Proposing a model to match virtual identity of Internet users with their true identity

Zahra Lotfi, Hassan Zafari, Rassol Roostae *

Abstract—**Identity is one of the important issues in the discussion of cyberspace. Identity is a complex global reality concept with various philosophical and practical implications. Fake identities are a major challenge in cyberspace which creates numerous problems for users of this field. In present study, verification methods in cyberspace have been firstly investigated. Then, the challenges of each method have been stated and finally, a new model proposed to counter fake identities. The results obtained from evaluations indicated the superiority of proposed method compare to the other methods.**

Keywords: **virtual identity, fake identity, identity verification, real identity**.

I. Introduction

Identity is a concept has been transmitted from real world to virtual world. Identity means individuals' personality and cognitive identifiers. In the real word, identity of individuals is defined by a combination of components such as name, surname name, father's name, place and date of birth and Id (birth certificate number) and so on. The number of involved components varies with magnitude of surrounding environment. As an example, identities of individuals are identified with their first name in a friendship group while the same individuals are identified with name, surname and Id number of student card in university. The important thing is that identity is unique and non-repetitive in the desired environment. Therefore, the number of identification (verification) identifiers can be changed.

II. Identity in cyberspace

Cyberspace has transformed the way of people interact. This space has facilitated human access to information, communication and led to development of relationships. The concept of identity has been changed along with this transformation and transforming offline status of real world into online status of virtual world. Various methods have been defined for expressing virtual identity. A username and password is one of the ways to represent a user's virtual identity. Each method of verifying identity in cyberspace has its own strengths and weaknesses which this matter is a detailed discussion itself.

III. Fake identity

Impersonation in virtual network is one of the most important issues in facing with cyberspace, so that individuals or organizations sometimes use fake identities to achieve specific goals. As an example, a person in a virtual network can take advantages of other people with a fake identity. Determining the true identity of individuals and matching their virtual identities with the real identity is a very important issue in social networks. Various strategies have been provided for this purpose that each of them has its own limitations and challenges.

- The method of providing information by the user and the assignment of the username and password is the simplest form of verification in which it is easy to spoof the title and identity.
- E-Mail verification is a solution in which user must insert an email at the time of registering, then confirm and activate his/her virtual identity through received message. In this method, fake identity can also be easily created due to the ability of user for creating fake email address.
- The method of identity verification by telephone is another way of identifying individuals in the virtual world. In this method, the user specifies his/her phone number on the site and receives the activation code by calling or message. This method also has some problems due to the possibility of registering virtual phone numbers or inaccurate information provided by individuals while buying phone numbers. However, websites such as Twitter with millions of members have made two-step security verification to increase the security of their users' information. Social networks have recently been targeted by cybercriminals and have created such a security system for non-hacking. Large networks like Facebook and Google that have switched over to Claude have already set up specific actions for users to log in. In a two-step system, users need to enter a special code sent to SSH in addition to writing the password in order to get the final confirmation. In this way, it is possible to block illegal inputs and limit the use of virtual single-use numbers, but it is not still possible to prevent impersonation in cyberspace. Because individuals can insert fake information when registering caller ID information in communications companies.
- Identity verification using OTP device (One Time Password) is a data protection method to prevent password theft. In this method, the password is generated using encryption approaches and valid only for one login. The most important advantage of using an OTP or one-time password

is the impossibility of stealing information by knowing the password. The need for this method is encryption device.

IV. The recommended model

In this section, a protocol has been proposed to anticipate and prevent impersonation in cyberspace based on mentioned problems. It would firstly better to specify the components required in this protocol.

- Identity verification server
- Well-known user in a social network
- Unknown user on all social networks
- Applicant user for registration on a social network
- User registered by verification server
- Collections of collaborative social networks

In the proposed model of present study, a server has been used to authenticate user's virtual identity. An organization must be the administrator of verification server so that each user receives a confirmation ID by signing up his/her real identity on the server and in virtual network queries, his/her virtual identity is authenticated with his/her real identity. If a user has a username in virtual network, the virtual network can match his/her real identity by submitting a request on the server.

As noted above, users of virtual environment are divided into three categories:

- The first group of users includes those with verification on the identity verification server. In present study, this group of users referred as class A.
- The users of class B includes those users are not registered on the verification server but identified on the social network as an honest person. The social network itself has adapted their virtual identities to their real identity through a different approach.
- Those users who are not members of the previous two classes but identified by at least one member of group A have been involved in class C.
- Those users who are not affiliated with any social networking sites and also are not registered on the verification server are called unknown users and involved in class D.

