ISSN: 0975-766X
CODEN: IJPTFI
Research Article

Available Online through
www.ijptonline.com

A NOVEL INTELLIGENT FRAMEWORK FOR SOCIAL MEDIA NETWORK TO CURTAIL THE ABUSED CONTENTS

Bala Sundara Ganapathy.N*[1], Mohana Prasad.K[2]
[1]PG Scholar, Computer Science and Engineering, Sathyabama University, Tamil Nadu, India.
[2]Associate Professor, Computer Science and Engineering Sathyabama University, Tamil Nadu, India.
Email: balabsg@gmail.com

Received on: 18-04-2017 *Accepted on: 06-06-2017*

Abstract

Social Media Networks (SMNs) are inclined to a heap of anomalies these days. There is no centralized control to screen and break down the exercises followed in the SMNs. Subsequently we have proposed a novel intelligent framework to diagnose and confine to the content being shared in the SMNs. As a part of the framework, we have proposed an algorithm namely CO for curtailing obscenities. We have used Natural Language processing to curtail the obscene text messages being shared in the SMNs. The usage pattern mining is useful to spot the user activities in the SMN and identify the abnormal user behavior using SVM classifier. The proposed framework is implemented as a model SMN called Satbook LTE, which successfully curtails the abused contents being shared in the form of text. The system diagnoses the content being shared through the SMN without affecting the privacy of the user.

Keywords: Social Media Network, Natural Language Processing, SVM classifier, Usage Pattern Mining, Abused Contents.

Introduction

Social media networks are de facto mechanism for all kinds of people. People access social media network as often as possible with a specific goal to satisfy their undertaking. A kid playing games, a teen surfing through various networks, an adult chatting with his/her friends, a student searching for subject related subtle elements, a maid paying the bill, an educator searching for a topic, a scientist uploading his/her invention, or a lawmaker passing on his/her perspectives may be dealing with his/her social media accounts. Social media networks provide a set of functionalities and features that are useful to common people. Social media networks concentrate to acquire a lot of customers for their network and they give a lot of freedom in maintaining the users' details. They haven't validated the users' profiles, more particularly their email ids, phone numbers and year related details like date of birth and work experience to avoid any conflict. Some of the SMN's (more specifically Google, Linked in), have started validating their users' email ids, and phone numbers nowadays. But they haven't validated the profiles thoroughly.

There is no restriction to access or to post obscene comments over the net and also sharing banned images and abused images over the net. Firewalls or other utility hardware and softwares or combination of these two are available to restrict the abused SMNs. But In spite of the most useful SMNs like Facebook, Whats App, Twitter, LinkedIn and many other SMNs that are available to the users provide the users with a heap of features to share the data. Therefore we tend to use these SMNs. These kinds of SMNs do not restrict the content being shared. It has necessitated a centralized control or framework to restrict the illegal content being shared. The proposed framework validates the user profiles, and restricts the abused content present in the user's message of type text. The other types of contents like image, clipart, audio, video, and other attachment types have been considered as a future work to be included in the framework.

Related Work on Social Media Network Problems

In Social Media Networks, many user accounts are duplicated and misused. The exploiter may post illegal contents in the SMNs, by using false identities. These duplicated anonymous users can be identified by using various techniques like Profile-Based anonymous User Identification, Content-Based anonymous User Identification, Network Structure-Based anonymous User Identification, and Cross-Platform Identification of anonymous users.

2.1 Profile-Based Anonymous User Identification

Numerous cogitations addressing mysterious user identification revolve around public profile attributes that includes profile name, gender, birthday, native place, living place, working place and profile image. Perito et al. [1], Liu et al. [2] computed the closeness of profile names and identified users using binary classifiers. Zafarani and Liu [3][4] designed a method to map individualities throughout different SMN platforms. Acquisti et al. [5] dealt the user identification task with a face recognition algorithm. Iofciu et al. [6] proposed an approach by evaluating the distance between user profiles.

Motoyama and Varghese [7] gathered attributes like education, occupation, etc. as sets of words and matched users by calculating the similarity of users. Similar studies across multiple platforms are also found in [8] [9] [10] [11], [12], [13]. Profile attributes provide principal information for user identification. In any case, a few attributes are replicated in SMNs, and are simply portrayed. Along these lines, exclusively profile-based methods have some confinements when they are employed in large scale SMNs.

