I J C T A, 8(5), 2015, pp. 1903-1910

An improved recommendation system for social networks

C.K. Shyamala*, Hemaashri S. and Swetha R.*****

Abstract: Social networking sites have become extremely popular in the past few years because of their extensible online connectivity and information sharing capabilities. The ability to communicate and connect online easily has attracted millions of users. Although this establishes high connectivity, privacy of user data is at stake due to various privacy threats such as sensitive data revelation, fake identities etc. We propose an automated dynamic grouping system to have better control over sensitive data that is shared online. This system analyses friends in social networking sites and categorizes them into best friends, normal friends and visitors. The categorization is based on various interaction parameters. The automated grouping system is enhanced using a recommendation system to make the system dynamic.

Index terms:Recommendation system, automated grouping, social networking, privacy, Sybil attack, privacy preservation, information revelation, sensitive data protection, closeness computation

1. INTRODUCTION

Social networking sites(SNS) have become a part of our daily lives because of their ability to establish and maintain relationships that may seem impossible in a real life scenario. The relationships in SNS like Facebook are categorized into: Friends, Friends of friends and not a friend. Friends of a user can access all the user's private data. The revelation of user's private data to friends of friends and not friends is decided by the privacy policy specified by the user. Friends in SNS are similar to friends in real life but there is a discrepancy in the former case. The sensitive information revealed by the user to close friends or family is also accessible to strangers in his friend list and hence the privacy of user data is at risk. To mitigate the attack on sensitive user data, an automated dynamic grouping system is modelled.[3] presents a grouping system which categorizes friends into best friends, normal friends and visitors with varied level of information access using various interaction parameters. Our system uses user interaction history to make the categorization. Best friends are given the benefit of complete information access to the user's data. Normal friends and visitors have access to less sensitive information. A recommendation system is used to enhance the performance of the grouping system. The recommendation system analyses a friend request and advises users to either accept or reject the request. The recommendations obtained are categorized into normal, good and exceeded. These recommendations are further used to categorize friends. The grouping system is enhanced to adapt to dynamic changes in SNS using the recommendation system.

Although the grouping system improves privacy in SNS, the system may be vulnerable to attacks such as mutual friend attack, whitewashing. [10] Sybil attack is one where a user has multiple identities and these identities are used to initiate malicious activities[8]. The system is resilient to the attacks, in particular

* Department of Computer Science and Engineering, Amrita School of Engineering-Coimbatore Amrita Vishwa Vidyapeetham (University), Amrita Nagar PO 641112, Coimbatore, India, *Email: ck_shyamala@cb.amrita.edu*

** Department of Computer Science and Engineering, Amrita School of Engineering-Coimbatore Amrita Vishwa Vidyapeetham (University), Amrita Nagar PO 641112, Coimbatore, India, *Email: hemaashris@gmail.com*

*** Department of Computer Science and Engineering, Amrita School of Engineering-Coimbatore Amrita Vishwa Vidyapeetham (University), Amrita Nagar PO 641112, Coimbatore, India, *Email: rswethaa@gmail.com*

Sybil attack. This paper evaluates the system for Sybil attacks. The proposed model identifies and removes of Sybil nodes present in the system.

The system improves privacy in SNS by automating a grouping system that controls sensitive information shared online. The paper is organized as follows. Section II vividly discusses the related work. Section III presents the model for automated grouping system. The results are discussed in section IV followed by conclusion in section V.

