

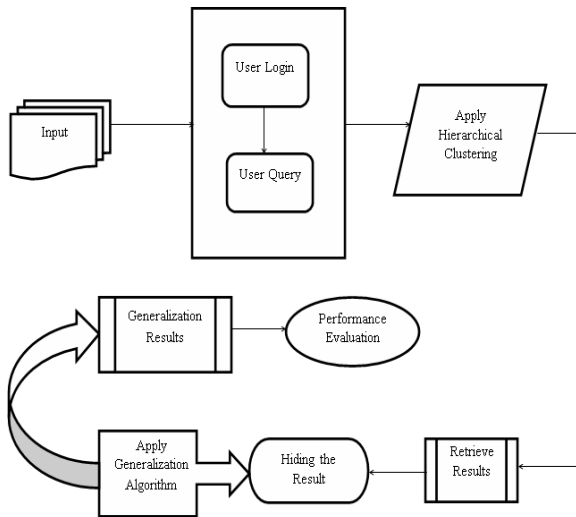




**b. Disadvantages**

1. It is not more users -specific information into consideration. The difficulties in doing this include integration of different information sources, modeling of the correlation between various information and the user's search behaviors, and efficiency concerns.
2. It does not design more sophisticated learning and ranking algorithms to further improve the performance of our system.

**3. System Architecture**



**Figure 1: System Architecture**

**4. Modules**

**A. Dataset preprocessing**

Most commonly a data set corresponds to the contents of a single statistical data matrix, or a single database table, where every column of the table represents a particular variable, and each row corresponds to a given member of the data set in question. The data set lists values for each of the variables, such as height and weight of an object, for each member of the data set. Each value is known as a datum. The data set may comprise data for one or more members, corresponding to the number of rows. This module, choose input dataset. Chosen dataset has been loaded into the database. After loading the dataset into the database, we can view the dataset. By using the string matching algorithm we filter out unwanted values in the dataset and it has been preprocessed and store into the database.

**B. User Login**

This is for user login page. In this module, users are entered by using the unique id and password. In this module, users are entered after registering. After registering each user has unique id. After login, user posts some queries which are based on our dataset which is loaded into the database.

**C. Query Searching and Search Results Retrieval**

In this module, user submits query. Based on the query, relevant results has been displayed and also based on the submitted query some history results are displayed. Based on the query and already posted queries, we can calculate

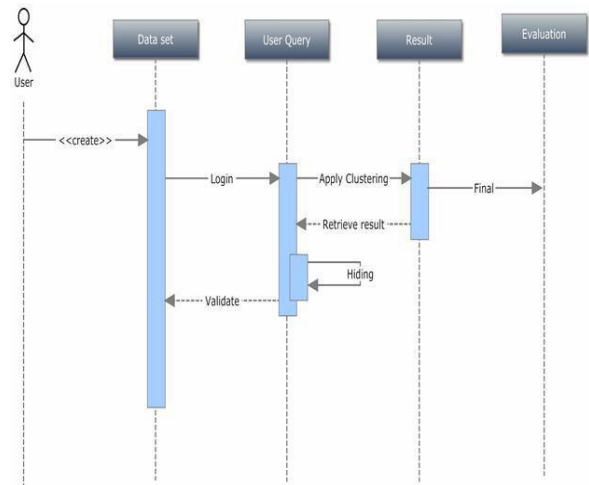
the similarity values between them. In that three types of similarity values has been estimated. From that, the result is retrieved which is based on the high relevant results by using the minimum range of similar values.

**D. Estimate Relevant Results**

In this module, user posts query and sub query also. Based on the query and sub query, estimate the results based n string matching. Based on the relevant results and total number of data's in the dataset, we can estimate the support values.

**E. Retrieve user profile in privacy manner**

In this module, adversaries to mine the history results means, only query time has been displayed. In this, other information such as query, query results, username are not displayed by using the background knowledge. First we generalize the table, and then suppress the values based on the generalized table. Generalized values are stored in the history results. When the adversaries' views the history result means, they can only view the generalized results. Finally, the performance can be evaluated by using the parameter such as time, cost and communicational and computational cost.



**Figure 2: Sequence Diagram**

**5. Existing System**

**A. Methodology of Existing System**

In the Existing Work, a client-side privacy protection framework called UPS for personalized web search was proposed. UPS could theoretically be adopted by any PWS that captures user profiles in a hierarchical taxonomy. The context allowed users to stipulate customized privacy requirements via the hierarchical profiles. In addition, UPS also performed online generalization on user profiles to protect the personal privacy without compromising the search quality. In this they proposed two greedy algorithms, namely GreedyDP and GreedyIL, for the online generalization. In this for query mapping process it has various steps to compute the relevant items.

Most works on anonymization focus on relational data where every record has the same number of sensitive attributes. There are a few works taking the first step towards anonymizing set-valued or transactional data where





