

User Perception of Level of Security on Social Networking Sites Using Fuzzy Logic

Eric Afful-Dadzie^{1*}, Arnošt Veselý²

¹Faculty of Applied Informatics, Tomas Bata University, Zlin, Czech Republic

²Department of Information Engineering, CULS, Prague, Czech Republic

ABSTRACT

Human nature often frowns on engaging or interacting with near strangers; however on online social media networks, this is largely ignored. There is an open interaction among both known users and loosely-connected users, and as a result, the normal social barriers against interacting with strangers are lowered. This rather careless openness has resulted in the rampant increase in cybercrime and identity theft worldwide, awaiting a potential privacy disaster in the near future. Since users raise concerns about the privacy and the security of social media sites, there is the need to evaluate the perception that users have of the security on social media sites. This paper presents a technique for evaluating user perception of level of security on social networking sites using fuzzy logic. The inputs to the system were fuzzy sets representing linguistic variables for information security evaluation goals of confidentiality, integrity and availability. The IF-THEN rules were constructed using the Mamdani fuzzy reasoning technique and the defuzzification technique was done using the centroid technique. The implementation of the design was carried out using the MATLAB Fuzzy logic tool box. Using three of the popular online social networking sites namely, Facebook, Twitter and LinkedIn the results show a system that can effectively be employed to evaluate user perception of Information Security.

Keywords: Fuzzy Logic; Social Networking Sites (SNSs); Information Security; User perception

1. INTRODUCTION

Social Networking sites (SNSs) have come to stay and are now an integral part of our lives. In recent years, participation in social networking sites has dramatically increased. Online social media services such as Facebook, Twitter, and LinkedIn allow millions of people to create online profiles and share personal information with vast networks of friends and sometimes unknowingly with strangers. However, while popularity is soaring for these SNSs and millions of users sign onto these sites on a daily basis, there have also been growing concerns about breach of security on these sites. Privacy issues and identity theft on social media sites are a huge concern. This phenomenon is attracting the attention of academic and industry researchers who are intrigued not just by the wide reach of audiences for these social networking sites but the increasing concerns of security risks posed to users.

* Corresponding author: Email: afful@fai.utb.cz

33 1.1. Overview of the Problem

34 As of June 2010, 22 percent of all the time online or one in every four and half minutes spent
35 online was social-through sharing, messaging, commenting, and blogging [1]. It is also
36 interesting to note that for the first time ever, social networks or micro-blogging sites are
37 visited by three quarters of global consumers who go on-line [2]. Brazil leads the world chat
38 with the highest percentage (86%) of internet consumers visiting a social networking site and
39 in the U.S. the total minutes spent on social networking sites has increased eighty-three
40 percent year-over-year. Facebook alone, as one of the major players of SNSs had by
41 September, 2012 reached one billion active users with each active user linked to an average
42 of 305 other users making it the second most visited website on the Internet [3, 4].

43
44 While there is no doubt about the range of opportunities for communication and real-time
45 exchange of all kinds of information offered by social networking sites, there has emerged
46 critical issues of concern about its privacy and security [5]. In most of the SNSs, there is very
47 little protection against copying of personal data from profiles and re-publishing the data
48 elsewhere [6] but one of the most important challenges of information sharing is how to
49 assure its security [7]. For example, in recent times, the reputation of social networking sites
50 has been hit by a number of incidents as reported by various media platforms [8,9,]. It is
51 therefore incumbent on SNSs to have clear policies regarding data protection so as to
52 deliver the same level of social privacy that exists face to face. As it is with any new tool or
53 application, it is always important to keep a close watch on its security especially when
54 majority of the population is actively involved in the use of such tool.

56 1.2. Information Security Evaluation Criteria

57 The widely accepted model or criteria for evaluating information security is the basic CIA
58 triad; standing for Confidentiality, Integrity and Availability. These three key criteria principles
59 are deemed fundamental to guaranteeing security in any information system. These criteria
60 have been applied across the whole subject of Security Analysis, from access to a user's
61 internet history to security of encrypted data across the internet [10]. Therefore the universal
62 classic definition of information security is brief and very simple: Information security is the
63 confidentiality, integrity, and availability of information [11]. By extension, if any one of the
64 three principles are violated or breached, it can have serious consequences for the parties
65 concerned be it an organization or the individual user of an information system. The
66 Information Technology Security Evaluation Criteria (ITSEC), a consortium of Information
67 Security experts from France, Germany, Holland and the United Kingdom, also employ
68 confidentiality, integrity and availability as the yardstick for evaluation of Information
69 Technology security [12]. The relationship among these factors however, has much
70 ambiguity and conflict [13, 14,15] such that it is reasonable and scientific to apply fuzzy
71 comprehensive evaluation method for evaluating security risk in an information technology
72 system such as an online social networking site. The CIA Triad variables are not
73 independent and sometimes oppose themselves as regards their use. For example [15]
74 explains"locking your data in a safe and throwing away the key may help confidentiality and
75 integrity but not availability".

77 1.3 Fuzzy Sets Theory Applications in Information Security Evaluation

78
79 The Fuzzy set theory approach, pioneered by Zadeh [16] was intended to deal with the issue
80 of uncertainties that are not statistical in nature. The approach has been widely used to
81 represent the uncertainties of real-life situations. The decade has witnessed a rapid growth
82 in the number and variety of applications using fuzzy set theory. In the field of computer
83 security, fuzzy set theory was used by [17] to assess quality performance of E-Banking
84 Security system. The research focused on the complex and dynamic nature of the various

factors that are considered in E-banking security assessment. They were convinced that fuzzy logic (FL) model presents an effective tool in assessing and evaluating e-banking security performance and quality. Fuzzy set theory is also applied in assessing online risk for distributed intrusion prediction and prevention systems [18]. The research illustrated how the design of fuzzy logic based on Distributed Intrusion Prediction and Prevention Systems (DIPPS) can be used to effectively assess online risk. Hierarchical Takagi-Sugeno Models is also used for online security evaluation Systems [19] where the risk assessment was carried out using an evolutionary algorithm to automatically design a Hierarchical Takagi-Sugeno fuzzy inference system. The hierarchical structure is evolved using Probabilistic Incremental Program Evolution (PIPE) with specific instructions. Authors [20] further on used a neuro-fuzzy learning method to optimize the performance of fuzzy risk models. The architecture of the developed hierarchical fuzzy inference system was however designed manually. Other works include information security risk assessment method based on Fuzzy Logic [21] and Network information security comprehensive evaluation using interval-valued Fuzzy mathematics [22].

2. METHODOLOGY

The methodological approach and design implementation selected for the evaluation of users' perception of security on online social networking sites by fuzzy set theory was done through the following stages: Figure 1 outlines the procedure.

