ISSN XXXX XXXX © 2018 IJESC

Research Article Volume 8 Issue No.2

Personalized Web Search Based On Client Side Ontology

K.Senthil Kumar[1], A.Abirami[2]
Assistant Professor[1], Student[2]
Department of Computer Applications
Karpagam College of Engineering, Coimbatore, India

Abstract:
Web search engines are valuable tools that are widely used to find specific information in the World Wide Web. Every user has a distinct background and a specific goal when searching for information on the Web. The goal of Web search personalization is to tailor search results to a particular user based on that user's interests and preferences. Personalized Web Search (PWS) is a general category of search techniques with the aim of better results for the search, which are designed for the individual needs of users. The user information must be collected, analysed, mined and used behind the issued query by the user. In this project we propose and build a personalized web search model for weblog data. The log contains useful information such as date, time, IP address and query.
We present this to personalized search that involves building models of user context as ontological profiles by assigning completely derived interest scores to existing concepts in domain ontology. The web search results from existing search engine are ranked based on the client's profile and the semantic evidence in an ontological user profile is used for elective presentation of the most relevant results to the user. Client's profile is stored as an ontology tree.

I. INTRODUCTION

The project entitled as "PERSONALIZED WEB SEARCH BASED ON CLIENTSIDE ONTOLOGY" is developed using PHP and MySQL. This project is developed for providing relevant results to the submitted query in a web search to the intended web users along with rank based on number of number of visitors, number of click through the data. We present an approach to personalized search which involves building models of user context as ontological profiles by assigning implicitly derived interest scores to existing concepts in domain ontology. The web search results from existing search engine are ranked based on the client's profile and the semantic evidence in an ontological user profile is used for elective presentation of the most relevant results to the user. Client's profile is stored as an ontology tree. Personalized web search (PWS) is a general category of search techniques. This project is aimed at providing better search results, which is created for individual user needs. Ontological concepts and methods in the computer field have been used for knowledge representation, knowledge sharing and reuse. Ontology concepts are utilized to represent user profiles. Ontologies provide a common understanding of topics for communication between system and users, and enable Web-based knowledge processing, sharing, and reuse between applications. Ontologies enable intelligent agents to gather Web information for users in knowledge-based Web gathering.

II. PROBLEM SPECIFICATION

Web search engine provides each user with more relevant information, several approaches were proposed for adapting search results according to each user's information need. Although there are several search engines currently present, it has been observed that they fails to capture user's preference and behaviour and hence the search results may or may not be related with the profile of the user.

Literature review
Most commercial search engines return roughly the same results to all users. However, different users may have different information needs even for the same query.

Architecture of personalized search model

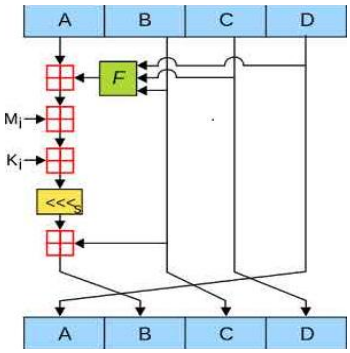
Personalized search system's outline is as follows:
User has to create his profile by signing up into the system. New user has to fill up the sign up form and submit it to the system to create username and password to access and login into the proposed system. Once the user login into the system, a home page will appear where the user has to enter a query into the search text box to get search results from the web. Query search using STRING SIMILARITY MATCHING ALGORITHM. Encrypt using the rijndael algorithm.

III. ALGORITHM

There are 3 algorithm used for this ontology method.
System Development is a series of operations performed to manipulate data to produce output from a computer system. In this module, users can login using the unique email id and password. The password is encrypted using MD5 algorithm. After login, user post some queries based on the dataset that is loaded into the database already. In registration module, user should register their details in signup form.

