Detecting Fake User Accounts on Different Social Media Networks

Sachin Ingle[1], Satish Borade[2], Sagar Awasare[3]

[1]Student, Information Technology, SIT,LONAVALA,
[2]Student, Information Technology, SIT,LONAVALA,
[3] Student, Information Technology, SIT,LONAVALA,

Abstract

Use of Social Networking is Increased Drastically. But every good thing have bad side like that, disadvantage of social network is fake user accounts. Anyone can create fake account easily on it and use it for bad purposes. We are here proposing a system that can help to detect fake accounts. This will find fake user from multiple Social Networking platform by matching their friends network, matching profile details and their writing styles. In this way, we are matching different user accounts on multiple social network platforms and finding matched account. In this way we can detect fake user. Along with this we are improving efficiency of previous works in this area.

Keywords- *Social Media Networks (SMN's), Cross Platform, Anonymous Identical Users, Friend Relationship User Identification. Network Based, Profile Based. Time Stamp. Friend Relation User Identification (FRUI).*

1.INTRODUCTION

Now a days use of internet is increased. with the use of internet, the term social media networks become popular. Everyone who use internet is well known about social media networks. Social media network is collection of many social networking websites. Social networking is platform, where a user of social network can express their point of view towards anything. Also with this user can share the information which they have with other peoples which are connected to user with social media networks. It provide full authority to express your opinion. It have advantage and disadvantage also[1]. Advantage of Social Media Networks is that, People can get connected to other peoples which are globally situated at far distances. Due to this they can share information about their cultures, can share thoughts on something, share global issues, may share solutions on some problems. Can share educational knowledge etc. important thing is that people can get social platform where they can share their problems globally[2].

Now we some bad sides/disadvantage of social media networks. Some people can use Social Media Sites for bad purpose like people create their accounts on social networking sites using fake information and can use it for doing bad things like spreading rumours about something. For pretending them as genuine user and becoming friend with any people that they even don‟t know them. Harassing someone. Black mailing peoples etc.

To work on these disadvantages we are proposing a system that can help to detect fake user from multiple social networking platforms and detect those users who have two accounts on different SMN‟s site with different name means same user creating multiple accounts but different name for hiding identity. Proposed system will detect fake accounts by matching users on one social networking platform with the user of another social networking platform. user matching will be done on the basis of user profile details, user friends network matching and user post matching.

In First step of implementation profile matching is done. In this two user profiles from Social network one and Social network two are matched with each other. Matching is done on the basis of information provided by the user at the time of registration like email, name mobile no. Date of birth and other fields included in profile. After profile matching is done successfully the extracted users are further used for next step. In next step further algorithm is applied like FRUI, in this step algorithm is applied on that users which are identified from profile based algorithm. In FRUI, users of social network one and two are matched based on the friend relation pattern, and those users are extracted which are having same friends relation and network of friends i.e. friends of friends network. After identifying and extracting users based on Friend relation in next step content matching is done. In content matching user posts are collected and analysed and words from user posts are matched like writing style of words like some user will write „the‟ as „d‟, „have‟ as „hve‟ and so on. This is called writing

style. Users are extracted based on this content matching further things are matched like on which post user will hit likes. Users login times to Social Sites etc. in this way Fake users are identified.

2.RELATED WORK IN THIS AREA

There are few works done by some peoples in this area on profile based, content based user identification, but our work is based on Friend Relation Based User Identification. In this FRUI user Identification algorithm user identification is based on degree of user match pairs. It is implanted to make previous network Structure based algorithm more Effective.[1][3][4][5]

Following Table show works done by different Peoples in this area.

Sr. No.	Name of Paper	Author	Year of Publishing
1	Connecting Corresponding Identities across Communities	Reza Zafarani and Huan Liu	2009
2	Network and profile based measures for usersimilarities on social networks	Cuneyt Gurcan Akcora, Barbara Carminati, Elena Ferrari	2011
3	Cross-Platform Identificationof Anonymous Identical Users in Multiple Social Media Networks	Xiaoping Zhou, Xun Liang	2016

3.working of the proposed system

Proposed System will work in the following way.

First User will create Accounts on Both Social Network Prototype. After successful accounts creation User will login to system. Then it will use all the provided features of site and post status . likes and comments on user posts.

Admin can login through provided admin panel then at very first, admin has to select the option of first algorithm which is to be applied on the social networks user. In this multiple users based on the profiles are matched. And the user with same information are extracted and they are identified as fake. In the next step option of second algorithm will be shown on next page, when admin clicks on this algorithm then network based algorithm is applied on the users accounts identified previously in profile based algorithm. Due to network structure based algorithm user pairs are matched and degree is calculates of user network structure.

