UDC 004

An example of fuzzy ontology usage for risk assessment and attack impact

Oleh Kozlenko[1]

[1] *National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine*

Annotation

The article discusses the use of fuzzy ontology for assessing risks and impacts of attacks in the field of information security. Fuzzy ontology, which is a formalized way of representing knowledge, offers effective solutions for processing complex and informal processes. The article substantiates the significance of fuzzy logic in structural analysis and presents an example of how new types of attacks influence the ontology. Key findings include the identification of risks associated with attacks through the application of fuzzy sets and entropy theory. The discussion highlights how these methods can enhance threat response and risk management in information systems.

Keywords: Cybersecurity, risk, fuzzy logic, ontology, threat analysis

Introduction

Ontologies represent a structured and formal way of organizing and presenting knowledge, facilitating communication and information exchange. Several ontologies have been created in the field of security, but there is still no unified and consistent model. Fuzzy ontology serves as an effective tool for processing various descriptions of ontologies created for a single domain. The apparatus of fuzzy mathematics is a key method for analysis when modeling structures and informal processes, including tasks in information security. Fuzzy set theory complements and extends classical set theory, and its counterpart in formal logic is fuzzy logic. The emergence of this new theory is driven by the need to describe processes, systems, and objects using fuzzy and approximate reasoning.

Theoretical Foundation of Fuzzy Ontology

Let X be a space of points (objects) with a common element x, then X = {z}. A fuzzy set A in X is defined by a membership function f_A(x), which maps each point in X to a real number in the interval [0, 1], indicating the "degree of membership" of x in A. The closer f_A(x) is to one, the higher the degree of membership of x in A. If A is a traditional set, its membership function can only take values of 0 and 1, where f_A(x) equals 1 or 0 depending on whether x belongs to A. Thus f_A(x) in this case defines the familiar functions of set A, which can significantly exceed 1.

The membership function of a fuzzy set bears some similarity to a probability function when X is a crisp set, but there are significant differences between these concepts that become clearer with the establishment of rules for combining membership functions and their fundamental properties. The concept of a fuzzy set is dynamic: it is empty only when its membership function equals zero for all x in X.

Introducing vagueness into the structure of ontologies helps eliminate ambiguities arising from discrepancies in user requirements specifications and descriptions of concepts. Creating a fuzzy ontology can be a challenging task, especially when done manually. Some classical ontologies, such as WordNet, exhibit various types of relationships between concepts, and in some cases, these relationships may be associated with degrees of the values of membership functions. The degree of membership can reflect the category of relations from which it originates. Another approach involves domain analysis and statistical study of terms to establish connections.

Ontologies require verification and adjustment by experts, even if their basic versions are generated automatically. This is necessary not only due to the limitations of automated generation but also because ontologies depend on specific applications and often contain pragmatic information. The main

goal of defining methods for constructing fuzzy ontologies is to address these issues. Reference [1] presents a fuzzy relational model of ontologies. Gottroy [2] focuses on knowledge extraction from databases using fuzzy rules to refine ontologies. However, his approach to this topic remains quite unclear, and formal semantics are lacking. Additionally, there are studies related to fuzzy OWL ontologies [1] and fuzzy reasoning in DL [3]. However, these approaches are still based on a "traditional" representation of knowledge grounded in logic, which we believe is inadequate for working with learned ontologies. Artificial intelligence (AI) methods, heuristic approaches [4], and similar paradigms [5] offer alternative frameworks but have not yet been connected to the mechanism of automatic knowledge acquisition in the real world. The literature also contains several descriptions of fuzzy logics, including fuzzy extensions of OWL [6].