MD5 ALGORITHM (MESSAGE DIGEST)
The MD5 algorithm is a widely used hash function producing a 128-bit hash value. Although MD5 was initially designed to be used as a cryptographic hash function, it has been found to suffer from extensive vulnerabilities.. Like most hash functions, MD5 is neither encryption nor encoding. It can be cracked by brute-force attack and suffers from extensive vulnerabilities as detailed in the security section below. The source code in RFC 1321 contains a "by attribution" RSA license. The abbreviation "MD" stands for "Message Digest." The security of the MD5 has

been severely compromised, with its weaknesses having been exploited in the field, most infamously by the Flame malware in 2012. The Despite this known vulnerability, MD5 remains in use. Md5 is an algorithm that is used to verify data integrity through the creation of 128-bit message digest from data input (which may be a message of any length) which is claimed to be as unique to that specific data as a fingerprint is to the specific individual. Message digest is a cryptographic hash function containing a string of digits created by a one-way encryption formula. Message digest are designed to protect the integrity of a piece of data or media to detect changes and alteration to any part of a message. Reversing MD5 is actually considered as malicious, there is a very small chance for reversing .so it is often used for password encryption. MD5 has been utilized in a wide mixture of security applications, and is additionally ordinarily used to check the trustworthiness of records.



PSEUDO CODE:

STEP 1: Initialized the state variables A, B, C, D.

STEP 2: It uses four functions to thoroughly goober the above state variables.

F(X,Y,Z) = (X & Y) | (~(X) & Z)

G(X,Y,Z) = (X & Z) | (Y & ~(Z))

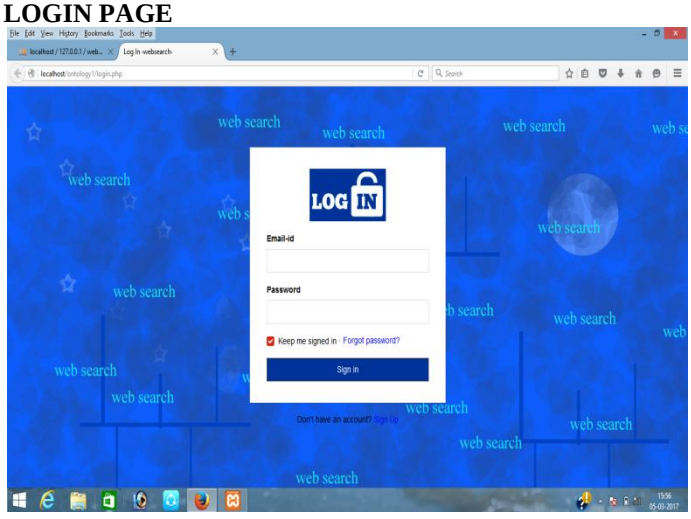
H(X,Y,Z) = X ^ Y ^ Z

I(X,Y,Z) = Y ^ (X | ~(Z)) Where &, |, ^, and ~ are the bit-wise AND, OR,

XOR, and NOT operator.

STEP 3: The transformation of the state variables from their initial state is achieved by using the state variable and input message along with the above functions. Then the message digest is generated and stored in the state variables A, B, C, D.

LOGIN PAGE



QUERY SEARCHING

In this module, user submits the query. Based on the query, relevant results will be displayed and based on the submitted query some history results will also be displayed. Here, the data is retrieved by using String Similarity Match (SSM) algorithm.

STRING SIMILARITY MATCHING ALGORITHM

In string similarity matching algorithm, the objective is to calculate the similarity between two strings and returns the number of matching characters in both strings. And also matches for short strings in long texts, in situations where a small number of differences is to be expected. For Client post query q to the server, server retrieves query results C to the clients. Results have been extracted by using the String similarity matching algorithm.

PSEUDO CODE

STEP 1: Given a pattern string P =P1, P2… Pn.

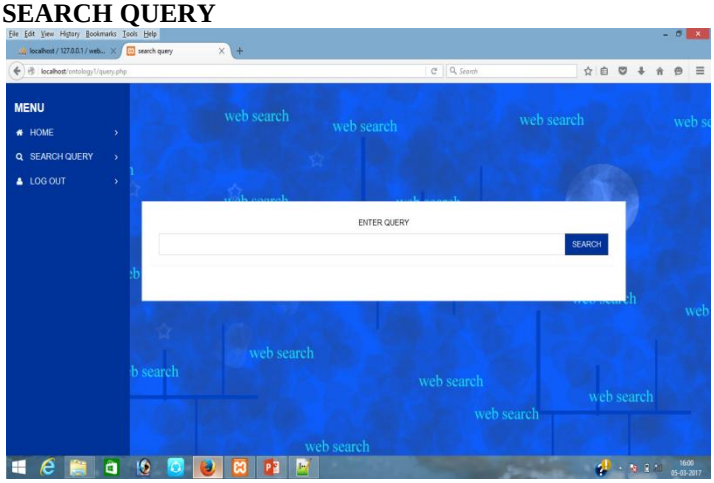
STEP 2: Text string T = T1 , T2….Tn.