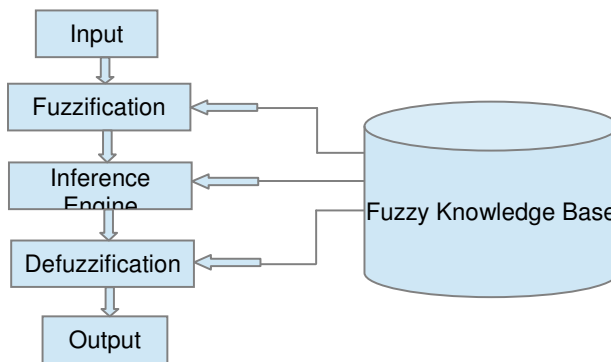


Fig. 1. Input to Output fuzzification-defuzzification process

2.1 Design Model of Linguistic Variables

The inputs to the system were confidentiality, integrity and availability. These criteria or linguistic variables are assumed to be of the same weight and a particular value is determined for each of them based on questions that are answered about a specific social networking site. Designing the fuzzy system requires that the different inputs (that is, confidentiality, integrity, and availability) are represented by fuzzy sets. The fuzzy sets are in turn represented by a membership function. The membership function used in this paper is the triangular membership function which is a three point function defined by minimum (α), maximum (β) and modal (m) values where ($\alpha \leq m \leq \beta$) as shown in figure 2.

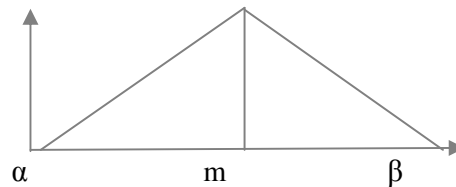


Fig. 2. Triangular Membership Function

2.2 Fuzzy Sets

The values of linguistic variables were represented with fuzzy sets defined by triangular membership functions [20]. The triangular membership function was chosen mainly because of its simplicity and appropriateness for this work [25]. Each linguistic variable takes 5 values considered to be an ideal choice because with more than 5 values the design becomes cumbersome. The level of confidentiality as a linguistic variable was defined on a set of membership functions of *not confidential*, *slightly confidential*, *confidential*, *very confidential* and *extremely confidential*. The level of integrity was also defined based on the scales of *very low*, *low*, *high*, *very high*, and *extremely high* while the level of availability was defined by the scales of *not often*, *rarely often*, *often*, *very often*, and *always available*. The levels defined above were based on a range definition with an estimated interval of [0-10]. The level of security, the output, is defined based on the scales of *not secure*, *slightly secure*, *secure*, *very secure*, and *extremely secure* within the range of [0 - 30]. A web survey on the CIA Triad was put on the three social networking sites for a period of 4 months. Some were also sent through emails to people. In all there were 829 respondents but 18 were incomplete bringing the number to 811. 376 of the respondents were between ages 19-25, 227 between 26-32, and 187 between ages 33-39. The rest were either below age 17 or above age 39. The respondents were from diverse backgrounds of culture and race. Based on the results from the web survey, the interval estimation method was used to define the ranges for each of the membership functions belonging to each of the three inputs and the output.

Table 1. Range of inputs for confidentiality

Name	Fuzzy Number
Not confidential	0,0,2.5
Slightly confidential	0,2.5,5
Confidential	2.5,5,7.5
Very confidential	5,7.5,10
Extremely confidential	7.5,10,10

Table 2. Range of inputs for integrity

Name	Fuzzy Number
Very Low	0,0,2.5
Low	0,2.5,5
High	2.5,5,7.5
Very High	5,7.5,10
Extremely High	7.5,10,10

Table 3. Range of Inputs for availability

Name	Fuzzy Number
Not often	0,0,2.5
Rarely often	0,2.5,5
Often	2.5,5,7.5
Very often	5,7.5,10
Always available	7.5,10,10

169
170

Table 4: Range of outputs for level of security

Name	Fuzzy Number
Not secure	0,0,7.5
Slightly secure	0,7.5,15
Secure	7.5,15,22.5
Very secure	15,22.5,30
Extremely secure	22.5,30,30

171
172
173
174
175

3. DESIGN OF THE FUZZY INFERENCE SYSTEM

3.1 Input variables

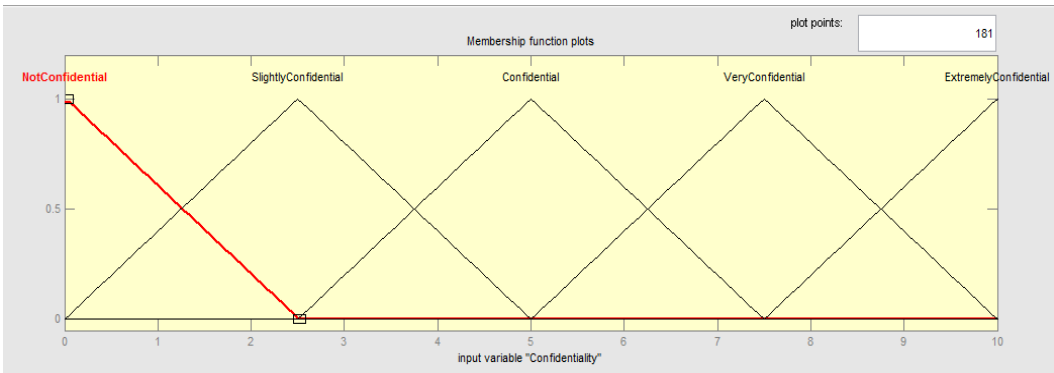
The input variables were the CIA TRIAD of confidentiality, integrity and availability that are used in the evaluation of information technology security. The same CIA criteria were deemed appropriate to be incorporated into a secure social networking site application system.

179
180
181

3.2 Membership functions for input variables

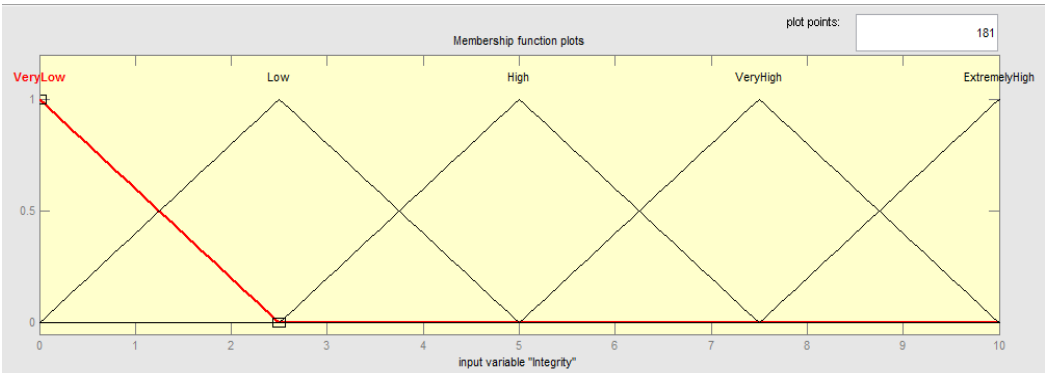
182
183
184
185
186
187

The inputs were defined on a domain interval of [0 - 10], based on the results from the survey using the interval estimation method. Again based on the results, the domain was then divided into 2N + 1 regions and to each region a membership function was attached. In this paper, the domain was divided into 5 regions (N =2). The regions were represented by triangular membership functions as exhibited in MATLAB below.