An example of the unknown attack's impact on the ontology structure

Let's consider how an unknown attack will affect the ontology presented in [7]. Suppose a new type of attack called Unknown appears. During the research, 150 cases of this type of attack were identified, of which 100 led to information leaks. The following cybersecurity measures have been identified as helping to contain an Unknown attack: web application testing, proprietary software development materials, regular updates and patches, incident management, incident roles, DLP systems, administrator access control, no sensitive data exposed, backup, encryption, employee information security training, employee screening, and collaboration with IS and management related to the following types of attacks:

- Attack on web applications (3)
- DOS (3)
- Insider threat (1)
- Various errors (2)
- Physical(2)
- Skimmers (0)
- Cyberespionage (2)
- Malware (2)
- POS (2)
- Coordination (2)
- Management readiness (2)
- CS adoption rate (2)
- Personnel security (2)

Unknown affects the confidentiality, integrity and availability of information and can have remote and local characters. The corresponding required characteristics to be added to the ontology will have the following values:

- G = 3
- SM = 14
- W = 2
- R_i = 2.176
- R_r = 2
- Wq = 2

The corresponding weight element of this attack will be:

$$F = W_q \cdot G \cdot W \cdot R_i / R_r = 13.056 \qquad (1)$$

The set P for the Unknown attack will then have the following elements: {Testing of web applications, Closure of materials for developed software, Updates and patches, Work with incidents, Roles in incidents, DLP - system, Control of administrators, No confidential data in open form, Backup, Encryption, IT training for collaborators, Verification of employees, Cooperation with IS department, Cooperation with management }, {0, 208, 416, 624, 832, 1040, 1248, 1456, 1664, 1872, 2080, 2288, 2496, 2704,2912}), Unknown, {Privacy, integrity, availability}, {0, 78, 156, 234}), Unknown, {Remote, local}, {0, 52, 104 }0).

The value of the set of connections R will have the following structure:

{(Unknown, Web application attack, medium), (Unknown, DOS, medium), (Unknown, Insider attacks, weak), (Unknown, Misc bugs, medium), (Unknown, Physical theft, medium), (Unknown , Skimmers, weak), (Unknown, Cyber espionage, medium), (Unknown, Criminal software, medium), (Unknown, POS, medium), (Unknown, Coordination, medium), (Unknown, Management readiness, medium), (Unknown, CS acceptance rate, medium), (Unknown, Personnel security, medium)}.

For example, we will use the ontology as a basis for risk determination using fuzzy sets and entropy theory [8]. The basis of this method is the assumption that the system can be in different states $\{S_1, S_2, S_3, \ldots, S_n\}$, and probability of P_i in state S_i[8]:

$$i = 1,2,\dots,n; 0 \le P_i \le 1; \sum_{i=1}^{n} P_i = 1 \quad (2)$$

Entropy will be [8]:

$$H = -\sum_{i=1}^{n} P_i \cdot ln(P_i) \quad (3)$$

Entropy for the weight of factor A_i will be [8]:

$$H_i = -\sum_{j=1}^{m} p_{ij} \cdot ln(p_{ij}) \quad (4)$$

When $P_i = 1/n$, then entropy value will be maxed out - $H_i = lnm$. Taking into account all the above, the determination of the value of entropy is of relative importance for A_i will be shown as:

$$e_i = \frac{-1}{lnm} \sum_{j=1}^{m} p_{ij} \cdot ln(p_{ij}) \quad (5)$$

After normalization, ϕ_i for every A_i will be as:

$$\phi_i = \frac{1 - e_i}{n - E} \quad (6)$$

Among the factors affecting the determination of the level of system risk, the most "unclear" are the impact on assets (asset impact), threat frequency (threat frequency) and the degree of vulnerability (severity of vulnerability). The use of fuzzy sets for system analysis requires [8]:

1. Determination of a set of system risk factors $U = u_1, u_2, u_3, \dots, u_n$
2. Experts assess the threat of risk factors for each element of the system and divide the results into m levels accordingly. As an example, $V = v_1, v_2, v_3, \dots, v_m$ - evaluation set for the above factors
3. Experts estimate U, V, and a fuzzy mapping function is constructed f:$U \to F(V)$,$V = v_1, v_2, v_3, \dots, v_m$, so $u_i \to f(u_i) = (p_{i1}, p_{i2}, \dots, p_m) \in F(V)$

Fuzzy connections will then look like this:

$$R_f \epsilon F(UxV), R_f(u_i, v_j) = f(u_i)(v_j) = p_{ij}$$

To obtain the final risk value, all values for each element of the system must be determined, as well as their respective weights $\phi = \phi_1, \phi_2, \phi_3, \dots, \phi_n$. The risk assessment for a specific element will look like this [8]:

$$R_a = \phi \cdot P_a \cdot U' \quad (7)$$

An identical procedure should be applied for the parameters threat frequency and severity of vulnerability.