STEP 3: Find a substring in Tj`,j = Tj`…..Tj in T, which, of all substrings of T, has the smallest edit distance to the pattern P.

STEP 4: Compare the user given word and word in database.

STEP 5: In that similarity calculation, extract the features in the dataset and display it.

SEARCH QUERY



ENCRYPTION

In this module, chosen dataset has been loaded into the database in encrypted form by using Rijndael Algorithm; we can view the query and IP Address field in encrypted form.

IV. RIJNDAEL ALGORITHM

Rijndael is the block cipher algorithm recently chosen by the National Institute of Science and Technology (NIST) as the Advanced Encryption Standard (AES). It supercedes the Data Encryption Standard (DES). Rijndael as the standard symmetric key encryption algorithm to be used to encrypt sensitive information. Rijndael is an iterated block cipher. The algorithm that they developed was designed as an easily understandable mathematical structure that can be broken down into simple components. Daemen and Rijmen write in their proposal to AES that Rijndael was designed based on the following three criteria

- Resistance against all known attacks;
- Speed and code compactness on a wide range of platforms;
- Design simplicity

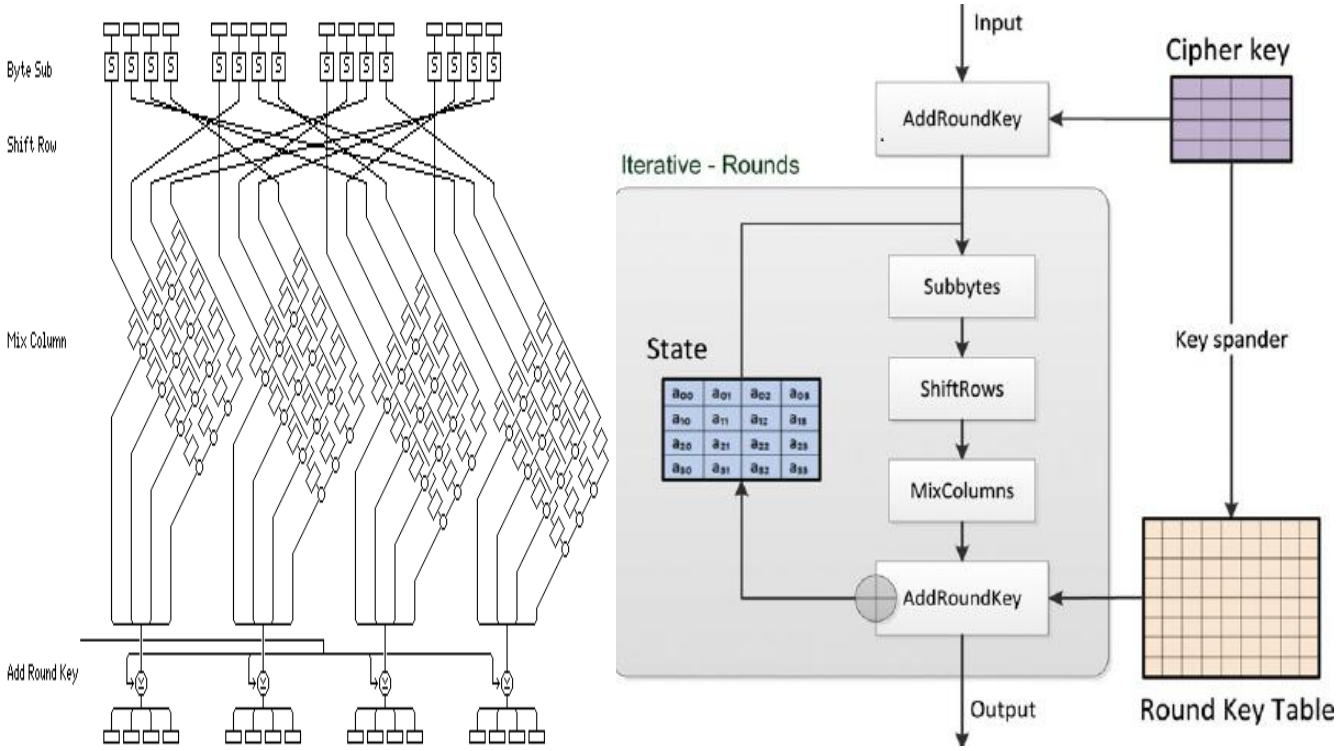


Figure.1. Rijndael Algorithm

ADVANTAGE[5]