2. RELATED WORK

Social networking is the current trend for communication and information sharing among online entities [23],[24] analyses the reason and the effects of increase in continuous usage of SNS by the users.[1], [20] analyses the lack of awareness of the users towards the problems related to privacy in SNS which results in privacy leaks.[22] proposes a method to quantify the privacy leaks in SNS. Response theory model is suggested here as a metric to measure the privacy of the user.[13] A number of methods have been analysed for improving privacy and this has led to the development of various anonymization techniques for SNS data. This issue of privacy leaks and privacy protection is addressed in [15][14]. Analyzing privacy protection in relationships, SNS has only 3 categorization of relationships: Friends, Friends of Friends and Not a Friend. The shortcoming of this model is the vast relationship categorization scheme.[2] presents a scheme of user created groups which acts as an overhead to the user. This issue is overcome in [26] which proposes various data mining methods and [6] proposes a friend group recommendation method to classify relationships. Though grouping of friends, based on their preferences exist,[25], [3] proposes an automatic grouping system to classify friends. Recommendation systems are the best source for analyzing the trust for a system. New recommendation systems have been developed for various applications in SNS. [4] proposes a system that analyses games played by the user and the user's connections in SNS to recommend games to the user. Media content recommendation system is proposed by [9] that analyses user preferences in media to make recommendations to the appropriate users. In SNS there is always a risk at securing private. As a step towards this [16] proposes a recommendation system, that safeguards the user's answers and provides recommendation based on weighted average of the results. [17] presents a serendipity recommendation system, that uses the vast Social networking data to recommend products or services to the user. There are systems that recommend friends to users. [19] proposes one such system where friends are recommended based on similarity factors and an approach based on hubs and authorities. For recommendation of a node to a node, [5] proposes a recommendation system for peer to peer systems. This system proposes a trust model that decreases malicious activity in a peer to peer system by giving recommendations based on the services provided by a peer. This recommendation system can be modelled to suit SNS as significant parameters like history, interactions, etc. can be used to make stronger recommendations to the users. Although the privacy issue is addressed, social networking data is prone to various attacks which is addressed in [21]. [11], [7] and [27] study attacks on online reputation systems and recommendation systems. The other attacks such as mutual friend attacks and neighborhood attacks which are addressed in [10],[12]. Amongst these attacks, Sybil attack is currently prominent in SNS which is discussed [8].

3. PROPOSED MODEL

3.1. Automated Dynamic Grouping System

In an attempt to enhance the privacy of users, grouping of friends is done based on various interaction parameters. This work aims at automating a grouping system that groups friends into best friends, normal friends and visitors by computing the closeness degree between the user and his/her friends. The automated grouping system suggested in [3] is adopted and enhanced to model a dynamic grouping system. The user's

closeness degree is computed using the concern degree and the interaction degree among his/her friends.The closeness degree computation is proceeded by adapting the model suggested in [3].

$$Cn(y)\ (1 * j) = 1/Cn(x)\ (1 * j) \quad (1)$$

where *Cn* represents the concern degree, which is a measure of various parameters such as the number of times the user visits the profile and the album of his/her friends, shares he/she makes and initiates conversation among his/her friends. *P* denotes the number of parameters considered for computation of concern degree. Let *F* be the total number of friends for a particular user.

$$Cn(1 * j) = Cn(y)\ (1 * j) * Cn(z)\ (i * j) \quad (2)$$

where for every user, the sum of parameters such as profiles visited, albums visited, shares made and initiated conversation for all the user's friends is calculated and stored ina *(1*P)* matrix as $Cn_{(x)}$ *(1*j)*, where *i* corresponds to the rows and j corresponds to the columns.The values of the parameters for the user's friends are stored in a *(P*F)* matrix as $Cn_{(z)}$ *(i*j)*.

$$In(1, j) = In(z)\ (i, j)/In(x)\ (i, j) \quad (3)$$

Where *In(1, j)* represents the interactions among user's friends. *In* is computed using the parameters such as number of posts made in each of his friend's wall, number of interactions, tags made and number of feedbacks obtained from each friend. For each user, all the values of these parameters are added and calculated separately for all his/her friends and stored as $In_{(z)}(i, j)$where *i* corresponds to the users and *j* corresponds to the friends. For each user, $In_{(x)}(i, j)$is calculated as the total interactions for each of his/her friends

Algorithm for closeness degree computation:

For each user *i*:

For each friend *j*,

Step 1:Obtain $Cn_{(y)}$ (1*j) value.