The process of proposed verification protocol is performed with requesting a user to register on a social network. Whenever a user submits a registration request on a given social network, he or she will be placed in one of three above classes based on his/her declared conditions. The decision process of provided protocol is in a way that if the user is involved in class A, the corresponding social network matches virtual identity of the user with his/her real identity by referring to verification server and this user is displayed with the real label. If a user is involved in class B, he/she admitted in new network and displayed with the real or intermediate label. Users of class C are identified with recommended label and admitted in social network by receiving the confirmation of a referral member. Finally, users of class D are not admitted in social network and their request is rejected.

V. Mathematical model of members of each class

The following parameters have been considered to determine the users' number of each class.

V. Mathematical model of members of each class	V. Mathematical model of members of each class
The following parameters have been considered to determine the users' number of each class.	The following parameters have been considered to determine the users' number of each class.
V. Mathematical model of members of each class	V. Mathematical model of members of each class
The following parameters have been considered to determine the users' number of each class.	The following parameters have been considered to determine the users' number of each class.
V. Mathematical model of members of each class	V. Mathematical model of members of each class
The following parameters have been considered to determine the users' number of each class.	The following parameters have been considered to determine the users' number of each class.
V. Mathematical model of members of each class	V. Mathematical model of members of each class
The following parameters have been considered to determine the users' number of each class.	The following parameters have been considered to determine the users' number of each class.
V. Mathematical model of members of each class	V. Mathematical model of members of each class
The following parameters have been considered to determine the users' number of each class.	The following parameters have been considered to determine the users' number of each class.

N_A is definitely zero at the moment of starting the proposed model. Therefore, all users must confirm their identity by referring to the identity verification organization. Hence, the number of visits to the relevant organizations is 2.5 billion in the worst case. But the maximum number of users of class B can be calculated as follow.

ISSN 2519-0814

$$N_B = N_{SUM} - N_{total}$$

$$N_B = 15\ BIL - 2.5\ BIL = 12.5\ BIL$$

The above amount indicates users who are easily verified by peer-to-peer sites.

The duration of the total user registration can be calculated based on the user logon rate to the social network and the total number of social network users.

$$Time = {}^{N_{SUM}}/_{R_E}$$

$$Time = {}^{15\ BIL}/_{10\ MIL} = 1500\ days$$

Now, the users' number of class C can be calculated given the total registration time and the percentage of users approved by existing users (referral users). It should be noted that all individuals should be logged in via the referrals after the first 20 days due to the rate of individuals' log in and percentage of individuals approved through the referrals.

$$T = \frac{{}^{R_E}/_{R_{REF}}}{R_E} = \frac{1}{R_{REF}}$$

$$T = \frac{1}{5\%} = 20$$

$$N_C = R_{REF} * R_E * (0 + 1 + 2 + 3 + \cdots + T) + R_E * (Time - T + 1)$$

$$N_C = 0.5(0 + 1 + 2 + 3 + \cdots + 20)MIL + 10 * 1479MIL$$

$$N_C = (0.5 * \frac{20 * 21}{2} + 10 * 1479)MIL$$

$$= 105 + 14790\ MIL = 14895MIL$$

The above results prove that out of 15 billion user names i.e. a maximum of fourteen billion eight hundred ninety five million usernames are approved through referrals. The flowchart of proposed protocol has been represented in Figure (1).