After this FRUI algorithm will be applied on the identified user accounts from step two i.e. network structure. Degree of UMP‟s are identified this is based on the user‟s friends network. Users are matched based on the friends at one level, two level and third level. And those users which having highest level of matching in network will be considered as fake accounts.

Along with this we are adding extra algorithms for identifying users based on the content they share, posts, comments and likes of users. Also we are adding login times in which users will mostly sign in to site. This can also help to identify fake user. This will improve efficiency of existing fake user detection algorithm.

This all algorithm can improve the results and system efficiency if used together in system. In this way system will work.

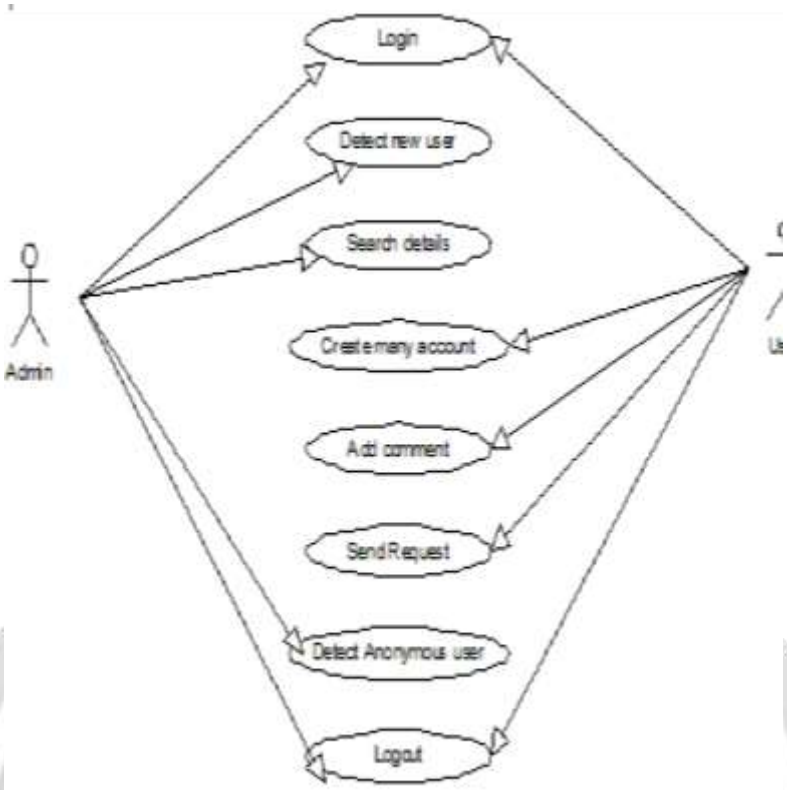

Fig. Use Case Diagram of Proposed System

4.SYSTEM ARCHITECTURE

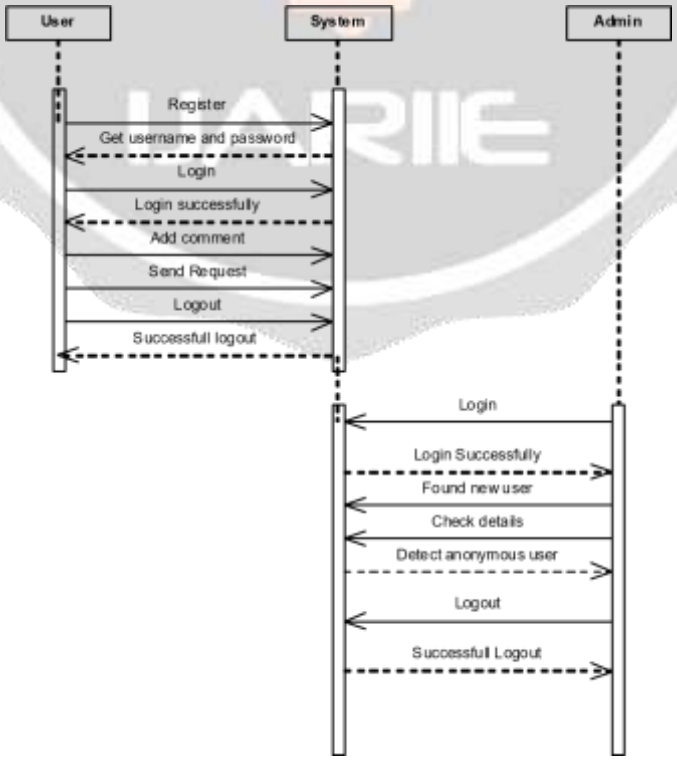

Fig. Sequence Diagram

5.Algorithm Used for working on system

Algorithm 1: FRUI

Input: SMN_A, SMN_B, Priori UMPs: PUMPs
Output: Identified UMPs: UMPs
1:**function** FRUI(SMN_A, SMN_B, PUMPs)
2: T = {}, R = dict(), S = PUMPs, L = [], max = 0, F_A = [], F_B = []
3: **while** S is 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: **foreach** $UMP_{A\text{-}B}(i, j)$ in S **do**
13: **foreach** UE_{Ai} in the unmapped neighbors of UE_{Ai} **do**
14: $F_A[i] = F_A[i] + 1$
15: **foreach** UE_{Bj} in the unmapped neighbors of UE_{Aj} **do**
16: $R[UMP_{A\text{-}B}(a, b)]$ += 1, $F_B[j] = F_B[j] + 1$
17: Add $UMP_{A\text{-}B}(a, b)$ to $L[R[UMP_{A\text{-}B}(a, b)]]$
18: **if** $R[UMP_{A\text{-}B}(a, b)]$ > max **do**
19: max = $R[UMP_{A\text{-}B}(a, b)]$
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 M_{ij} in C using (5)}
27: S = {un-Controversial UMPs in I}
28: **if** I == C **do**
29: **break**
30: **return** T

Identifier first calculate matrix R and identifies match degree. Then it iterates and identifies UMPs using function *g* until no UMP can be identified. In each iteration, once the UMPs are identified, the items are removed from the Candidate UMP list, and *R* is recalculated based on Proposition 2. The process is summarized in Algorithm 1.

Suppose that there are *s* Valid Priori UMPs in any iteration. Lines 4-11 in Algorithm 1 remove the identified UMPs and update the maximum match degree, and the time complexity costs $O(s) + O(min(vA, vB))=O(min(vA, vB))$, where *v*A and *v*B denote the numbers of the users in SMNA and SMNB, respectively. Lines 12-19 update the Candidate UMP list and the maximum match degree using Propositions1 and 2. As discussed above, the computational complexity is $O(sdAdB)$. Lines 20-29 identify identical users for the next iteration using (6). Normally, the max u(*M*) can be found in the candidate UMPs with the largest *Mij*, and the time complexity is no more than $O(min(vA, vB))$. In summary, the complexity of FRUI is $O(min(vA, vB))+ O(sdAdB) + O(min(vA, vB))\leq O(min(vA, vB)dAdB)$.Obviously, the com-plexity of FRUI is lower than $O((eA+ eB)dAdB)$ in NS[19], where *e*A and *e*B are the numbers of the edges in the two SMNs.