British Journal of Applied Science & Technology
3(4): 714-734, 2013

SCIENCEDOMAIN *international*
www.sciencedomain.org

User Perception of Security on Social Networking Sites Using Fuzzy Logic

Eric Afful-Dadzie[1*] and Arnošt Veselý[2]

[1]Faculty of Applied Informatics, Tomas Bata University, Zlin, Czech Republic.
[2]Department of Information Engineering, CULS, Prague, Czech Republic.

Authors' contributions

This work was carried out in collaboration between all authors. Author EAD designed the study, performed the analysis, managed the literature searches and wrote the first draft of the manuscript with guidance from author AV. Author EAD designed the mathematical models together with author AV. All authors read and approved the final manuscript.

Research Article

Received 15th February 2013
Accepted 28th March 2013
Published 2nd May 2013

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.

**Corresponding author: E-mail: afful@fai.utb.cz;*

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, whiles popularity is soaring for these SNSs and millions of users sign onto these sites on a daily basis, there are also growing concerns about breach of security on these sites. Privacy issues and identity theft on social media sites are huge concerns. The 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.

1.1 Overview of the Problem

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

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

1.2 Information Security Evaluation Criteria

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

1.3 Fuzzy Sets Theory Applications in Information Security Evaluation

The Fuzzy set theory approach, pioneered by Zadeh [16] was intended to deal with the issue of uncertainties that are not statistical in nature. The approach has been widely used to represent the uncertainties of real-life situations. The decade has witnessed rapid growth in the number and variety of applications using fuzzy set theory. In the field of computer security, fuzzy set theory was used by [17] to assess quality performance of E-Banking 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. METHODOLOGIES

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: Fig. 1 outlines the procedure.

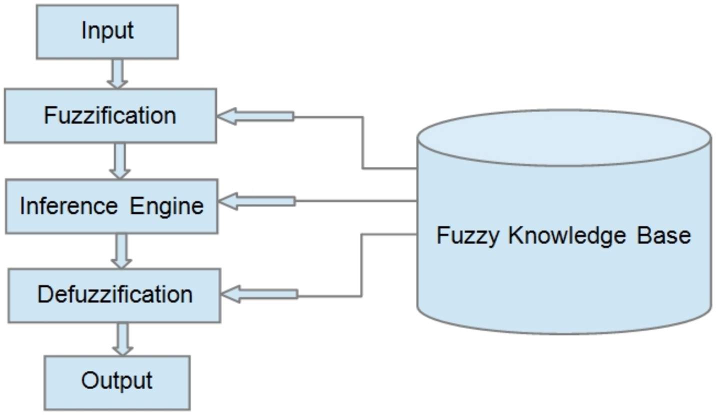


Fig. 1. Input to output fuzzification-defuzzification process

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 (α ≦ m≦ β) as shown in Fig. 2.

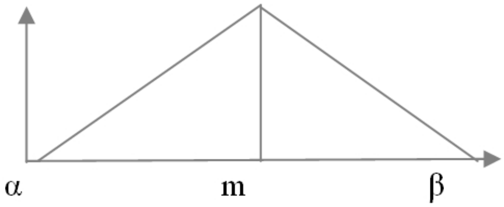


Fig. 2. Triangular membership function

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* whiles 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

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

3. DESIGN OF THE FUZZY INFERENCE SYSTEM

3.1 Input Variables

The input variables were the CIA TRIAD of confidentiality, integrity and availability which 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. Tables 1,2,3 and 4 show the ranges for each membership function for the input variables respectively.

3.2 Membership Functions for Input Variables

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 shown in Figs. 3, 4 and 5 respectively for confidentiality, integrity and availability in MATLAB.

