

Definition 4 ([44]) A Quasi IVFN-IOWAD operator of dimension n is a mapping $f : \text{IVFN}^n \times \text{IVFN}^n \times \text{IVFN}^n \to \mathbb{R}$ that has an associated weighting vector W of dimension n with $w_j \in [0, 1]$ and $\sum_{j=1}^{n} w_j = 1$, such that:

$$f(\langle U_1, A_1, B_1\rangle, \langle U_2, A_2, B_2\rangle, \dots, \langle U_n, A_n, B_n\rangle) \\ = g^{-1}\left(\sum_{j=1}^{n} w_j g(D_j)\right), \tag{3}$$

where D_j is the $d(A_i, B_i)$ value of the triplet $\langle U_i, A_i, B_i\rangle$ having the jth largest U_i and $g : \mathbb{R} \to \mathbb{R}$ is a continuous, strictly monotone function.

Theorem 1 *If f is an Quasi IVFN-IOWAD operator, then the following properties are satisfied:*

1. *f is commutative:*

$$f(\langle U_1, A_1, B_1\rangle, \langle U_2, A_2, B_2\rangle, \dots, \langle U_n, A_n, B_n\rangle) \\ = f(\langle U_1^{'}, A_1^{'}, B_1^{'}\rangle, \langle U_2^{'}, A_2^{'}, B_2^{'}\rangle, \dots, \langle U_n^{'}, A_n^{'}, B_n^{'}\rangle),$$

where $(\langle U_1^{'}, A_1^{'}\rangle, \langle U_2^{'}, A_2^{'}\rangle, \dots, \langle U_n^{'}, A_n^{'}\rangle)$ is any permutation of the arguments.

2. *f is monotone: if $d(A_i^1, B_i^1) \geq d(A_i^2, B_i^2)$ for all i, then*

$$f(\langle U_1, A_1^1, B_1^1\rangle, \langle U_2, A_2^1, B_2^1\rangle, \dots, \langle U_n, A_n^1, B_n^1\rangle) \\ = f(\langle U_1, A_1^2, B_1^2\rangle, \langle U_2, A_2^2, B_2^2\rangle, \dots, \langle U_n, A_n^2, B_n^2\rangle).$$

3. *f is idempotent: if $d(A_i, B_i) = d(A_j, B_j) = d, \forall i, j$, then*

$$f(\langle U_1, A_1, B_1\rangle, \langle U_2, A_2, B_2\rangle, \dots, \langle U_n, A_n, B_n\rangle) = d.$$

4. *f is bounded:*

$$\min_i \{d(A_i, B_i)\} \leq \\ f(\langle U_1, A_1, B_1\rangle, \langle U_2, A_2, B_2\rangle, \dots, \langle U_n, A_n, B_n\rangle) \leq \\ \max_i \{d(A_i, B_i)\}.$$

Proof The proofs are straightforward consequences of the definition and the arithmetic operations on interval-valued fuzzy sets, we only prove the boundedness. It can be proven by comparing the aggregated value to the minimum and maximum as follows:

$$\min_i \{d(A_i, B_i)\} = g^{-1}\left(g(\min_i \{d(A_i, B_i)\})\right) \\ = g^{-1}\left(\sum_{j=1}^{n} w_j g(\min_i \{d(A_i, B_i)\})\right) \leq g^{-1}\left(\sum_{j=1}^{n} w_j g(D_j)\right) \\ = f(\langle U_1, A_1, B_1\rangle, \langle U_2, A_2, B_2\rangle, \dots, \langle U_n, A_n, B_n\rangle)$$

and

$$\begin{aligned}\max_i \{d(A_i, B_i)\} &= g^{-1}\left(g(\max_i \{d(A_i, B_i)\})\right)\\ = g^{-1}\left(\sum_{j=1}^{n} w_j g(\max_i \{d(A_i, B_i)\})\right) &\geq g^{-1}\left(\sum_{j=1}^{n} w_j g(D_j)\right)\\ &= f(\langle U_1, A_1, B_1\rangle, \langle U_2, A_2, B_2\rangle, \ldots, \langle U_n, A_n, B_n\rangle).\end{aligned}$$

