Scientific Journal of Impact Factor (SJIF): 5.71

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 5, Issue 04, April -2018

IDENTIFICATION OF ANONYMOUS USER ACROSS DIFFERENT SOCIAL NETWORKS

Prateeksha Nagargoje[1], Pooja More.[2], Pooja Patil[3], Tanaya Shelar[4]

[1]*Department of Computer Engineering , Indira College of Engineering and Management, Pune*
[2]*Department of Computer Engineering, Indira College of Engineering and Management, Pune*
[3]*Department of Computer Engineering , Indira College of Engineering and Management, Pune*
[4]*Department of Computer Engineering , Indira College of Engineering and Management, Pune*

Abstract — *The previous couple of years have witnessed the emergence and evolution of a vivacious analysis stream on an outsized form of on-line Social Media Network (SMN)platforms. Recognizing anonymous, however identical users among multiple SMNs remains associate refractory drawback. Clearly, cross-platform exploration might facilitate solve several issues in social computing in each theory and applications. Since public profiles will be duplicated and simply impersonated by users with completely different functions, most current user identification resolutions, that chiefly target text mining of users' public profiles, square measure fragile. Some studies have tried to match users supported the situation and temporal order of user content still as literary genre. However, the locations square measure thin within the majority of SMNs, and literary genre is tough to tell apart from the short sentences of leading SMNs like S in a very Microblog and Twitter. Moreover, since on-line SMNs square measure quite cruciform, existing user identification schemes supported network structure aren't effective. The real-world friend cycle is very individual and nearly no 2 users share a congruent friend cycle. Therefore, it's a lot of correct to use a friendly relationship structure to research cross-platform SMNs. Since identical users tend to line up partial similar friendly relationship structures in several SMNs, we have a tendency to planned the Friend Relationship-Based User Identification (FRUI)algorithm. FRUI calculates a match degree for all candidate User Matched Pairs (UMPs), and solely UMPs with prime ranks square measure thought-about as identical users. we have a tendency to additionally developed 2 propositions to boost the potency of the algorithmic program. Results of intensive experiments demonstrate that FRUI performs far better than current network structure-based algorithms*

Keywords- *Cross-Platform, Social Media Network, Anonymous Identical Users, Friend Relationship, User Identification*

I. INTRODUCTION

In the last decade, many sorts of social networking sites have emerged and contributed vastly to massive volumes of real-world information on social behaviors. Twitter 1, the most important micro blog service, has over 600 million users and produces upwards of 340 million tweets per day [1]. Sina Microblog2, the first Twitter-style Chinese micro blog web site, has a lot of than500 million accounts and generates run out a hundred million tweets per day [2].

Due to this diversity of on-line social media networks (SMNs), folks tend to use totally different SMNs for various functions. for example, Ren Ren 3, a Facebook-style however antonymous SMN, is employed in China for blogs, whereas Sina Micro blog is employed to share statuses (Fig.1). In different words, each existent SMN satisfies some user wants. In terms of SMN management, matching anonymous users across different SMN platforms will give integrated details on every user and inform corresponding laws, like targeting services provisions. In theory, the cross-platform explorations enable a bird's-eye read of SMN user behaviors. However, nearly all recent SMN-based studies focus on one SMN platform, yielding incomplete information. Therefore, this study investigates the strategy of crossing multiple SMN platforms to color a comprehensive image of those behaviors.

Nonetheless, cross-platform analysis faces various challenges. As shown in Fig.1, with the expansion of SMN platforms on the net, the cross-platform approach has integrated numerous SMN platforms to form richer information and a lot of complete SMNs for social computing tasks. SMN users type the natural bridges for these SMN platforms. the first topic for cross-platform SMN analysis is user identification for various SMNs. Exploration of this subject lays a foundation for additional cross-platform SMN analysis..

II. LITERATURE SURVEY

According to literature survey after studying different IEEE paper, collected some related papers and documents some of the point discussed here:

1) How unique and traceable are user nemeses'?

AUTHORS: D. Perito, C. Castelluccia, M.A. Kaafar, and P. Manils,

This paper explores the chance of linking users profiles solely by viewing their usernames. The intuition is that the chance that 2 usernames talk to constant physical person powerfully depends on the "entropy" of the username string itself. Our experiments, supported crawls of real net services, show that a major portion of the users' profiles will be joined exploitation their usernames. To the simplest of our data, this is often the primary time that usernames area unit thought of as a supply of knowledge once identification users on the web.