- Rijndael's advantages based on implementation aspects, simplicity of design, variable block length and extensions.
- Rijndael's implementation is very flexible since it can be used with varying key sizes and block sizes.
- It is also possible to change the sequence of some steps in Rijndael without affecting the cipher and the cipher has a simple and elegant structure.
- It does not hide its structure by using complex components. Instead, it benefits from the advantages gained by the use of simple components in a well defined structure.

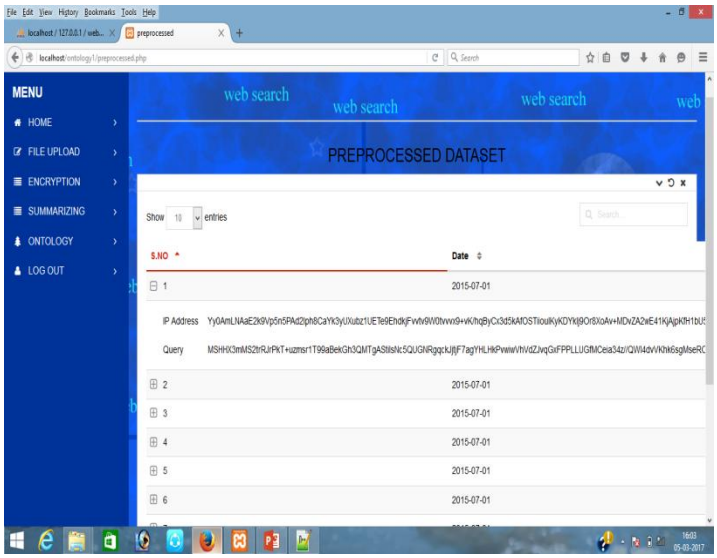
DISADVANTAGE[5]

- Rijndael can be subject to standard techniques of differential and linear cryptanalysis.
- It is also weak to an attack called the "Square attack" that is based on the way matrix multiplication works. However, in practical terms, this attack is not sufficient to compromise the security of the Rijndael algorithm.
- Rijndael is also limited by its inverse cipher. Though the encryption is suited for many applications and is quite fast, the inverse cipher takes more code.

PSEUDO CODE:

STEP 1: An initial Round Key addition (perform an Add Round Key
step (XOR ing with the block).
STEP 2: Nr-1 Rounds (Nr is the number of rounds).
STEP 3: Final Round
Round(State,RoundKey) {
ByteSub(State);
ShiftRow(State);
MixColumn(State);
AddRoundKey(State,RoundKey);