Note 1 One special case of this definition is the generalized IVFN-IOWAD operator, where $g(x) = x^{\alpha}$, $\alpha \in \mathbb{R}$, and it takes the following form:

$$\left(\sum_{j=1}^{n} w_j D_j^{\alpha}\right)^{\frac{1}{\alpha}}.$$

Definition 5 ([44]) An IVFN-IOWAD operator of dimension n is a mapping $f : \mathbb{R}^n \times \text{IVFN}^n \times \text{IVFN}^n \to \mathbb{R}$ that has an associated weighting vector W of dimension n with $w_j \in [0, 1]$ and $\sum_{j=1}^{n} w_j = 1$, such that:

$$f(\langle u_1, A_1, B_1\rangle, \langle u_2, A_2, B_2\rangle, \ldots, \langle u_n, A_n, B_n\rangle) = \sum_{j=1}^{n} w_j D_j, \tag{4}$$

where D_j is the $d(A_i, B_i)$ value of the triplet $\langle u_i, A_i, B_i\rangle$ having the jth largest u_i, where u_i is the order inducing variable and A_i, B_i are the argument variable represented in the form of IVFN's.

Example 1 To illustrate the definition, we will calculate the OWA- distance of trapezoidal-shaped IVFN's (the upper and lower fuzzy numbers are trapezoidal fuzzy numbers) choosing $g(x) = x$, which is a special case of the definition, an IVFN-IOWAD operator. The upper and lower fuzzy numbers can be represented as $A^L = (a, b, \alpha, \beta)$ and $A^U = (c, d, \theta, \tau)$ respectively, where $[a, b]$ and $[c, d]$ stand for the central intervals, (α, β) and (θ, τ) denotes the left and right width of the fuzzy numbers. The mean value of a triangular IVFN can be expressed as

$$E(A) = \frac{a+b}{4} + \frac{c+d}{4} + \frac{\beta-\alpha}{12} + \frac{\tau-\theta}{12}.$$

In the example, we suppose that new data is available (concerning a potential intrusion) and it is compared to two previous intrusion cases stored in the database. One expert evaluates the cases based on three criteria, and this evaluation will be used to calculate the distance between the new observation and the two stored cases. The expert's evaluation is described in Table 1.

Table 1 The evaluation of the expert for the different cases

Criteria	New case	Case 1	Case 2
C_1	$A_1^U = (4, 6, 2, 2)$	$B_1^{1,U} = (3, 6, 1, 1)$	$B_1^{2,U} = (5, 5, 3, 3)$
C_1	$A_1^L = (4, 5, 1, 1)$	$B_1^{1,L} = (4, 5, 1, 1)$	$B_1^{2,L} = (5, 5, 2, 2)$
C_2	$A_2^U = (8, 10, 3, 3)$	$B_2^{1,U} = (7, 8, 3, 4)$	$B_2^{2,U} = (9, 11, 2, 2)$
C_2	$A_2^L = (9, 9, 2, 1)$	$B_2^{1,L} = (7, 7, 2, 4)$	$B_2^{2,L} = (10, 11, 2, 2)$
C_3	$A_3^U = (2, 4, 1, 1)$	$B_3^{1,U} = (5, 7, 2, 2)$	$B_3^{2,U} = (4, 6, 1, 1)$
C_3	$A_3^L = (3, 4, 1, 1)$	$B_3^{1,L} = (6, 7, 2, 1)$	$B_3^{2,L} = (5, 6, 2, 1)$

The corresponding order inducing variables for the criteria (we use crisp values in this example) are $u_1 = 5, u_2 = 7, u_3 = 2$. The weights are defined as $W = (0.4, 0.2, 0.4)$. The aggregation can be calculated as

$$f_i(\langle 5, A_1, B_i^1\rangle, \langle 7, A_2, B_i^2\rangle, \langle 2, A_3, B_i^3\rangle) = 0.4|E(A_2 - B_i^2)| \\ + 0.2|E(A_1 - B_i^1)| + 0.4|E(A_3 - B_i^3)|.$$

for $i = 1, 2$. The obtained values are $f_1 = 1.78$ and $f_2 = 1.35$, which indicate that the new case is more similar to Case 2 from the database, as the distance between these two instances is smaller than the distance between the new case and Case 1.

To use the distance measure for obtaining similarities, we need to normalize the aggregated values by dividing by the factor

$$\sup \{x \in \cup_i (supp(A_i) \cup supp(B_i)\} - \inf \{x \in \cup_i (supp(A_i) \cup supp(B_i)\}$$

and compute $s = 1 - d$, where d stands for the Quasi IVFN-IOWAD operator.