188

Fig. 3. Membership function for confidentiality



189

Fig. 4. Membership function for Integrity

* Corresponding author: Email: afful@fai.utb.cz

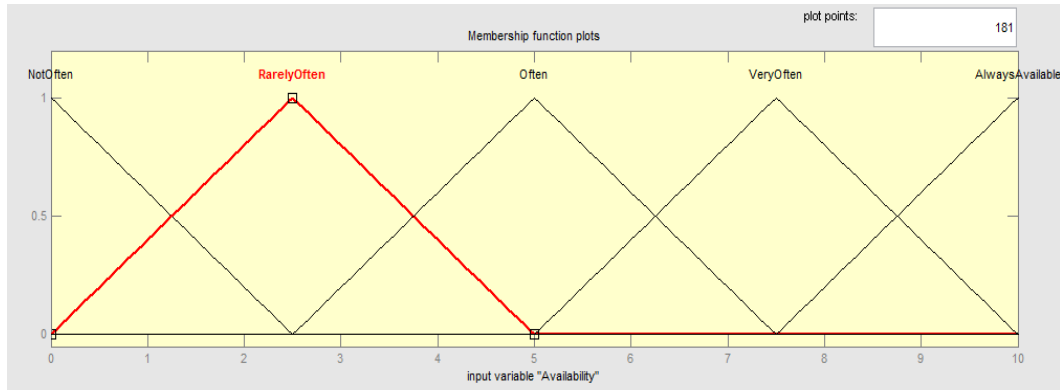


Fig. 5. Membership function for Availability

3.3. Membership functions for Output Variables

The output domain interval was estimated to be [0 - 30]. The domain interval was further divided into $2N + 1$ regions and to each region, a membership function was attached. For example level of security (the output) is divided into 5 regions ($N = 2$) represented by *not secure*, *slightly secure*, *secure*, *very secure*, and *extremely secure* as the fuzzy sets.

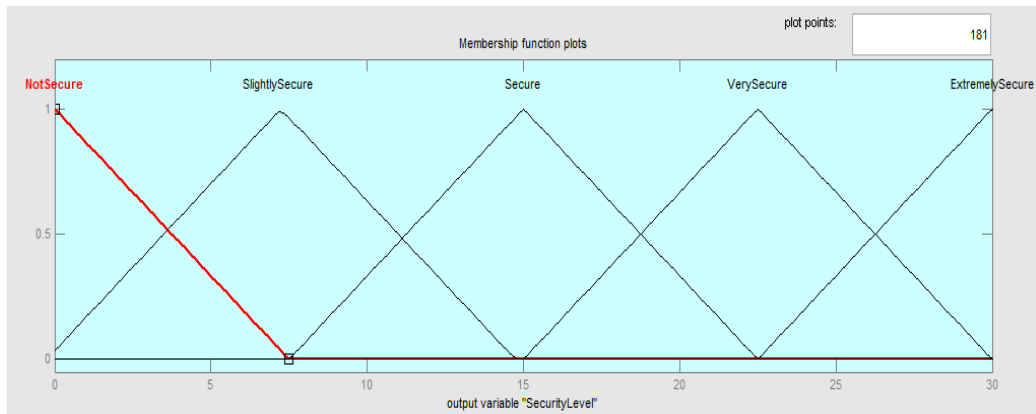


Fig. 6. Membership function Level of Security

3.4 Formulating rules and populating the rule base.

The rules were built based on intuitive knowledge of the relationships between the variables. The rules were formulated so as to reflect the relationships between any possible relations of the input variables to the output variable. The rules in this work reflected the relationships between the levels of confidentiality, integrity and availability and also the level of security. To determine the overall security risk level for each social networking site, the rule base needs $5^3 = 125$ rules since there were five linguistic values and three linguistic variables (Confidentiality, Integrity and availability). A sample of the rule base used to construct the overall knowledge base is shown in Table 5 for different linguistic values. The levels of Confidentiality, Integrity, and Availability were used in the antecedent of rules and the level of security as the consequent of rules. A fuzzy rule is a conditional statement in the form: IF X_1 is A_1 and X_2 is A_2 and X_3 is A_3 THEN y is B where X_1 , X_2 , X_3 and y are linguistic variables and A_1 , A_2 , A_3 and B are linguistic values determined by fuzzy sets on universe of discourses X_1 , X_2 , X_3 and Y .

214

Table 5. The Sample of the Rule base

RULE r	IF Confidentiality is	AND Integrity is	AND Availability is	THEN Security Level is
1	Not Confidential	Very Low	Not Often	Not Secure
2	Not Confidential	Very Low	Rarely Often	Not Secure
3	Not Confidential	Very Low	Often	Not Secure
4	Not Confidential	Very Low	Very Often	Slightly Secure
5	Not Confidential	Very Low	Always Available	Slightly Secure
6	Not Confidential	Low	Not Often	Slightly Secures
7	Not Confidential	Low	Rarely Often	Slightly Secure
8	Not Confidential	Low	Often	Slightly Secure
9	Not Confidential	Low	Very Often	Slightly Secure
10	Not Confidential	Low	Always Available	Slightly Secure
11	Not Confidential	High	Not Often	Slightly Secure
12	Not Confidential	High	Rarely Often	Secure
13	Not Confidential	High	Often	Secure
14	Not Confidential	High	Very Often	Secure
15	Not Confidential	High	Always Available	Secure

215

216 4. APPLICATION OF FUZZY TECHNIQUE

217

218 In order to construct a fuzzy rule-based assessment for the evaluation of users' perception of
 219 security risk on social networking sites, an online questionnaire was designed mainly based
 220 on the CIA triad of Confidentiality, Integrity and Availability. In all there were seven social
 221 network security based questions under each Linguistic variable to help form users'
 222 perception of security on each of the selected social networking sites which were Facebook,
 223 Twitter and LinkedIn respectively. The link for the survey was pasted on all the three social
 224 networking sites for users to respond. Other users were emailed with the link to respond to
 225 the survey. The survey questionnaire was online for approximately three months. In
 226 representing users' perceptions as a fuzzy membership function, the interval estimation
 227 method was used. The interval estimation generates more suitable results for continuous
 228 measurements. Participants understand and represent their opinions more easily using
 229 interval estimation. Often an interval estimation method for constructing fuzzy membership
 230 functions is the most appropriate and is commonly used [23, 24]

231

232 4.1 Fuzzy Aggregation using Weighted Average.

