ISSN 2395-1621

# Personalized Web Search

[#1]Harshali Kudle, [#2]Shilpa Goski, [#3]Varsha Taru, [#4]Prof Deepak Uplaonkar

[1]harshalikudle@gmail.com
[2]shilpagoski4@gmail.com
[3]varshataru17@gmail.com
[4]uplaonkar@gmail.com

[#1234]Department of Computer Engineering,JSPM NTC Pune,India

**ABSTRACT**

**Personalized web search (PWS) is used to improving the quality of various search services on the Internet. We study privacy protection in PWS applications and it display reran ked result on Google search engine according to user interest. For finding user interest personalized web search uses user's tweets or comments or likes on social sites. For example:- Facebook, tweeter.We propose a PWS framework is also called as UPS(User customizable privacy preserving search). We present ANN (Artificial Neural Network) algorithm for runtime generalization. The experimental results also reveal that ANN is very efficiency.**

**Keywords— Privacy protection, personalized web search, utility, risk, profile**

ARTICLE INFO

**Article History**

Received 21st March 2016

Received in revised form :

23rd March 2016

Accepted : 25th March 2016

**Published online :**

**28th March 2016**

## I. INTRODUCTION

The web search engine is useful for ordinary people which is provide useful information on the web. sometimes it return the all possible result to the users but many times search engines return irrelevant results that do not meet their real intentions. personalized web search (pws) providing better search results for individual user. in pws, user information has to be collected and analyzed then re-ranked the results according to user interest.PWS is divided into two types namely click-log-based methods and profile-based. The click-log based methods are simple and they impose bias to clicked pages in the user's query history. In Profile-based methods improve the search result from user profile history. Profile-based methods can be potentially effective for almost all sorts of queries.UPS is divided into two types namely Non trusty search engine server and number of clients. Each client or users accessing the search service on the internet. UPS framework work in a two phases first one is online phases and second is a offline phase. In offline phase all hierarchical user profile is constructed and in online phase handled all the queries. The main aim of the UPS is protecting the privacy of every individual user profile.

.

## II. MOTIVATION

Our project is based on profile based pws.it means in profile based searching history of user cannot be saved.it used only single session.it improve the search quality with the personalization utility of the user profile. On the other hand , in the existing system click log based was used. In that searching history of user was saved.

## III. GOAL

In our project, Within fraction of second Google search engine display the result according to the user's.

## IV. OBJECTIVE

It will overcome the problem of function overloading. Our main aim is to remove the redundant and irrelevant features. This is the main object of our system code.

## V. SYSTEM ARCHITECTURE

Our system work in a two phases namely training phase and search phase. In training phase we have to log in to the social networking site such as a tweeter then download all the tweets. After that parse that tweets then apply the TF (Term Frequency- is nothing but it displays how many times one word is repeated in single document). Next step we have to perform i.e categorized the tweet according to domain.

 

In search phase we have to fired queries on Google search engine then fetch the result from Google then following operations are performed on that queries that is parse and term frequency.

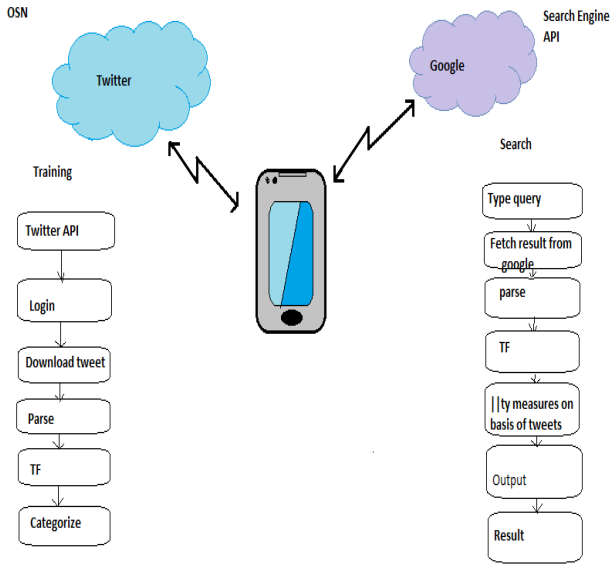


FIG1: SYSTEM ARCHITECTURE

## VI. STATEMENT OF SCOPE

For future work, we will try to use more number of domains and sub-domain like sport, entertainment, health etc.

## VII. MAJOR CONSTRAINTS

If the training of ANN is not properly done then it will not generate the proper result.
We will use limited number of domains.

## VIII. SOFTWARE REQUIREMENT SPECIFICATION

Language: Java J2SE and JDK: J2SE (Java 2 Standard Edition) Java would be the required as language for development of the project. JDK is the development kit used to compile java programs.
IDE: NetBeans: Just like visual studio provides development environment for VB and .Net, NetBeans provides an integrated development environment (IDE) for Java.
Client-Server Architecture using Serialized Objects: In case the project needs client-server communication this is how it is handled in java.
GUI - AWT& SWING are used for GUI design.

## IX. EQUATIONS

- TF(Term frequency) for single document=Total occurrences of word/Total word
- IDF(Inverse Document Frequency) for document=log2(Total document /doc containing word)
- TF*IDF

## X. ALGORITHM

In our project we are using the ANN Algorithm.ANN stand for Artificial Neural Network.Below fig shows that every neuron is connected to each neuron.