6 The Financial Institution Ontology

With the ambition to illustrate how fuzzy ontologies and similarity measures could be used for intrusion detection purposes, we created a simple fuzzy ontology. Although intrusion methods are seldom limited to a specific context, our ontology was adapted to fit relevant risks associated with financial institutions. This ontology was then used as the base for creating a simple application, showing a practical example on how fuzzy ontologies can aid in intrusion detection by generating the risk for certain intrusions to occur. In Sect. 6.1 we demonstrate this by presenting a couple of scenarios, pointing out where the fuzzy ontology could contribute, Sect. 6.2 presents the structure of the ontology whereas the technical parts of the application are presented more in Sect. 6.3.

6.1 Intrusion Scenarios

As to clarify the context and functions of the application proposed, we describe a couple of scenarios, showing how the fuzzy ontology could aid in detecting possible intrusions.

Scenario 1.
The first scenario addresses a malware attack, presumably from the widely used Zeus malware. In this scenario, the advice generated by the system is assessed and combined by human experts, whereas a final result is produced, based on both the ontology result and the expert assessments.

As the surveillance system notices an abnormal behaviour, the recorded values are processed by our proposed intrusion detection system. This generates a result showing how likely the detected abnormality is an intrusion attempt. In other words, this example would display the following result:

Value 1 is 95 % similar to Zeus_Intrusion_nr45
Value 2 is 59 % similar to Zeus_Intrusion_nr32
Value 3 is 67 % similar to Gauss_Intrusion_nr2
Value 4 is 85 % similar to Zeus_Intrusion_nr45
Value 5 is 75 % similar to Zeus_Intrusion_nr5

*It is **Highly** likely that the detected intrusion is a **Zeus-based malware**.*

The values represent different measures relevant to the behaviour of the intrusion. Regarding the Zeus Malware, they could represent: amount of hazardous .php files detected; amount of hazardous .exe files detected; amount of functions reporting a malfunction. The detected files are compared with different lists containing hazardous files frequent in different types of intrusions.

A human expert would then asses the results generated by the ontology. The expert does also have the option to see not only the most similar case, but also the whole list of generated similarities. In this case, the expert could notice that Zeus_Intrusion_nr45 and Zeus_Intrusion_nr47 had a 65 and 64 % similarity to value 4, respectively. This aids in the experts' decision making, making it possible to embrace the whole picture and decide in favour of the proposed analysis. The defence systems would then take the appropriate measures, being more efficient, as the intrusion method is likely to be known.

Scenario 2.
The second scenario is assumed to be a denial-of-service attack (Dos), conducted with the purpose of overloading the institutions online system, creating chaos that would consume both time and money to be sorted out. The number of Dos attacks has increased lately, with the main goal of punishing the target by making their online system crash. This scenario excludes the experts, as Dos require immediate action and therefore can not wait for human input.

Online systems can easily define what the normal range of data traffic is, using historical data, and also defining when the crucial limits are reached. One could use fuzzy interval values to model when the values are closing in on the critical limits. In other words, we can use linguistic terms, such as: low risk, medium risk and high risk for indicating how close to the critical limit the amount of data traffic is. This means that one can observe even small risks, where several slightly suspicious factors (which would not have been noticed in a non-fuzzy system) together can detect possible intrusions. Risks or changes that otherwise would have been unnoticed.

For example, a bank usually registers 1000 logins per hour in their online banking system, the record is 1500 and the minimum 500. In other words, it usually moves between 500 and 1500. By using type-2 fuzzy sets, one can define that if the value goes over 1500, it is considered to be "Highly trafficked" and as it reaches closer to 2000, it becomes more and more "Critical". However, the system does not need to shut down if the number of logins exceeds a critical limit; if several similar measures are starting to reach a critical level, the system can conclude that a possible attack is occurring. The fuzzy ontology then can be used to define what kind of Dos attack is most likely taking place and adjust the counter measures according to that knowledge, for instance, by quickly shutting down the system before it crashes and wait for human maintenance persons to make the final decision. In this way one could avoid costly maintenance conducted after a real crash has happened.

