**International Journal of Engineering Sciences & Emerging Technologies, Dec. 2013.**
**ISSN: 2231 – 6604 Volume 6, Issue 3, pp: 368-372 ©IJESET**

# ATTRIBUTES SIMILARITIES SUPPORTS PROFILE MATCHING IN SOCIAL NETWORK

Asmita Joshi, J. S. Sodhi
Amity University, Noida

***ABSTRACT***
*In the current environment through social network people can find others and make their own network. An Individual user can have multiple accounts of social networking sites. In This paper we proposed an algorithm for profile matching in social networks, which helps to identify a particular person who has multiple social networking accounts and map his/her profile's attribute values with others in the same network to make a search of friends easier.*

***KEYWORDS:*** *Social networking site, Profile matching*

## I. INTRODUCTION

Over the years, social networking among 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 users' profiles differently. A social networking site is an online service, platform, or site that focuses on building and reflecting of social networks or social relations among people, who, for example, share interests and/or activities. A social network service consists of a representation of each user his/her social links, and a variety of additional services. Most social network services are web-based and provide means for users to interact over the Internet. 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. In this approach we use the methods of entity matching for different profile based on their same attribute value.

## II. RELATED WORK

**Elie, Richard, Albert** [1] they proposed a framework that is able to discover the biggest possible number of profiles that refer to the same person .This approach assigned weights manually or automatically to the attributes, string and semantic similarity metrics were used to compare attribute values. Aggregation functions were used for data fusion and for decision making.

**Vosecky J., Hong D. and Shen V.Y.[**13,12] They Proposed a matching technique in which each user profile is represented as a vector consisting of the values of individual profile fields (e.g., name, date of birth, etc.). The comparison between any two vectors consists of two phases. In the first phase, the algorithm calculates a similarity score between corresponding vector fields using an appropriate string matching function for each field, resulting in a similarity vector. In the second phase, a weighting vector is applied to the similarity vector to calculate the overall similarity.

# III. PROPOSED APPROACH

Social site helps peoples to expand their networking. A person can use multiple social accounts to make their networking strong between professional friends or school and child hood friends. Through our proposed idea we are introducing profile matching of a single person between two social networks and person can find people like him/her, those have similar features in their profile.

Our approach will move in two direction, in the first part of our research we match profile of a single person between two networks and in the second approach system will make task simple to find friends in the social network it will find appropriate matched profile of friends in the single social networking site.

In our proposed approach we prepare data set from basic attributes exist in profiles of users. In the data set we collect attribute values that are used in profile building over a network. These values are the values that shows the basic information about a persons .This data set will used in both approaches.

Our first task will match the common attribute values between two social sites and second will find the friends over the network if they have one of the common values for at least one attribute. These attributes can be refined as per the use in our approach.

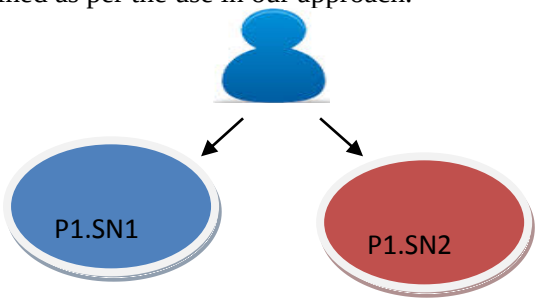


**Fig.1** Person P1 has two accounts on SN1(Social Network) and SN2

We collect the values of attributes like name, school name, college name, employer, home town, friends etc. This data set will be considered for our both approaches. The values of the attributes are different as per the profiles of different people. We refined the attributes for processing. Following table shows the respective data set field in our experiment.

We match the profile of one person over the two sites and make apply rule based classification algorithm to decide whether this two profiles belongs to a single person. Similarities between two attribute's values are calculated using similarity functions of name similarity.

