

# Identifying User by Collaborative Statistics: A Review

Rasika Retar<sup>1,\*</sup>

Pratiksha Dhoble<sup>2,\*</sup>

Abhishek Chauhan<sup>3,\*</sup>

Swati Gourkhede<sup>4,\*</sup>

Bhavna Sonkusre<sup>5,\*</sup>

Prof. Pradeep O. Balbudhe<sup>6,#</sup>

\*Student of Bachelor of Engineering,

#Assistant Professor,

Department Of Computer

Engineering,,

Suryodaya College Of Engineering &

Technology, Nagpur, Maharashtra,

India

**Abstract** - Today's we have many as n numbers of online services. Those services use by most of users. Every user has their own unique behaviour or pattern of uses. The users behaviour or we can say usage pattern which can we are going to give a feat by tracking there usage of online services for identifying the user statistically. We are using a powerful algorithm which can help us to identify the user by time, location, as worldwide and which give a output with the help of histogram as a graph.

As user identities of both the database which a database of source one and source two as mention above, then this is a insignificant. A common requirement in our database is to analyzing for identifying user by a feat statistics of their data which we are going to actually work on.

**Keywords** - ICS system, Web User, Data Flow Diagram, Algorithm.

## I. INTRODUCTION

A common requirement in our database is to analyzing for identifying user by a feat statistics of their data which we are going to actually work on. Now a day we have many applications, which can help us to access some information about the user those using the database or we can say source and the same information about the users we can get of the users from the another database or we can say source, and the requirement is to match the particular information from the first database of information from database that according to same hidden user.

As user identities of both the database which a database of source one and source two as mention above, then this is a insignificant. Although, as we say about some applications, the identities of the users which are unknown either in the first database or in the second database or in both; thus this type of situation, the requirement become significant. for ex. The both database might contain information about the users with there a location, city statistics measure over the time period which we calculated distinct. Another example of the online services for the

matching problem the database which collected from the online services which are mention above at the accurate time period.

## II. EXISTING SYSTEM

The studies of this paper which we are using or our subject for study is matching the user those who's have been clearly participated and is to overcome the several problem which occurs in other different communities. For these several problem we present comprise of our approach with related problems of several areas and highlight our contributions. Our work is closely related to de-anonymization concept which is studied in the our reference paper. These concept is closely related to the directly work on the our work with the datasets.

## III. PROPOSED SYSTEM

We resolve all the drawbacks of existing system in our proposed system. The proposed system helps in saving the time and making information flow easy giving beneficial reports.

## A. System Design

The system design deals with data flow diagram & detailed flow graph of the Identifying User By Matching Statistics.

### a) Data Flow Diagram

A Data Flow Diagram (DFD) is a graphical representation of the “flow” can be used to represent a system in terms of the input data to the system, various processing carried out on these data, and the output data is generated by the system.

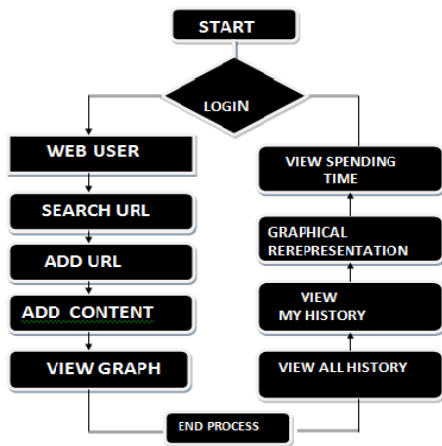


Fig. 1 Data Flow Diagram

A DFD represents flow of data in a system. Data flow diagrams are commonly used for problem analysis. Movement of information through the different level or processes in the system is shown in Data Flow Diagram of Fig. 1.

### b) Detailed Flow Graph

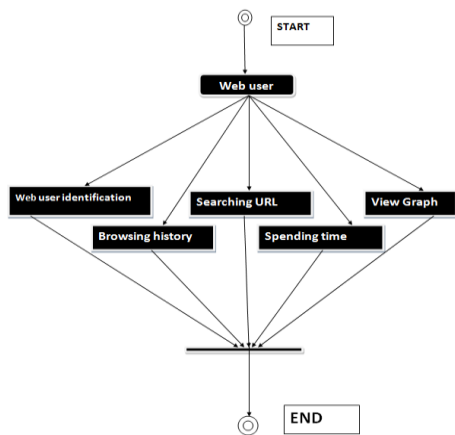


Fig. 2 Detailed Flow Graph

In the above flow diagram we can say the fully process of our system which is identifying user by matching statistics.

In the above flow diagram we see the various components of our system. The main component of our system which we are gone implemented. Our process start with login from registration; the registration makes our system secure. The web user follows many processes those are included with the search url, add url, add content and view graph.

#### i. Web user

The web user is for entering the information regarding new person who are login through the registration, controlling all the database like any changes regarding to the name, phone numbers, location ,url etc. The web user also controls the faulty registration like entering a new registration. The web user also have their own standard for the making or working with all this database. Its controls all the activity for the generating the database in the various of the requirement as the including the registration. The webuser will check the all the information about the i.e. person updates, url updates, url content update etc. The web user has the highest level of priority in the our identification of user bt collaborating system.

#### ii. Web user identification

While the person going to be registries and get there username as well as their password they have choose the option of the web user for them to identified as a web user . if the are once entered in the web user for there particular it help to them to see all these functionality about the web user these are browsing history, searching url ,spending time, view graph.

#### iii. Browsing history

In the web user as the user done his /her registration, the registered person get on the login page. And add their user name and password and select the web user option which is giving in the dropdown arrow. Then they see the

next page in which they have to search for the url. And adding content for the those particular url which is entered the same. This all information help us to generating the database for the browsing history in the web user for the working on it.