6.2 The Fuzzy Ontology

The ontology was created with Protégé [32], the main modelling software for OWL ontologies. Fuzzy datatypes where added to the ontology using the Fuzzy OWL plug-in [7]. The plug-in is an important step towards including fuzzy logic in OWL and making fuzzy logic available for general users. The intrusion risks included in the ontology were collected from different computer security companies and reports, e.g. from The Kaspersky Lab[2] and S2sec.[3]

Figure 1 shows a overview of the ontology. The ontology is structured by fuzzy classes according to the intrusion type, e.g. **Social_Engineering** and **Malware_and _Viruses**. Each of these general classes have more specified subclasses, such as: **Phisihing** and **Win32.** . The subclasses are populated with individuals, representing specific intrusions, such as the famous Zeus malware and previous recorded intrusion attempts. All the individual instances have a set of recorded values or behaviours showing how the intrusion was conducted. Using similarity measures these values are compared with the new intrusions.

[2] http://www.kaspersky.com/

[3] http://www.s21sec.com/

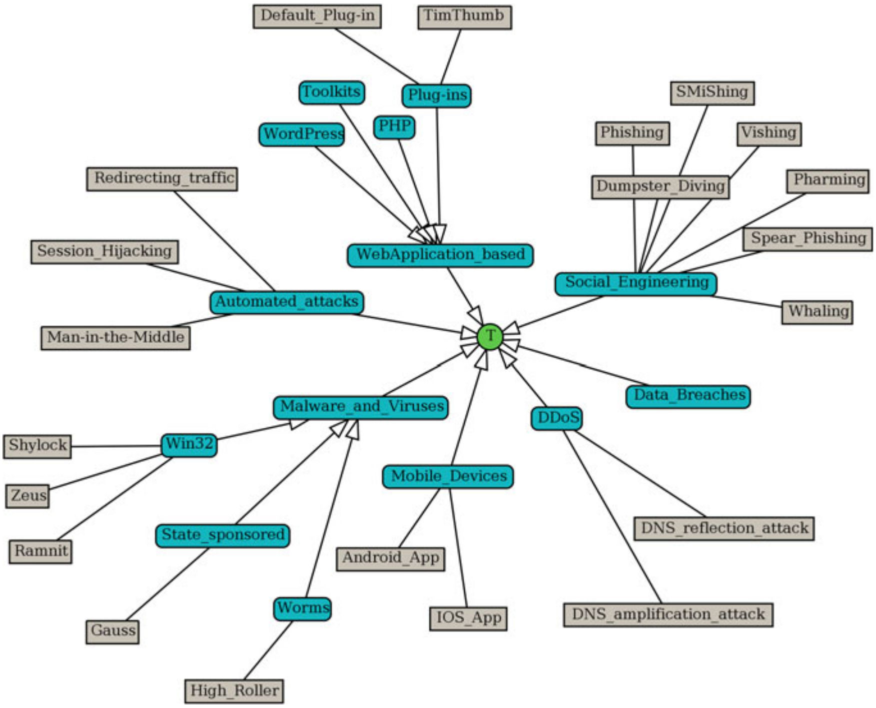
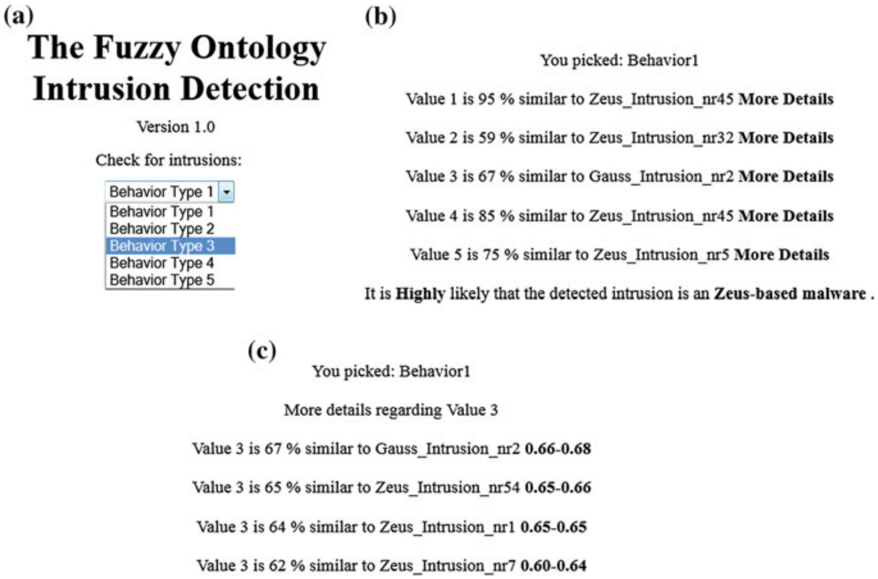