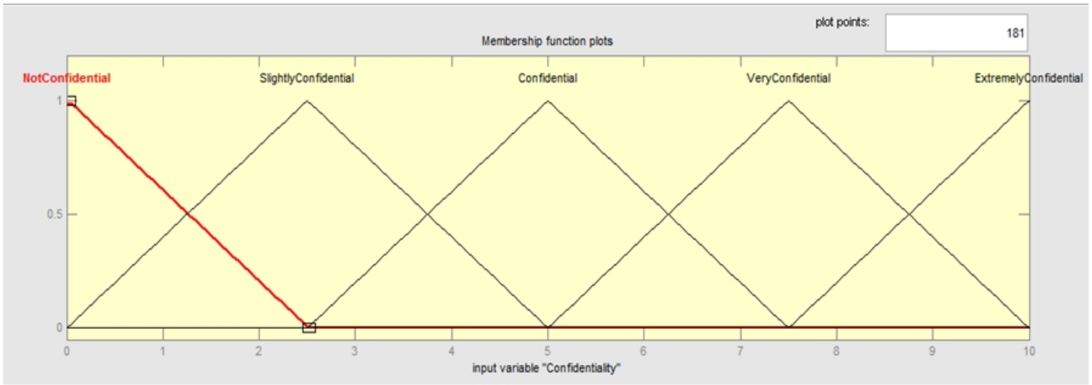


Fig. 3. Membership function for confidentiality

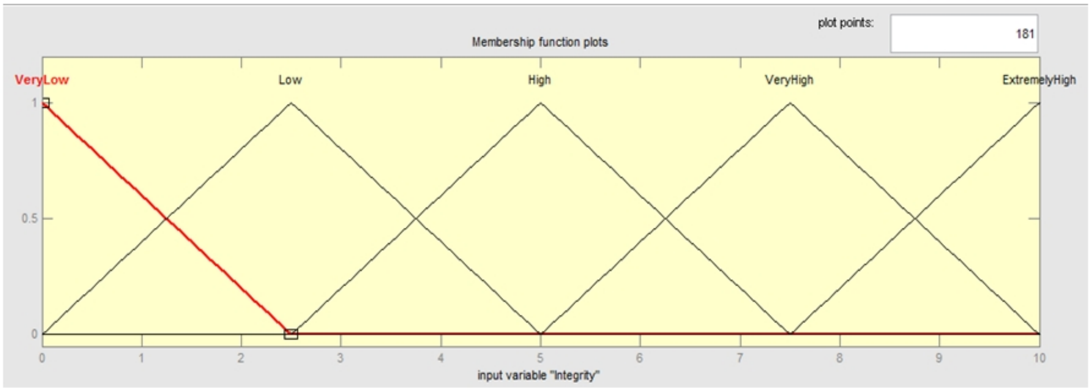


Fig. 4. Membership function for integrity

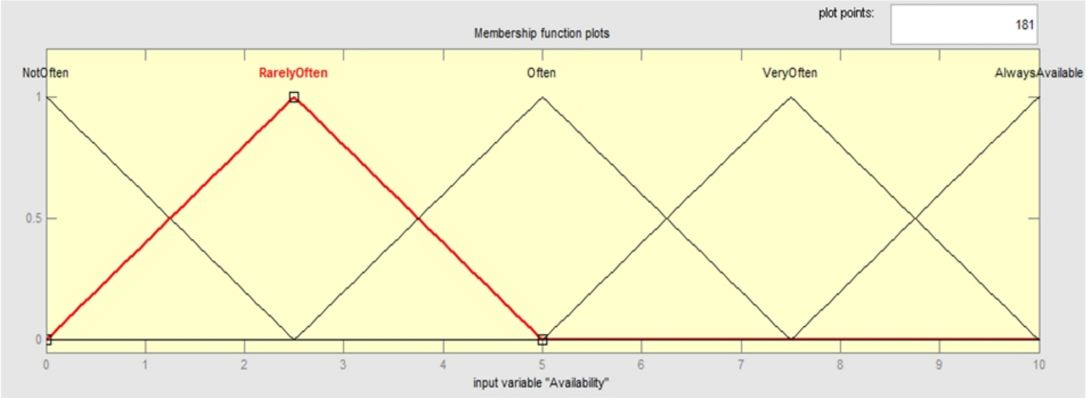


Fig. 5. Membership function for availability

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. The 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 shows the triangular membership functions for the output variable as modeled in MATLAB.