2.2 Content-Based Anonymous User Identification

Content-Based User Identification solutions seek to distinguish the users based on the times and locations that they post their contents, in addition to the writing style of the content. Zheng et al. [16] proposed a framework for authorship identification using the writing style of online messages and classification techniques. Almishari and Tsudik [17] proposed connecting users across different SMNs by exploiting the writing style of the authors. Kong and Zhang [18] proposed Multi-Network Anchoring (MNA) to map users. They calculated the combined similarities of user's social, spatial, temporal and text information in different SMNs, and examined a stable matching problem between two sets of user accounts. Goga et al. [19] exploited the geo-location attached to users' posts, the timestamp of posts, and users' writing style to address user identification tasks. Geo-location appears to have forceful features for user recognition. However, this information is often sparse in SMNs, since only a small portion of users are willing to post their locations. Although writing style solutions perform well in scenarios involving long content, these techniques are not applicable to SMNs such as Twitter and Sina Microblog, in which short sentences are posted mostly.

2.3 Network Structure-Based Anonymous User Identification

Network structure-based studies [20], [21], [22] on user identification across multiple SMNs are used to recognize identical users solely by user network structures and seed, or priori, identified users. These works had similar workflow, finding seed users first, then using these seed users to recursively propagate information through networks and extend sets of mapped nodes. The task on user identification is closely related to the deanonymization problem [23] for privacy-preserving social network analysis, which re-identifies the individuals with the online published SMN datasets. Zhou and Pei analyzed the neighborhood attacks of de-anonymization and proposed privacy preservation approaches using k-anonymity and l-diversity [24], [25], [26], [27]. Since cross-platform user identification is similar to the de-anonymization task, it can be applied to address the de-anonymization problem.

The joint use of profile information, user behaviors hidden content and network structures may lead to better results. Jain and Kumaraguru [14], [15] developed Finding Nemo, a method that matches Facebook and Twitter accounts. However, this text-based network search method has low accuracy and high complexity in terms of user identification, as it can recognize only the text with the same nickname when searching the friend sets of friends [10], [11]. Bartunov et al. [21] integrated profiles with a network structure using a Conditional Random Fields model and obtained better user identification results. Network structure-based user identification is a hard nut to crack, and can be used to identify only a portion of identical users. NS, the first network structure-based user recognition algorithm across SMNs, can carry out user recognition tasks by using the network structure only and can identify 30.8 percent identical users in a ground-truth dataset [20].

2.4 Cross-Platform Identification of Anonymous User

The real-world friend cycle is highly individualistic and virtually no two users share a congruent friend cycle. Xiaoping Zhou et al. [28] proposed the Friend Relationship-Based User Identification (FRUI) algorithm for identifying anonymous identical users in cross-platform. FRUI calculates a match degree for all candidates, User Matched Pairs (UMPs), and only UMPs with top ranks are considered as identical users. FRUI performs much better than current network structure-based algorithms, but identifying anonymous users across multiple SMNs is a challenging work. Therefore, only a portion of identical users with different nicknames can be recognized with this method. These methods are complementary and not mutually exclusive, since the final decision may rely on human user's involvement.

The information-age.com [29] website analyzed about various abused contents being shared in reputed SMNs like Flickr, Facebook, Twitter and so on. In our previous work [30], we have analyzed about various risk factors present in the SMN's, and we have proposed some of the mitigation techniques to overcome the problems encountered in the SMNs. There is no previous study for restricting abused content being shared in various SMN's. Therefore the proposed methodology is a novel Social Media Network framework, which is used to identify and restrict the anonymous users and the abused content being shared.

Problem Definition of a Novel Intelligent Framework for Social Media Network

The proposed framework consists of three modules. There is no restriction on the number of modules defined. Depending on the requirements and restrictions applied on the content being shared, the modules have been included in to the framework. We have taken into account the text contents which are the main components being shared. We have also identified the anonymous users through profile validation and usage pattern mining. Based on these a framework has been evolved and is shown in figure1. The modules defined in the framework are:

