RAMA UNIVERSITY

www.ramauniversity.ac.in

FACULTY OF EGINEERING AND TECHNOLOGY

Soft Computing

LECTURE -35

Umesh Kumar Gera
Assistant Professor
Computer Science & Engineering

OUTLINE

- Genetic algorithm based Internet Search Techniques
- Important process in web information retrieval framework
- Components of an Information Retrieval System
- Clustering of Document and Terms
- Steps of TSP using GA
- Multiple Choice Question
- References

Genetic algorithm based Internet Search Techniques

People use search engines to find information they desire with the aim that their information needs will be met. Information retrieval (IR) is a field that is concerned primarily with the searching and retrieving of information in the documents and also searching the search engine, online databases, and Internet. Genetic algorithms (GAs) are robust, efficient, and optimized methods in a wide area of search problems motivated by Darwin's principles of natural selection and survival of the fittest. This discussion describes information retrieval systems (IRS) components also looks at how GAs can be applied in the field of IR and specifically the relevance of genetic algorithms to internet web search. Finally, from the proposals surveyed it turns out that GA is applied to diverse problem fields of internet web search.

GENETIC ALGORITHM BASED INTERNET SEARCH TECHNIQUES

Important process in web information retrieval framework

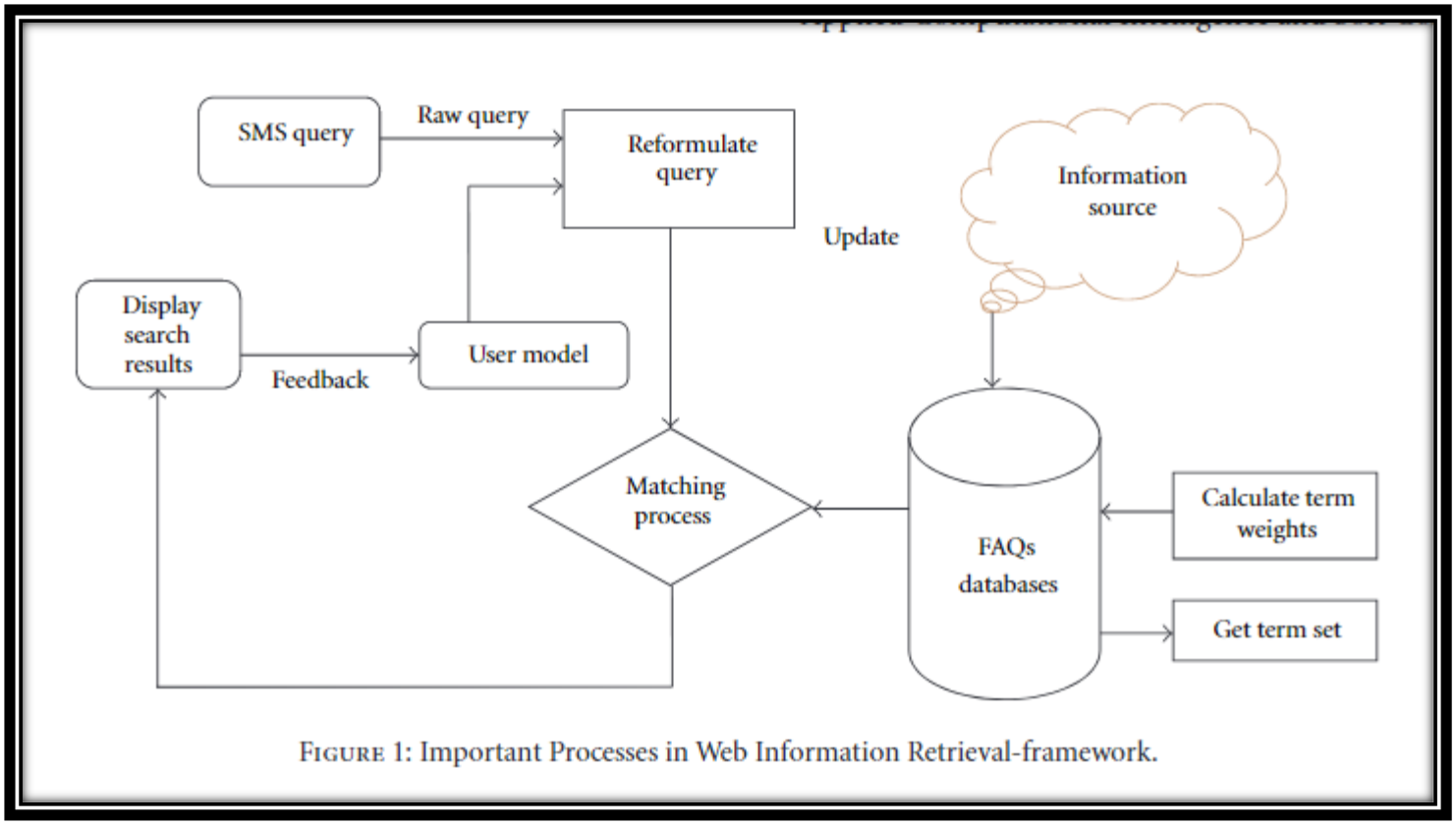

FIGURE 1: Important Processes in Web Information Retrieval-framework.

Components of an Information Retrieval System

There are two major components of IRS system:

1. **Document databases**

   The document databases stores the databases related to documents and the representations of their information contents based on TFIDF approach. An SMS-query keyword term is also associated with this component which automatically generates a representation for each document by extracting the frequency of the SMS-query keyword terms from the document contents.

2. **Reformulated query processing system**

   There formulated query processing System consists of two sub-systems: Searching-Matching Unit and Displaying-Ranking Unit. Searching unit allows user to search the documents from the document database, and matching unit does a comparison of all documents against the user's query.

Clustering of Document and Terms

In this field, two approaches have been applied for obtaining user-oriented document clusters.

The main features of the GA are as follows.

1. **Representation scheme:**

   Two different coding schemes are considered to include division-assignment and separator methods

2. **Initial population:**

   The first generation of the chromo-somes depends on the chosen coding, and the rest of individual are randomly Generated.

3. **Operators:**

   Each operator has an application probability associated and is selected spinning the roulette. Different crossover and mutation operators are used.

4. **Fitness function:**

   a measure of the relative entropy and Pratt's measure are two proposals adopted.

MULTIPLE CHOICE QUESTION

1. Which of the following is false in the case of a spanning tree of a graph G?

a) It is tree that spans G

b) It is a sub graph of the G

c) It includes every vertex of the G

d) It can be either cyclic or acyclic

2. Every graph has only one minimum spanning tree.

a) True

b) False

3. Consider a complete graph G with 4 vertices. The graph G has ____ spanning trees.

a) 15

b) 8

c) 16

d) 13

4. Which of the following is not the algorithm to find the minimum spanning tree of the given graph?

a) Boruvka's algorithm

b) Prim's algorithm

c) Kruskal's algorithm

d) Bellman–Ford algorithm

5. Which of the following is false?

a) The spanning trees do not have any cycles

b) MST have n – 1 edges if the graph has n edges

c) Edge e belonging to a cut of the graph if has the weight smaller than any other edge in the same cut, then the edge e is present in all the MSTs of the graph

d) Removing one edge from the spanning tree will not make the graph disconnected

REFERENCES

- https://www.javatpoint.com/artificial-neural-network-hopfield-network
- https://www.geeksforgeeks.org/traveling-salesman-problem-using-genetic-algorithm/