2) Connecting corresponding identities across communities
AUTHORS: R. Zafarani and H. Liu,

One of the foremost fascinating challenges within the space of social computing and social media analysis is that the questionable community analysis. A accepted barrier in cross-community (multiple website) analysis is that the disjunction of those websites. during this paper, our aim is to supply proof on the existence of a mapping among identities across multiple communities, providing a way for connecting these websites. Our studies have shown that easy, nonetheless effective approaches, that leverage social media's collective patterns are often used to seek out such a mapping. The utilized ways with success reveal this mapping with sixty six accuracy

3) Connecting users across social media sites: a behavioral-modeling approach
AUTHORS : R. Zafarani and H. Liu,

Social media is enjoying a vital role in our existence. folks typically hold varied identities on completely different social media sites. User-contributed internet information contains numerous data that reflects individual interests, policy making and alternative behaviors. To integrate these behaviors data, it's useful to spot users across social media sites. This paper focuses on the challenge of characteristic unknown users across completely different social media sites. a technique to relate user's identities across social media sites by mining users' behavior data and options is introduced. the strategy has 2 key parts. the primary part distinguishes completely different users by analyzing their common social network behaviors and finding robust opposing characters. The second part constructs a model of behavior options that helps to get the distinction of users across social media sites. the strategy is evaluated through 2 experiments on Twitter and Sina Weibo. The results of experiments show that the strategy is effective.

4) Privacy in the age of aug-mented reality,"
AUTHORS: A. Acquisti, R. Gross and F. Stutzman,

We investigate the practicability of mixing in public on the market we have a tendency tob a pair of.0 knowledge with ready-made face recognition computer code for the aim of large-scale, machine-controlled individual re-identification. 2 experiments illustrate the flexibility of characteristic strangers on-line (on a chemical analysis website wherever people shield their identities by mistreatment pseudonyms) and offline (in a public space), supported photos created in public on the market on a social network website. a 3rd proof-of-concept experiment illustrates the flexibility of inferring strangers' personal or sensitive data (their interests and social insurance numbers) from their faces, by combining face recognition, data processing algorithms, and applied mathematics re-identification techniques. The results highlight the implications of the convergence of face recognition technology and increasing on-line self-disclosure, and also the emergence of "personally predictable" data, or PPI. They raise questions about the longer term of privacy in associate "augmented" reality world during which on-line and offline knowledge can seamlessly mix.

5) I seek you: searching and matching individuals in social networks
AUTHORS: M. Motoyama and G. Varghese,
An online user joins multiple social networks so as to en- joy die rent services. On every joined social network, she creates Associate in Nursing identity and constitutes its 3 major dimensions specifically professional lupus erythematous, content and affiliation network. She for the most part governs her identity formulation on any social network and thus will manipulate multiple aspects of it. With no international identifier to mark her presence unambiguously within the on-line domain, her on-line identities stay unlinked, isolated and difficult to go looking. Earlier analysis has explored the higher than mentioned dimensions, to go looking Associate in Nursing link her multiple identities with an assumption that the thought of dimensions are least disturbed across her identities. However, majority of the approaches are restricted to exploitation of 1 or 2 dimensions. we tend to build a rest decide to deploy Associate in Nursing integrated system Finding Memo that uses all the 3 dimensions of Associate in Nursing identity to go looking for a user on multiple social networks. The system exploits a glorious identity on one social network to go looking for her identities on alternative social networks. we tend to check our system on 2 preferred and distinct social networks { Twitter and Facebook. we tend to show that the integrated system offers higher accuracy than the individual algorithms. we tend to report experimental endings within the paper

III. PROPOSED SYSTEM

We proposed the FRUI algorithm. Since FRUI employs a unified friend relationship, it is apt to identify users from a heterogeneous network structure.
Unlike existing algorithms, 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.