233

234 One of the most common aggregation operator often found in literature is the weighted
 235 average (WA) also known as the weighted mean. It is similar to an arithmetic mean where
 236 instead of each of data points contributing equally to the final average, some data points
 237 contribute more than others. There are weighted versions of other means such as the
 238 weighted geometric mean (WGM) and the weighted harmonic mean (WHM). There is also
 239 the ordering weighted average (OWA).

240

241

242

243

244

245 4.2 Triangular Fuzzy Numbers (TFN)

246 A triangular fuzzy number X (TFN X) having minimum value X1, modal value X2 and
247 maximum value X3 is written as (X1, X2, X3).

248 Let TFN X = (X1, X2, X3) (1)

249 And TFN Y = (Y1, Y2, Y3) (2)

250 Then the sum of X and Y is (X1+Y1, X2+Y2, X3+Y3) (3)

251 In this paper, respondents were to choose between a series of statements on the
252 ordinal/interval scale the one they judge most appropriate and it is argued that the choice of
253 score is, in effect, a judgement between 3 indicator statements. Thus, for example,
254 respondents rate the level of confidentiality of the dating history or intimate secrets they
255 submit with friends on social media on the following scale:

Name	Fuzzy number	
■ Not Confidential	0, 0, 2.5	▼
■ Slightly Confidential	0, 2.5, 5	▼
■ Confidential	2.5, 5, 7.5	▼
■ Very Confidential	5, 7.5, 10	▼
■ Extremely Confidential	7.5, 10, 10	▼

256 Fig. 7. Inputs for confidentiality

257
258 In this interpretation, a respondent who judges “very confidential” to be the appropriate score
259 makes a constrained choice in the range where 5 is the minimum value, 7.5 is the modal
260 value and 10 the maximum. (To think of it another way, respondents must consider which of
261 the five hypotheses, *Not confidential*, *Slightly confidential*, *Confidential*, *Very confidential* and
262 *Extremely confidential* best represent their judgement of the situation.) In extracting the fuzzy
263 scores on a range of 10, the descriptor “*Not confidential*” corresponds to a triangular fuzzy
264 number (0, 0, 2.5). Similarly, the descriptor “*very confidential*” also corresponds to (5, 7.5,
265 10), and so on.

275 4.3 Evaluation using weighted average

276 To summarize the answers modeled by the fuzzy numbers put to many users, the weighted
277 average was used. For summarizing the answers that were put to n users we used the
278 average of fuzzy numbers representing users’ answers.

$$280 (s_1, s_2, s_3) = \frac{1}{n} \sum_{i=1}^n (X_1^i, X_2^i, X_3^i) \quad (4)$$

281 It is clear that we may also consider this average to be a weighted sum of the obtained
282 answers. For example, assuming that n users responded to a question concerning
283 confidentiality and that:

- 284 n_1 users answered Not confidential (represented with fuzzy number(0,0,2.5))
- 285 n_2 users answered Slightly confidential (represented with fuzzy number(0,2.5,5))
- 286 n_3 users answered Confidential (represented with fuzzy number(2.5,5,7.5))
- 287 n_4 users answered Very confidential (represented with fuzzy number(5,7.5,10))
- 288 n_5 users answered Not confidential (represented with fuzzy number(7.5,10,10))

289 Then the summarized weighted average for such a question would be the fuzzy number:

$$290 (s_1, s_2, s_3) = \frac{1}{n} (n_1(0,0,2.5) + n_2(0,2.5,5) + n_3(2.5,5,7.5) + n_4(5,7.5,10) + n_5(7.5,10,10)) \quad (5)$$

291

4.4 Center of gravity

To determine perceived confidentiality, integrity and availability, users responded to 7 different questions (see Appendix A). Further on, seven summarized fuzzy numbers are derived as shown in Table 6 and subsequently in Table 7 and 8. To represent these 7 fuzzy numbers by crisp value, the centroid method was used in this instance. The center of gravity was evaluated as follows:

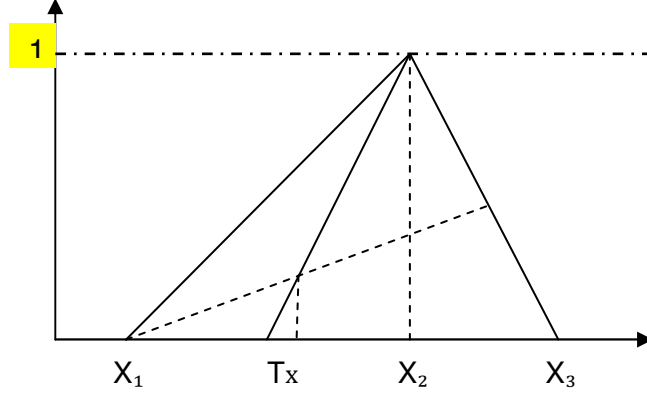


Fig. 8. Center of gravity T_x of one triangle (X_1, X_2, X_3)

$$T_x = \frac{(X_1 + X_2 + X_3)}{3}$$

(6)

For center of gravity T_x of 2 triangles (X_1, X_2, X_3) and (Y_1, Y_2, Y_3) we get

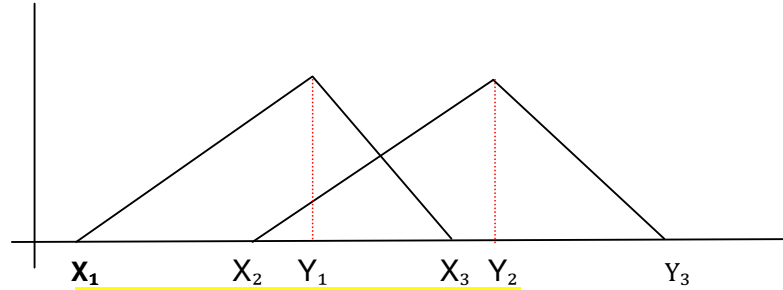


Fig. 9. Center of gravity of 2 triangles

T_x is the center of gravity of (X_1, X_2, X_3) and

T_y is the center of gravity of (Y_1, Y_2, Y_3)

$$T_c \left((X_3 - X_1) \frac{1}{2} + (Y_3 - Y_1) \frac{1}{2} \right) = T_x (X_3 - X_1) \frac{1}{2} + T_y (Y_3 - Y_1) \frac{1}{2} \quad (7)$$

$$T_c = \frac{1}{((X_3 - X_1) + (Y_3 - Y_1))} [T_x (X_3 - X_1) + T_y (Y_3 - Y_1)] \quad (8)$$