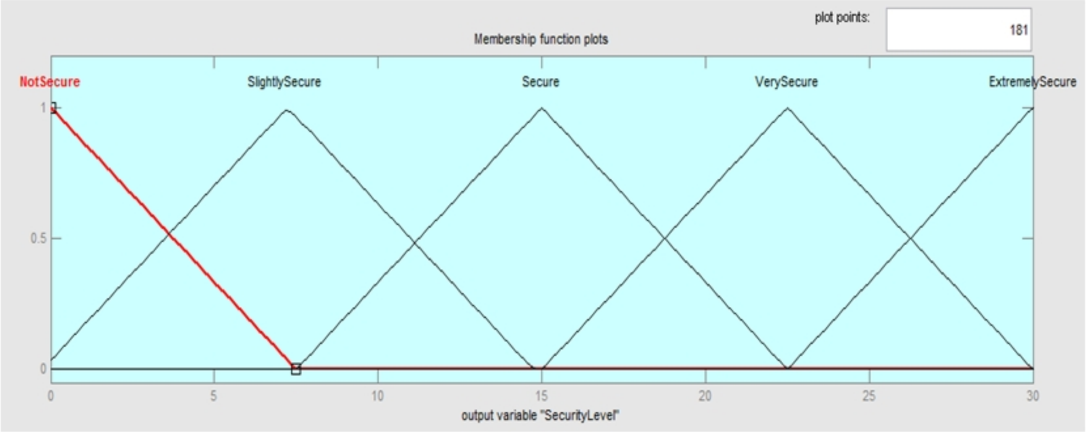


Fig. 6. Membership function level of security

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 among the levels of confidentiality, integrity and availability to the level of security. To determine the overall security 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.

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

4. APPLICATION OF FUZZY TECHNIQUE

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

4.1 Fuzzy Aggregation using Weighted Average

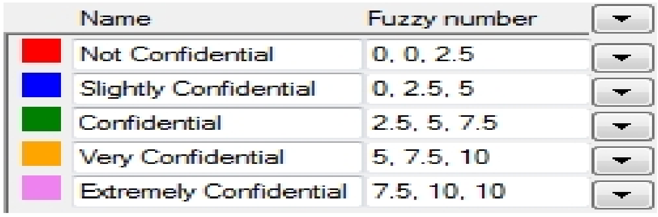
One of the most common aggregation operator often found in literature is the weighted average (WA) also known as the weighted mean. It is similar to an arithmetic mean where instead of each of data points contributing equally to the final average, some data points contribute more than others. There are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

4.2 Triangular Fuzzy Numbers (TFN)

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

Let TFN X = (X1, X2, X3) (1)
And TFN Y = (Y1, Y2, Y3) (2)
Then the sum of X and Y is (X1+Y1, X2+Y2, X3+Y3) (3)

In this paper, respondents were to choose between a series of statements on the ordinal/interval scale the one they judge most appropriate and it is argued that the choice of score is, in effect, a judgement between 3 indicator statements. Thus, for example as shown in Fig. 7 for the linguistic variable confidentiality, respondents rate the level of confidentiality of the dating history or intimate secrets they 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

Fig. 7. Linguistic variable confidentiality

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

4.3 Evaluation using Weighted Average

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

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

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

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

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

$$(s_{1,} s_{2,} s_3) = \frac{1}{n}(n_1(0,0,2.5) + n_2(0,2.5,5) + n_3(2.5,5,7.5) + n_4(5,7.5,10) + n_5(7.5,10,10)) \quad (5)$$

4.4 Center of Gravity

To determine perceived confidentiality, integrity and availability, users responded to 7 different questions (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. Figs. 8 and 9 illustrate the centers of gravity for one triangle and two triangles respectively.