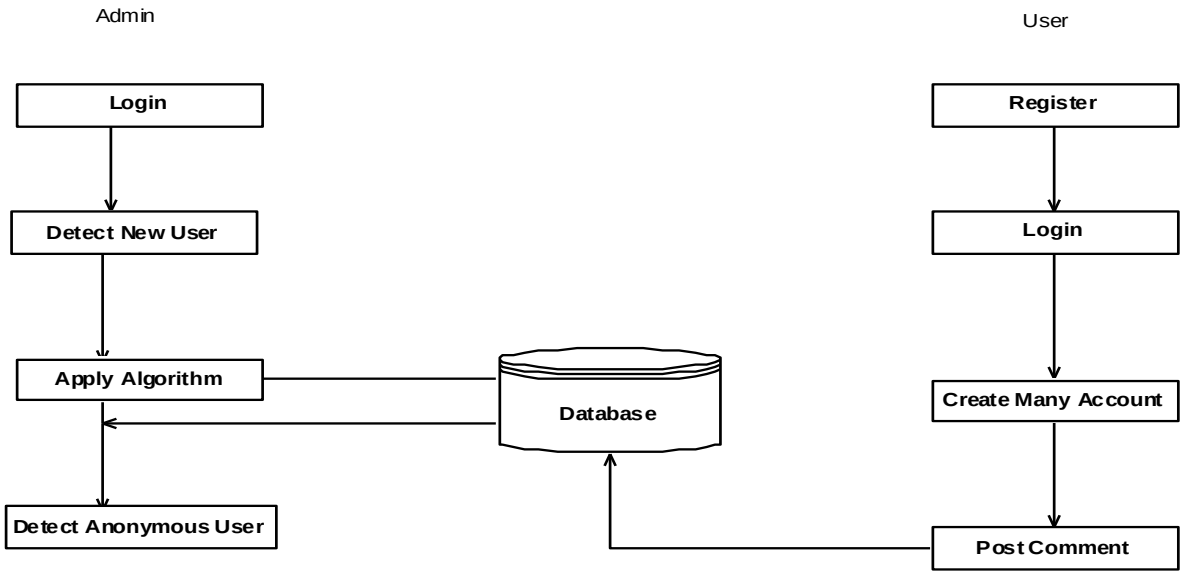


Fig 1. Block Diagram

IV. MATHEMATICAL MODULE

Let S is the Whole System Consist of

S= {I, P, O}

I = Input.

I = {U, Q, D}

U = User

U = {u1,u2….un}

Q = Query Entered by user

Q = {q1, q2, q3…qn}

D = Dataset

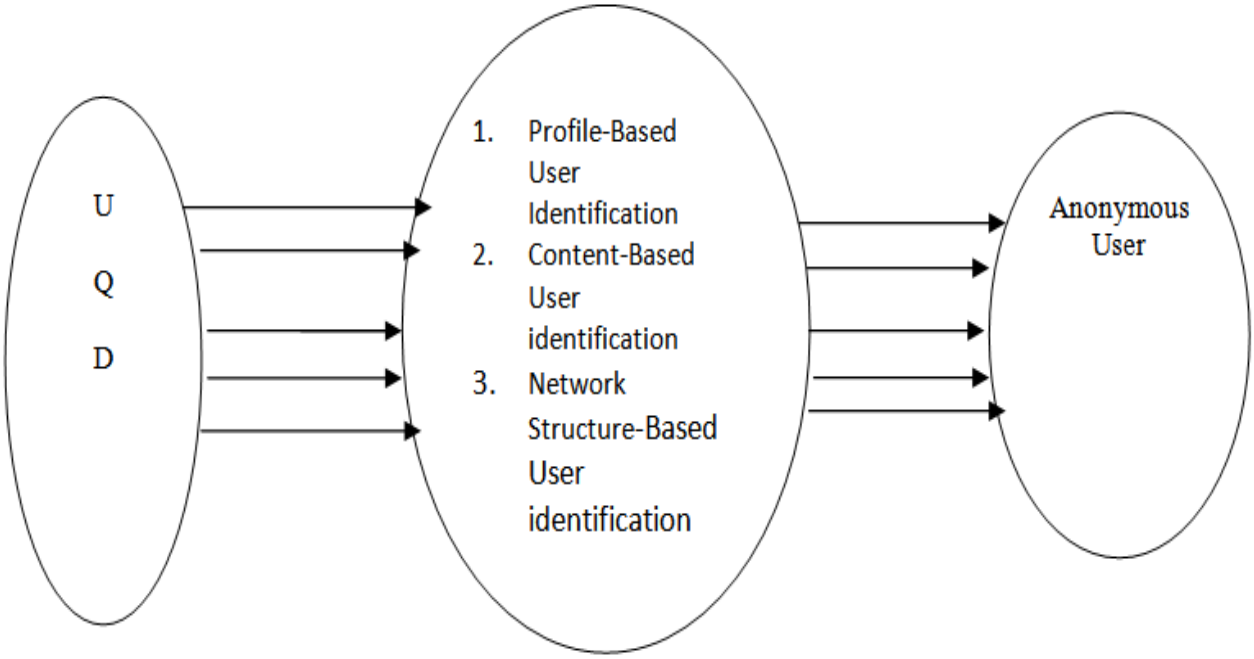
P = Process:

Step1: Social network creation.

Step2: User will register to particular social network for creating an account.