1. Profile Validation
2. Usage Pattern Mining
3. Curtailing Obscenities

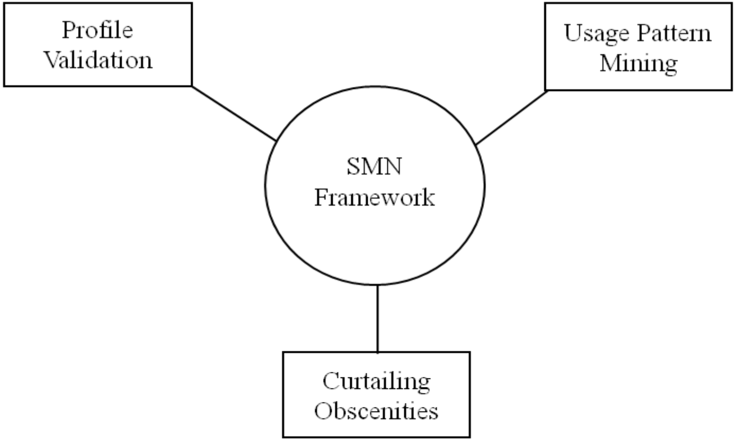


Figure 1. A Framework for Social Media Network

3.1 Profile Validation

The user of SMN has to register by providing the email ID, phone number, date of birth, education, experience etc. The email id and phone number have to be verified primarily by sending pin number to the corresponding accounts. Date related information provided by the user has been verified by correlating age with the education and experience given in the date fields. For example user has to complete the schooling within the age of 18 approximately, under graduation after the age of 21, post graduation after 23. Similarly work experience may include the age after 18. The school, college and working organization names and their locations have been verified and their interrelationship also has been validated.

The user profile validation has been further improved by posting the photo of the user. The users are insisted to provide the image of their own through webcam or selfie cameras. And country specified Social Security Number is also helpful to know the identity of the user. The process of validating the profile information has been simplified through this. If the user identities are known to the SMNs, they may not misbehave or misuse the networks.

3.2 Usage Pattern Mining

Profile similarity check method checks the text content of the profile being duplicated. Graph method is used for identifying the duplicate users by calculating degree. Semantic based approach is also difficult to confirm the duplicate user, since users may use different ids for different purposes. For example users may maintain official id and personal ids separately in a single SMN.

We can further improve the findings of duplicate users by identifying the usage patterns. Duplicate users may not use the social networks at a time with multiple ids. The suspected users' profile has been monitored and it has been verified through usage timings and IP address of the accessing system. Multiple ids with the same IP address of the system are considered for usage pattern verification. The malicious user's usage patterns are different from the normal users. They may keep on searching the new people regardless of their own profile. The location of the user may be same though the user uses multiple ids. Using longitude and latitude, the user's exact location can be identified, if the user enables to view his location details. If so it can be blocked after verification.

Some of the abnormal usage patterns are:

- The user always searching for the new friends.
- Sending friendly requests to befriend to many unknown users to whom the profile mat not match with the users' profile.
- Sharing unauthorized data's like photos, text messages, commenting others, asking bank account number, asking to join the unknown group, advertisements, asking for the details of the family members, etc.,

We have been keeping three list namely suspected user list, normal user list and malicious user list. If the user posts harmful messages and images, the user will be included directly to the malicious users' list. If the user sends friendly requests and surff frequently through the profiles of the various users, he has to be posted in suspected user list. Their activities are closely monitored and are verified for confirmation. Based on their behavior they have been posted either to the normal user or malicious users list. This malicious users list has to be intimated to the corresponding social media network for further follow up. It is their responsibility to suspend the malicious users' account. The process of usage pattern mining is shown in figure 2.

