HUMAN ASPECT PROFILE ANALYSIS FOR PRIVACY PRESERVING

MUVITHA.A [1], SOWMIYA.K [2], THEERTHAGIRI.K [3]

MRS.LEELA.V (AP\SR.GR) [4]

Department of Information Technology

Velalar College of Engineering and Technology

Erode

a.muvithadharan@gmail.com[1], sowmiyakarke18@gmail.com[2]

theerthagiri97@gmail.com[3]

leelabtechit@gmail.com[4]

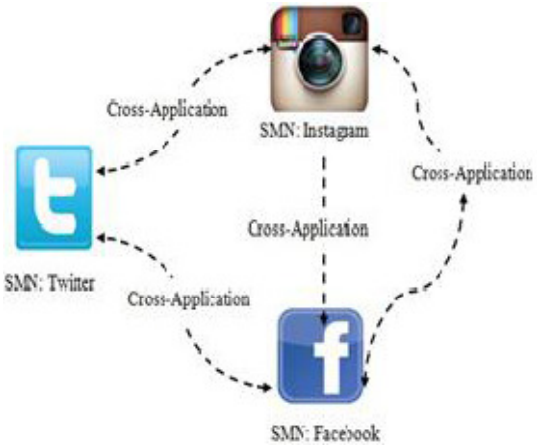
ABSTRACT:

Now-a-days, social networking has become an important part of the online activities on the web. Currently, many social networking websites use different ways to store and display different information to interact with more number of users. The personal details might be collected from the illegal person and they misuse those information. Then, introduce a new metric called profile matching technique to identify and deduct the same profiles in the dataset. An FRUI(Friend Relationship User Identification) algorithm issued to identify the same profiles in a dataset. The two profiles describe the same physical user (e.g.The Facebook, homepage,etc...)Are the same. Anonymous user tends to create more than one account in social network. As a result the user prefers different social networking websites which exchanges the friend's information.In this paper we proposed a method to identify the users duplicate accounts in the different social media networks by using profile matching technique.

INTRODUCTION:

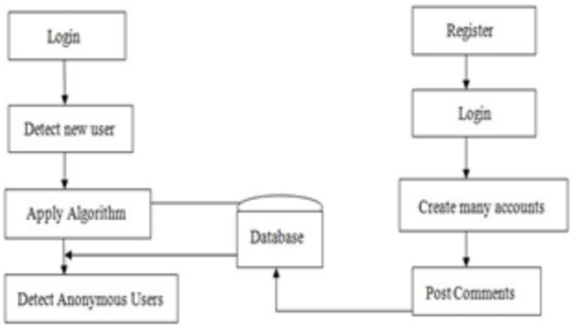
Over the year's social networking has become more and more popular. It is a way to make connections, not only people with us, but with friends whom we know but are not met several years. Social Networking is a way that helps many people feels as though they belong to a community. Social networks offer to users interesting means and ways to connect, communicate, and share information with other members within their platforms. However those Sites have currently different structures and they represent user's profile differently. Social networking sites allow users to share ideas, activities, events, and interests within their individual networks. The main types of social networking services are those that contain category places (such as former

school year or classmates) mean to connect with friends recommendation system linked to trust. An individual user can have multiple accounts of social networking sites. In this paper we proposed a method for profile matching technique in social media network, which helps to identify a particular person who has multiple social networking accounts in the same network to make a search of friends easier.



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 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 FRUI requires no control parameters.



LITERATUREREVIEW:

1. MAINTAINING AN ONLINE BIBLIOGRAPHICAL DATABASE, THE PROBLEM OF DATA QUALITY:

In this work, Cite Seer and Google-Scholar are huge digital libraries which provide access to (computer) science publications. Both collections are operated like specialized search engines, they crawl the web with little human intervention and analyze the documents to classify them and to extract some metadata from the full texts. On the other hand there are traditional bibliographic data bases like INSPEC for engineering and public method for medicine. For the field of computer science the DBLP service evolved from a small specialized bibliography to a digital library covering most subfields of computer science. The collections of the second group are maintained with massive human effort. On the long term this investment is only justified if data quality of the manually maintained collections remains much higher than that of the search engine style collections. In this paper we discuss management and algorithmic issues of data quality.

