

Cross System: Identifying Anonymous User Using Multiple Social Media Network

^{#1}Ms. Rohini N Khedkar, ^{#2}Prof. Sujata Bhairnallykar, ^{#3}Prof. Vaishali Jadhav



¹rohnikhedkar10@gmail.com

²bsujata5@rediffmail.com

³vaishalijadhav_2000@yahoo.com

^{#1}M.E Student, Dept. of Computer Engineering,

^{#2}Asst. Prof, Dept. of Computer Engineering

^{#3}Assot Prof, Dept. of Information Technology.

Saraswati College of Engineering, Mumbai University

ABSTRACT

In today's generation, everyone uses online social network for their different purpose. In-between different social media networks some personal information is shared to all on social media network. That shared information is misused by fake (anonymous) users. Fake user collect the information, misuse it as blackmail to user, create good relation to get their belief and then they make some social crimes. This happens day by day. To solve this problem cross platform exploration may help in social computing in both theory and real time. Some technique used to identify the anonymous user focused on public profile attribute, location and timing of user content as well as writing style. Existing user identification based on network structure are not effective. To identify anonymous user system proposed approach for friend relationship based user identification algorithm .also use profile based identification and content based identification technique. Proposed System test on two social media networks-Facebook and LinkedIn.

Keywords: Cross Platform, FRUI, SMN Profile and content based Identification technique.

ARTICLE INFO

Article History

Received: 1st August 2018

Received in revised form :

1st August 2018

Accepted: 3rd August 2018

Published online :

11th August 2018

I. INTRODUCTION

In these year, many types of social networking sites have emerged and contributed immensely to large volumes of real-world data on social behaviors. Twitter 1, the largest microblog service has more than 600 million users and produces upwards of 340 million tweets per day [2]. Sina Microblog 2, the primary Chinese microblog social media website, has more than 500 million accounts and generates well over 100 million tweets per day [4]. Due to this diversity of online social media networks (SMNs), people tend to use different SMNs for different purposes. For sharing information to friends, to be in contact with friends, to share new ideas to others on social media. Similarly for instance, Ren Ren 3, is used in China for posting blogs on social media, while Sina Microblog is used to share statuses [4]. In other words, every existent social media network satisfies some user needs. Facebook also used for sharing status, different thoughts, posting blogs and also to keep connected with all friends [3].

Now day's peoples having lots of accounts on different social media networks of a same person. Who are using the

lots of accounts on different social media they are using those accounts for different purpose. Like getting personal information of some people, making frauds using it and for blackmailing purpose. Now fake users count on social media network is going on increasing continuously.

In terms of SMN management, matching anonymous users across different SMN platforms can provide integrated details on each user and inform corresponding regulations, such as targeting services provisions. In theory, the cross-platform explorations allow a bird's eye view of SMN user behaviors [1]. However, nearly all recent SMN based studies focus on a single SMN platform, due to that getting incomplete data. Therefore, this study investigates the strategy of cross-system multiple social media network platforms to paint a comprehensive picture of the behaviors of the different users. With the growth of SMN platforms on the Internet, the cross-platform approach has merged various SMN platforms to create richer raw data and more complete SMNs for social computing tasks. SMN users form the natural bridges for these SMN platforms. The

primary topic for cross-platform SMN research is user identification for deferent SMNs. Exploration of this topic lays a foundation for further cross-platform SMN research. SMN connections fall into two categories: single following connections and mutual following connections. Single following connections are also called following relationships or following links. If user A follows user B, then user A and user B have a following relationship (single way fans in which one knows the other, but not vice versa). Following relationships are common in microblogging SMNs, such as Twitter and Sina Microblog. Likewise, mutual following connections are called friend relationships [1]. In microblogging SMNs, a friend relationship refers to the mutual following relationships between two users. In most other SMNs, such as Facebook, Ren Ren and WeChat, a friend relationship forms only if a friend request is sent by one user and Confirmed by the other user. Friend relationships are difficult to fake by malicious users, and therefore react real-world relationships much better. Due to their reliability and consistency, friend relationships are more robust in user identification tasks. Moreover, since unified friend relationships are formed, our algorithm can also be applied to SMNs with a heterogeneous network structure, such as Twitter and Facebook.

II. RELATED WORK

In the today's generation, the social life of everyone has become associated with online social network. These site have made a drastic change in the way we pursue our social life. People tend to use different SMNs for different purposes for sharing ideas, thoughts, and posting blogs and connected to each other. Because of different social media networks some personal information is shared to all on social media network. That shared information is misused by fake users. Fake user collect the information, misuse it as blackmail to user, create good relation to get their belief and then they make some social crimes

In [5] D. Perito, C. Castelluccia, M.A. Kaafar, and P. Manils, This paper explores the possibility of linking user's profiles only by looking at their usernames. The intuition is that the probability that two usernames refers to the same physical person strongly depends on the entropy of the username string itself. **"How Unique and Traceable Are Username?"**