ENCRYPTION



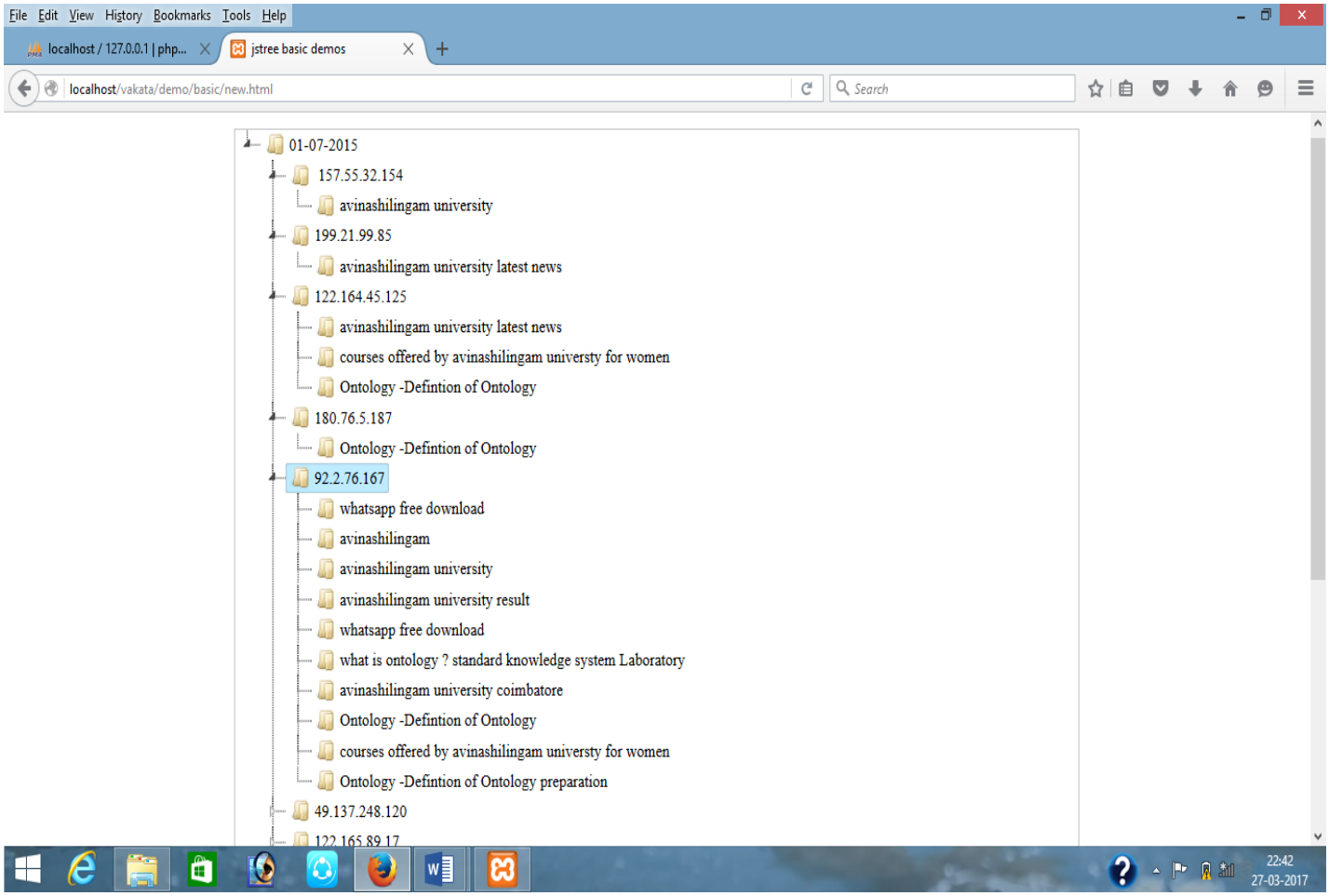
SUMMARIZING

In this module, again pre-processed the dataset which will displayed query with count based on number of frequent visitor, number of click through the data.

DESIGNING ONTOLOGY

In this module, creating a ontology tree based on client side profile using jquery frame. It includes fields like date, IP address, and query.

ONTOLOGY TREE



V. CONCLUSION

Personalized web search is a better way to improve web search quality. Although personalized search was unclear, whether personalization is constantly effective on different queries for different users and under different search context, the proposed framework has overcome this limitation. It requires users to grant the server full access to personal information on the Internet, which violates the user's privacy. In this project, it proposes to provide fast and relevant search are personalized using user profile and also protect the user's privacy where the data are encrypted and stored at server. So it can protect web server from web attacks like URL manipulation attack, terror attacks. The objective of this work is to improve the performance and also to reduce the query execution time. With this progression, we have created an ontology based on client profile which assists for analysis of user's behavior.

VI. REFERENCES

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

[2]. Lidan Shou, He Bai, Ke Chen and Gang Chen, 2014, "Supporting Privacy Protection in Personalized Web Search", IEEE. 26(2).

[3]. Jayanthi. J, Dr. K. S. Jayakumar, Sruthi Surendran, 2011, "Generation of Ontology Based User Profiles for Personalized Web Search", 978-1-4244-8679-3/11 IEEE.

[4]. Fang Liu, Clement Yu, Senior Member, Weiyi Meng, Member, IEEE., 2004., "Personalized Web Search for Improving Retrieval Effectiveness", IEEE Transactions on Knowledge and Data Engineering, 16(1).

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

WEBSITES

[6]. https://www.researchgate.net

[7]. https://www.academia.edu

[8]. https://www.ijarcce.com

[9]. https://www.w3schools.com

[10]. https://www.php.net