Fig. 1 The structure of the fuzzy ontology

6.3 *F.I. Application*

Using the programming language Java, an application, retrieving information from the OWL ontology, was constructed [8]. Java makes it possible to connect the structure with other techniques, for instance HTML, meaning that one can run the application online. The application functions in the following way:

- The user decides among a couple of pre-defined example treats that have been "registered". The function of registering and comparing the possible intrusion would be automatic in a real world intrusion detection system (Fig. 2a).
- The chosen intrusion is modelled with interval type-2 fuzzy sets. The previously stored intrusions are retrieved from the ontology and the similarities to the current intrusion are computed.
- The results of the computation, i.e. how likely the detected abnormally is an intrusion and in that case which previous intrusion it resembles, is presented to the user, see Fig. 2b.
- The user has the possibility to view also other similar intrusions, screenshot shown in Fig. 2c, offering, for instance, for the experts that should make the final decision an opportunity to get a more comprehensive picture of the situation.



You picked: Behavior1

Value 1 is 95 % similar to Zeus_Intrusion_nr45 **More Details**

Value 2 is 59 % similar to Zeus_Intrusion_nr32 **More Details**

Value 3 is 67 % similar to Gauss_Intrusion_nr2 **More Details**

Value 4 is 85 % similar to Zeus_Intrusion_nr45 **More Details**

Value 5 is 75 % similar to Zeus_Intrusion_nr5 **More Details**

It is **Highly** likely that the detected intrusion is an **Zeus-based malware** .

You picked: Behavior1

More details regarding Value 3

Value 3 is 67 % similar to Gauss_Intrusion_nr2 **0.66-0.68**

Value 3 is 65 % similar to Zeus_Intrusion_nr54 **0.65-0.66**

Value 3 is 64 % similar to Zeus_Intrusion_nr1 **0.65-0.65**

Value 3 is 62 % similar to Zeus_Intrusion_nr7 **0.60-0.64**

Fig. 2 Example of a fuzzy intrusion detection user interface: **a** The initial choice, **b** First retrieved results, **c** More detailed results

It has to be acknowledged that the functions in the application are only basic, however, the structure of the application and the techniques it builds on, e.g. OWL, fuzzyDL [7], Java and HTML to make it suitable for extending and combining with numerous other applications.

7 Conclusion

Intrusion detection is becoming more and more essential to handle the risks associated with network activities. New intrusion detection systems should be capable of preventing organisation not only from an increasing numbers of attacks but also from more and more sophisticated intrusion strategies. One promising solution would be to exercise expert insights in the detection process. As experts have accumulated an extensive knowledge of their field, in many cases they can point out some irregularities which can indicate anomalies that otherwise could not be identified by automatic intrusion detection systems and would result in significant losses.

In many cases, tacit knowledge of experts can be expressed only using linguistic (imprecise) terms. In our proposal the combination of fuzzy logic and ontologies can transform expert knowledge into a systematic description processable by computational methods. The fusion of type-2 fuzzy ontologies and similarity measures to identify possible causes of intrusions provides benefits to organisations that are not achievable by other methods.

To support the results of the ontology and provide additional information that can be essential to identify anomalies, expert opinions expressed in terms of linguistic information and modelled by interval-valued fuzzy numbers are employed. As numerous intrusions can occur at the same time, the proposed system can provide estimations on the seriousness of different activities in terms of potential losses. Based on this, the decision makers can assign the limited available resources in a way that is optimal: by assigning more resources to the more serious cases, the potential loss can be minimized. Our proposal can be extended by incorporating more detailed database of intrusions for testing purpose and also by using different types of similarity measures.

