International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue 17
December 2017

# Profile Matching Of Anonymous Users in Multiple Social Media Networks

**Cheemaladinne Kondaiah**

M.Tech Student, Dept. of Computer Science and Systems Engineering,
AU College of Engineering (A),
Visakhapatnam, India

***Abstract:***

The SMN proves to be the best platform for information retrieval. However, identifying unknown and identical users on multiple social media application is still an unsolved problem. People use different social media for different purpose; the idea of integrating multiple social media application can take the research a step forward. The main idea of this project is to identify alias and identical accounts by merging multiple SMN in order to get complete information about a particular user. In social media networks, profile details of one user can be used by others to create account with original user identity or the original user may have multiple accounts in multiple social media sites. Discovery of multiple accounts that belong to the same person is an interesting and challenging work in social media analysis. Profiles, contents and network structures can be used for user identification in social media sites. In this project, we develop a methodology Friend Relationship - Based User Identification (FRUI) algorithm for mapping individuals on cross application SMN's. The friend cycle of every individual differs therefore, accuracy of our result will be maintained if we use friend list as a key component to analyse cross application social media networks. We also focus on using two more methods to improve efficiency of our algorithm. Our study has shown that FRUI is effective to analyze and de -anonymize social media.

**Keywords:** Cross-Application, Social Media Network, Anonymous Identical users, Friend Relationship based User Identification.

## 1. INTRODUCTION

Today, most of the people use social media sites. It is obvious that people tend to use different social media application for different purpose. Facebook, is a for-profit corporation and most popular social media application in the world, has more than 1.7 billion users. Twitter is an online social networking service that enables users to send and read short 140-character messages called "tweets". At the second half of 2016 the number of registered users was more than 313 million users. Registered users can read and post tweets, but those who are unregistered can only read them. Instagarm is a mobile photo sharing network which has reached 500 million of active users in the month of September 2015.

Every social media network is famous for its distinct features, for e.g. facebook is used to connect with people all over the world and exchange their thoughts through messaging. Twitter provides microblog service where people tweet or share their opinion. So we can conclude that every existing social media application is build to satisfy some user needs.

To analyze user's profile, we will require complete knowledge about the user. Single application social media network provides us with incomplete information which degrades the accuracy to analyze the user as anonymous or not. The idea of cross-application social media network can be applied here. In this paper we present a method Friend List based User Identification (FRUI) for mapping individuals on

cross-application social media network. Our proposed system uses the friend list and additional information of user available on different social media to calculate the better results. Profile contains different information (public posts, friends, photos personal information). As private data is not possible to retrieve we collect public posts of user on different social media

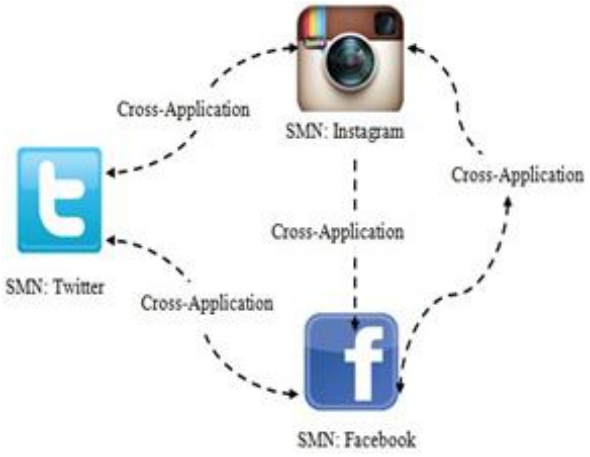


Fig. 1 Cross-application research to merge a variety of SMN's

## II. RELATED WORK

Xiangnan Kong, et al [4] proposed inferring anchor links across multiple social networks by a technique called multi-network anchoring. The proposed Multi-Network Anchoring (MNA) method performs other baseline methods. Multiple social networks can provide different types of information about the users. By explicitly consider the users complete data within the networks. It shows that by incorporating the one-to-one constraint in the inference process can improve the performance of anchor link prediction. "Inferring Anchor Links across Multiple Heterogeneous Social Networks".

