

A Survey on Semantic Search in E-Commerce Website Ranking

Kaviya.M¹, Siddique Ibrahim S. P², Kirubakaran

¹ PG Scholar, ^{2,3} Assistant Professor, Department of Computer Science and Engineering.

Kumarguru College of Technology, Coimbatore, TamilNadu, India

ABSTRACT

In this speedy life everybody need an immediate answers for their searches in websites .In this scenario semantic search engines will be helpful. It deals with the actual meaning of the queries. Searched by keyword and do not understand its meaning are some reasons why the traditional search engine is not suitable anymore. Semantic web search technologies (like Google Knowledge Graph)is emerging into the answer engine market to transform the unstructured data into structured data. The proposed search engine is user-friendly and gives right-to-the-point result .The E-commerce industry has grown rapidly and everyone started using it, like online shopping (flip kart). Even though it is the emerging trend, it has some defects like comparison of price between different E-commerce websites, compare weakness and strength between the competitors. In this research, Semantic and neural based page ranking algorithm is used to overcome this defects.

Keywords : Semantic, Ontology, Crawler.

I. INTRODUCTION

A. Semantic Web

A web search engine like Google is software used to search for information on the World Wide Web (WWW) [1]. The search results are usually displayed in a line of results often referred to as search engine results pages (SERPs)[5]. The information is a mix of texts, images, videos and other files.The tremendous growth in the volume of data or the information lead the traditional search engines to get the answers syntactically correct but large in amount [7].

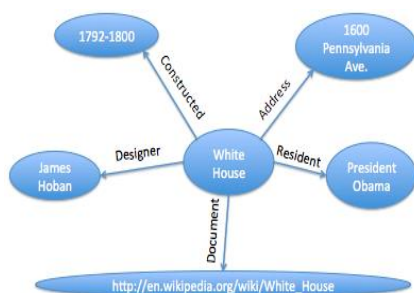


Figure 1: Example of semantic search

The diagram in Figure 1 shows an overview of a semantic search, where every keyword is treated as a thing, instead of a string . Semantic search tries to improve search result by knowing searcher’s idea and the contextual meaning of terms to generate more relevant results. Here, all the nodes are interconnected in a semantic way such that the semantic search is capable of answering the questions by just looking at attributes instead of the documents containing the keywords. The Semantic Web is made of a set of technologies, tools and Standards that form the basis of an infrastructure to support the vision of a Web associated with meaning [2].

B. E-Commerce Ranking

The overall objective of this paper is to improve the E-commerce website ranking process by developing a SNEC page ranking and it’s implemented in the form of an automated tool to assist the customer while carrying out E-commerce transactions [6]. In this research, a semantic and neural based mathematical approach to deal with various ranking problem is discussed and this algorithm will reduce the use of web dictionary, previously spend time statistics.

II. PROPOSED SYSTEM

A. Semantic Search Engine

A semantic search engine should consist of the following components: (1) Ontology development, (2) Ontology Crawler, (3) Ontology Annotator, (4) Web crawler, (5) Performing semantic search, (6) Query builder, and (7) Query pre-processor [4]. The ontology's are developed to make (1) shares understanding of structure between people, (2) Reuse of domain knowledge, (3) Domain assumptions explicit, (4) separate domain knowledge from the operational knowledge, and (5) To analyze domain knowledge.