In [6] Reza Zafarani proposed connecting user identities across communities by a process called link analysis algorithm. The relationship between usernames selected by one person in different social media network, and on some of the web regarding usernames and communities **"Connecting Corresponding Identities across Communities."**

In [7] M. Motoyama and G. Varghese proposed searching and matching individuals in social networks. An online user joins multiple social networks in order to enjoy different services. On each joined social network, she creates an identity and constitutes its three major dimensions namely profile, content and connection network. She largely governs her identity formulation on any social network and therefore can manipulate multiple

aspects of it." **I Seek You: searching and matching individuals in social media network."**

In [8] Paridhi Jain, *et al* proposed finding nemo: Searching and resolving identities of users across online social networks using algorithm called profile search.. **"Finding Nemo: Searching and Resolving Identities of Users Across Online Social Media Network"**.

In [9] Reza Zafarani, *et al* proposed connecting users across social media sites: A behavioral-modeling approach by an algorithm called learning algorithm. The proposed behavioral modeling approach exploits information redundancy due to these behavioral patterns. An alternative solution addressing the age verification problem by exploiting the nature of social media and its networks. The information available on all social media sites (usernames) to derive a large number of features that can be used by supervised learning to connect users across sites [8]. Users often exhibit certain behavioral patterns when selecting usernames. It includes analyzing these possibilities and discovering features indigenous to specific sites, beyond those constricted to usernames, and incorporating them into MOBIUS for future needs. **"Connecting Users across Social Media Sites: A Behavioral-Modeling Approach"**.

In [11] Tereza Iofciu¹, Peter Fankhauser¹, Fabian Abel², Kerstin Bischoff this paper study those questions and investigate whether users can be identified across social tagging systems. Authors combine two kinds of information: their user ids and their tags. Author's introduce and compare a variety of approaches to measure the distance between user profiles for identification [11]. **"Identifying Users Across Social Tagging Systems."**

In [12] Oana Goga, Daniele Perito, Howard Lei, Renata Teixeira, and Robin Sommer proposed Large-scale Correlation of Accounts across Social Networks. Organizations are increasingly mining the personal data users generate as they carry out much of their day-to-day activities online. Authors perform a large scale study that assesses a range of correlation approaches for matching accounts between five popular social networks: Twitter, Facebook, Google+, Myspace, and Flickr. **"Large-scale Correlation of Accounts across Social Networks"**

III. PROPOSED SYSTEM

Sharing of personal information on different SMN, lots of cybercrime, robberies, murders are going to happen. This all happen by fake user (anonymous user). To detect anonymous users System proposed approach of Friend Relationship based User Identification algorithm (FRUI). In this system to detect anonymous user FRUI applied first. Then outcome of FRUI get applied profile based and content based algorithm. Finally that anonymous user will be detected. Since FRUI employs a Unified friend Relationship, it is to identify users from a heterogeneous network structure. FRUI chooses candidate matching pairs from currently known identical users rather than unmapped ones. This operation reduces computational complexity, since only a very small portion of unmapped users are

involved in each iteration. Moreover, since only mapped users are exploited, our solution is scalable and can be easily extended to online user identification applications.

ALGORITHM 1: FRUI

Input: SMNA, SMNB, Priori UMPs: PUMPs

Output: Identified UMPs: UMPs

```

1: function FRUI (SMNA, SMNB, PUMP)
2: T= { }, R= dist ( ), S=PUMPs, L= [ ], max=0, FA= [ ],
   FB= [ ]
3: while Sis not empty do
4: Add S to T
5: if max>0 do
6: Remove S from L[max]
7: while L[max] is empty
8: max=max-1
9: if max= =0 do
10: return UMPs
11: Remove UMPs with mapped UE from L[max]
12: for each UMPa~B (i, j) in S do
13: for each UEAi in the unmapped neighbors of UEAi do
14: FA[i] =FA[i] +1
15: for each UEAb in the unmapped neighbors of UEAj do
16: R [UMPA~B (a, b)] +=1, FB[j] =FB[j] +1
17: Add UMPA~B (a, b) to L[R [UMPA~B (a, b)]]
18: if R [UMPA~B (a, b)]>max do
19: max=R [UMPA~B (a, b)]
20: m=max, S= { }
21: while s is empty do
22: Remove UMPs with mapped UE from L[max]
23: C = L[m], m = m - 1, n = 0
24: S = {un-Controversial UMPs in C}
25: while S is empty do
26: n = n + 1, I = {UMPs with top n Mij in C using (5)}
27: S = {un-Controversial UMPs in I}
28: if I == C do
29: break
30: return T
    
```

3.1 System Architecture

The basic architectural block diagram for proposed model is as shown in the figure 1.1: In this System model considers two types of entities: Admin and user. Admin will have the access of both social media network (Facebook and LinkedIn) which we are using in our system. First Admin

log in both social media network. User first register on social media network then login .user entry goes into the data base. Then user create many account and these entry goes into the database .admin detect the new user .on this user admin apply the algorithm such as friend relationship based user identification, profile based, content based, network based. After applying these algorithm admin detect anonymous user.