Reza Zafarani [5] 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. The most social media network preserve the anonymity of users by allowing them to freely select usernames instead of their real names and the fact that different websites allows different username and security systems. If there exists a mapping between usernames across different social media networks and the real identities behind them, then connecting social media application across the web becomes an easy task. "Connecting Corresponding Identities across Communities".

Paridhi Jain, et al [6] proposed identifying users across multiple online social media application by a technique called identity search algorithms. They introduced two novel identity search algorithms based on content and network attributes and search algorithm based on prole attributes of a user that exploiting multiple identity search methods, a new technique to identify users unlike existing methods (e.g., similar name) and thus, increases the efficiency of finding correct matching users across social media applications. In this work, they attempt to understand if search methods based on an identity's content and network attributes, along with search methods based on an identity's prole International Journal of Engineering Science and Computing, March 2016 2739 http://ijesc.org/ attributes. "Identifying Users Across Multiple Online Social Networks".

Nitish Korula, et al [7] proposed an efficient algorithm by merging learning algorithms for social networks. A deeper understanding of the characteristics of a user across different networks helps to construct a better portrait of her, which can be used to serve personalized content or advertisements to the best of our knowledge, it has not yet been studied formally and no rigorous results have been proved for it. Even if certain

behaviour can be observed in several networks, there are still serious problems because there is no systematic way to combine the behaviour of a specific user across different social networks and because some social relationships will not appear in any social network. For these reasons, identifying all the accounts belonging to the same individual across different social services is a fundamental step in the study of social science. "An Efficient Reconciliation Algorithm for Social Networks".

## III.PROPOSED SYSTEM

We developed a Friend Relationship Based User Identification algorithm (FRUI). FRUI assumes every user has a unique friend circle; this is used to identify users across multiple social applications. Unlike existing algorithms [4], [6], [7], FRUI chooses a 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. In contrast with current algorithms [4], [6], [7], FRUI requires no control parameters.

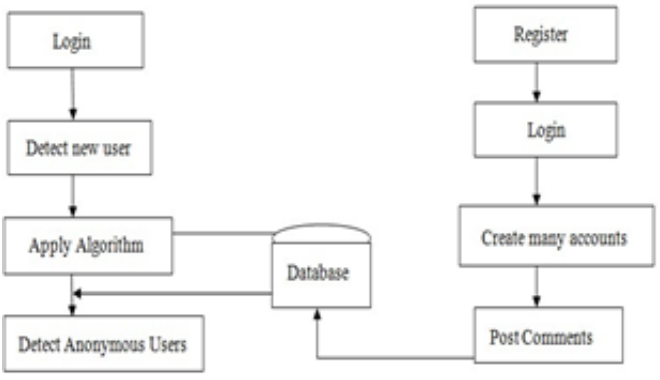


Fig. 2 System Architecture

## IV. IMPLEMENTATION

In this paper we use Friend Relationship-Based User Identification (FRUI) to match a degree of all candidates User Matched Pairs (UMP), and only UMP's with top ranks are considered as identical users. We also scrutinize the identical profiles and find out the common attributes to improve the accuracy of our algorithm.