## A. LCS Name Similarity

The longest common sub-string (LCS) repeatedly finds and removes the longest common sub-string in the two compared strings up to a minimum length. A similarity measure can be calculated by dividing the total length of the common sub-strings by the average lengths of the two original strings. We use the LCS algorithm to calculate the similarity between the two users' full names. [9, 12]

## B. N-Gram Name Similarity

An *n*-gram similarity between two strings is calculated by counting the number of *n*-grams in common(i.e.-grams contained in both strings) and dividing by either the number of *n*-grams in the shorter string (called Overlap coefficient), or the number of *n*-grams in the longer string or the average number of *n*-grams in both strings. We used 2-grams and 3-grams to calculate the similarity between the two users' full names. [9, 12]

Initial similarity functions match the two string values if they are similar then we use rule based technique for decision, that will help to make decision that two profiles belongs to the same person. The matched attributes are used to generate rules. Strongest rules are the rules that are greater or equal to the threshold values of the rules should be true. They give a good result. In the first part we proposed an algorithm for matching two profiles belongs to the same person.

**Table 1:** Show Common Attributes of Profile

| |
|---|
| Name |
| School Name |
| Graduate university /College |
| Post graduate university /College |
| Employer |
| Family |
| Employer |
| Mutual friends |
| Email ID |
| Home Town |
| Language |
| Interest |

## C. Proposed Algorithm

**Input:** P1 and P2 are the two profiles, a1, a2….. Attributes of the profiles.
**Begin**
1. We input the value for a1, a2,.an for profile P1
And Profile P2.
2. Find the Similarity (P1.a1, P2.a1)
3. If Match is Found in the attribute values.
3. Then We Increase the counter according to matched values, this counter value is used to generate rules
4 Show the matched values.
5. Now we assign 1 to each match value output and 0 to each unmatched value (These similar valued attributes are used to classify that two profiles are belong to same person).
6. Rules are generated to classify the data for each 1's or matched values.
6 Integer value for each rule repeatedly assign to independent variable V
7. Profiles are similar if values (total 1's)>= threshold values of rules to be true then it shows that two profile are belongs to single person.
**End**
Above described process will find that two profiles exist at two social sites belongs to one person.
Next phase of our proposed approach is to build some suggestions that help people to find their friends on a social networking site. 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 schemas and they represent users' profiles differently.[1,2,3] It's the simple act of expanding the number of people you know by meeting your friends' friends, their friends' friends and so on. Social networking sites give you a great chance to keep in touch with your old friends, keep tabs on current friends and family, and to create new online friendships with people who share similar interests. Our idea is to expand your network that shows the list of the people those have similar attributes with you in their profile.
They may be your school friends, college friends, colleagues etc.

### A. Social Network structure for networking

Social network structure has special features that features are common between two profiles

### B. Mutual Friends

This feature represents the number of mutual friends of two users. Mutual friends are counted by the number of friends with identical names in both cycles of friends. [2]

### C. Mutual Friends of Friends

This feature represents the number of mutual friends of friends of two users [12, 13]

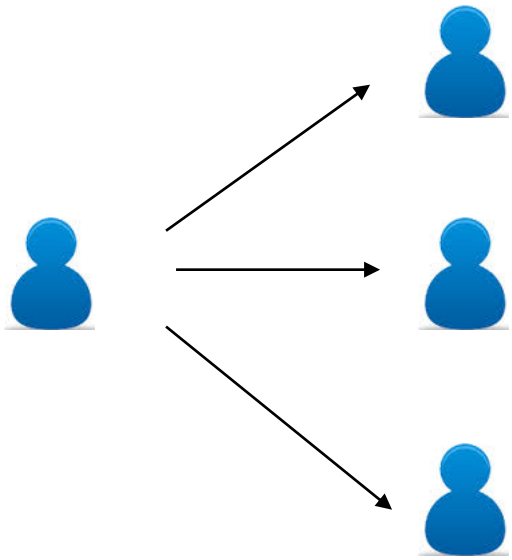

**Fig.2** Find a friend in the network

### D. Approach to find friends in the same network

Our model is based on the similarities of mutual friends that are described in the profile attributes values. These values take as an entity values. The system will match the entity values of profiles. If in general approach two profiles are share some similar entity values then it results shows two person may know each other. We extract attributes values from both the profiles and find similarities between both the values of attributes. This process shows that we find the profile of his/her friend but decision is must to confirm that both profiles of friends so we take a common attribute from the profile which has common value like Common friend names.
This gives the strength to our result. We assign this value to the variable K for each value of attribute k will increase. The improved value of K shows two profiles are of people who know each other. The value of K is use to rank the suggested list.

## IV. CONCLUSION

In this paper our approach is to identify profile of single person between two social networks and also find the friends for the same person in his social network. This algorithm is based on attribute values which are used to build a profile on the network. We extract this value and match with existing string matching algorithm and decision algorithm that justify our approach.