It can also similarly be done as in the case of 2 triangles with 7 triangles representing the 7 sets of questions under each of the linguistic variable as below:

Center of gravity T_c of 7 triangles $(X_1^1, X_2^1, X_3^1), \dots, (X_1^7, X_2^7, X_3^7)$

338 The center of gravity T_c of fuzzy numbers $(X_1^1, X_2^1, X_3^1), \dots (X_1^7, X_2^7, X_3^7)$ can be
 339 evaluated as:

$$340 \quad T_c \sum_{i=1}^7 (X_3^i - X_1^i) = \sum_{i=1}^7 T_i (X_3^i - X_1^i) \quad (9)$$

$$341 \quad T_c = \frac{\sum_{i=1}^7 T_i (X_3^i - X_1^i)}{\sum_{i=1}^7 (X_3^i - X_1^i)} \quad (10)$$

342 where

343 T_i ; $i=1, \dots, 7$ are centres of gravity

344
 345 This method of aggregation is used for all the three linguistic variables (confidentiality,
 346 integrity and availability). The three inputs representing the three variables are then fed into
 347 the fuzzy logic tool box to generate the appropriate output of the level of security on each
 348 social networking site namely Facebook, Twitter and LinkedIn.

349
 350 For illustration purposes, the table results for the linguistic variables; confidentiality, integrity
 351 and availability for Facebook is shown below. The same was done for the Twitter and
 352 LinkedIn.

353
 354 Table 6. Aggregation of responses of Confidentiality on Facebook
 355 **Question:** *In general, how would you rate the level of confidentiality of the following*
 356 *information you submit with friends on Facebook.*
 357

Information	Not confidential	Slightly confidential	Confidential	Very confidential	Extremely confidential	Weighted average
Dating history	0	2	135	81	3	3.46,5.96,8.43
Financial info (things you buy, where you buy from)	4	4	94	118	1	3.77,6.22,8.71
Gossips between friends	4	156	59	1	1	0.71,3.18,5.67
Intimate secrets	3	42	72	103	1	3.18,5.64,8.13
Lifestyle related (eg. Photos, blogs, history)	36	105	55	24	1	1.20,3.30,5.78
Professional/work related information	33	80	79	28	1	1.56,3.70,6.18
Religious/Political beliefs	77	76	3	64	1	1.51,3.14,5.63

358 Center of Gravity $T_c = 4.52$

359
 360
 361
 362
 363

364 Table 7. Aggregation of responses of Integrity on Facebook
365 **Question:** *Rate the level of authenticity of the origin of the message and files you receive*
366 *from the following people on Facebook*
367

Information	Very Low	Low	High	Very High	Extremely High	Weighted Average
Close Friends	2	2	52	160	5	4.37,6.85,9.30
Co-workers	1	56	10	148	6	3.67,6.23,8.58
Family members	1	16	75	126	3	3.80,6.28,8.76
Friends	1	48	27	139	6	3.65,6.14,8.57
People who live far away from you	1	52	109	46	13	2.71,5.20,7.55
Strangers (people you have never met before)	3	60	10	127	21	3.70,6.16,8.42
Friends of your friends	2	2	114	102	1	3.63,6.10,8.60

368 Center of Gravity $T_c = 6.10$

369 Table 8. Aggregation of responses of Availability on Facebook
370 **Question:** *How often do you have ready access to the following information on Facebook?*

Information	Not Often	Rarely often	Often	Very often	Always Available	Weighted average
Message history	1	3	6	43	168	6.74,9.23,9.76
Chat history	2	2	36	102	79	7.02,9.51,9.93
The website itself	1	1	1	34	184	6.74,9.23,9.76
Intimate secrets	1	2	1	1	216	7.36,9.85,9.90
Lifestyle related (photos, blogs)	1	1	34	71	144	5.85,8.35,9.56
Professional/ Work related information	3	2	10	105	101	5.91,8.38,9.74
Profile Information	1	1	1	41	177	6.95,9.34,9.93

371 Center of gravity $T_c = 8.53$

372 4.5 Summary of Results

373 Table 9: Final value of aggregated responses

	Confidentiality	Integrity	Availability
Facebook	4.52	6.10	8.53
Twitter	6.30	7.42	8.21
LinkedIn	3.98	2.70	7.70

376 4.6 Implementation Procedure in MATLAB

377 The final result for each of the linguistic inputs derived after aggregating the responses from
378 the well-constructed online social networking sites security questions were fed into MATLAB
379
380

to derive the final output (security risk level) for each of the three selected social media sites. The inputs were supplied through the graphical user interface called rule viewer.

4.6.1 MATLAB Fuzzy Inference System (FIS) Editor

The fuzzy inference system editor in figure 7 shows the summary of the fuzzy inference system. In the editor, there is shown the mapping of the inputs to the system and to the output. The input variables were respectively *confidentiality*, *integrity* and *availability*. The output was *security level* while the rules were constructed using the Mamdani fuzzy reasoning and the defuzzification technique was done using the centroid technique. The mamdani method was chosen over Sugeno it well suitable for human input like this research and generally has broad acceptance and applicability[25].