2.NEAR OPTIMAL HASHING ALGORITHMS FOR APPROXIMATE NEAREST NEIGHBOR IN HIGH DIMENSIONS:

In this work, we present an algorithm for the c-approximate nearest neighbor problem in a d-dimensional Euclidean space, achieving query time of O (dn1/c2+o (1)) and space O (Dn + n 1+1/c2+o (1)). This almost matches the lower bound for hashing-based algorithm recently obtained .We also obtain a space-efficient version of the algorithm, which uses dn + n logo(1) n space, with a query time of dnO(1/c2). Finally, we discuss practical variants of the algorithms that utilize fast bounded-distance decoders for the Leech Lattice.

3. PRIVACY PRESERVING RECORD LINKAGE VIA GRAM PROJECTIONS:

In this work, Record linkage has been extensively used in various data mining applications involving sharing data. While the amount of available data is growing, the concern of disclosing sensitive information poses the problem of utility Vs privacy. In this paper, we study the problem of private record linkage via secure data transformations. In contrast to the existing techniques in this area, we propose a novel approach that provides strong privacy guarantees under the formal framework of differential privacy. We develop an embedding strategy based on frequent variable length grams mine in a private way from the original data. We also introduce personalized threshold for matching individual records in the embedded space which achieves better linkage accuracy than the existing global threshold approach. Compared with the state-of-the-art secure matching schema, our approach provides formal, provable privacy guarantees and achieves better scalability while providing comparable utility.

4. SOME METHODS FOR BLINDFOLDED RECORD LINKAGE:

In this work, a simple but effective algorithm for matching adult patients seen at more than one site in a multi-site de identified registry is described. In a data set of 19,000 records a derived match variable consisting of a 2-character prefix from both first and last name combined with date of birth has 97% sensitivity; by contrast, an anonym zed identifier based on the patients' full names and date of birth has sensitivity of only 87%.

5.LEARNING TO MATCH AND CLUSTER LARGE HIGH-DIMENSIONAL DATA SETS FOR DATA INTEGRATION:

In this work, Part of the process of data integration is determining which sets of identifiers refer to the same real-world entities. In integrating databases found on the Web or obtained by using information extraction methods, it is often possible to solve this problem by exploiting similarities in the textual names used for objects in different databases. In this paper we describe techniques for clustering and matching identifier names that are both scalable and adaptive, in the sense that they can be trained to obtain better performance in a particular domain. An experimental evaluation on a number of sample datasets shows that the adaptive method sometimes performs much better than either of two non-adaptive baseline systems, and is nearly always competitive with the best baseline system.

6. EFFICIENT ROBUST PRIVATE SET INTERSECTION:

In this work, Computing Set Intersection privately and efficiently between two mutually mistrusting parties is an important basic procedure in the area of private data mining. Assuring robustness, namely, coping with potentially arbitrarily misbehaving (i.e., malicious) parties, while retaining protocol efficiency (rather than employing costly generic techniques) is an open problem. In this work the first solution to this problem is presented.

MODULE DESCRIPTION:

1. DATA PREPROCESSING: In this module pre-processor is intended to secure however many Priori Anonymous users as would be prudent. There is no regular approach accessible to get Anonymous users between two SMNs. Indicated strategies must be planned by SMNs. It can be received by application, e.g., email address, screen name, URL, and so forth. An email deliver seems, by all accounts, to be a special element for every record, and can be utilized to gather Priori Anonymous users. The email locations to discover indistinguishable clients among various SMNs with the proposed algorithm.

2. FRUI PROFILE ANALYZER:

Profile retriever: It is used to extract profiles having the same FRUI value from the initial set of profiles. This can be done using a LSH or by accessing a dataset of profile provided locally. It is important to note that crawling profiles from social network is a difficult task due to social site protection policy.

Weight assignment: It is used to assign manually or automatically each attribute in the user profile to a weight as indicated.

Profile matcher: It returns the decision whether the two compared profiles are the same or not. This decision, done via a decision making algorithm, is computed using the weighted similarity scores.