#### iv. Searching URL

The searching the url is the target we achieving for the particular web site address. Which is we helping for the various types of their particular information that going to be we are searching as well as it helps us to generate the database with respect to the quires, as well as updating the database for their contribution for the working.

The searching url along with the one work is there for us is the add content about the searching url. Is it is very necessary to adding the contents for the particular web site which are we going ton to search it help us to generate the database as with the information about the every url i.e if we search the web site as [www.scet.in](http://www.scet.in) after searching we have to add something details about the scet such as what is sect?, sect location etc for our database.

#### v. Spending time

As we are see above that there are we are going to register and select the web user option for the controlling all these operation then we are selecting the browsing history along with the searching url and adding the contents.

These all procedure is recorded with the particularly time period. Such as how many times we are spending the web site as example given above. Its a way to calculated all the time which is going to we are spending with these website,

The web sites are those which are we can search in the searching the url as well as which are save in our database and with their content.

#### vi. View graph

At end we can see the after all these process the output for browsing history, searching url, spending time etc.

We are getting all these information with the help of there graphically presentation as we can say there are we making the database are of the two types . the one database is compared with the second database graphically and we can get the output graphically for the same .

But all of the sudden its is very in fraction of time we can get the graphically output with the graph according to the searching the url , browsing history , web user can get the output but the most important with the respect of the time.

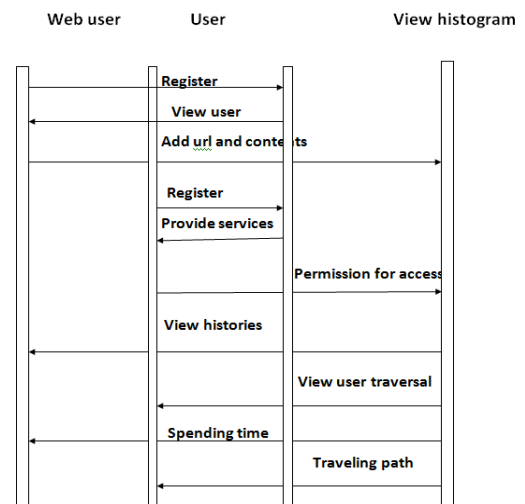


Fig.3 Working of web user

## IV. ALGORITHM

The step by step representation of our program is known as the algorithm. In this system the our algorithm is the heart our all the system . as we know that all the procedure we are making with the web user actually we are creating the database. The database the real database we are on working state. For the combining all the database which are separated by the time period and collecting the information with the help of our browsing history , searching url, spending time.

### A. Collaborative Filtering

Collaborative filtering algorithm we are using in this system for making automatic predictions (filtering) according to a behaviour of users by collecting the use of

their own priority or tested of their information from many users.

The collaborative filtering algorithm which works on the two subgroups, or we can say the two database which we are making in the system.

The collaborative filtering approach is that if a person A has the same point of view as person B on an issue, A is more likely to have B's point of view on a different issue than that of a randomly chosen person.

In the general or popular way, the collaborative filtering algorithm is the process of filtering for the information or patterns using technique including collaboration along multiple agent view points, data source as database, etc.

Collaborative filtering methods have been applied to many different types of data such as in mineral exploration, monitoring data.

## V. CONCLUSION

In the system we which we are implemented in this paper where we have study the requirement for user identification from the statistics of their online uses of services. Accordingly, we are given an data mining concept with database in the form of histograms which accordingly belonging to database of the users and another independent database of histogram which generated by the same users which are using online services. We are spracing in the system the accuracy of our collaborative filtering algorithm.

Our proposed system which can be implemented with the help of minimum weight maximal collaborative algorithm on a complete weighted graph and higher accuracy than heuristics based method with the three databases of different parameters or we can say different natures.

The identification of users via matching collaborative filtering algorithm can something result in higher accuracy than existing system based more complicated data patterns or models.

## REFERENCES

- [1] Farid M. Naini, Jayakrishnan Unnikrishnan, Patrick Thiran and Martin Vetterli, "Where you are is who you are: user identification by matching statistic", (volume: 11, issue: 2), 05 November 2016.
- [2] Joonseok Lee, Mingxuan Sun, Guy Lebanon, "A Comparative Study of Collaborative Filtering Algorithms." 2012.
- [3] Reena Pagare, Shalmali A. Patil, "Study of Collaborative Filtering Recommendation Algorithm- Scalability Issue" international journal of computer Application(0975-8887) Volume 67-No. 25, April 2013.
- [4] Peng Peng, Hu Bin, "A Collaborative Filtering Recommendation Based on User Characteristics and Time Weight[J]", Journal of Wuhan University of Technology, vol. 31, no.3, pp. 24-28, 2009.
- [5] Huang Huang, Wei Wei, Ye Ye et al., "Collaborative Filtering Algorithm Based on User property and item Category[J]", Computer & Digital Engg., vol. 40, no.10, pp.5-7, 2012.
- [6] Zha Zha, Li Li, Xu Guqiong, "An Optimised Collaborative Filtering Algorithm Based on Combined Similarity[J]", Computer Applications and Software, vol. 31, no.12, pp 323-328, 2014.
- [7] Wu Wu, Wang Haoran, "Collaborative Filtering Algorithm Using User Background Information[J]", Computer Application, vol. 28, no. 11, pp .58-60, 2008.
- [8] P B K kantar, L Rokach, F Ricci et al., "Recommender Systems handbook[M]", Springer, 2011.
- [9] L Lu, M Medo, C H Yeung et al., "Recommender systems[J]", Physics Reports, Vol. 519, no. 1, pp. 1-49, 2012.