Step3: Admin will login to the system.

Step3: Admin Model
Step4:System will detect an anonymous user
Step 5: The algorithm repeats steps 2 to 4 until certain terminating conditions are fulfilled, such as a pre-defined number of iterations.



V. SYSTEM IMPLEMENTATION

- **Profile-Based User Identification**

Several studies addressing anonymous user identification have centered on public profile attributes, as well as screen name, gender, birthday, town and profile image.
A screen name is that the publicly needed profile feature in most SMNs. it's been wide explored as the way to acknowledge users across completely different SMNs. Peritoet al. [3] calculated the similarity of screen names and known users victimization binary classifiers. Similarly, Liu et al. [4] matched users in Associate in Nursing unattended approach victimization screen names. Za farani and Liu [5] projected a way to map identities across completely different SMN platforms, by trial and error confirming many hypotheses. On prime of this work, they[6]further developed a user mapping methodology by modeling user behavior on screen names. Among public profile attributes, the profile image is another feature that has received extended study. Acquisitive al. [7] self-addressed the user identification task with a face recognition rule. though each screen name and profile image will determine users, they can not be applied to giant SMNs. this can be as a result of some users could have identical screen name and profile pictures. as an example, several users have the screen name "John Smith" on Facebook.

- **Content-Based User Identification**

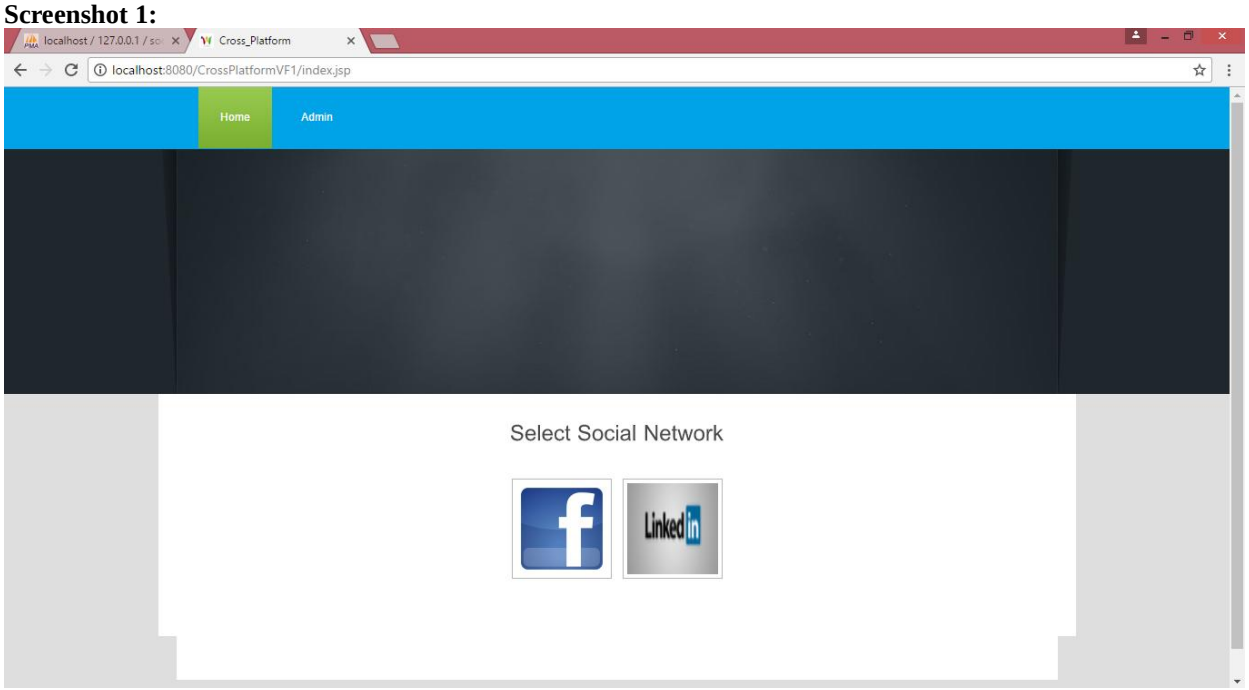
Content-Based User Identification solutions commit to acknowledge users supported the days and locations that users post content,as well because the literary genre of the content.
Zhenget al. [18] projected a framework for authorship identification victimization the literary genre of on-line messages and classification techniques.Almishari and Tsudik projected linking users across totally different SMNs by exploiting the literary genre of authors.Kong and Zhang projected Multi-Network Anchoring(MNA)to map users. They calculated the combined similarities of user's social, spatial, temporal and text data in numerous SMNs, and examined a stable matching drawback between 2 sets of user accounts.Goga et al. [21] exploited the geo-location hooked up to users' posts, the timestamp of posts, and users' literary genre to deal with user identification tasks.