$Cn_{(y)}$ (1*j) = 1/ $Cn_{(x)}$ (1*j)

Step 2:Obtain $Cn_{(z)}$(i*j) value

Step 3:Obtain *Cn*(1*j) value from $Cn_{(y)}$(1*j) and $Cn_{(z)}$(i*j).

Cn(1, *j*) = $Cn_{(y)}$(1*j) * $Cn_{(z)}$(i*j).

Step 4:Obtain $In_{(z)}(i, j)$ and $In_{(x)}(i, j)$

Calculate *In*(1, *j*) value

Step 5: *In*(1, *j*)= $In_{(z)}(i, j)$ / $In_{(x)}(i, j)$.

Calculate closeness degree.

Close(*d*)(1, *j*) = *Cn*(1, *j*) + *In* (1, *j*).

Best friend: close(*d*) > *a* + 2*s*

Normal friend: *a*-*s* <close(d) < *a* + 2*s*

Visitor: close(d) < *a*-*s*

Figure 1:Closeness degree computation algorithm

The algorithm in Figure 7 illustrates the computation of closeness degree. A standard value for best friend, normal friend and casual friend is calculated using the closeness degree computation. The standard value denotes the range that decides each category of friends. Standard values for categorization of friends is computed as *a+2s* and *a-s* where *a* denotes the average and *s* denotes the standard deviation of the closeness value *close(d)(1, j)* of a particular user's friends. This computation results in 3 categories of friends and each type of friend is assigned a value. The friend values are: Best friend: 15, Normal friend: 10, Visitor: 5. The closeness computation algorithm results in categorization of friends. To make the categorization dynamic, a recommendation system for peer to peer systems proposed in [5] is adapted to suit social networking and is used along with closeness computation algorithm.Let *p* and *q* be nodes in the social networking site. *p* gives friend request to *q* and *r* is the mutual friend to *p* and *q*. *q* requests recommendation from all the mutual friends between *q* and *p*. Friend requests are analyzed based on the recommendations obtained. Three parameters are used to evaluate recommendation. The parameters used are:

1. Satisfaction – friend value of p to r (output of closeness computation algorithm).
2. Weight – No of times (profile viewed + album viewed + shared)/history between p and r.
3. Fading Effect – No of interactions/history between p and r.

Table 1
Parameters of Recommendation system

Notation	*Description*
p, *q*	Nodes in social networking sites
R	Mutual friends between p and q
Rc_i	Recommendation from every mutual friend *r*
Rth_{high}	Upper Threshold value
Rth_{low}	Lower Threshold value
E_{cnt}	Count of exceeded recommendations
N_{cnt}	Count of normal recommendations
G_{cnt}	Count of good recommendations

Algorithm-recommendation system:

Step1: Obtain the mutual friends(*n*) between *p* and *q*.

Step2: For every mutual friend *R* obtain the required parameters

Rc_i = Satisfaction + weight + fading Effect

Step 3: For every *n*

If ($Rc_i > Rth_{high}$) $E_{cnt} = E_{cnt}+1$

If ($Rc_i < Rth_{low}$) $N_{cnt} = N_{cnt}+1$

If ($Rc_i <= Rth_{high}$ and $Rc_i >= Rth_{low}$)

$G_{cnt} = G_{cnt}+1$

Step 4: If (Ncnt > = $n/2 + 1$)

Recommend user to reject request.

For every $Rc_i > Rth_{high}$

If (Fvalue = 15) Fvalue = 10

If (Fvalue = 10) Fvalue = 5

Else Recommend user to accept request

Figure 8: Algorithm for recommendation system

The system obtains recommendation from every mutual friend. The recommendations are characterized as normal (Rc less than threshold), good (Rc value between the threshold range) and exceeded (Rc greater than the threshold) recommendations. The count of each type of recommendation is calculated and majority quorum is used to accept or reject friend requests. In the latter case, for every exceeded recommendation the friend level is reduced for every friend who gave the exceeded recommendation. The algorithm in Figure 8 illustrates the computation of the recommendations obtained. This system analyses friend requests and based on the recommendation quality, changes the friend level to accommodate dynamic changes in the system.