After determining the weight parameters and all the arguments, the risk value R can be found using the formula [8]:

$$R = k_1 \cdot R_a + k_2 \cdot Rt + k_3 \cdot R_v, \quad k_1 + k_2 + k_3 = 1 \quad (8)$$

where *k* is the weight of the evaluation element.

Consider the situation with an organization working with a database. The risk assessment of the information system is based on the definition of entropy and fuzzy sets and will use the appropriate calculation model.

Table 1

Correspondence of the lexical value of the risk level to its quantitative value

Quantification of risk	Risk level
0.0 < ϕ < 0.2	Minor impact
0.2 < ϕ < 0.4	Low impact
0.4 < ϕ < 0.6	Medium impact
0.6 < ϕ < 0.8	High impact
0.8< ϕ < 1.0	Severe impact

Let's assume that the basic analysis, which is given in the example, identified such threats in the system:

$$S = \langle DOS, Malware, Variouserrors, Insiderthreat \rangle.$$

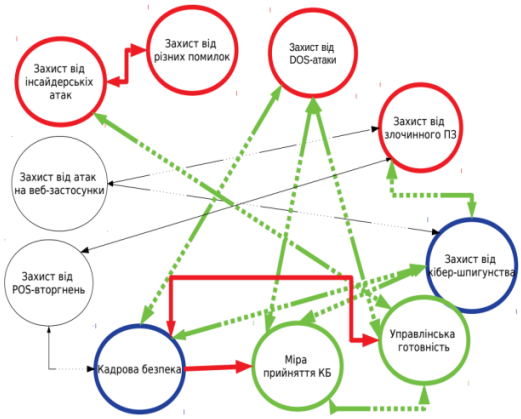
Applying the created fuzzy ontology from [7], we conclude that the connection in this situation is incomplete. Therefore, we will supplement the elements in the set S from the ontology [7] with medium and high connections. The following elements will be added to the resulting set:

- Cyberespionage,
- Personnel management,
- Security measure,
- Management readiness.

The relationship between the elements of Security Level and Management Readiness are high categories, so we can exclude them.

The final set will consist of the following elements: *DOS, Malware, Various errors, Insider threat, Cyberespionage*. We will use this set for risk analysis.

We will determine the expert assessment of each of the elements in 3 areas: Asset, Threat, Vulnerability. For example, lets take 5 experts. The values of the estimates are shown in Table 2:



Pic. 1*: Ontology for the risk assessment example*

Table 2

Expert evaluations for the example of DOS attacks

	Experts	Work with incidents	Roles in incidents	DLP-system
Asset	1	0.2	0.4	0.2
	2	0.1	0.6	0.5
	3	0.1	0.3	0.2
	4	0.2	0.1	0.6
	5	0.4	0.5	0.6
Threat	1	0.1	0.6	0.4
	2	0.1	0.2	0.6
	3	0.4	0.2	0.4
	4	0.4	0.6	0.4
	5	0.1	0.1	0.2
Vulnerability	1	0.1	0.1	0.1
	2	0.3	0.5	0.2
	3	0.4	0.4	0.4
	4	0.1	0.6	0.2
	5	0.5	0.1	0.5

Let's determine the value of e (3) for each of the experts for each direction. The results are shown in Table 3.

$$e_i = \frac{1}{ln(m)} \cdot \sum_{j=1}^{m} r_{ij} \cdot ln(r_{ij}) for i \in 1, n \qquad (9)$$

Table 3

The e value is for an example of DOS attacks

Asset	e value	Threat	e value	Vulnerability	e value
1	0.54	1	0.46	1	0.42
2	0.4956	2	0.94	2	0.526
3	0.51	3	0.827	3	0.68
4	1.201	4	0.54	4	0.39
5	0.701	5	0.57	5	1.07

According to formula (10), Table 4 shows the results of determining the entropy index:

$$E = \sum_{i=1}^{n} e_i \qquad (10)$$

Table 4

The value of E for an example of DOS attacks

Set	E value
Asset	3.45
Threat	3.348
Vulnerability	3.109

The weight indicators will be determined according to formula (6). The results are shown in Table 5.