- **Network Structure-Based User Identification**

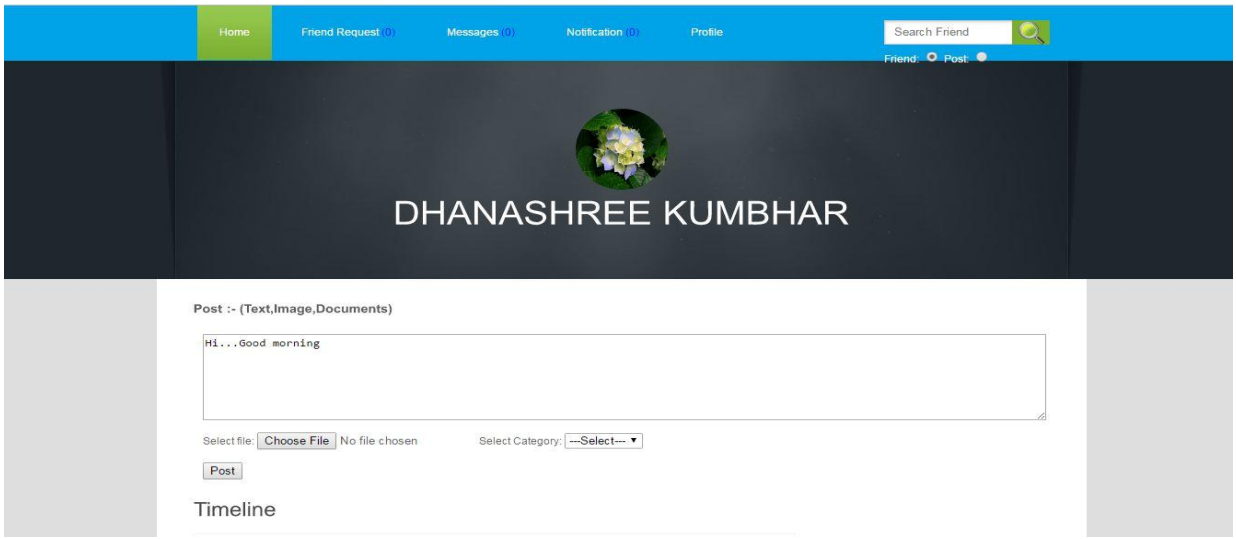
Network structure-based studies on user identification across multiple SMNs area unit accustomed acknowledge identical users entirely by user network structures and seed, or priori identified users. As shown higher than, network-based user identification poses many major challenges, with few studies to create on.
To address this drawback, Bartunovet al. proposed associate approach supported conditional random fields referred to as Joint Link-Attribute (JLA).JL A considered each profile at-tributes and network properties. To analyze privacy and obscurity, Narayanan and Shmatikov developed NS, based mostly entirely on constellation.