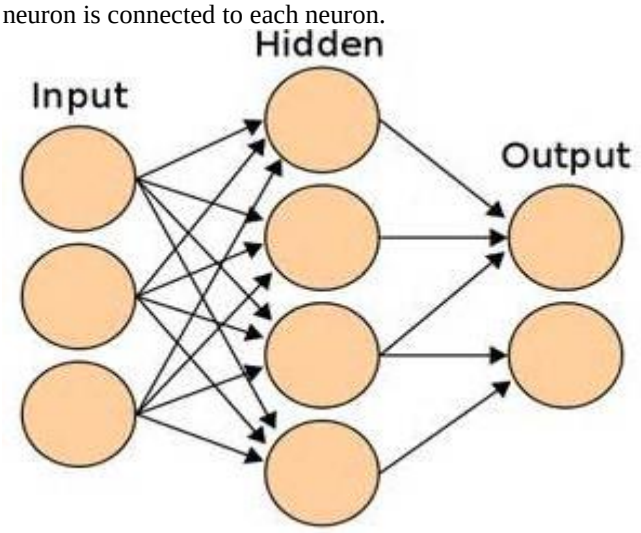


FIG2: ARTIFICIAL NEAURAL NETWORK

1. Input Layer
2. Hidden Layer
3. Output Layer

Artificial neural networks are generally presented as systems of interconnected neurons which exchange messages between each other. Input are always Normalize."

- ANN contain 2 phases:
1. Training Phase
2. Recognization Phase

In training phase, we have to get Input and Output also. Training phase is also called as BPNN. It contains two types: Feed Forward and Feed Backward.
In recognization phase, we have to only give Input and ANN provide the Output. The recognization phase contains only Feed Forward.Neurons work by processing information. They receive and provide information in form of spikesan artificial neural network is composed of many artificial neurons that are linked together according to a specific network architecture. The objective of the neural network is to transform the inputs into meaningful outputs.

## XI. CONCLUSION

In this project Personalized web search (PWS) is used to improve the quality of various search services on the Internet. Privacy preserved PWS methods are used to protect the disclosure of personal information in search process. User customizable Privacy-preserving Search (UPS) framework is used to support privacy in search process. The UPS scheme is enhanced with attack resistant methods. Personalization utility is high in the personalized web search scheme. The system reduces the generalization risk levels. The system increases the attack control rate. Priority based user profile construction process is supported by the system

 

## XII. ACKNOWLEDGEMENT

I would like to express my sincere gratitude towards my guide Prof. Deepak Uplaonkar for the help, guidance an encouragement in the development of this methodology. They supported me with scientific guidance, advice and encouragement, and were always helpful and enthusiastic and this inspired me in my work. I have benefitted from numerous discussions with guide.

## REFRENCES

[1] Z. Dou, R. Song, and J.-R.Wen, "A Large-Scale Evaluation and Analysis of Personalized Search Strategies," Proc. Int'l Conf. World Wide Web (WWW), pp. 581-590, 2007.

[2] J. Teevan, S.T. Dumais, and E. Horvitz, "Personalizing Search via Automated Analysis of Interests and Activities," Proc. 28th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), pp. 449-456, 2005.

[3] M. Spertta and S. Gach, "Personalizing Search Based on User Search Histories," Proc. IEEE/WIC/ACM Int'l Conf. Web Intelligence (WI), 2005.

[4] B. Tan, X. Shen, and C. Zhai, "Mining Long-Term Search History to Improve Search Accuracy," Proc. ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD), 2006.

[5] K. Sugiyama, K. Hatano, and M. Yoshikawa, "Adaptive Web Search Based on User Profile Constructed without any Effort from Users," Proc. 13th Int'l Conf. World Wide Web (WWW), 2004.

[6] X. Shen, B. Tan, and C. Zhai, "Implicit User Modeling for Personalized Search," Proc. 14th ACM Int'l Conf. Information and Knowledge Management (CIKM), 2005.

[7] X. Shen, B. Tan, and C. Zhai, "Context-Sensitive Information Retrieval Using Implicit Feedback," Proc. 28th Ann. Int'l ACM SIGIR Conf. Research and Development Information Retrieval (SIGIR), 2005.

[8] F. Qiu and J. Cho, "Automatic Identification of User Interest for Personalized Search," Proc. 15th Int'l Conf. World Wide Web (WWW), pp. 727-736, 2006.

[9] J. Pitkow, H. Schu¨ tze, T. Cass, R. Cooley, D. Turnbull, A. Edmonds, E. Adar, and T. Breuel, "Personalized Search," Comm. ACM, vol. 45, no. 9, pp. 50-55, 2002.

[10] Y. Xu, K. Wang, B. Zhang, and Z. Chen, "Privacy-Enhancing Personalized Web Search," Proc. 16th Int'l Conf. World Wide Web (WWW), pp. 591-600, 2007.

[11] K. Hafner, Researchers Yearn to Use AOL Logs, but They Hesitate, New York Times, Aug. 2006.

[12] .A. Krause and E. Horvitz, "A Utility-Theoretic Approach to Privacy in Online Services," J. Artificial Intelligence Research, vol. 39, pp. 633-662, 2010.

[13] J.S. Breese, D. Heckerman, and C.M. Kadie, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," Proc. 14th Conf. Uncertainty in Artificial Intelligence (UAI), pp. 43-52, 1998.

[14] P.A. Chirita, W. Nejdl, R. Paiu, and C. Kohlschu¨ tter, "Using ODP Metadata to Personalize Search," Proc. 28th Ann. Int'l ACM SIGIR Conf. Research and Development Information Retrieval (SIGIR), 2005.

[15] .A. Pretschner and S. Gauch, "Ontology-Based Personalized Search and Browsing," Proc. IEEE 11th Int'l Conf. Tools with Artificial Intelligence (ICTAI '99), 1999.

 