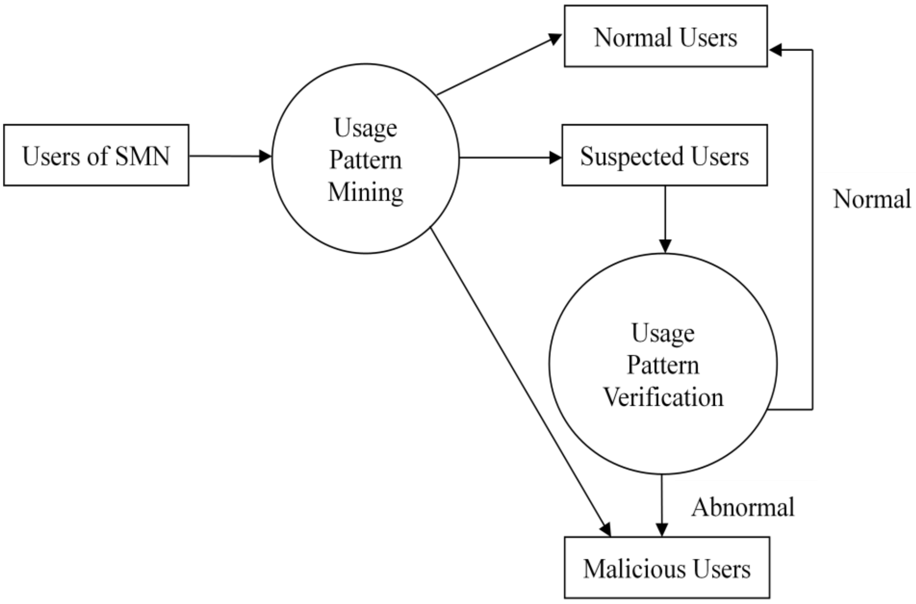


Figure 2. The process of Usage Pattern Mining

3.3 Curtailing Obscenities

This module is used for restricting the obscenities present in the content being shared. Most of the contents being shared in the social media network are textual contents. The users in sharing their views or conveying their feedbacks or giving their opinions and views have to express them through text content. Therefore the text messages are the important component involved in social media networks. The content being shared may include many worth full feedbacks. Some users they may express their views abruptly without any hesitation. This negative feedback also consider for improving the discussion. But bad words usage has to be strictly prohibited in the social media networks. We should have a social responsibility to not share the indecent content usage in social media network. But things become worse nowadays. People are using unwanted, unnecessary and abused words in their contents. Therefore the proposed framework has a mechanism of restricting those contents being shared.

The framework has the ability to monitor the content being shared in the social media network. The natural language processing and Bayesian Network algorithm has been used to monitor the content being shared. We have proposed an algorithm called CO for Curtailing Obscenities as shown in Algorithm 1.

Algorithm 1. CO

```
Input: Dictionary, W[]
Output: Curtail W
function CO (Dictionary, W)
    PC=0, NC=0
    if W[i-1] ε Dictionary.IW and W[i] ε Dictionary.PW do
        PC = PC + H
    elseif W[i-1] ε Dictionary.DW and W[i] ε Dictionary.PW do
        PC = PC + L
    elseif W[i] ε Dictionary.PW do
        PC = PC + M
    elseif W[i-1] ε Dictionary.IW and W[i] ε Dictionary.NW do
        NC = NC + H
    elseif W[i-1] ε Dictionary.DW and W[i] ε Dictionary.NW do
        NC = NC + L
    elseif W[i] ε Dictionary.NW do
        NC = NC + M
    compute PV using PC and NC
```

```
if PV > T do
    delete W
else
    return W
```