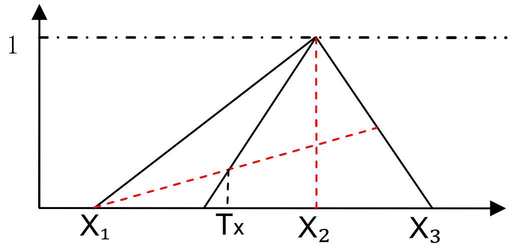


Fig. 8. Center of gravity T_x of one triangle (X_1, X_2, X_3)

$$T_x = \frac{(x_1 + x_2 + x_3)}{3} \quad (6)$$

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

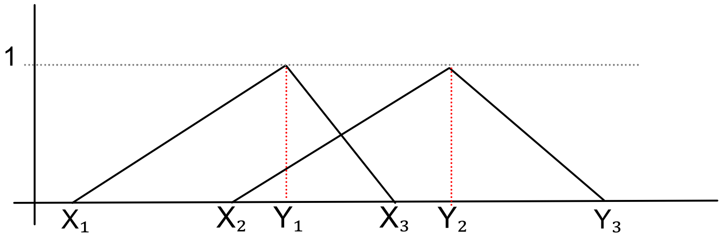


Fig. 9. Center of gravity of 2 triangles

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)

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

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

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 $(X1^1, X2^1, X3^1)$,...$(X1^7, X2^7, X3^7)$

The center of gravity T_c of fuzzy numbers $(X1^1, X2^1, X3^1)$,...$(X1^7, X2^7, X3^7)$ can be evaluated as:

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

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

where

T_i ; i=1, …,7 are centers of gravity

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

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

Table 6. Aggregation of responses of confidentiality on Facebook

Question: *In general, how would you rate the level of confidentiality of the following information you submit with friends on Facebook*.

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

Center of Gravity T_c ***= 4.52***

Table 7. Aggregation of responses of integrity on Facebook

Question: *Rate the level of authenticity of the origin of messages and files you receive from the following people on Facebook*

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

Center of Gravity T_c ***= 6.10***

Table 8. Aggregation of responses of availability on Facebook

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

Center of gravity T_c ***= 8.53***

4.5 Summary of Results

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

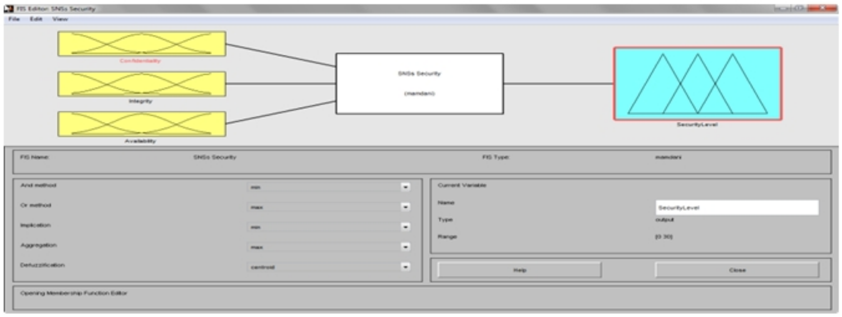
Table 9 summarizes the crisp values from the aggregated responses from users for the input variables.

4.6 Implementation Procedure in MATLAB

The final result for each of the linguistic inputs derived after aggregating the responses from the well-constructed online social networking sites security questions were fed into MATLAB 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 Fig. 10 shows the summary of the fuzzy inference system. In the editor, is shown the mapping of the input variables to the output. The input variables were respectively *confidentiality*, *integrity* and *availability*. The output was *security level* whiles 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 because it is well suitable for human input like this research and generally has broad acceptance and applicability [25].