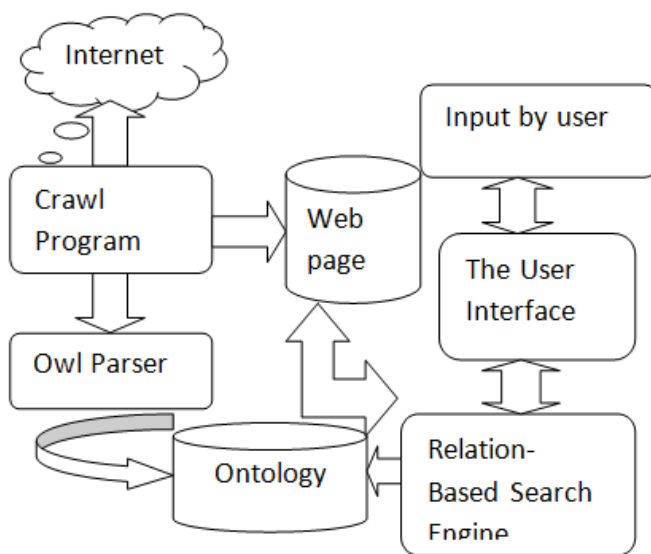


Figure 2 : Semantic Search Engine

An ontology crawler is used to crawl through the Web page to find new ontology's and populate the database with them. The ontology's crawled will not be directly dumped into the database. Ontology translator, will translate the crawled data into the database tables. Ontological annotations identify real-world and relations that characterize the entities' attributes and role in their textual context. Web crawler is an automatic program (also called a "robot") which inquire the World Wide Web, and often used to build automated indexes for the Web, allowing users to do keyword searches for Web documents[4]. (A) Ontology's are usually created in plain text format (.OWL or .DAML). Ontology translator translates them into relational database tables. Ontology Crawler finds new ontology's on the Web and adds them to the ontology library. (B) Users use Ontology Annotator to annotate their Web pages with this ontology's. (C) Web

Crawler crawls the Web to find Web pages explained with this ontology's and builds knowledge base from the instances of these ontology's. (D) Users construct search queries with the help of Query Builder and sent to the Inference Engine after pre-processing by Query Pre-processor [3]. (E) Inference Engine carries out reasoning on these search queries. The results are finally sent to the user to be viewed on the web page as shown in fig[2].

The Use of Semantic Search Engine improves the results of searches in two ways, a) traditional search results uses a list of a document/web pages and b) the search phrase in research searches typically denotes real-world concepts. The advantages over the traditional search are

- A semantic search makes it easy to locate relevant information with respect to the user's interest and the user need not waste time in reading unwanted websites.

B. Ranking Mechanism

In this research work, the retrieved candidate is processed from search engine or manually enters the URL using Profiling and Dictionary Implementation Module to implement web log pre-processing in order to remove incomplete entries, data cleaning, and removal of stem words and finally implementation of a web dictionary. This web dictionary and candidate webpage is further passed to content priority module to apply web content mining to check for relevancy of web page and hence to determine its priority as well as the removal of irrelevant webpage. Then webpage would be passed to time spent priority module which will determine priority of candidate web page using time stamp of the creation of web pages as well as previous user time spent statistic that searched for the same product. This module will assign highest priority to the web pages, now it is passed to Semantics recommendation module which in turn is responsible for identification of user session from various navigation pattern profiles using longest common subsequence algorithm and determination of ontology class to avoid wrong interpretation of user search query[6]. The overall priority of E-commerce website is finally determined by using supervised back-propagation neural network. This network categorise E-commerce website at different layers: Input layer, Hidden layer and output layer as shown in fig[3]. Input layer accepts five inputs: content priority, Time spent

priority, Feedback about the website, recommended semantics and biased input. The actual output of the network is compared to the threshold value set by human volunteer. The simplified neural design of proposed system is shown in fig:3

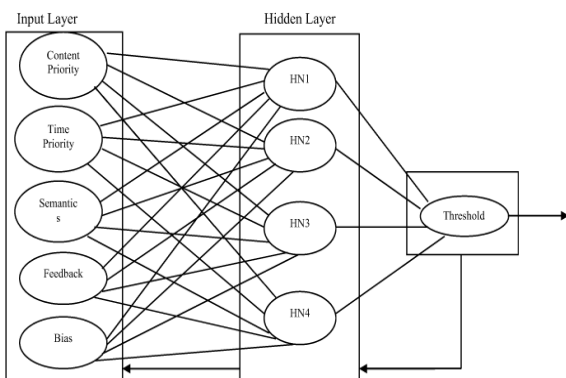


Figure 3. Network categories layers

SNEC page ranking algorithm

SNEC is defined as Semantic and Neural based E-Commerce Page Ranking Algorithm as shown in fig:4

MODULE 1:

Step 1: Accept search string from user.

Step 2: Remove stem words from search string

Step 3: Record navigation sequence pattern in user profile database.

Step 4: search the web documents using search engine.

Step 5: Determine the minimum and maximum length among the various words of search engine.