References

1. Acampora, G., Loia, V.: Using FML and fuzzy technology in adaptive ambient intelligence environments. Int. J. Comput. Intell. Res. **1**(1), 171–182 (2005)
2. Anderson, J.P.: Computer security threat monitoring and surveillance. Technical Report James P. Anderson Company, Fort Washington, Pennsylvania (1980)
3. Apel, M., Bockermann, C., Meier M.: Measuring similarity of malware behavior. In: IEEE 34th Conference on Local Computer Networks (LCN 2009), pp. 891–898 (2009)
4. Atannasov, K.: Intuitionistic Fuzzy Sets: Theory and Applications. Physica-Verlag, New York (1999)
5. Berners-Lee, T.: Semantic web on XML. http://www.w3.org/2000/Talks/1206-xml2k-tbl/ (2000)
6. Bobillo, F.: Managing vagueness in ontologies. PhD Thesis, University of Granada, Spain (2008)
7. Bobillo, F., Straccia, U.: Fuzzy ontology representation using OWL 2. Int. J. Approximate Reasoning **52**(7), 1073–1094 (2011)
8. Bobillo, F., Straccia, U.: Aggregation operators for fuzzy ontologies. Appl. Soft Comput. **13**(9), 3816–3830 (2013)
9. Botha, M., von Solms, R.: Utilising fuzzy logic and trend analysis for effective intrusion detection. Comput. Secur. **22**(5), 423–434 (2003)
10. Burillo, P., Bustince, H.: Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets. Fuzzy Sets Syst. **78**(3), 305–316 (1996)
11. Bustince, H.: Indicator of inclusion grade for interval-valued fuzzy sets. application to approximate reasoning based on interval-valued fuzzy sets. Int. J. Approximate Reasoning **23**(3), 137–209 (2000)
12. Carlsson, C., Fullér, R., Mezei J.: Project selection with interval-valued fuzzy numbers. In: IEEE 12th International Symposium on Computational Intelligence and Informatics (CINTI), pp. 23–26 (2011)
13. Catania, C.A., Garino, C.G.: Automatic network intrusion detection: current techniques and open issues. Comput. Electr. Eng. **38**(5), 1062–1072 (2012)
14. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: a survey. ACM Comput. Surv. (CSUR) **41**(3), 15 (2009)
15. Chen, S.-J., Chen, S.-M.: Fuzzy risk analysis based on measures of similarity between interval-valued fuzzy numbers. Comput. Math. Appl. **55**(8), 1670–1685 (2008)
16. Chen, S.-M., Chen, J.-H.: Fuzzy risk analysis based on similarity measures between interval-valued fuzzy numbers and interval-valued fuzzy number arithmetic operators. Expert Syst. Appl. **36**(3), 6309–6317 (2009)