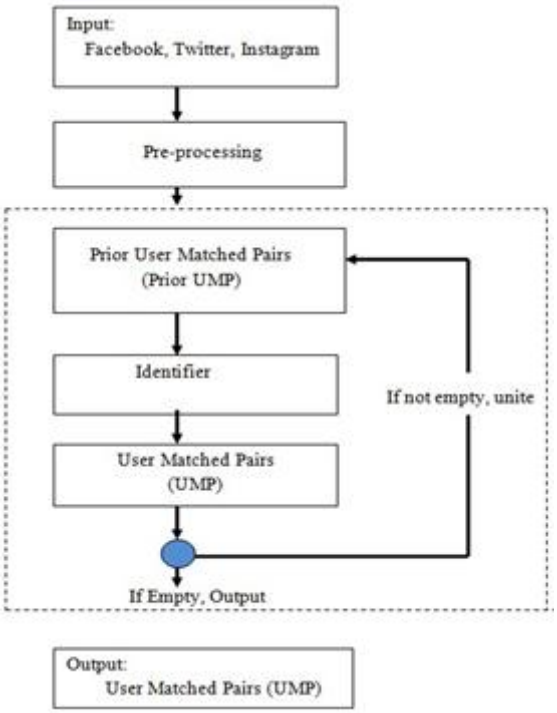


Fig. 3 Network structure based user identification

The network structure based user identification first obtains a Prior UMP'S through pre-processor, and then identifies more UMPs through the Identifier in an iteration process.

**Definition 1** (SMN) An SMN is defined as SMN= {U, C, I}, where U, C and I denote the users, connections and interactions among users, respectively.

**Definition 2** (User Matched Pair). Two social media networks SMNA and SMNB, if UEAi and UEBj belong to the same user in real-life, which is denoted as Ψ, then we hold that UEAi and UEBj match on Ψ, and they compose a User Matched Pair UMPΨ. UMP Ψ can also be expressed as UMPA~B (i, j) or UMP (UEAi, UEBj), equivalently.

Definition 3 (Priori UMP). Priori UMPs are UMPs given before user identification work is executed. Priori UMPs are used as the condition to identify more UMPs.

**Phases of the system:**

i. The six phases of the system are:
ii. Network Structure Based User Identity
iii. User Matched Pair
iv. User Identification
v. Friend Relationship Based User Identification (FRUI)Algorithm
vi. Notification of identical accounts & anonymous user

**FRU ALGORITHM**

Input :SMNA,SMNB,PrioriUMPs: PUMPs
Output:Identified UMPs:
Function FRUI(SMNA,SMNB,PUMPS)
T={}, R=dict(),s=PUMPs,L=[],max=0,FA=[],fb[]
While s is not empty do
Add s to T
If max >0 do
Remove s fromL[max]
While L[max]is empty
Max=max-1
If max==0 do
Return UMPs
Remove UMPswith mapped UE from L[max]
Foreach UMPA-B(I,j) in S do
Foreach UEAa in the unmapped neighbors of UEAi do
FA[i]=FA[i]+1
Foreach UEAb in the unmapped neighbors of UEAjdo
R[UMPA-B(a,b)]+=1,FB[j]=FB[j]+1
Add UMPA-B(a,b)to L[R[UMPA-B(a,b)]]
If R[UMPA-B(a,b)]>max do
Max=R[UMPA-B (a,b)]
M=max,S={}
While S is empty do
Remove UMPs with mapped UE from L[max}
C=L[m],m=m-1,n=0
S={uncontroversial UMPs in C}
While S is empty do
N=n+1,I={UMPs with top n Mij in C
S={ uncontroversial UMPs in I}
If I==C do
Break
Return T

**CONCLUSION**

Our study addresses the intractable problem of unknown user identification across SMN applications and offers an innovative solution. We will also use an algorithm friend relationship-based algorithm called FRUI. To improve the accuracy of FRUI, we described two propositions and addressed the complexity. We expect the result that the network structure can accomplish important user identification work. Our FRUI algorithm is simple, yet efficient, and performed much better than NS, the existing state-of-art network structure-based user identification solution. FRUI is extremely suitable for cross-application tasks when raw text data is sparse, incomplete, or hard to obtain due to privacy settings. In addition, our solution can be easily applied to any SMNs with friendship structure, including Twitter, Facebook and Instagram. It can also be extended to other studies in social computing with cross- application problems. Since only the adjacent users are involved in each iteration process, our method is scalable and can be easily applied to large datasets and online user identification applications. Identifying unknown users across multiple SMNs is challenging work. Therefore, only a portion of identical users with

different nicknames can be recognized with this method. This study will take the research a step forward. Other user identification methods can be applied simultaneously to study multiple social media application.