3. SOCIAL MEDIA USER MATCHED PAIR:

In this module, we methodically talk about our answer for the client ID issue by utilizing clients' companions, and create two recommendations to enhance the productivity of our calculation. The identifier discovers Anonymous users utilizing associations among clients and Priori Anonymous users. This algorithm is used to provide the details which have been deleted in the add friend history. Inside this application we have a separate tab called history and this provides the details which have been deleted in the conversation tab.

4. DATA ANALYSIS AND ANONYMOUS USER IDENTIFICATION MODULE:

Input: When the user logs into the system he has to provide the username, password and friends on available list. When he performs all these actions the user successfully logs into the system.

For a new user of this application the user has to register with all the details such as username, password and selecting one level of friend relationship predict and answering it and then giving a password to access the user panel. After giving all these details the user gets successfully logs into the system. Once we get login into the application the user has access to different tabs using user profile, add friend, friend list.

Profile: In this user can perform the actions such as edit his profile by changing the

details which he have been at the time of registration.

Friends: In this user can have access to three different tabs such as Friend List, and Notifications. The New Friends provides sending friend request that are registered with this application. Notification tab has the notifications of all the other users who have provided us the friend request. Friend List provides the list with all the friends we have in this application.

Logout: Once you finished all your actions by clicking on this Logout tab it will take out of this application.

ALGORITHM FOR FRUI:

Input: SMNA, SMNB, Priori UMPs: 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 UMP switch mapped UE from L[max]
For each UMPA-B (I,j) in S do
Foreach UEA a 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

OUTPUT:

SMN(A)	SMN(B)
Ashwin	Ashwini
Kaviya	Kavin
Praveen	Praveena
Swathi	Soundhar

CONCLUSION:

In this paper, we addressed the issue of providing social network operations and functionalities. In this work, we proposed a framework for user profile matching in social networks. This framework is able to discover the biggest possible number of profiles that refer to the same physical user that existing approaches are able to detect. In our work, we proposed a method to identify the users duplicate accounts in different social networks by using this profile matching technique. The result of the experimentation showed improvements compared to other classical method. As a

future work, we are planning to further explore and propose more interesting social operations and functionalities.

REFERENCES:

[1] D. Balzarotti, M. Cova, and G. Vigna, "Clear Shot: Eavesdropping on Keyboard input from video," inProc. IEEE Symp. Security Privacy, May 2008, pp. 170–183.

[2] J. Bandeau, "The science of guessing: Analyzing an anonym zed corpus of 70 million passwords," inProc. IEEE Symp. Security Privacy, May 2012, pp. 538–552.

[3] J. Bandeau, C. Harley, P. C. van Outshot, and F. Stajano,"The quest to replace passwords: A framework for comparative evaluation of Web authentication schemes," inProc. IEEE Symp. Security Privacy, May 2012, pp. 553–567.

[4] J. Bandeau, S. Preibusch, and R. Anderson, "A birthday present every Eleven wallets the security of customer-chosen banking pins," in Financial Cryptography and Data Security. Berlin, Germany: Springer, 2012.

[5] J. Brainerd, A. Jules, R. L. Rivest, M. Sydlo, and M. Yung, "Fourth-factor authentication: Somebody you know," inProc. ACM CCS, 2006,pp. 168–178.

[6] P. Cao, H. Li, K. Nahrastedt, Z. Kalbarczyk, R. Ayer, and A. J. Slagell,"Personalized password guessing: A new security threat," inProc. ACM Symp. Boot camp Sci. Security, 2014, p. 22.

[7] C. Castelluccia, A. Chaabanes, M. Demuth, and D. Perito. (Apr. 2013)."When privacy meets security: Leveraging personal information for password cracking."

[8] A. Das, J. Bandeau, M. Caesar, N. Brasov, and X. Wang, "The tangled Web of password reuse," inProc. NDSS, 2014, pp. 23–26.

[9] D. Davis, F. Montrose, and M. K. Reiter, "On user choice in graphical password schemes," inProc. USENIX Security, 2004, p. 211.

[10] X. de C. de Carnavalet and M. Mannan, "From very weak to very strong: Analyzings password-strength meters," inProc. NDSS, 2014, pp. 23–26.

[11] S.Ramkumar , G.Emayavaramban, A.Elakkiya, "A Web Usage Mining Framework for Mining Evolving User Profiles in Dynamic Web Sites", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 4 (7), pp.889-894, Aug – 2014.