## REFERENCES

[1] Elie Raad, Richard Chbeir, Albert Dipandla“ User Profile matching in social network”
[2] C. Zhou, H. Chen, and T. Yu, “Learning a probabilistic semantic model from heterogeneous social networks for relationship identification,” in Tools with Artificial Intelligence, IEEE International Conference on,vol. 1. Los Alamitos, CA, USA: IEEE Computer Society, 2008, pp.343–350.
[3] L. Ding, L. Zhou, T. Finin, and A. Joshi, “How the semantic web is
being used: An analysis of FOAF documents,” in Proceedings of the
38th Annual Hawaii International Conference on System Sciences, Big Island, HI, USA, 2005, pp. 113c–113c.
[4] P. Mika, “Flink: Semantic web technology for the extraction and analysis of social networks,” JOURNAL OF WEB SEMANTICS, vol. 3, pp. 211—223, 2005.
[5] J. Golbeck and M. Rothstein, “Linking social networks on the web with FOAF: a semantic web case study,” in Proceedings of the 23rd national conference on Artificial intelligence - Volume 2. Chicago, Illinois:AAAI Press, 2008, pp. 1138–1143.

[6] S. Bortoli, H. Stoermer, P. Bouquet, and H. Wache, "Foaf-o-matic -solving the identity problem in the foaf network," in Proceedings of the Fourth Italian Semantic Web Workshop (SWAP2007), 2007.
[7] M. Rowe and F. Ciravegna, "Disambiguating identity through social circles and social data," in Collective Intelligence Workshop ESWC 2008,Tenerife, Spain, 2008.
[8] L. Shi, D. Berrueta, S. Fernndez, L. Polo, S. Fernndez, and A. Asturias,"Smushing RDF instances: are alice and bob the same open source developer?" in ISWC2008 workshop on Personal Identification and Collaborations: Knowledge Mediation and Extraction (PICKME 2008),2009.
[9] P. Bouquet, H. Stoermer, and B. Bazzanella, "An entity name system (ENS) for the semantic web," in The Semantic Web: Research and Applications, 2008, pp. 258–272.
[10] J. J. Jiang and D. W. Conrath, "Semantic similarity based on corpus statistics and lexical taxonomy," International Conference Research on Computational Linguistics (ROCLING X), Sep. 1997.
[11] A. ElSayed, H. Hacid, and D.-A. Zighed, "A multisource context dependent semantic distance between concepts," in 18th International Conference on Database and Expert Systems Applications (DEXA 07),Regensburg, Germany, ser. Lecture Notes inComputer Science, R. Wagner,N. Revell, and G. Pernul, Eds., vol. 4653. Springer, September 2007, pp. 54–63.
[12] Olga Peled, Michael Fire,Lior Rokach and Yuval Elovici: "Entity Matching in Online Social Networks"
[13] Vosecky J., Hong D., and Shen V.Y., "User identification across multiple social networks", In Proc. of First International Conference on Networked Digital Technologies, 2009.
[14] Veldman I., "Matching Profiles from Social Network Sites", Master'sthesis, University of Twente, 2009.
[15] Carmagnola F., Osborne F., and Torre I., "User data distributed on thesocial web:How to identify users on different social systems andcollecting data about them", Proc. of HetRec 2010, USA
[16] Iofciu T., Fankhauser P., Abel F., and Bischoff K., "Identifying usersacross social tagging systems", ICWSM 2011.
[17] Narayanan A., and Shmatikov V., "Robust De-anonymization of LargeSparse Datasets", IEEE Symposium on Security and Privacy, 2008.
[18] A. Hogan, "The Expert Finder corpus 2007 for the benchmarking and development of Expert-Finding systems," in First International ExpertFinder Workshop, 2007.

## AUTHORS

**J. S. Sodhi** Currently working as Head-IT & CIO (Assistant Vice President) with AKC Data Systems (An Amity Group and AKC Group Company).Doctorate in Information Security Management and Certified Security Compliance Specialist (CSCS).His area of interest is IT Management.

**Asmita Joshi** is working as a asst. prof in CSE dept. in CET-IILM-AHL Greater Noida, UP, India. She is pursuing M.tech in CS from Amity University, Noida.