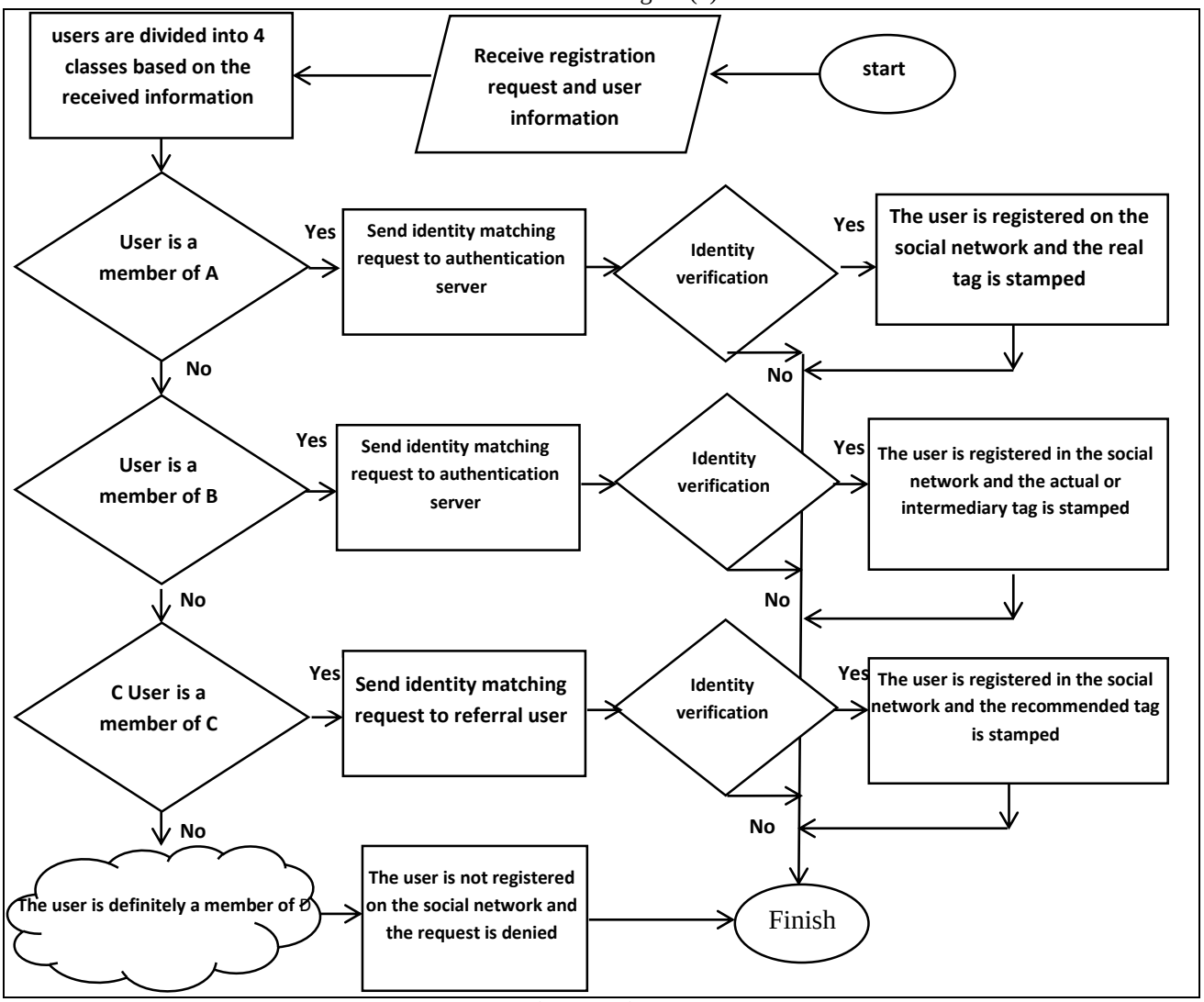

Figure 1. the flowchart of proposed protocol

ISSN 2519-0814

As it can be seen from the above flowchart, the known users are more likely to subscribe to social networks and verify their identities. The important point of this plan is the possibility of using distributed systems for the verification server so that as like the DNS server, it can be distributed according to the location or country of the verification server. By this, the reliability and distribution of load are also provided for this server.

Collaborative social networks that contribute to the identity validation of Class B users increase the load distribution and error tolerance in the identity verification system. In this way, the volume of requests sent to the identity verification system can be reduced by performing identity verification via a social network as like a peer to peer (P2P) system and also has an alternative role for error tolerance.

Figure (2) shows the communication architecture of proposed idea. U1 user is placed in class A, because it has been registered in identity verification server and it would be possible for him to register on social network with the confirmation of social network from verification server. U2 user is familiar to U1 user, so his request for verification is sent to U1 as a referral user by the social network. U4 user is familiar with another social network and its verification issued by the relevant social network and placed in class B. Finally, user U3 is placed in class D and his request denied because he is unknown.