Step 6: search the time database

MODULE 2:

Step 7: Eliminate all those web pages which does not match the searched criteria.

MODULE 3:

Step 8: Assign priority to the web pages.

Step 9: Update the time database.

MODULE 4:

Step 10: Identify navigation session by computing user search query with each of the search query present in user profile database.

MODULE 5:

Step 11: Normalize all the priority inputs.

Step 12: Determine error rate.

Step 13: Display all the retrieved web pages in decreasing order of their corresponding priority ranking. The overall design of proposed system is shown in fig:4

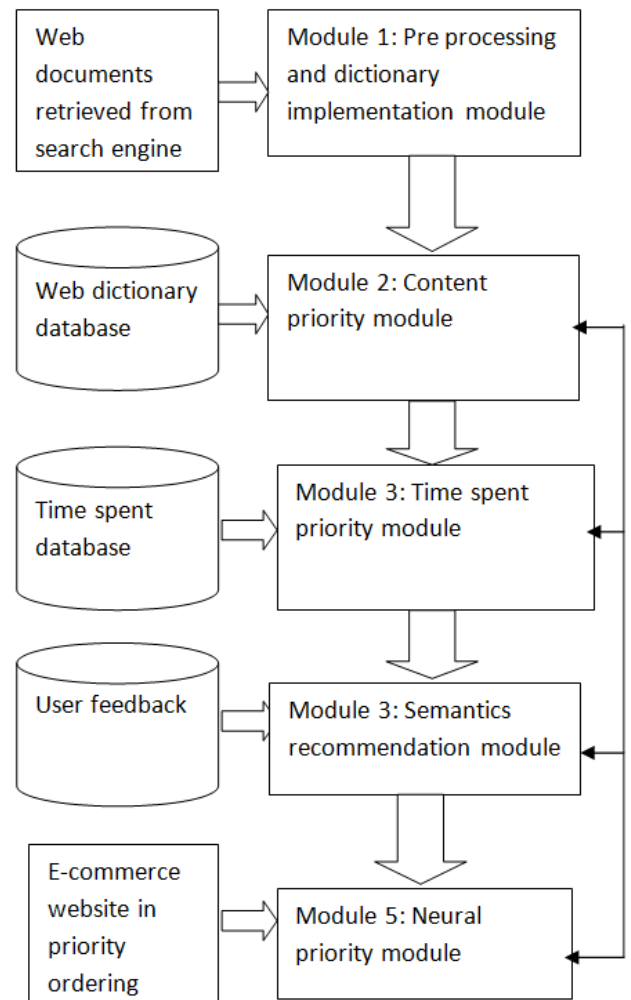


Figure 4: Design of proposed system

III. CONCLUSION

This research work presents a semantic and neutral based novel approach for E-commerce websites priority ranking with respect to specific product search query. The ranking provided will help the website designer for analysing the competitive ranking. By analyzing the research, it can be concluded that the proposed method can ensure the correct rate and the recall rate of product review corpus is high. But compared with other similar methods, also achieve a higher performance because of the introduction of semantic information with deep features.

IV. REFERENCES

- [1]. "Size of World Wide Web," <http://www.worldwidewebsite.com> (accessed on 1/15/16)
- [2]. "semanticsearch", https://en.wikipedia.org/wiki/Semantic_search

- [3]. “Google Semantic-Search (Knowledge Graph),” https://en.wikipedia.org/wiki/Knowledge_Graph (accessed on 1/15/16).
- [4]. Junaidah Mohamed Kassim and Mahathir Rahmany.,”Introduction to Semantic Search Engine”in international conference, august 2009.
- [5]. Venkatesh Mabbu ,Abu Asaduzzaman ,and Muhammad F. Mridha .,“A Novel Semantic Knowledge Engine Using Automated Knowledge Extraction from World Wide Web” in international conference 2016.
- [6]. Neha verma, Dheeraj malhotra, monika malhotra and jatinder singh., ”E-commerce website ranking using semantic web mining and neural computing” in ICACTA-2015.
- [7]. Anusree.ramachandran , R.Sujatha., “Semantic search engine”: A survey, School of Information Technology and Engineering,VIT University.