FIS Name: SNSs Security

FIS Type: mamdani

And method: min

Or method: max

Implication: min

Aggregation: max

Defuzzification: centroid

Current Variable

Name: SecurityLevel

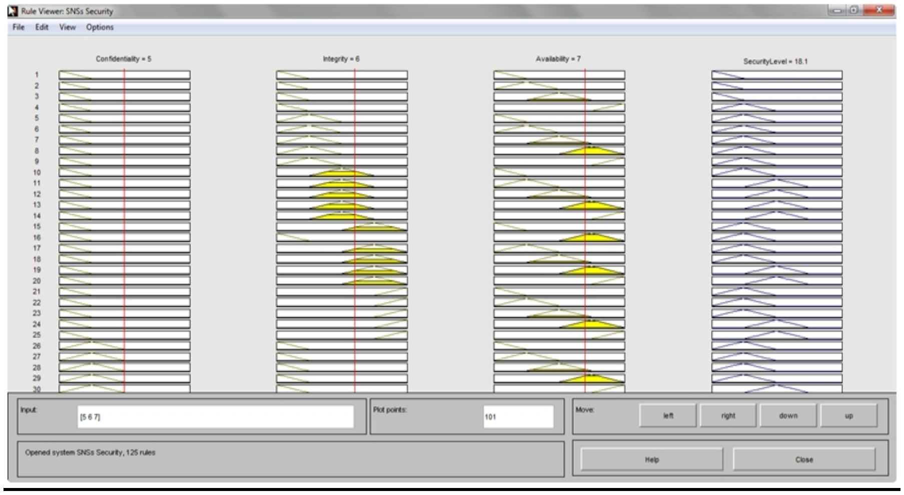
Type: output

Range: [0 30]

Opening Membership Function Editor

Fig. 10. MATLAB FIS editor

4.6.2 The MATLAB rule viewer



Confidentiality = 5

Integrity = 6

Availability = 7

SecurityLevel = 18.1

Input: [5 6 7]

Plot points: 101

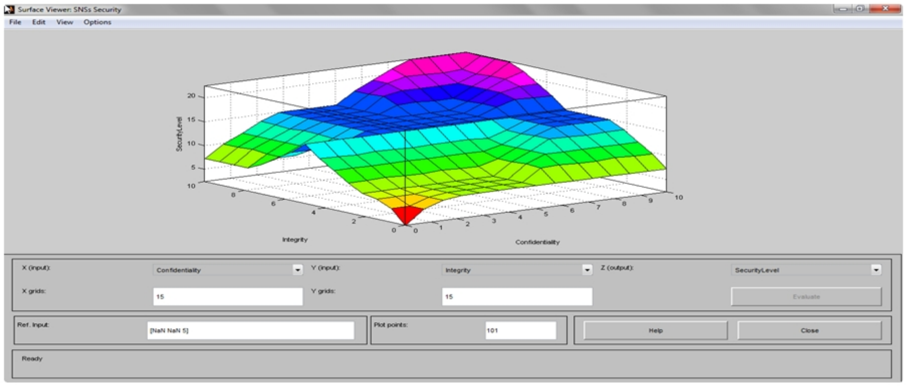
Opened system SNSs Security, 125 rules

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 Fig. 11, the input values for the variables confidentiality, integrity and availability are respectively (5,6,7) and the corresponding output (security level) is 18.1, as shown at the top of the corresponding graphs. The result for each of the input variables is specified at the top of the section corresponding to them, so also is the output variable.

4.6.3 The surface viewer

The MATLAB surface viewer as shown in Fig. 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.



X (input): Confidentiality

Y (input): Integrity

Z (output): SecurityLevel

X grids: 15

Y grids: 15

Ref. Input: [NaN NaN 5]

Plot points: 101

Ready

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:

a) if *x* is greater or equal to 15, then the value of the membership function is

$$secure(x)=(22.5-x)\ )/(22.5-15)=(22.5-x)/7.5 \tag{11}$$

b) if *x* is less than 15, then value of membership function is

$$secure(x)=(x-7.5)/(15-7.5)=(x-7.5)/7.5 \tag{12}$$

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. Figs. 13 and 14 show the security levels for Facebook based on the membership functions secure and very secure respectively.