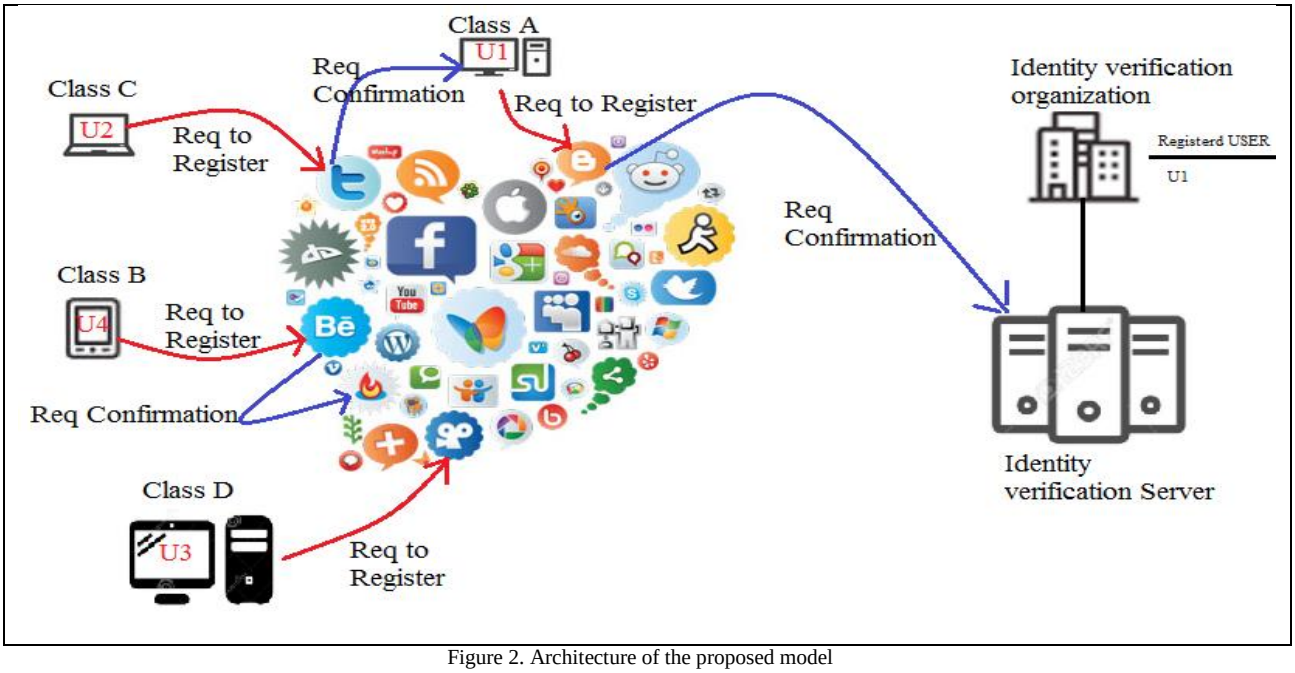

Figure 2. Architecture of the proposed model

Another issue taken into account in proposed model is classification of users into four classes with different labels which among them, three classes can login on social networks and the fourth class is unknown for virtual space and cannot register on social network unless verifies his identity on identity verification server. This classification allows the creation of arbitrary limits for different classes according to social networking policies. As an example, it would be possible to determine the number of users recommended by a known user based on the status of referral user.

VI. Evaluation of proposed model

In this section, the cost of impersonating has been evaluated and ease and difficulty of each method of impersonating have been investigated in the form of following table.

Table 1. investigating the steps and costs necessary for impersonating in each method

Difficulty level of each method	Cost and prerequisite for impersonation	Required steps	Verification method
X	-----	Register (sign up) on the relevant site	Password-based
XX	Email	Register on the relevant site and create an e-mail	Email-based
XXX	Phone (only once in use)	Register on the relevant site and	Phone-based (offline)

		activate with a phone	
XXXX	Phone (every time to log in)	Register on the relevant site and receive a password with a phone	Two steps with phone (online)
XXXXX	Having encoder device	Register on the relevant site and receive the password from the OTP machine	OTP device
XX	Sign up at the source site	Register on the relevant site and activate with the peer system	With Collaborative (peer-to-peer) System
XXXXX	The presence of a referral user who helps the applicant to impersonate	Register on the relevant site and verifying the identity via related server, peer site or the referral user	The recommended method

In Table (1), the steps of identity verification in each method have been firstly identified. Then, the costs of impersonating as well as ease and difficulty of it have been investigated for each method. As it can be seen, the proposed method and the OTP method have the most difficult mechanism for counterfeiting. In OTP method, the applicant needs to register in person in the organization where the encoder device is shipped. In this way, it's very difficult to impersonate. Also in the recommended method, the users of class (A) are registered in identity verification server by in person referring to the identity verification organization. The users of class B have also taken the steps other social networks and finally, the users of class C should receive a confirmation from a referral user. In this way, it's very difficult to impersonate.