Table 5

The value of φ for an example of DOS attacks

Set	φ value	Set	φ value	Set	φ value
Asset	0.29	Threat	0.326	Vulnerability	0.301
	0.32		0.034		0.250
	0.31		0.104		0.167
	-		0.275		0.318
	0.130				
	0.193		0.259		-0.038

The total weight index for each direction Asset, Threat, Vulnerability according to the formula will have values of 0.428, 0.480, 0.0912.

At a value of k for all directions, 1/3 the risk value for DOS will look like this:

$$R = kR_a + kR_t + kR_v = 0.3282,$$
$$R_i = \varphi_i \cdot P_i \cdot W_i^T$$

Similarly to the point about DOS attacks, all calculations in all other elements of the ontology are shown in Tables 6 – 13.

Table 6

Experts evaluations for Malware

Asset		Evaluation										
	1	0.023	0.119	0.087	0.064	0.069	0.197	0.169	0.205	0.013	0.054	
	2	0.139	0.104	0.011	0.038	0.138	0.135	0.139	0.032	0.135	0.129	
	3	0.074	0.180	0.137	0.019	0.094	0.154	0.183	0.053	0.001	0.105	
	4	0.092	0.137	0.133	0.008	0.066	0.040	0.127	0.126	0.143	0.128	
	5	0.017	0.037	0.090	0.131	0.163	0.079	0.164	0.090	0.096	0.135	
Threat		Evaluation										
	1	0.042	0.057	0.114	0.175	0.074	0.102	0.160	0.050	0.152	0.072	
	2	0.062	0.103	0.061	0.177	0.079	0.138	0.088	0.156	0.028	0.108	
	3	0.060	0.061	0.020	0.130	0.106	0.122	0.126	0.171	0.149	0.055	
	4	0.091	0.098	0.076	0.167	0.189	0.223	0.052	0.050	0.034	0.021	
	5	0.164	0.137	0.029	0.133	0.143	0.044	0.157	0.028	0.138	0.026	
Vulnerability		Evaluation										
	1	0.101	0.097	0.149	0.110	0.123	0.084	0.049	0.038	0.093	0.156	
	2	0.153	0.181	0.072	0.094	0.024	0.078	0.121	0.157	0.045	0.076	
	3	0.194	0.251	0.098	0.119	0.134	0.035	0.005	0.018	0.137	0.009	
	4	0.040	0.028	0.108	0.186	0.043	0.058	0.234	0.101	0.202	0.000	
	5	0.111	0.014	0.172	0.160	0.150	0.003	0.097	0.156	0.058	0.078	

The next stage is the determination of the value of E. The results of the calculations are shown in Table 7.

Table 7

E and e values for Malware

	e	E
Asset	1.304	6.990
	1.325	
	1.454	
	1.578	
	1.331	
Threat	1.334	6.543
	1.328	
	1.334	
	1.260	
	1.288	
Vulnerability	1.415	6.448
	1.341	
	1.197	
	1.193	
	1.302	

Table 8

Value of indices and risk for Malware

Asset	φ	W	R
1	0.153	0.128	
2	0.163	0.106	
3	0.228	0.123 0.095	0.102
4	0.290	0.016	
5	0.166	0.096 0.116	
Threat	φ	0.149	R
1	0.217	0.036 0.134	
2	0.212		
3	0.216		0.095
4	0.168		
5	0.186		
Vulnerability	φ		R
1	0.287		0.101
2	0.235		
3	0.136		
4	0.134		
5	0.209		