17. Chiang, H.-S., Tsaur, W.: Mobile malware behavioral analysis and preventive strategy using ontology. In: IEEE Second International Conference on Social Computing (SocialCom), pp. 1080–1085 (2010)
18. Dai, S.-Y., Fyodor, Y., Kuo, S.-Y., Wu, M.-W., Huang Y.: Malware profiler based on innovative behavior-awareness technique. In: IEEE 17th Pacific Rim International Symposium on Dependable Computing (PRDC), pp. 314–319 (2011)
19. Dengfeng, L., Chuntian, C.: New similarity measures of intuitionistic fuzzy sets and application to pattern recognitions. Pattern Recogn. Lett. **23**(1), 221–225 (2002)
20. Dickerson, J.E., Juslin, J., Koukousoula, O., Dickerson, J.A.: Fuzzy intrusion detection. In: IEEE 9th joint IFSA World Congress and 20th NAFIPS International Conference, vol. 3, pp. 1506–1510 (2001)
21. Feng, Z.-Q., Liua, C.-G.: On similarity-based approximate reasoning in interval-valued fuzzy environments. Informatics **36**, 255–262 (2012)
22. Gorzałczany, M.: A method of inference in approximate reasoning based on interval-valued fuzzy sets. Fuzzy Sets Syst. **21**(1), 1–17 (1987)
23. Grzegorzewski, P.: Distances between intuitionistic fuzzy sets and/or interval-valued fuzzy sets based on the hausdorff metric. Fuzzy Sets Syst. **148**(2), 319–328 (2004)
24. Hendler, J.: Agents and the semantic web. Intell. Syst. **16**(2), 30–37 (2001)
25. Horridge, M., Krötzsch, M., Parsia, B., Patel-Schneider, P., Rudolph, S.: OWL 2 web ontology language, primer. W3C Working Group (2009)
26. Hua, J., Bapna, S.: The economic impact of cyber terrorism. J. Strateg. Inf. Syst. **22**(2), 175–186 (2013)
27. Huang, H.-D., Acampora, G., Loia, V., Lee,C.-S., Hagras, H., Wang, M.-H., Kao, H.-Y., Chang J.-G.: Fuzzy markup language for malware behavioral analysis. In: On the Power of Fuzzy Markup Language, pp. 113–132. Springer (2013)
28. Huang, H.-D., Acampora, G., Loia, V., Lee, C.-S., Kao, H.-Y.: Applying FML and fuzzy ontologies to malware behavioural analysis. In: IEEE International Conference on Fuzzy Systems, pp. 2018–2025 (2011)
29. Huang, H.-D., Lee, C.-S., Wang, M.-H., Kao, H.-Y.: IT2FS-based ontology with soft-computing mechanism for malware behavior analysis. Soft Comput. **18**(2), 267–284 (2014)
30. Hung, S.-S., Liu, D.S.-M.: A user-oriented ontology-based approach for network intrusion detection. Comput. Stan. Interfaces **30**(1–2), 78–88 (2008)
31. Hwang, C.-M., Yang, M.-S.: New similarity measures between interval-valued fuzzy sets. In: Proceedings of the 15th WSEAS International Conference on Systems, pp. 66–70 (2011)
32. Knublauch, H., Fergerson, R., Noy, N., Musen, M.: The Protégé OWL plugin: an open development environment for semantic web applications. The Semantic Web-ISWC **2004**, 229–243 (2004)
33. Lau, A., Tsui, E., Lee, W.: An ontology-based similarity measurement for problem-based case reasoning. Expert Syst. Appl. **36**(3, Part 2):6574–6579 (2009)
34. Lawry, J.: A framework for linguistic modelling. Artif. Intell. **155**(1–2), 1–39 (2004)
35. Leder, F.S., Martini, P.: Ngbpa next generation botnet protocol analysis. In: Emerging Challenges for Security, Privacy and Trust, pp. 307–317. Springer (2009)
36. Lee, C., Wang, M., Hagras, H.: A type-2 fuzzy ontology and its application to personal diabetic-diet recommendation. IEEE Trans. Fuzzy Syst. **18**(2), 374–395 (2010)
37. Li, W., Tian, S.: An ontology-based intrusion alerts correlation system. Expert Syst. Appl. **37**(10), 7138–7146 (2010)
38. Li, Y., Olson, D.L., Qin, Z.: Similarity measures between intuitionistic fuzzy (vague) sets: a comparative analysis. Pattern Recogn. Lett. **28**(2), 278–285 (2007)
39. Liang, Z., Shi, P.: Similarity measures on intuitionistic fuzzy sets. Pattern Recogn. Lett. **24**(15), 2687–2693 (2003)
40. Liao, Y., Vemuri, V.R.: Using text categorization techniques for intrusion detection. In: USENIX Security Symposium, vol. 12 (2002)
41. Liu, W.: Research of data mining in intrusion detection system and the uncertainty of the attack. In: International Symposium on Computer Network and Multimedia Technology, pp. 1–4 (2009)