VI. RESULT AND DISCUSSIONS

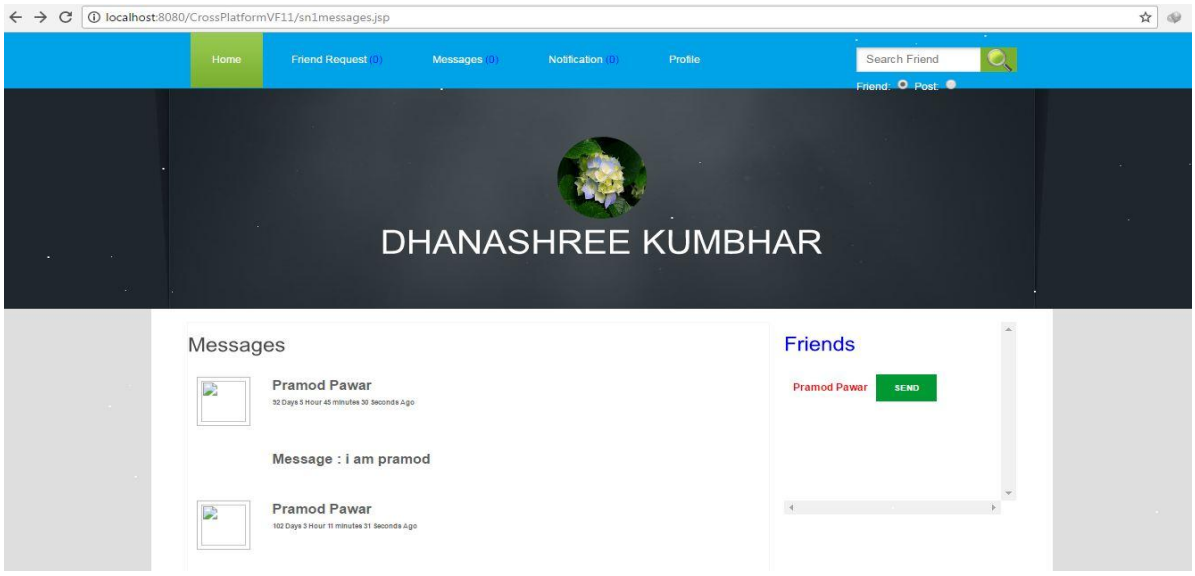
1. **Screenshot 1:**



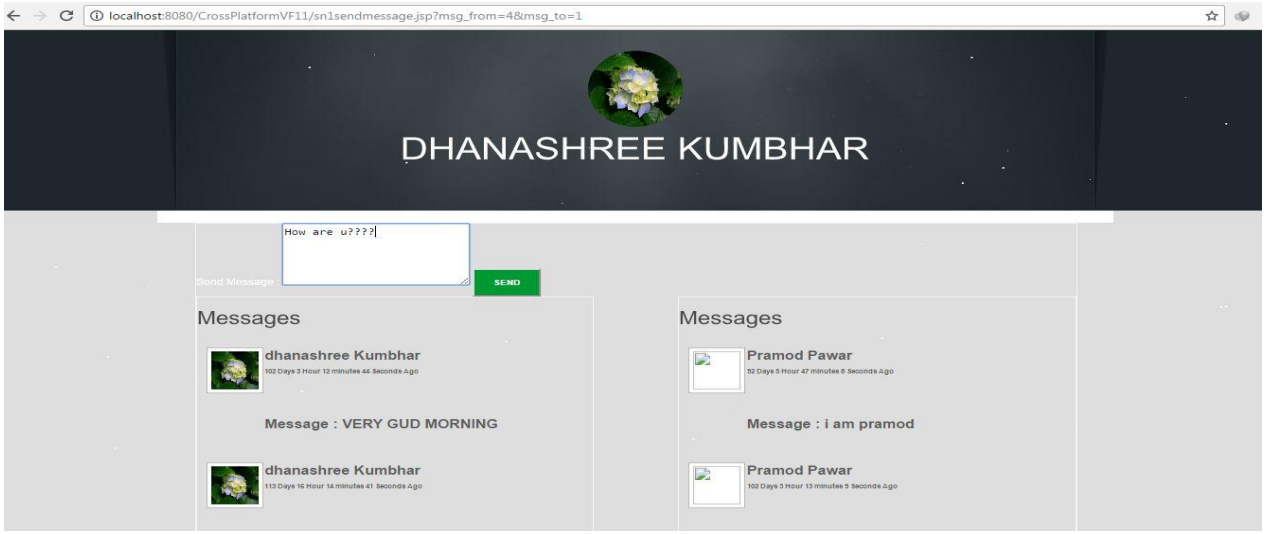