The resulting risk value for Malware will be:
R = 0.098

Table 9
Experts evaluations for Various errors

Asset	Evaluation		
1	0.038	0.233	0.729
2	0.453	0.290	0.257
3	0.505	0.216	0.279
4	0.111	0.503	0.387
5	0.290	0.002	0.708
Threat	Evaluation		
1	0.066	0.193	0.741
2	0.057	0.630	0.313
3	0.516	0.221	0.263
4	0.222	0.373	0.405
5	0.372	0.514	0.114
Vulnerability	Evaluation		
1	0.463	0.504	0.033
2	0.646	0.075	0.279
3	0.567	0.017	0.416
4	0.465	0.244	0.291
5	0.688	0.220	0.092

Table 10
The value of indices and risks for Various errors

	e	E	f	W	R
Asset	0.947	5.677	-0.078	0.428 0.480 0.091	0.064
	0.643		-0.527		
	0.685		-0.465		
	0.531		-0.692		
	2.870		2.761		
Threat	1.066	3.277	-0.038		0.371
	0.372		0.364		
	0.665		0.194		
	0.685		0.183		
	0.488		0.297		
Vulner-ability	0.450	3.632	0.402		0.420
	0.745		0.186		
	1.292		-0.214		
	0.690		0.227		
	0.454		0.399		

The resulting risk value for Various errors:
R = 0.24

Table 11
Experts evaluations for for Insider threat
(A – Asset, T- Threat, V-Vulnerability)

A	1	2	3	4	5	6	7
1	0.148	0.261	0.221	0.004	0.040	0.144	0.182
2	0.117	0.055	0.078	0.196	0.127	0.263	0.164
3	0.157	0.013	0.198	0.003	0.220	0.189	0.220
4	0.110	0.043	0.109	0.167	0.195	0.171	0.205
5	0.058	0.222	0.269	0.043	0.115	0.162	0.131
T	1	2	3	4	5	6	7
1	0.318	0.022	0.047	0.151	0.245	0.023	0.194
2	0.254	0.049	0.239	0.031	0.110	0.187	0.132
3	0.126	0.162	0.186	0.167	0.161	0.025	0.174
4	0.261	0.069	0.136	0.167	0.103	0.104	0.161
5	0.135	0.208	0.032	0.230	0.057	0.152	0.187
V	1	2	3	4	5	6	7
1	0.144	0.071	0.185	0.168	0.099	0.163	0.171
2	0.341	0.082	0.080	0.045	0.002	0.222	0.228
3	0.027	0.101	0.102	0.104	0.143	0.211	0.312
4	0.125	0.084	0.105	0.087	0.208	0.140	0.252
5	0.229	0.097	0.045	0.214	0.221	0.159	0.036

	e	E	f	W	R
Asset	1.598	6.814	0.330	0.045 0.111 0.118 0.160 0.290 0.125 0.152	0.139
	1.097		0.054		
	1.565		0.312		
	1.130		0.071		
	1.423		0.233		
Threat	0.977	5.772	-0.029		0.131
	1.402		0.521		
	1.166		0.215		
	1.143		0.186		
	1.083		0.108		

Vulnerability	1.194	5.547	0.355		0.147
	1.023		0.041		
	1.097		0.178		
	1.172		0.314		
	1.061		0.112		

The resulting risk value for Insider Attack:
R = 0.1356

Table 12
Experts evaluations for Cyberespionage

Asset	1	2	3	4	5	6	7
1	0.256	0.141	0.244	0.087	0.185	0.047	0.040
2	0.095	0.205	0.210	0.122	0.207	0.094	0.067
3	0.178	0.114	0.163	0.164	0.077	0.112	0.191
4	0.198	0.179	0.065	0.116	0.182	0.065	0.194
5	0.086	0.110	0.059	0.164	0.206	0.176	0.199
Threat	1	2	3	4	5	6	7
1	0.204	0.030	0.144	0.101	0.199	0.161	0.161
2	0.402	0.060	0.128	0.013	0.096	0.049	0.252
3	0.323	0.152	0.079	0.081	0.166	0.079	0.119
4	0.196	0.195	0.090	0.127	0.218	0.163	0.011
5	0.008	0.037	0.107	0.273	0.275	0.047	0.252
Vulner-ability	1	2	3	4	5	6	7
1	0.192	0.092	0.192	0.056	0.264	0.184	0.020
2	0.104	0.087	0.201	0.170	0.201	0.126	0.111
3	0.126	0.201	0.030	0.304	0.120	0.036	0.183
4	0.186	0.142	0.203	0.186	0.267	0.009	0.008
5	0.120	0.177	0.133	0.055	0.245	0.043	0.227