3.2. Detection and prevention of Sybil attack

The system aims at preserving privacy of the user by grouping of friends but the system is still prone to attacks. This system analyses the detection and prevention of Sybil attacks. Sybil attack is an outcome of multiple pseudonymous identities created by a user to perform malicious activity. To mitigate the presence of Sybil nodes, Sybil Identification using Connection Threshold(SICT) and Sybil Identification using

Connection Threshold and Frequency of visits(SICTF) algorithms in [8] are adapted to suit the system. The Sybil nodes can be detected by analyzing the connections between users and his/her friends. Connections denote the number of times the user gets acquainted to other users. Friends denote the current relationships of the user. The Sybil attack is detected using the following parameters:

Length: l = number of friends/number of connections

Frequency: f = number of times profile visited/number of friends

Relation: r = ((number of connections-number of friends)/(number of friends))

To detect a Sybil node, the three conditions to be satisfied are:

Relation >1, Length < 0.5, Frequency < 0.

The Sybil nodes are identified and removed from the system. Thereby attack resilience is achieved by Sybil identification and detection.

4. SIMULATION AND RESULTS

The system aims at categorizing friends into best friends, normal friends and casual friends based on their relationships and interactions. The interaction history and the closeness degree of the user's friends is stored and updated. The SNS is simulated for 500 users, where the closeness degree is computed for each of the user's friends and they are grouped in various categories using closeness computation algorithm. To enhance the privacy of the user, the system analyses the friend requests obtained from a new node and provides recommendation for the new node, from the other users. The dishonest recommendation provided by some of the nodes are also identified and their reputation is reduced. Figure:1 illustrates the initial state of the system. The friends are categorized for various users. The relationship status and interactions made by each user keeps changing. The closeness degree varies based on the interactions made and the degree of concern for each of the user's friends. The categorization of friends is made after 100 interactions. The database tracks the dynamic changes made in the categorization of friends, as shown in Figure: 2

Figure:3 depicts the categorization of friends for node #1, which illustrates the initial state of the system in which friends are grouped as best friends, normal friends and casual friends.The interactions get updated in the system and the relationship status of each friend varies. Figure: 4 depicts the categorization of friends in the updated system.

The system is evaluated to detect the presence of Sybil users. The malicious activity performed by the Sybil nodes are identified. The proposed algorithm provides high accuracy rates and hardly fails to detect