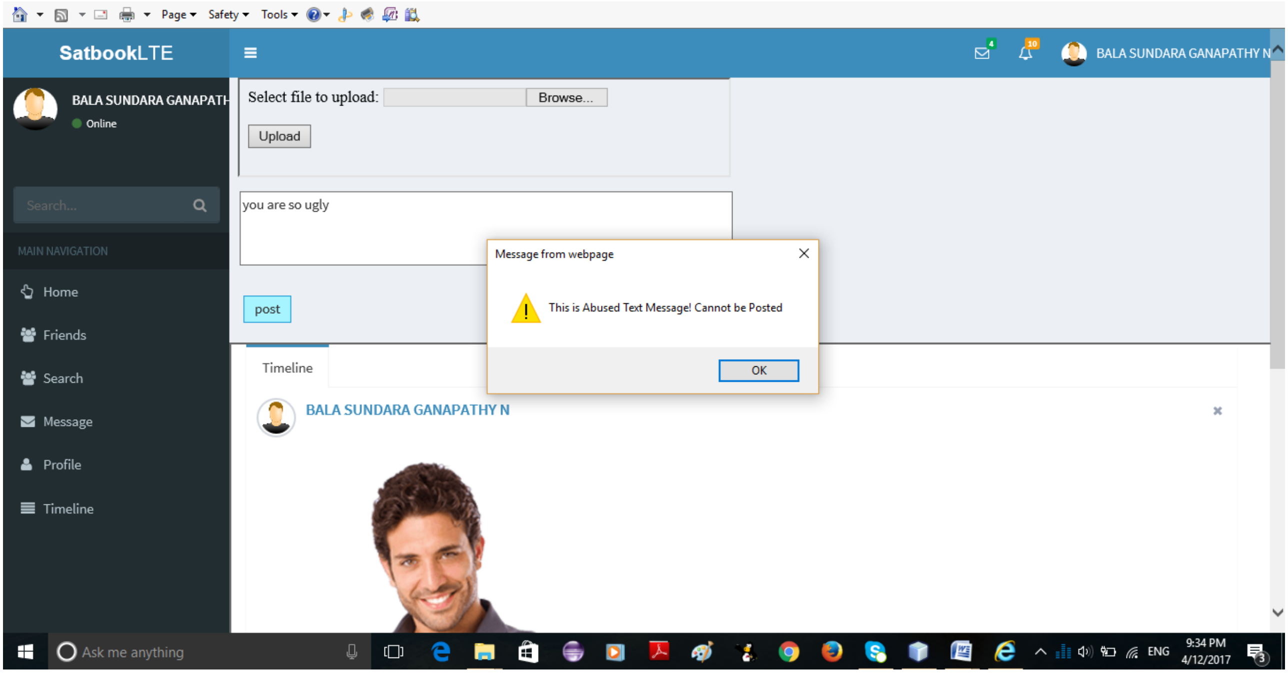
In CO algorithm, we have used a bunch of words approach, in which a list of positive (PW) and negative (NW) words/phrases are used to check the text content of the message being shared. Words that increase or decrease the severity of the meaning is coded as incrementing words (IW) and decrementing words (DW) in order to improve the findings of obscenities. The words/phrases are classified as positive words, negative words, positive with incrementing words, positive with decrementing words, negative with incrementing words and negative with decrementing words. According to the nature of words, the weights have been assigned as higher, moderate, and lower weights (HW, MW and LW respectively). Highly sensitive abused words may be assigned with higher weights (HW).The cumulative weights have been considered for restricting the content being shared. The probability value (PV) is computed using PC and NC. If the PV is greater than the threshold value (T), the content being shared will be restricted. Otherwise the text message has been shared in the SMN's. The combination of different words which forms a sentence that leads to wrong meaning is considered as a second stage of monitoring to be proposed. The synonym of the sentences has been evaluated then restricted if it gives wrong meaning. Ontology inference service has to be used to verify the content of the message according to the meaning. In the third stage, users may feed their text in their colloquial language. These colloquial terms have been restricted in sharing. A word or phrase being shared has been checked with the WordNet dictionary and other dictionary services, which are helpful to identify the colloquial term being used by the users. The framework formed in this way provides three stage protecting mechanism to curtail obscenities contents that are shared in the social media networks.

Experimental Studies

The proposed framework has been implemented through sample social media network called Satbook LTE. The system has functionalities that are very similar to the popular social media network. User can register with their email id or phone number in Satbook LTE. Once the user authentication is verified, they can share text messages. User profile information is verified through email id or phone number. The verification code will be send to the corresponding email id or phone number. On successful verification, the user's account is enabled for use. Until then user accounts has been disabled. We have validated other profile information moderately and the usage pattern mining is also in proposed level. Most importantly the content being shared in SMN's is validated in full-fledged manner.

4.1 Application for enforcing the Curtailment of Obscenities

The user messages that are posted in their page have been monitored through natural language processing. The privacy of the data posted has been maintained since the process is automated and there is no human intervention during the processing. The system calculates the weighted average of the message being posted. The system restricts messages that are unusual or abnormal as shown in the figure 3. The messages that are below the threshold value have been posted. The framework validates each and every message being posted. In this way the abnormal or abused messages have been restricted in posting and sharing in the social media networks. If we follow these strategies in all social media networks, we can resolve many problems encountered in the social media networks as discussed in our paper [30] entitled "Risk factor analysis of social media network".



Select file to upload:

you are so ugly

Message from webpage

This is Abused Text Message! Cannot be Posted

Figure 3. SatbookLTE with Restricting Obscenities.

To validate the performance of the CO, we have conducted the experiment with fifty messages of both normal and abnormal categories. Each message is successfully validated by the system and the system has restricted the abused messages from posting. The figure 4 denotes the status of the message during posting. The threshold value being set in the system for identifying the abused message is 0. The messages with weights greater than the threshold are allowed to be posted. This is shown in green star points in figure 4. The red star points denote the abused messages that are being restricted by the system.