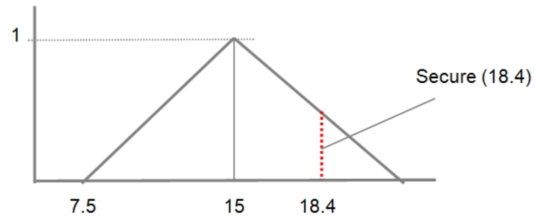


Fig. 13. Facebook: membership function secure(X)

$$\frac{1}{secure(18.4)} = \frac{(22.5-15)}{(22.5-18.4)} = \frac{7.5}{4.1} \longrightarrow \quad secure(18.4) = \frac{4.1}{7.5} = 0.55 \tag{13}$$

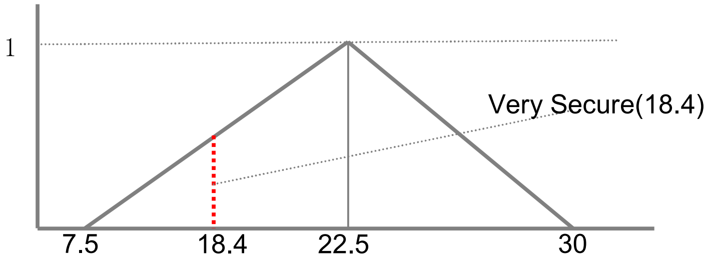


Fig. 14. Facbook: Membership function *very secure(x)*

$$\frac{1}{Verysecure(18.4)} = \frac{(22.5-15)}{(18.4-15)} = \frac{7.5}{3.4} \longrightarrow \quad Verysecure(18.4) = \frac{3.4}{7.5} = 0.45 \tag{14}$$

5.2 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. Figs. 15 and 16 show the security levels for Twitter based on the membership functions secure and very secure respectively.

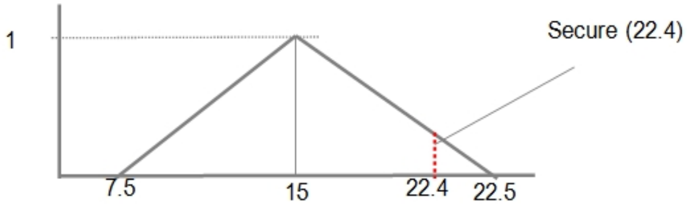


Fig. 15. Twitter: membership function *secure(x)*

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

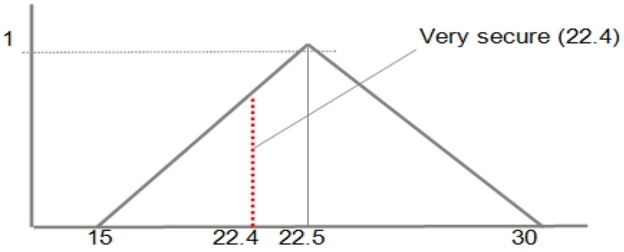


Fig. 16. Twitter: membership function very *secure(x)*

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

5.3 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 shows the security level for Facebook based on the membership functions secure.

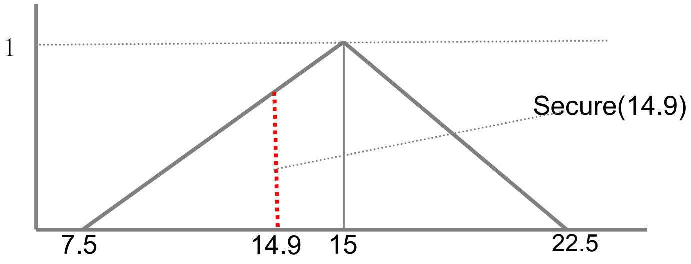


Fig. 17. LinkedIn: Membership function *secure(x)*

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

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	-

Table 10 gives a summary of the research in terms of how the final values were arrived at for the variables involved. 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. Whiles 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 two other media sites.

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 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.