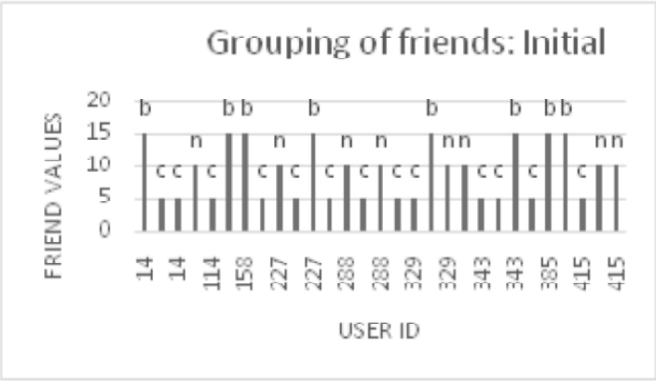

Figure1: Initial grouping of friends

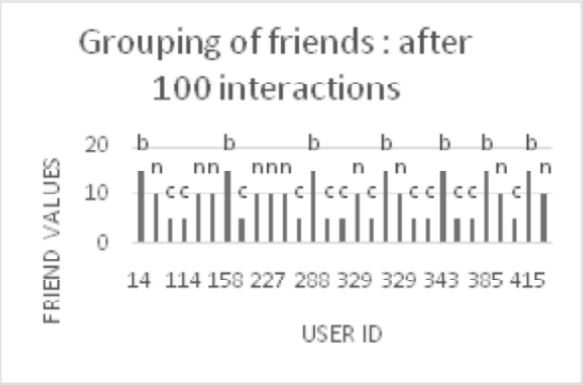

Figure2: Automated grouping of friends after 100 ... interactions

Key: Friend values:b-Best friend, n-Normal friend, c-Casual friend

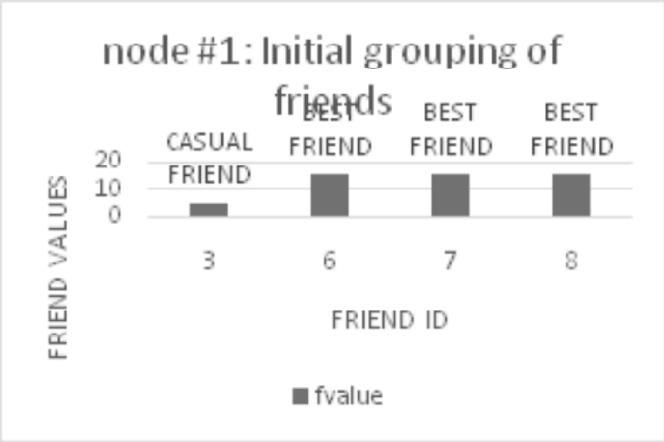

Figure 3: Initial Grouping of friends for node #1

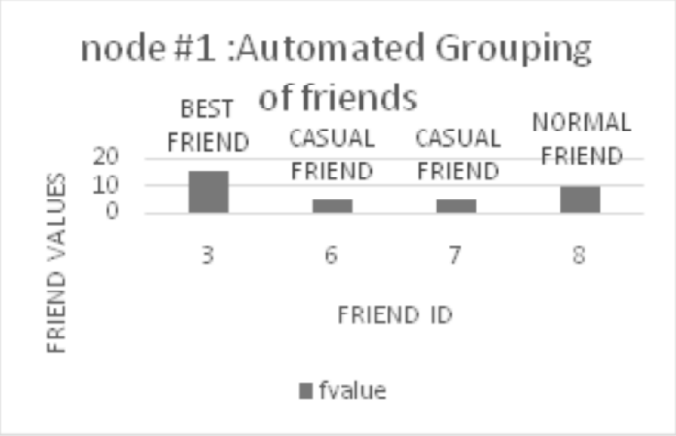

Figure4: Automated Grouping of friends for ... node #1

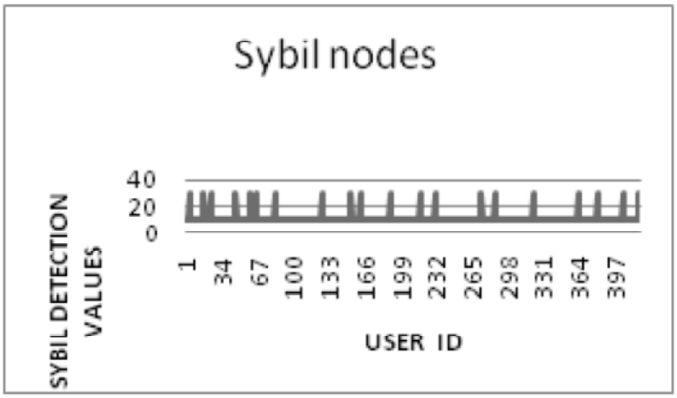

Figure 5: Total Sybil nodes in the system

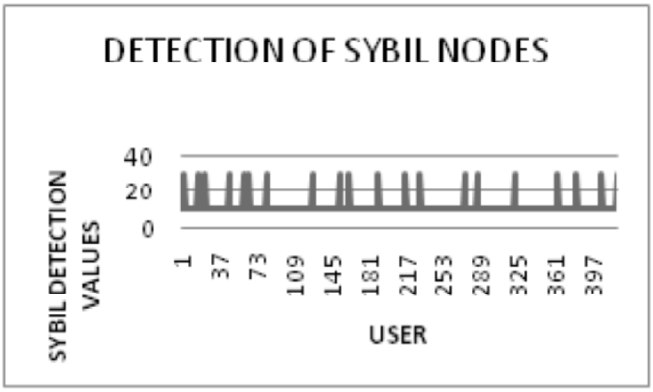

Figure6: Detection of the Sybil nodes in ... the system

Total number of Sybil nodes present in the system: 20
Total number of Sybil nodes identified:18
Accuracy rate: 90%

the presence of Sybil users. The results shown, explains the detection rates of the Sybil users identified. Figure: 5 portrays all the Sybil nodes present in the system.The proposed model gives 90% accuracy as shown in Figure:6. The identified Sybil nodes are removed from the system.

5. CONCLUSION

In this paper, we propose an automated dynamic grouping system for SNS. Our primary goal of automating a privacy enhancing system that protects sensitive user data is achieved. The automated dynamic grouping system consists of a grouping system and a recommendation system. The grouping system categorizes friends in SNS based on user interactions. Recommendation system assesses friend requests based on recommendation obtained from other users and this recommendation quality is also used to evaluate the categories of friends. The system also analyses and detects the presence of Sybil attack. Our system groups friends and suggests access levels to user data based on the categorization of friends. Further future research would enhance the privacy of the system by implementing sensitive data analysis. The sensitive data analysis identifies data such as user location, feelings, etc. These data are given varied level access to each type of friend. The future work is concentrated on sensitive data analysis and data access levels for these grouping of friends. Sensitive data analysis is in progress.