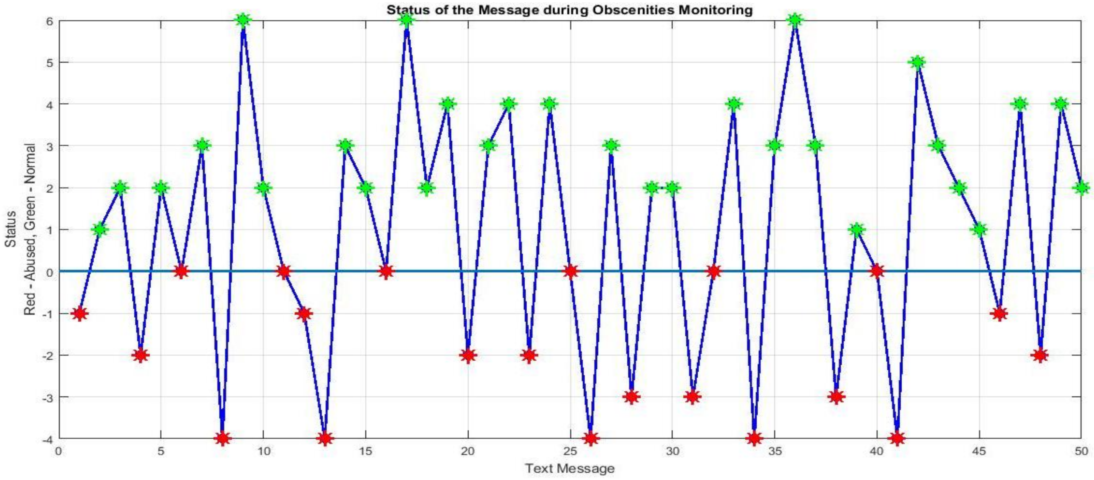


Figure 4. Status of the Message during Obscenities Monitoring.

Conclusion and Future Enhancement

The proposed framework is very much useful for the social media network users. Especially the countries that have steeped in their conservative culture are adapted to the framework, preserving their social importance and the values of their culture. The framework includes the methodology for validating the user profiles, user's usage pattern mining and restricting the abused text messages. In future the extension of this to social media network to restrict all other content types such as images, clipart, portable document format, video and audio is to be considered.

References

1. D. Perito, C. Castelluccia, M. A. Kaafar, and P. Manils, “How unique and traceable are usernames?” in Proc. 11th Int. Conf. Privacy Enhancing Technol., 2011, pp. 1–17.
2. J. Liu, F. Zhang, X. Song, Y. I. Song, C. Y. Lin, and H. W. Hon, “What's in a name?: An unsupervised approach to link users across communities,” in Proc. 6th ACM Int. Conf. Web Search Data Mining, 2013, pp. 495–504.
3. R. Zafarani and H. Liu, “Connecting corresponding identities across communities,” in Proc. 3rd Int. ICWSM Conf., 2009, pp. 354–357.
4. R. Zafarani and H. Liu, “Connecting users across social media sites: a behavioral-modeling approach,” in Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2013, pp. 41–49.
5. A. Acquisti, R. Gross, and F. Stutzman, “Privacy in the age of augmented reality,” in Proc. Nat. Acad. Sci., 2011, pp. 36–53, Available: https://www.usenix.org/legacy/events/sec11/tech/slides/ acquisti.pdf
6. T. Iofciu, P. Fankhauser, F. Abel, and K. Bischoff, “Identifying users across social tagging systems,” in Proc. 5th Int. AAAI Conf. Weblogs Social Media, 2011, pp. 522–525.