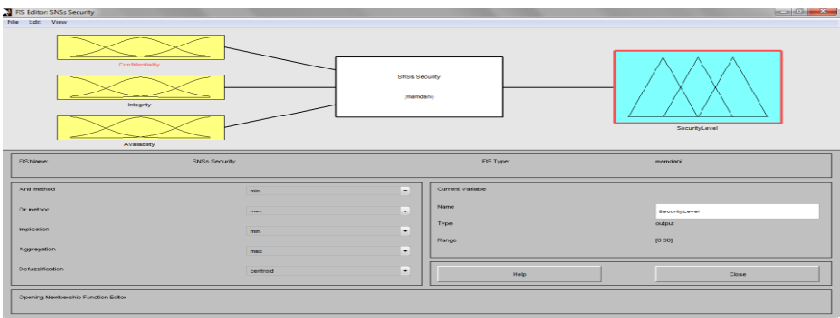


Fig. 10. MATLAB FIS editor

4.6.2 The MATLAB Rule Viewer

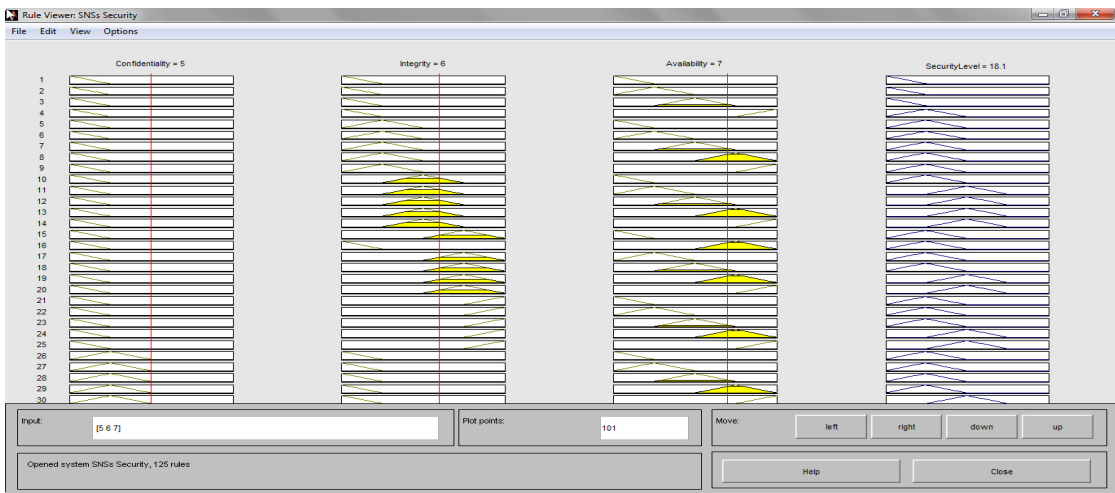


Fig. 11. MATLAB Rule Viewer

The three input variables (confidentiality, integrity and availability) were fed into the system. The appropriate input corresponds to the weighted averages of the user responses in the questionnaire for each of the input variables followed appropriately by their center of gravity. For example, in the figure 8, the input values for the variables confidentiality, integrity and availability are respectively [5, 6, 7] and the corresponding output (security level) is 18.1, which is shown at the top of the corresponding graphs. The input for each of the input variables is specified at the top of the section corresponding to them, so also is the output variable.

* Corresponding author: Email: afful@fai.utb.cz

4.6.3 The Surface Viewer

The MATLAB surface viewer as shown in figure 12 is a 3-D graph that shows the relationship between the inputs and the output. The output (security level) is represented on the Z-axis while 2 of the inputs (Confidentiality and Integrity) are on the x and y axes and the other input (Availability) is held constant. The surface viewer shows a plot of the possible ranges of the input variables against the possible ranges of the output.

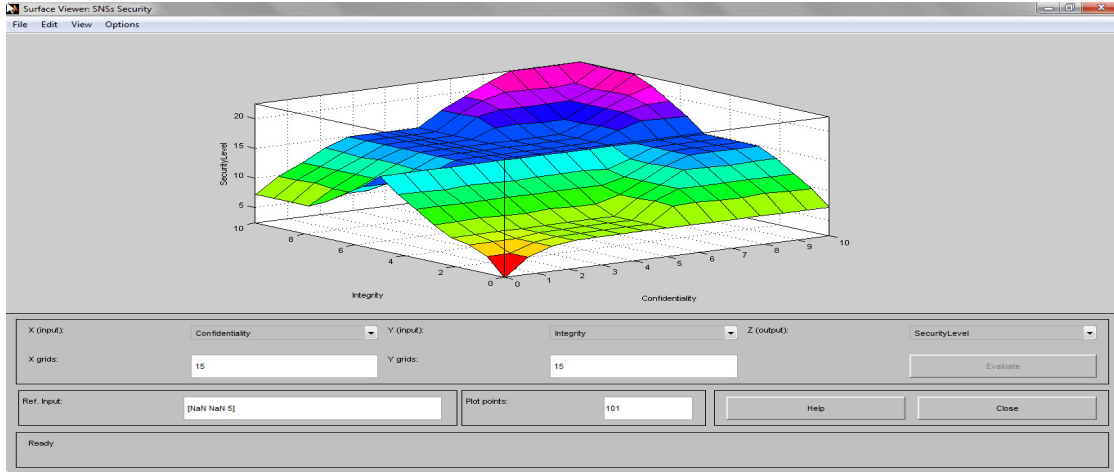


Fig. 12. MATLAB Surface Viewer

5. EVALUATION

The output from the fuzzy system is a crisp number whose value is not intuitive. We therefore interpreted it also in terms of used notions *not secure*, *slightly secure*, *secure*, *very secure*, and *extremely secure*. These notions were modeled with fuzzy sets and then evaluated the membership functions of the fuzzy sets that represent these notions. For example linguistic variable value *secure* is represented with the triangular membership function (7.5, 15, 22.5). It can be seen that for crisp value x the following holds:

$$\text{a) if } x \text{ is greater or equal to 15, then the value of the membership function is } \text{secure}(x) = (22.5 - x) / (22.5 - 15) = (22.5 - x) / 7.5 \quad (11)$$

$$\text{b) if } x \text{ is less than 15, then value of membership function is } \text{secure}(x) = (x - 7.5) / (15 - 7.5) = (x - 7.5) / 7.5 \quad (12)$$

5.1 User Security Perception of Facebook

Facebook was rated according to user responses with 4.52 as the score for confidentiality, 6.10 for integrity and 8.53 as the score for availability. This produced a crisp output of 18.4 representing the security level out of the set range of 30. This value of 18.4 shows that Facebook is 55% secure and 45% very secure.

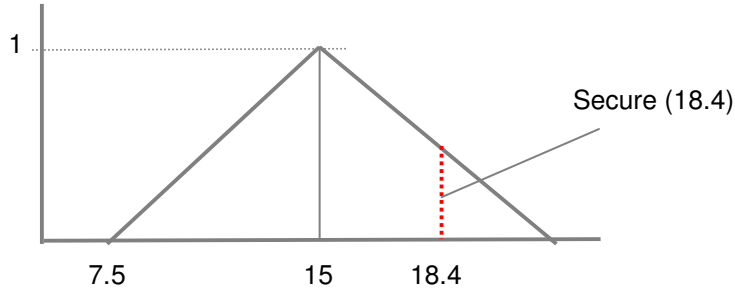


Fig. 13 Facebook: Membership function *secure*(X)

$$\frac{1}{\text{secure}(18.4)} = \frac{(22.5 - 15)}{(22.5 - 18.4)} = \frac{7.5}{4.1} \longrightarrow \text{secure}(18.4) = \frac{4.1}{7.5} = 0.55$$

(13)