References

[1] Nagle, F., & Singh, L. (2009, July). Can friends be trusted? Exploring privacy in online social networks. In *Social Network Analysis and Mining, 2009. ASONAM'09. International Conference on Advances in* (pp. 312-315). IEEE.

[2] Ur, B., & McGrath, R. Grouping Friends for Access Control in Online Social Networks.

[3] Qian, C., Xiao, X., Chen, S., & Wang, X. (2013, January). Grouping friends to improve privacy on Social Networking Sites. In *Conference Anthology, IEEE* (pp. 1-6). IEEE.

[4] Aranda, J., Givoni, I., Handcock, J., & Tarlow, D. (2007), An online social network-based recommendation system. *Toronto, Ontario, Canada*.

[5] Can, A. B., & Bhargava, B. (2013), Sort: A self-organizing trust model for peer-to-peer systems. *Dependable and Secure Computing, IEEE Transactions on*, *10*(1), 14-27.

[6] Zhe, Z., & Li, Z. (2012, May). A method of visualizing friends relations and recommending groups in online social network. In *Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on* (pp. 836-839). IEEE.

[7] Saini, N. K., Sihag, V. K., & Yadav, R. C. (2014, February). A reactive approach for detection of collusion attacks in P2P trust and reputation systems. In *Advance Computing Conference (IACC), 2014 IEEE International* (pp. 312-317). IEEE.

[8] Samuel, S. J., &Dhivya, B. (2015, March). An efficient technique to detect and prevent Sybil attacks in social network applications. In *Electrical, Computer and Communication Technologies (ICECCT), 2015 IEEE International Conference on* (pp. 1-3). IEEE.

[9] Seth, A. (2008), Design of a Social Network Based Recommender System for Participatory Media Content. *Manuscript, University of Waterloo*.

[10] Sun, C., Yu, P. S., Kong, X., & Fu, Y. (2013, December). Privacy Preserving Social Network Publication Against Mutual Friend Attacks. In*Data Mining Workshops (ICDMW), 2013 IEEE 13th International Conference on* (pp. 883-890). IEEE.