2. **Screenshot 2:**



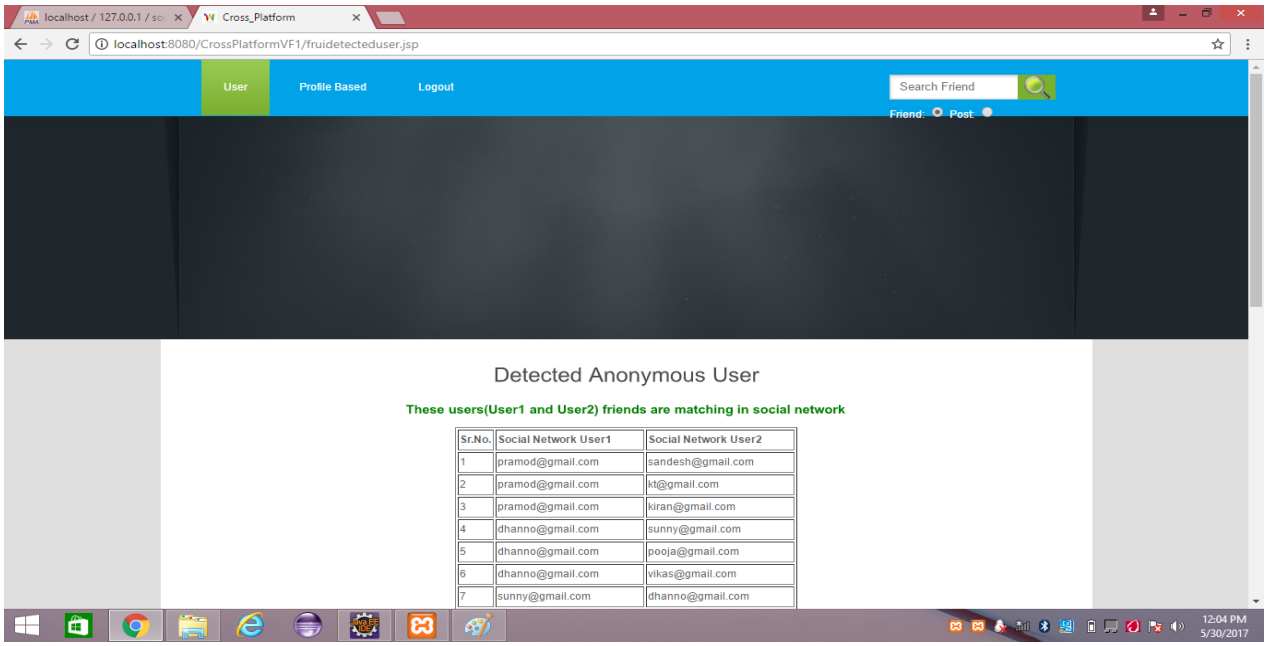
3. **Screenshot 3:**



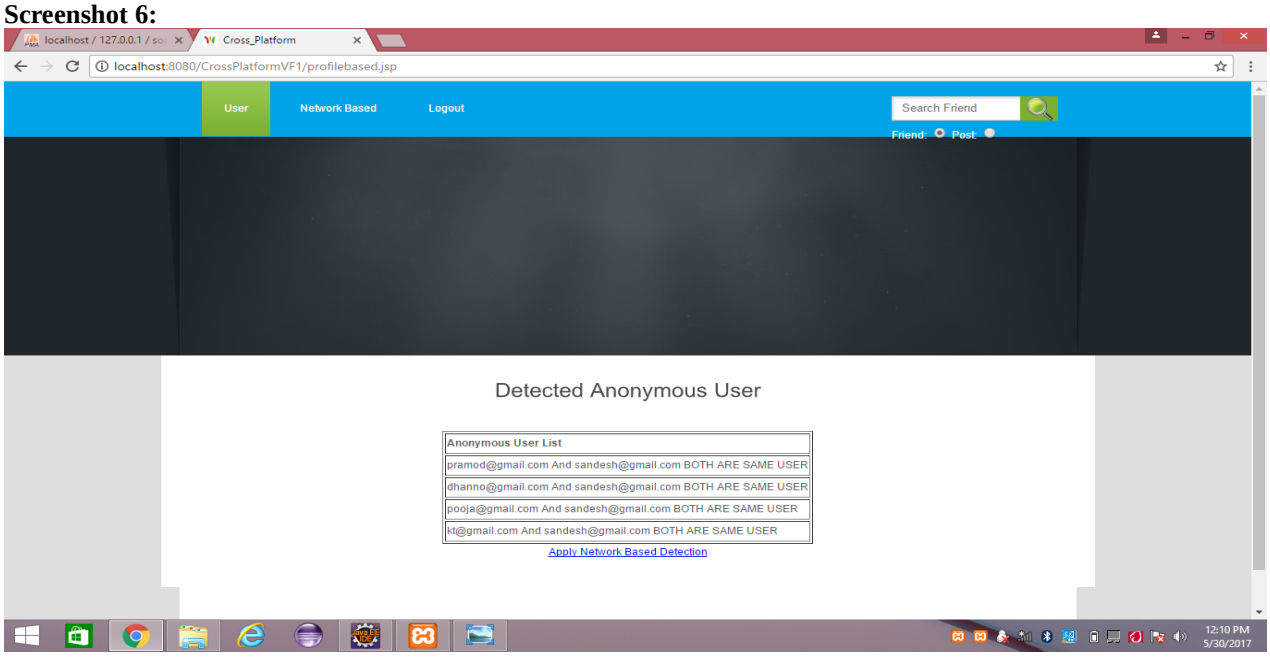
4. Screenshot 4:



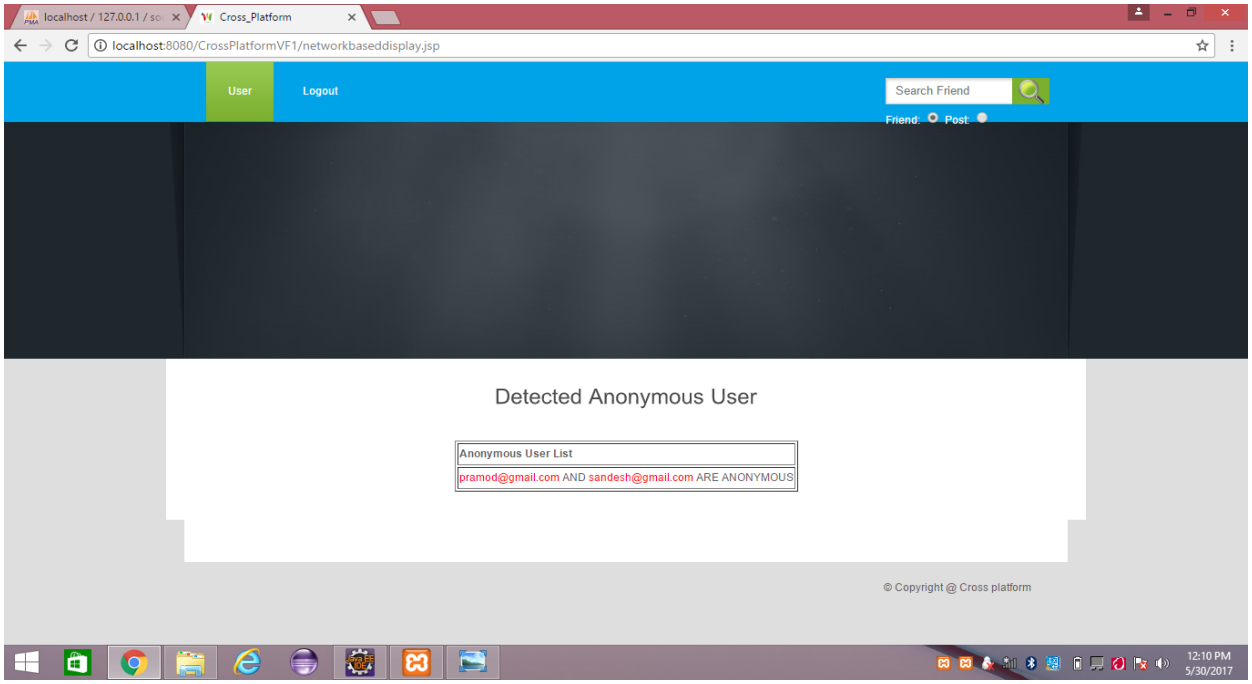
5. Screenshot 5:



6. Screenshot 6:



7. Screenshot 7:



VII. CONCLUSION

This study addressed the matter of user identification across SMN platforms associated offered an innovative answer. As a key side of SMN, network structure is of predominate importance and helps resolve de-anonymization user identification tasks. Therefore, we have a tendency to plan a consistent net-work structure-based user identification answer. We have a tendency to conjointly develop a completely unique friend relationship-based algorithmic program known as FRUI. To enhance the potency of FRUI, we have a tendency to de-scribed 2 propositions and addressed the quality. Finally, we have a tendency to verify our algorithmic program in each artificial net-works and ground-truth networks.

Results of our empirical experiments reveal that net-work structure will accomplish necessary user identification work. Our FRUI algorithmic program is straightforward, nevertheless economical, and performed far better than NS, the present state-of-art network structure-based user identification answer. In eventualities once raw text information is distributed, incomplete, or arduous to get because of privacy settings, FRUI is extraordinarily suitable for cross-platform tasks.

Moreover, our resolution may be simply applied to any SMNs with friendship networks, including Twitter, Face-book and Foursquare. It can even be extended to different studies in social computing with cross-platform issues like targeted promoting information retrieval, collaborative filtering, and sentiment analysis and additional. Additionally, since solely the Adjacent Users area unit concerned in every iteration method, our technique is ascendible and can be simply applied to large datasets and on-line user identification applications.

Identifying anonymous users across multiple SMNs is difficult work. Therefore, solely some of identical users with completely different nicknames may be recognized with this technique. This study engineered the inspiration for more studies on this issue. Ultimately, it's our hope that a final approach may be developed to spot all identical users with completely different nicknames. Different user identification strategies may be applied at the same time to look at multiple SMN plat-forms. These strategies area unit complementary and not mutually exclusive, since the ultimate call might trust human user's involvement. Therefore, we propose exploitation these strategies synergistically and considering strengths and weaknesses for the most effective results.

REFRENCES

[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 6thACM 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.