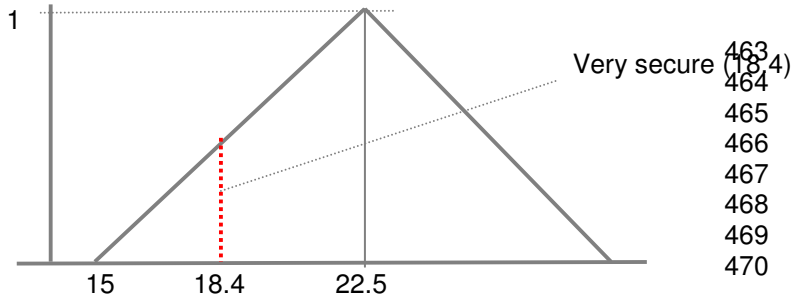


Fig. 14 Facebook: Membership function *very secure*(x)

$$\frac{1}{\text{Verysecure}(18.4)} = \frac{(22.5 - 15)}{(18.4 - 15)} = \frac{7.5}{3.4} \longrightarrow \text{Verysecure}(18.4) = \frac{3.4}{7.5} = 0.45$$

(14)

5.1 User Security Perception of Twitter

Twitter scored inputs of 6.3, 7.42 and 8.21 for confidentiality, integrity and availability respectively. The crisp output was 22.4. This value corresponded to 1% secure and 99% very secure.

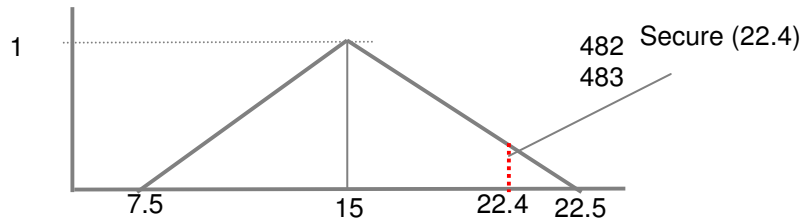


Fig. 15 Twitter: Membership function *secure*(x)

$$\frac{1}{\text{secure}(22.4)} = \frac{(22.5 - 15)}{(22.5 - 22.4)} = \frac{7.5}{0.1} = 75 \longrightarrow \text{secure}(22.4) = \frac{1}{75} = 0.01$$

(15)

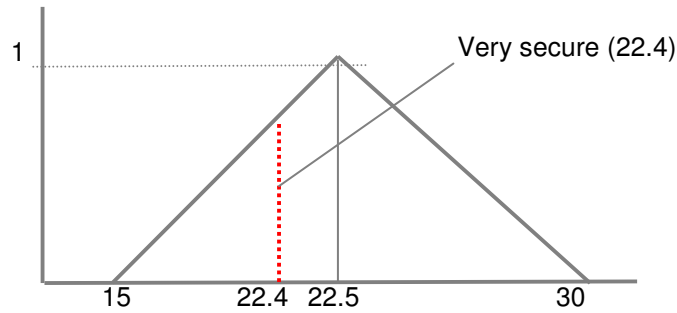


Fig. 16 Twitter: Membership function *very_secure(x)*

$$\frac{1}{\text{Verysecure}(22.4)} = \frac{(22.5 - 15)}{(22.4 - 15)} = \frac{7.5}{7.4} \longrightarrow \text{Verysecure}(22.4) = \frac{7.4}{7.5} = 0.99 \quad (16)$$

5.1 User Security Perception of LinkedIn

LinkedIn was judged with scores of inputs of 3.98, 2.70 and 7.70 for confidentiality, integrity and availability respectively. The crisp output was 14.9. This value corresponded to 1% slightly secure and 99% secure.

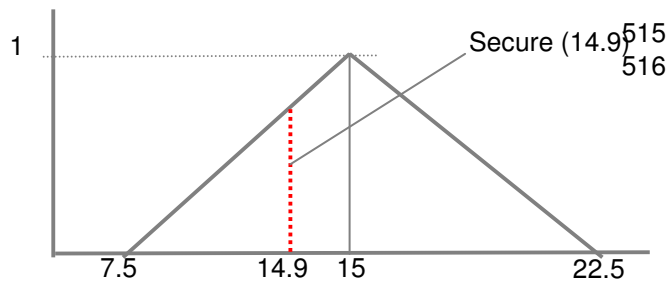


Fig. 17 LinkedIn: Membership function *secure(x)*

$$\frac{1}{\text{secure}(14.9)} = \frac{(15 - 7.5)}{(14.9 - 7.5)} = \frac{7.5}{7.4} \longrightarrow \text{secure}(14.9) = \frac{7.4}{7.5} = 0.99 \quad (17)$$

Table 10. Evaluation of Variables

SNSs	Variable Inputs	Crisp Output	Slightly Secure	Secure	Very Secure
Facebook	[4.52, 6.10, 8.53]	18.4	0	0.55	0.45
Twitter	[6.30, 7.42, 8.21]	22.4	0	0.01	0.99
Linkedin	[3.98, 2.70, 7.70]	14.9	0.01	0.99	-

The research finds that users consider Twitter *very secure* and LinkedIn *secure*. Facebook lies between secure and *very secure* on a range of level of security. In can be inferred therefore, from the research that, Twitter seems to meet the expectations of users of a more

robust security than the others. While the focus of the paper was to develop an appropriate methodology that captures the views of users to be incorporated into future security designs aimed at improving security on social media, the ranking on who best meets users' expectations on security on social media sites could also in part, help instill some healthy competition among the media sites even though admittedly, Facebook has more active users than the rest of the media.

6. CONCLUSION

Users are by far the main building block of any online social networking site and therefore their security and privacy should be of utmost concern to managers of these social media sites. One user complaint, user perception, is an important element when considering the concepts of social networking security. The perception processes of humans cannot be analyzed and assessed by a binary approach or in a simple quantitative way. The human thought process is subjective, imprecise and complicated, and human perception usually uses a linguistic approach, as opposed to a numerical approach, to classify, describe, or "value" a system.

In addition, user perception of security risk on social networking site is solely affected by an individual evaluator's needs and requirements of what would make him or her secured on a social networking site. In this paper, a fuzzy system was implemented using fuzzy logic theory to evaluate user perception of security risk on social networking sites. Facebook, Twitter and LinkedIn were used as case studies for this research. Employing MATLAB and its associated fuzzy logic toolbox to design the Fuzzy Inference System, an overall user perception of security risk on SNSs were realized.

ACKNOWLEDGEMENT

We would like to express our profound gratitude to all those who took time to respond to the survey questionnaire helping to provide the data to develop the model for evaluating users' perception of security on social networking sites.

COMPETING INTERESTS

Authors hereby declare that no competing interests exist as far as this paper is concerned.