[6] After applying FRUI then content based identification is done. Along with this user login time will be used and applied to identify the users which logs on to particular time period on system.[6]

6.Experimental Results

Experiments Performed on Dummy Dataset and the result was improved than previous work. Due to this degree of fake identical user identification is increased by some value.
Experimental Results will be increased as the dataset increased.

7.Problem Statement of proposed System

Increase the degree of fake user identification by means of some extra work on the previous/existing work with some new idea. Adding some extra user identification criteria in existing system.

8. What Extra Criterion/Work will be done in Proposed System?

Extra user identification criteria is that, in previous profile based and network structure based algorithm is used for fake identical user identification but here we are using Friend Relation User Identification along with Text

mining of User posts, comments and likes on SMN "s sites. It will help to identify writing style of user. And we will use login time of different users in to system. To identify identical users. At last we will integrate all previous and new work together to improve overall efficiency and accuracy of system.

Future Scope

System can be made more accurate by adding some security to users while account creation. Government ID proof can be added as mandatory field while creation of accounts along with this OTP can be used while account creation. This will help in restricting the user from fake account creation.

CONCLUSION

Our main goal behind this proposed system is that to improve identical fake user identification more accurate by means of using extra criterion of user identification like content based, FRUI based, login time based.

ACKNOWLEDGMENT

We would like to thanks our Guide, for helping us during this research work every time to solve our queries .Also with them we would like to thanks other teachers of my department and my friends who helped me during this research work.

REFERENCES

1.www.google.co.in
2.www.wikipedia.org
3.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.
4.R. Zafarani and H. Liu, "Connecting corresponding identities across communities," *Proc. of the 3rd International ICWSM Conference*, pp. 354-357, 2009.
5.Cuneyt Gurcan Akcora, Barbara Carminati, Elena Ferrari, "Network and profile based measures for user similarities on social networks", IEEE IRI, 2011
6.Xiaoping Zhou, Xun Liang, "Cross-Platform Identification of Anonymous Identical Users in Multiple Social Media Networks",ieee,2016