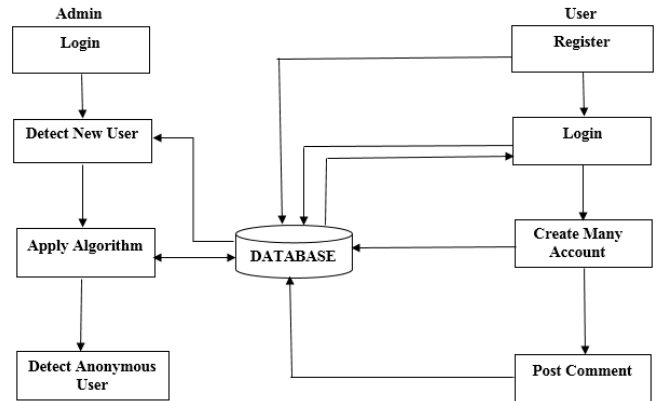


Fig 1.1 Proposed System Architecture

Following technique used in proposed system.

1) Profile based technique: -

Identity of a user on a social network includes a set of profile attributes, which gives basic information about the user such as username, name, location, gender, description etc. If the user does not demonstrate any active obfuscation and does not create altogether a different identity, it is likely that she re-uses certain profile attributes' value, on the social networks she joins. If the user demonstrates such behavior, profile attributes can be used to find her identity on other social networks. Profile Search method explores "Profile" dimension of a user's identity and exploits profile attributes as linking attributes. To make comparisons between any two identities using linking attributes, it is essential to have same set of attributes publicly available for both identities profile based user identification have focused on public profile attribute, including screen name, age ,birth date, Gender, and profile images.

2) Content based technique: -

An identity of a user on a social network includes the content that she creates or is shared with her .Content-Based User Identification solutions attempt to recognize users based on the times and locations that users post content, as well as the writing style of the content.

3) Network based technique: -

Network is an important dimension of a user's identity on a social network. It is a shared identity of a user build with the involvement of other users .if other users leak their identity on any other social network, it is likely that the user's identity also gets leaked. Network Search algorithm explores the possibility of a user's identity leak via her network attribute. Network structure-based studies on user identification across multiple SMNs are used to recognize identical users solely by user network structures and seed or priori, identified users.

IV. CONCLUSION

This paper address the problem of user identification across multiple SMN plat-forms and offered an innovative solution. As a key aspect of SMN, network structure is importance and helps to resolve user identification tasks problem. The proposed system for a friend relationship based algorithm called Friend Relationship based User Identification, calculates a match degree for all candidate User Matched Pairs (UMPs), and only UMPs with top ranks are considered as identical users. In scenarios when raw text data is sparse, incomplete, or hard to obtain due to privacy settings, FRUI is extremely suitable for cross-platform tasks.

V. REFERENCE

- [1] Xiaoping Zhou, Xun Liang, Senior Member, IEEE, Haiyan Zhang, Yuefeng Ma "Cross-Platform Identification of Anonymous Identical Users in multiple social media network. ". Citation information: DOI 10.1109/TKDE.2016.2485222, IEEE Transactions on Knowledge and Data Engineering.
- [2] Wikipedia, "Twitter, "<http://en.wikipedia.org/wiki/Twitter>. 2014.
- [3] Wikipedia, "Facebook" <https://en.wikipedia.org/wiki/Facebook>.
- [4] Xinhuanet, "Sina Microblog Achieves over 500 Million Users," http://news.xinhuanet.com/tech/201202/29/c_122769084.htm. 2014.
- [5] D. Perito, C. Castelluccia, M.A. Kaafar, and P. Manils, "How unique and traceable are usernames?" Privacy Enhancing Technologies (PETS'11), pp. 1-17, 2011.
- [6] R. Zafarani and H. Liu, "Connecting corresponding identities across communities," Proc. of the 3rd International ICWSM Conference, pp. 354-357, 2009.
- [7] M. Motoyama and G. Varghese, "I seek you: searching and matching individuals in social networks," Proc. of the 11th international workshop on Web Information and Data Management (WIDM'09), pp. 67-75, 2009.
- [8] P. Jain and P. Kumaraguru, "Finding Nemo: searching and re-solving identities of users across online social networks," arXiv preprint arXiv: 1212.6147, 2012.
- [9] R. Zafarani and H. Liu, "Connecting users across social media sites: a behavioral-modeling approach, " Proc. of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'13), pp.41-49, 2013.
- [10] 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," Proc. of the 6th ACM international conference on Web search and data mining (WDM'13), pp. 495-504, 2013.
- [11].T. Iofciu, P. Fankhauser, F. Abel, and K. Bischoff, "Identifying users across social tagging systems," Proc. of the 5th International AAAI Conference on Weblogs and Social Media, pp. 522-525, 2011.
- [12] O. Goga, D. Perito, H. Lei, R. Teixeira, and R. Sommer, "Large-scale Correlation of Accounts across Social Networks," Technical report, 2013.