7. REFERENCES

- [1] Nielsen Online. Social Networks/Blogs Now Account for One in Every Four and a Half Minutes Online. 2010;05:06. Accessed 9 September 2012. Available: <http://blog.nielsen.com/nielsenwire/global/social-media-accounts-for-22-percent-of-time-online/>
- [2] Nielson Online news Release. Time Spent on Facebook up 700 Percent. 2009; 02:06 Accessed 13 September 2011. Available: http://www.nielsen-online.com/pr/pr_090602.pdf
- [3] Facebook. Statistics-Facebook. 2012;15:09. Accessed 8 July 2012. Available: <http://newsroom.fb.com/Key-Facts>

- 578 [4] Alexa-Web Information Company. Facebook.com Site Info. 2011;08:08. Accessed 7
579 September 2012. Available: <http://www.alexa.com/siteinfo/facebook.com>
- 580 [5] Donath J. Signals in Social Supernets. Journal of Computer-Mediated Communication
581 2007;13 (1):231–251.
- 582 [6] Boyd D. Facebook's Privacy Trainwreck: Exposure, invasion, and social convergence.
583 The International Journal of Research into New Media Technologies 2008a; 14 (1):13–
584 20. DOI: 10.1177/1354856507084416
- 585 [7] Srivatsa M, Agrawal D, Reidt S. A Metadata Calculus for Secure Information sharing.
586 Proceedings of the 16th ACM conference on Computer and Communications Security
- 587 [8] Chiaramonte P, Martinez E. Jerks In Space. The New York Post, pp. 6. 2006.
- 588 [9] Hass N. In Your Facebook.com. The New York Times, : 30-31. 2006;08:01. Accessed 10
589 July 2012. Available: [http://www.nytimes.com/2006/01/08/education/edlife/facebook.html?](http://www.nytimes.com/2006/01/08/education/edlife/facebook.html?pagewanted=all)
590 [pagewanted=all](http://www.nytimes.com/2006/01/08/education/edlife/facebook.html?pagewanted=all)
- 591 [10] Bhaiji Y. Network Security Technologies and Solutions. A comprehensive all-in-one
592 reference for Cisco network security. 1st ed. Cisco Press; 2008.
- 593 [12] Dahal K. Loan Risk Analyzer based on Fuzzy Logic. Proceedings of IEEE International
594 Conference on E-Technology, E-Commerce and E-Service. 2005:363-366
- 595 [12] ITSEC. Information Technology Security Evaluation Criteria. Department of Trade and
596 Industry, London, 1991;12:06. Accessed 28 November 2012. Available:
597 https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Zertifizierung/ITSicherheitskriterien/itsec-en_pdf.pdf?__blob=publicationFile
598
- 599 [13] Wilson K. Conflicts Among the Pillars of Information Assurance. IT Professional, IEEE
600 Computer Society press, 2012. 14(6). DOI: 10.1109/MITP.2012.24
- 601 [14] Conrad E, Misenar S, Eldman J. CISSP Study Guide: Information and Security
602 Governance and Risk Management. 1st ed. Elsevier; 2010.
- 603 [15] CSEC. IT Security Risk Management: A lifecycle Approach. Annex 4-Profile 1. Nov;
604 2012.
- 605 [16] Zadeh L. A. Fuzzy Sets. Information and Control, 1965;8(3):338-353
- 606 [17] Hossain M.A, Thabatah F, Dahal K. Intelligent Quality Performance Assessment for E-
607 Banking Security using Fuzzy Logic. Fifth International Conference on Information
608 Technology: New Generations. 2008: 420 - 425. DOI: 10.1109/ITNG.2008.154
- 609 [18] Abraham A, Knapskog S. Fuzzy online risk assessment for distributed intrusion
610 prediction and prevention systems. Tenth International Conference on Modeling and
611 Simulation, IEEE Computer Society press.2008: 1-12. DOI: 10.1109/UKSIM.2008.30
- 612 [19] Abraham A, Grosan C, Liu H, Chen Y. Hierarchical Takagi-Sugeno Models for Online
613 Security Evaluation Systems, In Fifth International Conference on Information
614 Assurance and Security, IEEE Computer Society press. 2009: 687-692. DOI:
615 10.1109/IAS.2009.348

- 616 [20] Jang J.S, Sun C.T. Neuro-Fuzzy and Soft Computing. A Computational Approach to
617 Learning and Machine Intelligence. 1st ed. Prentice Hall; 1997.
- 618 [21] Yang Y, Zhou Y. Fuzzy Logic based Method for Network Information Security Risk
619 Assessment. Conference on Internet Technology and Applications (iTAP), 2011: 1 - 4
- 620
- 621 [22] Chen D, Li, X. A Comprehensive Evaluation Model of Network Information Security
622 Based on Interval-valued Fuzzy Mathematics. International Conference on Computer
623 Science and Service System. 2012: 777 – 780. DOI: 10.1109/CSSS.2012.199
- 624 [23] Juang C. H, Huang X.H, Shiff S.D. Determination of weight criteria for decision making
625 by the fuzzy eigenvector method. Civil Engineering System, 1992; 9(1): 1-16.
- 626 [24] Lee D, Pietrucha M.T, Sinha S.K. Application of Fuzzy Logic to Evaluate Driver
627 Perception of Variable Message Signs. Transportation Research Record 1937 TRB,
628 National Research Council, Washington, D.C. 2005; Vol. 1937:96-104.
- 629 [25] Sivanandam S.N, Sumathi S, Deepa S.N. Introduction to Fuzzy Logic using MATLAB.
630 1st ed. Springer;2007

631 Appendix A

632 Questionnaire for evaluating Confidentiality

633 **Question:** In general, how would you rate the level of confidentiality of the following
634 information you submit with friends on (Facebook, Twitter,
635 LinkedIn)?
636
637

Information	Not confidential	Slightly confidential	Confidential	Very confidential	Extremely confidential
Dating history					
Financial info (things you buy, where you buy from)					
Gossips between friends					
Intimate secrets					
Lifestyle related (eg. Photos, blogs, history)					
Professional/work related information					
Religious/Political belief					

649
650
651

Questionnaire for evaluating Integrity

Question: Rate the level of authenticity of the origin of the message and files you receive from the following people on (Facebook, Twitter, LinkedIn)

Information	Very Low	Low	High	Very High	Extremely High
Close Friends					
Co-workers					
Family members					
Friends					
People who live far away from you					
Strangers (people you have never met before)					
Friends of your friends					

652
653
654
655
656
657

Questionnaire for evaluating Availability

Question: How often do you have ready access to the following information on your chosen (Facebook, Twitter, LinkedIn)?

Information	Not Often	Rarely often	Often	Very often	Always Available
Message history					
Chat history					
The website itself					
Intimate secrets					
Lifestyle related (photos, blogs)					
Professional/ Work related information					
Profile Information					

658