VII. Implementation prerequisites

The tools used to create identity verification methods represent the hardware or software that is used to implement the method. Whatever these requirements are less, the work is easier and more convenient.

Table 2. System requirements for each method

The required tools	Verification method
Registration(Sign-up) server	Password-based
Registration (Sign-up) server and the system of activation link generator and send email	Email-based
Registration server and the system of sending activation code via telephone	Phone-based (offline)
Registration server and the system of sending activation code via telephone	Two steps with phone (online)
Registration server and OTP device to each user and one-time password check system	OTP device
Registration server and the system of verification via peer networks	With Collaborative (peer-to-peer) System
Registration server, verification server, ability to communicate with the identity verification server and peer site or the referral user	The recommended method

In the term of required tools to implement each of the methods, it is clear that the method of receiving username and password is the simplest method and has the least requirements. But in the term of difficulty in impersonating, only two methods of OTP and the proposed method have better conditions. Therefore, the system requirements of these two methods were only comprised. The proposed method requires only an identity verification server. While in OTP method, it is required that each has an encoder device in addition to the password server. Therefore, the recommended method is superior in this evaluation.

VIII. Operational over-load

Another parameter for evaluating new methods and protocols is operational over-load which has been investigated for proposed method. In Table (3), the operational over-load of each method has been identified based on the sent messages, code processing and so on. As it can be seen, the operational over-load of the proposed method is lower compared to the main competitor i.e. OTP method, because the OTP method needs a password inquiry in each login.

Table 3. the operational over-load of each method

operational over-load	Verification method
Processing the registration form	Password-based
Processing the registration form Sending email activation	Email-based

ISSN 2519-0814

Processing the registration form Send code by phone activation	Phone-based (offline)
Processing the registration form Send code by phone (per login) Activation (per login)	Two steps with phone (online)
Processing the registration form Attend the organization and receive the device Generate encryption with the device (per login) One-time password processing by the server (per login)	OTP device
Processing the registration form Submit the request to the peer system Identifying and matching the identity and sending results in the peer system	With Collaborative (peer-to-peer) System
Processing the registration form Presence in the organization and identity registration Submitting a request to the identity verification server or peer system or referral user	The recommended method

IX. Conclusion

In the present study, identity verification via an identity verification server or peer site or referral user has been provided with the classification of users. In the investigations, proposed idea was superior to other methods in terms of criteria such as impersonation cost, tools needed to implementation and operational over-load. In the other sides, some characteristics of proposed model such as peer system make it possible to utilize from P2P structure to identity authentication as well as provide error tolerance in the model along with load distribution. It should be noted that error tolerance and load distribution capabilities are also provided by providing identity verification server as a distributed system.

References

[1] Stryker, Sheldon, and P. J. Burke. (2000). The past, present, and future of an identity theory. *Social psychology quarterly*. pp. 284-297.

[2] Bilge, Leyla, et al. "All your contacts are belonged to us: automated identity theft attacks on social networks", Proceedings of the 18th international conference on World wide web ACM, (2009).

[3] Stets, Jan E., and Peter J. Burke. (2003). A sociological approach to self and identity. *Handbook of self and identity*. pp. 128-152.

[4] Burke, Peter J., and Donald C. Reitzes. (1991). An identity theory approach to commitment. *Social psychology quarterly*. pp. 239-251.

[5] Ibarra, Herminia, M. Kilduff, and W. Tsai. (2005). Zooming in and out: Connecting individuals and collectivities at the frontiers of organizational network research. *Organization science*. 16(4). pp. 359-371.

[6] Metzger, J. Miriam, A. J. Flanagin, and R. B. Medders. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*. 60(3). 413-439.

[7] T. Mike, D. Wilkinson, and S. Uppal. (2010). Data mining emotion in social network communication: Gender differences in MySpace. *Journal of the Association for Information Science and Technology*. 61 (1). pp. 190-199.

[8] https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/.

[9] https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/.