## FUTURE SCOPE

The use of social media application is increasing day by day.the misuse of social media will also go on increasing. Our project can prevent the user from being cheated. The identical accounts on different social media networks will be available which will help the user to get complete information. The anonymous activity on social media will be reduced to a great extent. As our proposed system is based on cross-application, if the system needs to be used in real time it may require the databases of all the social media application involved. Our FRUI algorithm uses friendship structure; in order to use this algorithm on other social media application it must have the friendship structure identifying unknown users across social media network is a challenging task. Therefore, only a portion of identical users with matching nicknames can be recognized with this method. Our system will work only if the users are registered on multiple social media sites (Facebook, Instagram, and Twitter). If the user is not registered on other SMN's our system may produce false results.

## REFERENCES

[1] Wikipedia, "Twitter, " http://en.wikipedia.org/wiki/Twitter. 2014.

[2] Xinhuanet, "Sina Microblog Achieves over 500 Million Users," http://news.xinhuanet.com/tech/2012-02/29/c_122769084.htm. 2014.

[3] D. Perito, C. Castelluccia, M.A. Kaafar, and P. Manils, "How unique and traceable are usernames?," *Privacy Enhancing Technol- ogies (PETS'11)*, pp. 1-17, 2011.

[4] 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.

[5] R. Zafarani and H. Liu, "Connecting corresponding identities across communities," *Proc. of the 3rd International ICWSM Con- ference*, pp. 354-357, 2009.

[6] 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. [7] A. Acquisti, R. Gross and F. Stutzman, "Privacy in the age of aug- mented reality," *Proc. National Academy of Sciences*, 2011.

[8] 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.

[9] M. Motoyama and G. Varghese, "I seek you: searching and matching individuals in social networks," *Proc. of the 11th inter- national workshop on Web Information and Data Management (WIDM'09)*, pp. 67-75, 2009.

[10] O. Goga, D. Perito, H. Lei, R. Teixeira, and R. Sommer, "Large-scale Correlation of Accounts across Social Networks," Tech- nical report, 2013.

[11] K. Cortis, S. Scerri, I. Rivera, and S. Handschuh, "An ontology-based technique for

online profile resolution," *Social Informatics*, Berlin: Springer, pp. 284-298, 2013.

[12] F. Abel, E. Herder, G.J. Houben, N. Henze, and D. Krause, "Cross-system user modeling and personalization on the social web," *User Modeling and User-Adapted Interaction*, vol. 23, pp. 169-209, 2013.

[13] O. De Vel, A. Anderson, M. Corney, and G. Mohay, "Mining e-mail content for author identification forensics," *ACM Sigmod Record*, vol. 30, no. 4, pp. 55-64, 2001.

[14] E. Raad, R. Chbeir, and A. Dipanda, "User profile matching in social networks," *Proc. Of the 13th International Conference on Network-Based Information Systems (NBiS'10)*, pp.297-304, 2010.

[15] J. Vosecky, D. Hong, and V.Y. Shen, "User identification across multiple social networks," *Proc. Of the 1st International Confer- ence on Networked Digital Technologies*, pp.360-365, 2009.

[16] P. Jain, P. Kumaraguru, and A. Joshi, "@ i seek 'fb. me': identify- ing users across multiple online social networks," *Proc. of the 22nd International Conference on World Wide Web Companion*, pp. 1259-1268, 2013.

[17] P. Jain and P. Kumaraguru, "Finding Nemo: searching and re-solving identities of users across online social networks," arXiv preprint arXiv:1212.6147, 2012.

[18] R. Zheng, J. Li, H. Chen, and Z. Huang, "A framework for au- thorship identification of online messages: writing- style fea- tures and classification techniques," *J. of the American Society for Information Science and Technology*, vol. 57, no. 3, pp. 378-393, 2006.

.

M.Tech Student, Dept. of Computer Science and Systems Engineering,
AU College of Engineering (A), Visakhapatnam, India