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.
4. Alexa-Web Information Company. Facebook.com Site Info. 2011;08:08. Accessed 7 September 2012. Available: http://www.alexa.com/siteinfo/facebook.com.
5. Donath J. Signals in Social Supernets. Journal of Computer-Mediated Communication 2007;13(1):231–251.
6. Boyd D. Facebook's Privacy Trainwreck: Exposure, invasion, and social convergence. The International Journal of Research into New Media Technologies 2008a;14(1):13-20. DOI: 10.1177/1354856507084416.
7. Srivatsa M, Agrawal D, Reidt S. A Metadata Calculus for Secure Information sharing. Proceedings of the 16th ACM conference on Computer and Communications Security
8. Chiaramonte P, Martinez E. Jerks In Space. The New York Post. 2006;6.
9. Hass N. In Your Facebook.com. The New York Times: 30-31. 2006;08:01. Accessed 10 July 2012. Available: http://www.nytimes.com/2006/01/08/education/edlife/facebooks.html? pagewanted=all.
10. Bhaiji Y. Network Security Technologies and Solutions. A comprehensive all-in-one reference for Cisco network security. 1st ed. Cisco Press; 2008.
11. Dahal K. Loan Risk Analyzer based on Fuzzy Logic. Proceedings of IEEE International Conference on E-Technology, E-Commerce and E-Service. 2005:363-366
12. ITSEC. Information Technology Security Evaluation Criteria. Department of Trade and Industry, London, 1991;12:06. Accessed 28 November 2012. Available: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Zertifizierung/ITSicherheitskriterien/itsec-en_pdf.pdf?_blob=publicationFile.
13. Wilson K. Conflicts among the Pillars of Information Assurance. IT Professional, IEEE Computer Society press, 2012. 14(6). DOI: 10.1109/MITP.2012.24
14. Conrad E, Misenar S, Eldman J. CISSP Study Guide: Information and Security Governance and Risk Management. 1st ed. Elsevier; 2010.
15. CSEC. IT Security Risk Management: A lifecycle Approach. Annex 4-Profile 1. Nov; 2012.
16. Zadeh L. A. Fuzzy Sets. Information and Control. 1965;8(3):338-353.
17. Hossain MA, Thabatah F, Dahal K. Intelligent Quality Performance Assessment for E-Banking Security using Fuzzy Logic. Fifth International Conference on Information Technology: New Generations. 2008;420-425. DOI: 10.1109/ITNG.2008.154.
18. Abraham A, Knapskog S. Fuzzy online risk assessment for distributed intrusion prediction and prevention systems. Tenth International Conference on Modeling and Simulation, IEEE Computer Society press. 2008;1-12. DOI: 10.1109/UKSIM.2008.30.
19. Abraham A, Grosan C, Liu H, Chen Y. Hierarchical Takagi-Sugeno Models for Online Security Evaluation Systems, In Fifth International Conference on Information Assurance and Security, IEEE Computer Society press. 2009;687-692. DOI: 10.1109/IAS.2009.348.
20. Jang JS, Sun CT. Neuro-Fuzzy and Soft Computing. A Computational Approach to Learning and Machine Intelligence. 1st ed. Prentice Hall; 1997.

21. Yang Y, Zhou Y. Fuzzy Logic based Method for Network Information Security Risk Assessment. Conference on Internet Technology and Applications (iTAP), 2011:1-4
22. Chen D, Li, X. A Comprehensive Evaluation Model of Network Information Security Based on Interval-valued Fuzzy Mathematics. International Conference on Computer Science and Service System. 2012;777–780. DOI: 10.1109/CSSS.2012.199.
23. Juang CH, Huang XH, Shiff SD. Determination of weight criteria for decision making by the fuzzy eigenvector method. Civil Engineering System. 1992;9(1):1-16.
24. Lee D, Pietrucha MT, Sinha SK. Application of Fuzzy Logic to Evaluate Driver Perception of Variable Message Signs. Transportation Research Record 1937 TRB, National Research Council, Washington, D.C. 2005;1937:96-104.
25. Sivanandam SN, Sumathi S, Deepa SN. Introduction to Fuzzy Logic using MATLAB. 1st ed. Springer; 2007.

Appendix A

Questionnaire for Evaluating Confidentiality

Question: In general, how would you rate the level of confidentiality of the following information you submit with friends on (Facebook, Twitter, Linkedin)?

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					

Questionnaire for Evaluating Integrity

Question: Rate the level of authenticity of the origin of messages 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					

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					

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=226&id=5&aid=1318