[11] Sun, Y. L., & Liu, Y. (2012), Security of Online Reputation Systems: The evolution of attacks and defenses. *IEEE Signal Process. Mag.*, *29*(2), 87-97.

[12] Zhou, B., & Pei, J. (2008, April). Preserving privacy in social networks against neighborhood attacks. In *Data Engineering, 2008. ICDE 2008. IEEE 24th International Conference on* (pp. 506-515). IEEE

[13] Shishodia, M. S., Jain, S., &Tripathy, B. K. (2013, August). GASNA: greedy algorithm for social network anonymization. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (pp. 1161-1166). ACM.

[14] Al Hasib, A. (2009), Threats of online social networks. *IJCSNS International Journal of Computer Science and Network Security*, *9*(11), 288-93.

[15] Wang, Z., Liao, J., Cao, Q., Qi, H., & Wang, Z. (2015), Friendbook: a semantic-based friend recommendation system for social networks. *Mobile Computing, IEEE Transactions on*, *14*(3), 538-551.

[16] Hoens, T. R., Blanton, M., & Chawla, N. V. (2010, August). A private and reliable recommendation system for social networks. In *Social Computing (SocialCom), 2010 IEEE Second International Conference on* (pp. 816-825). IEEE.

[17] Zhou, L. (2009), Trust based recommendation system with social network analysis. In *2009 International Conference on Information Engineering and Computer Science* (pp. 1-4).

[18] Chiu, Y. S., Lin, K. H., & Chen, J. S. (2011, December). A social network-based serendipity recommender system. In *Intelligent Signal Processing and Communications Systems (ISPACS), 2011 International Symposium on* (pp. 1-5). IEEE.

[19] Carullo, G., Castiglione, A., & De Santis, A. (2014, September). Friendship Recommendations in Online Social Networks. In *Intelligent Networking and Collaborative Systems (INCoS), 2014 International Conference on* (pp. 42-48). IEEE.

[20] Babour, A., & Khan, J. I. (2014, August). Tweet Sentiment Analytics with Context Sensitive Tone-Word Lexicon. In *Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 01* (pp. 392-399). IEEE Computer Society.

[21] Fire, M., Goldschmidt, R., & Elovici, Y. (2014), Online Social Networks: Threats and Solutions. *Communications Surveys & Tutorials, IEEE*, *16*(4), 2019-2036.

[22] Nguyen-Son, H. Q., Nguyen, Q. B., Tran, M. T., Nguyen, D. T., Yoshiura, H., & Echizen, I. (2012, August). Automatic anonymization of natural languages texts posted on social networking services and automatic detection of disclosure. In *Availability, Reliability and Security (ARES), 2012 Seventh International Conference on* (pp. 358-364). IEEE.

[23] Akaichi, J. (2013, September). Social Networks' Facebook'Statutes Updates Mining for Sentiment Classification. In *Social Computing (SocialCom), 2013 International Conference on* (pp. 886-891). IEEE.

[24] Jiang, X., Du, R., & Ai, S. Z. (2011, August). An empirical study of users' continued usage of social networking service website. In *Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), 2011 2nd International Conference on* (pp. 105-108). IEEE.

[25] Javed, Y., & Shehab, M. (2012, August). How do facebookers use friendlists. In *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012)* (pp. 343-347). IEEE Computer Society.

[26] Tanbeer, S. K., Jiang, F., Leung, C. K. S., MacKinnon, R. K., & Medina, I. J. (2013, August). Finding groups of friends who are significant across multiple domains in social networks. In *Computational Aspects of Social Networks (CASoN), 2013 Fifth International Conference on* (pp. 21-26). IEEE.

[27] Liu, Y., Sun, Y., & Yu, T. (2011, October). Defending multiple-user-multiple-target attacks in online reputation systems. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on* (pp. 425-434). IEEE.