42. Lukasiewicz, T., Straccia, U.: Managing uncertainty and vagueness in description logics for the semantic web. Web Semantics: Science, Services and Agents on the World Wide Web **6**(4), 291–308 (2008)
43. McCulloch, J., Wagner, C., Aickelin, U.: Extending similarity measures of interval type-2 fuzzy sets to general type-2 fuzzy sets. In: IEEE International Conference on Fuzzy Systems, pp. 1–8 (2013)
44. Mezei J., Wikström, R.: OWAD operators in type-2 fuzzy ontologies. In: Proceedings of the 2013 Joint IFSA World Congress NAFIPS Annual Meeting, number ISBN: 978-1-4799-0347-4, pp. 848-853 (2013)
45. Mitchell, H.: On the dengfeng-chuntian similarity measure and its application to pattern recognition. Pattern Recogn. Lett. **24**(16), 3101–3104 (2003)
46. Mitchell, H.B.: Pattern recognition using type-II fuzzy sets. Inf. Sci. **170**(2), 409–418 (2005)
47. Ning, P., Xu, D.: Learning attack strategies from intrusion alerts. In: Proceedings of the 10th ACM Conference on Computer and Communications Security, pp. 200–209. ACM, (2003)
48. Park, W.H.: Risk analysis and damage assessment of financial institutions in cyber attacks between nations. Math. Comput. Model. **58**(11–12), 18–45 (2012)
49. Riccardi, M., Oro, D., Luna, J., Cremonini, M., Vilanova, M.: A framework for financial botnet analysis. In: eCrime Researchers Summit (eCrime), pp. 1–7 (2010)
50. Riccardi, M., Pietro, R.D., Palanques, M., Vila, J.A.: Titans revenge: detecting Zeus via its own flaws. Comput. Networks **57**(2):422–435 (2013) (Botnet Activity: Analysis, Detection and Shutdown.)
51. Sharma, A., Pujari, A.K., Paliwal, K.K.: Intrusion detection using text processing techniques with a kernel based similarity measure. Comput. Secur. **26**(7–8), 488–495 (2007)
52. Simmonds, A., Sandilands, P., van Ekert, L.: An ontology for network security attacks. In: Applied Computing, pp. 317–323. Springer (2004)
53. Szmidt, E., Kacprzyk, J.: On measuring distances between intuitionistic fuzzy sets. Notes on IFS **3**(4), 1–13 (1997)
54. Szmidt, E., Kacprzyk, J.: Distances between intuitionistic fuzzy sets. Fuzzy Sets Syst. **114**(3), 505–518 (2000)
55. Tafazzoli, T., Sadjadi, S.H.: Malware fuzzy ontology for semantic web. Int. J. Comput. Sci. Network Secur. **8**(7), 153–161 (2008)
56. Tajbakhsh, A., Rahmati, M., Mirzaei, A.: Intrusion detection using fuzzy association rules. Appl. Soft Comput. **9**(2), 462–469 (2009)
57. Turksen, I., Zhong, Z.: An approximate analogical reasoning schema based on similarity measures and interval-valued fuzzy sets. Fuzzy Sets Syst. **34**(3), 323–346 (1990)
58. Undercoffer, J., Joshi, A., Pinkston, J.: Modeling computer attacks: an ontology for intrusion detection. In: Recent Advances in Intrusion Detection, pp. 113–135. Springer, (2003)
59. Wagener, G., Dulaunoy, A., et al.: Malware behaviour analysis. J. Comput. Virol. **4**(4), 279–287 (2008)
60. Wang, C., Entropy, AQu: similarity measure and distance measure of vague soft sets and their relations. Inf. Sci. **244**, 92–106 (2013)
61. Wang, G., Hao, J., Ma, J., Huang, L.: A new approach to intrusion detection using artificial neural networks and fuzzy clustering. Expert Syst. Appl. **37**(9), 6225–6232 (2010)
62. Wu, D., Mendel, J.: A comparative study of ranking methods, similarity measures and uncertainty measures for interval type-2 fuzzy sets. Inf. Sci. **179**(8), 1169–1192 (2009)
63. Wu, D., Mendel, J.M.: A vector similarity measure for linguistic approximation: interval type-2 and type-1 fuzzy sets. Inf. Sci. **178**(2), 381–402 (2008)
64. Xu, Z.: Some similarity measures of intuitionistic fuzzy sets and their applications to multiple attribute decision making. Fuzzy Optim. Decis. Making **6**(2), 109–121 (2007)
65. Xu, Z., Chen, J.: Ordered weighted distance measure. J. Syst. Sci. Syst. Eng. **17**(4), 432–445 (2008)
66. Zadeh, L.A.: Fuzzy logic = computing with words. IEEE Trans. Fuzzy Syst. **4**(2), 103–111 (1996)

67. Zeng, W., Guo, P.: Normalized distance, similarity measure, inclusion measure and entropy of interval-valued fuzzy sets and their relationship. Inf. Sci. **178**(5), 1334–1342 (2008)
68. Zeng, W., Li, H.: Relationship between similarity measure and entropy of interval valued fuzzy sets. Fuzzy Sets Syst. **157**(11), 1477–1484 (2006)
69. Zhang, C., Fu, H.: Similarity measures on three kinds of fuzzy sets. Pattern Recogn. Lett. **27**(12), 1307–1317 (2006)
70. Zhang, H., Zhang, W.: Inclusion measure and similarity measure of intuitionistic and interval-valued fuzzy sets. In: Proceedings of the 2007 International Conference on Intelligent Systems and Knowledge Engineering (ISKE2007) (2007)
71. Zheng, G., Wang, J., Zhou, W., Zhang, Y.: A similarity measure between interval type-2 fuzzy sets. In: International Conference on Mechatronics and Automation (ICMA), pp. 191–195 (2010)
72. Zheng, G., Xiao, J., Wang, J., Wei, Z.: A similarity measure between general type-2 fuzzy sets and its application in clustering. In: IEEE 8th World Congress on Intelligent Control and Automation (WCICA), pp. 6383–6387 (2010)