7. M. Motoyama and G. Varghese, “I seek you: searching and matching individuals in social networks,” in Proc. 11th Int. Workshop Web Inf. Data Manage., 2009, pp. 67–75.
8. O. Goga, D. Perito, H. Lei, R. Teixeira, and R. Sommer, “Large-scale correlation of accounts across social networks,” University of California at Berkeley, Berkeley, California, Tech. Rep. TR-13-002, 2013.
9. K. Cortis, S. Scerri, I. Rivera, and S. Handschuh, “An ontology based technique for online profile resolution,” in Proc. 5th Int. Conf. Social Informat., 2013, pp. 284–298.
10. F. Abel, E. Herder, G. J. Houben, N. Henze, and D. Krause, “Cross system user modeling and personalization on the social web,” User Model. User-Adapted Interaction, vol. 23, pp. 169–209, 2013.
11. O. De Vel, A. Anderson, M. Corney, and G. Mohay, “Mining e-mail content for author identification forensics,” ACM Sigmod Rec., vol. 30, no. 4, pp. 55–64, 2001.
12. E. Raad, R. Chbeir, and A. Dipanda, “User profile matching in social networks,” in Proc. 13th Int. Conf. Netw.-Based Inf. Syst., 2010, pp. 297–304.
13. J. Vosecky, D. Hong, and V. Y. Shen, “User identification across multiple social networks,” in Proc. 1st Int. Conf. Netw. Digital Technol., 2009, pp. 360–365.
14. P. Jain, P. Kumaraguru, and A. Joshi, “@ i seek ‘fb. me’: Identifying users across multiple online social networks,” in Proc. 22nd Int. Conf. World Wide Web Companion, 2013, pp. 1259–1268.
15. P. Jain and P. Kumaraguru, “Finding Nemo: searching and resolving identities of users across online social networks,” arXiv preprint arXiv:1212.6147, 2012.
16. R. Zheng, J. Li, H. Chen, and Z. Huang, “A framework for authorship identification of online messages: Writing style features and classification techniques,” J. Amer. Soc. Inf. Sci. Technol., vol. 57,no. 3, pp. 378–393, 2006.
17. M. Almishari and G. Tsudik, “Exploring linkability of user reviews,” in Proc. 17th Eur. Symp. Res. Comput. Security, 2012, pp. 307–324.
18. X. Kong, J. Zhang, and P. S. Yu, “Inferring anchor links across multiple heterogeneous social networks,” in Proc. 22nd ACM Int. Conf. Inf. Knowl. Manage., 2013, pp. 179–188.
19. O. Goga, H. Lei, S. H. K. Parthasarathi, G. Friedland, R. Sommer, and R. Teixeira, “Exploiting innocuous activity for correlating users across sites,” in Proc. 22nd Int. Conf. World Wide Web, 2013, pp. 447–458.
20. A. Narayanan and V. Shmatikov, “De-anonymizing social networks,” in Proc. IEEE 30th Symp. Security Privacy, 2009, pp. 173–187.

21. S. Bartunov, A. Korshunov, S. Park, W. Ryu, and H. Lee, “Joint link-attribute user identity resolution in online social networks,” in Proc. 6th SNA-KDD Workshop, 2012.
22. N. Korula and S. Lattanzi, “An efficient reconciliation algorithm for social networks,” arXiv preprint arXiv:1307.1690, 2013.
23. L. Backstrom, C. Dwork, and J. Kleinberg, “Wherefore art thou r3579x?: Anonymized social networks, hidden patterns, and structural steganography,” in Proc. 16th Int. Conf. World Wide Web, 2007, pp. 181–190.
24. B. Zhou and J. Pei, “Preserving privacy in social networks against neighborhood attacks,” in Proc. 24th IEEE Int. Conf. Data Eng., 2008, pp. 506–515.
25. B. Zhou and J. Pei, “The k-anonymity and l-diversity approaches for privacy preservation in social networks against neighborhood attacks,” Knowl. Inf. Syst, vol. 28, no. 1, pp. 47–77, 2011.
26. K. Liu and E. Terzi, “Towards identity anonymization on graphs,” in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2008, pp. 93–106.
27. X. Ying and X. Wu, “Randomizing social networks: A spectrum preserving approach,” in Proc. SIAM Int. Conf. Data Mining, 2008, pp. 739–750.
28. Xiaoping Zhou, Xun Liang, Senior Member, IEEE, Haiyan Zhang, and Yuefeng Ma, “ Cross-Platform Identification of Anonymous Identical Users in Multiple Social Media Networks” IEEE Transactions On Knowledge And Data Engineering, Vol. 28, No. 2, February 2016 pp. 411-424
29. http://www.information-age.com/child-sexual-abuse-social-networks-123462234/
30. Bala Sundara Ganapathy N, and Mohana Prasad K, “Risk Factors Analysis of Social Media Networks,” Research Journal of Pharmaceutical, Biological and Chemical Sciences, March–April 2017, RJPBCS 8(2), pp. 1940 – 1946.