Table 13
The value of indices and risks for Cyberespionage

	e	E	f	W	R
Asset	1.254	5.925	0.275	0.210	0.155
	1.228		0.247	0.103	
	1.184		0.199	0.149	
	1.134		0.145	0.162	
	1.124		0.134	0.296	
				0.078	
Threat	1.181	5.490	0.370	0.003	0.158
	1.158		0.321		
	1.123		0.250		
	1.096		0.196		
	0.933		-0.137		
Vulner-ability	1.239	5.712	0.335		0.164
	1.202		0.283		
	1.035		0.049		
	1.042		0.059		
	1.195		0.274		

The resulting risk value for Cyberespionage:
R = 0.1573

Overall results for all elements:

Table 14
Risk value for security measures

Measure	Risk evaluation
DOS	0.328
Malware	0.098
Various errors	0.240
Insider threat	0.138
Cyberespionage	0.157

The value of the final weights is given in the table:

Table 15
Weights of measures

Measure	Weight
DOS	0.002
Malware	0.369
Various errors	0.069
Insider threat	0.184
Cyberespionage	0.008

The final value of the risk, based on the value of the weights of the elements taken from the ontology will be

$$R = \sum_{i=1}^{k} W_i * R_i = 0.08$$

Conclusion

The use of the proposed methods and calculations in work [7], along with the fuzzy ontology presented in the same study, can significantly enhance the computation of risk values using fuzzy ontologies. The example demonstrated that the calculations are relatively straightforward but require expert values, which is one of the drawbacks of this approach. The impact of an unknown attack on the ontology structure can aid in responding to Zero-Day attacks if the potential effects of such attacks on related elements of the protection system are known. At the same time, the risk assessment showed that the approach proposed in work [7] can significantly improve the understanding of the risks present in a system with fuzzy relationships, which are characteristic of information security systems.

References

[1] Fenz, S., Ekelhart, A. "Formalizing Information Security Knowledge," in the International Symposium on Information, Computer, and Communications Security (ASIACCS ’09), New York, pp. 183-194.

[2] Gonzalez, C., Ben-Asher, N., Oltramari, A., Lebiere, C. "Cognitive Models of Cyber Situation Awareness and Decision Making," in Cyber Defense and Situational Awareness, A., Wang, C., Erbacher, R. Kott, Ed.: Springer, 2014, vol. 62

[3] Cure O. Merging expressive spatial ontologies using Formal Concept analysis with uncertainty considerations. Methods for Handling Imperfect Spatial Information. Spinger-Verlag. 2010; 256:188–209.

[4] Chen RC, Bau CT, Yeh VJ. Merging domain Ontologies based on the WordNet system and Fuzzy Formal Concept Analysis techniques. Applied Sot Computing. 2011 Mar; 11(2):1908–23.

[5] Ganter B, Stumme G. Creation and merging of Ontology top-levels. International conference on Conceptual structures; 2003 Jul; 2746. p. 131–45.

[6] Quan, T. T., S. C. Hui, A.C.M. Fong, and T.H. Cao. 2006. “Automatic Fuzzy Ontology Generation for Semantic Web.” IEEE Transactions on Knowledge and Data Engineering, 18(6): 842-856.

[7] Побудова нечіткої онтології для аналізу системи захисту інформації в ІТС - О.В. Козленко - Безпека інформації, 2018, ISSN 2225-5036 DOI: https://doi.org/10.18372/2225-5036.24.12973

[8] Yu Fu, Xiao-ping Wu, Qing Ye, Xi Peng, "An Approach for Information Systems Security Risk Assessment on Fuzzy Set and Entropy-Weight," Acta Electronica Sinica, vol.38, no.7,pp.1489-l494, 2010.