

Design of Fuzzy Keyword Search Engine

Shilpa Sagar Motghare^{Å*} and G. P. Bhole^Å

^ÅDepartment of Computer Engg. & IT, VJTI, Mumbai, India 400019

Accepted 10 July 2014, Available online 01 Aug 2014, Vol.4, No.4 (Aug 2014)

Abstract

This paper addresses to the design of fuzzy keyword search engine. In a fuzzy keyword search engine, in which if person try to search for the word gravity but by mistake he type graweety then also it show the result of gravity. In a fuzzy search engine it not rarely matter that exact word spelling, because fuzzy logic work on the partial truth, if some words of the search word is also matches it show the result. In fuzzy keyword search engine search for the synonyms word and also handle a singular plural condition. Now a day's fuzzy logic is used in a various application.

Keywords: Fuzzy search engine, Fuzzy logic

1. Introduction

Now a day's Internet is booming around, everyone uses internet, search for what they want and hope for the result they are expecting. Early days when there is exact search engine if there is a person don't know what exact he want to search he won't be able search what he want but in a fuzzy search engine it is not the case. Fuzzy search engine is work on a fuzzy set which has a degree of membership. This framework propose a mainly a general idea about the fuzzy search engine.

In this paper first part contain an introduction and second part contain a history of fuzzy logic is define, in third section different parameter of fuzzy logic is defined like 1) fuzzy set, 2) fuzzy interval, 3) membership function in fuzzy logic, membership of fuzzy logic means that the element of fuzzy set is belong to the fuzzy interval or not. 4) true value in the fuzzy set, section four define a development of fuzzy logic, in section 5 we define a flow chart of fuzzy engine, in sixth section the conclusion and future work the fuzzy search engine is defined.

Fuzzy search engine work on a small amount of data but give the very sufficient result, in fuzzy search the approximate word is also get search, in fuzzy search the if we enter a singular word then in document if its plural word is present then also it show result. In our search engine we can enter a many word to search, store it in an array and search those word line by line in a search document. Now a day's fuzzy logic is used in various applications like washing machine, pressure cooker, iron, refrigerator etc. A fuzzy search is a process that locates web pages that are likely to be relevant to a search argument even when the argument does not exactly correspond to the desired information.

2. History of Fuzzy logic

*Corresponding author: Shilpa Sagar Motghare

Aristotle 'The Father of Logic' has invented a two value logic i.e. true and false but he also accepted that his logic will not apply for the all the future event. In a 20th century Jan Łukasiewicz and Alfred Tarsk together formulated logic on a true value. In 1932 Hans Reichenbach formulated logic of many true values where $n \rightarrow \infty$. Kurt Gödel research that Intuitionistic logic has not has finite true value and defined a system of Gödel logics. Fuzzy set were Introduced by the Lotfi Asker Zadeh and Dieter Klaua in 1965. In 1970 research group were forms in Japan, in the 1974s Ebrahim Mamdani design a fuzzy logic for steam engine i.e. first fuzzy logic controller. In 1977 Dubois applied fuzzy set in comprehensive study of traffic condition. In NASA they have used Fuzzy Logic. First code of Fuzzy search engine was written in 1998 in TREX search engine which is combination of many logics including fuzzy logic.

3. Different Parameters in Fuzzy Logic

3.1 Fuzzy Set

Fuzzy set is set of that element which has a degree of membership. In a classical set theory, the set theory have a binary value [1 0]. Fuzzy set is used in condition in which information is not precise, incomplete or huge. A fuzzy set is a pair of (U, m) where U is a set and m is an $m=U \rightarrow [0, 1]$.

3.2 Fuzzy Interval

Fuzzy interval is not having a specific value it varies upon different situation. In Fuzzy interval fuzzy number/element is present, which have some specific degree of membership.

3.3 Membership Function

Membership Function is used to defined a membership of an element which is available in fuzzy set. Suppose F is

taken as a fuzzy set and then its membership function is denoted by μ_F and if we have to find out the membership degree of an element f in set F which is denoted by $\mu_F(f)$ if it 1 means it is fully member of a the fuzzy set, 0 means it is not at all member of the fuzzy set and value in between 1 and 0 means it is partially a member of a fuzzy set. Usually the fuzzy membership value lie in between [0 1] are then called [0 1]-value membership functions.

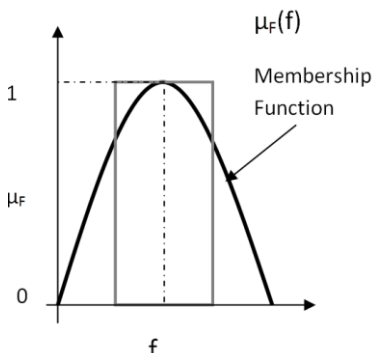


Fig. 1 Membership Function of Fuzzy Set

3.4 True Value

The True Value in a fuzzy keyword search engine has no fix range it depend on different scenario, different condition. For example there is 100ml of glass is there out of that 40ml is full of water, some might say it is half full, some might apply a concept if the glass's water is below 50ml then it is empty it depend on the observer perspective, in our fuzzy search engine we provided a true value i.e. tolerance level below the tolerance level the result fail and above the true range the success is given.

4. Development of Fuzzy Search Engine

We have implemented fuzzy search engine in .net framework and use a fuzzy logic for the searching a document from the browse the folder and select a file from which we want to search a word.

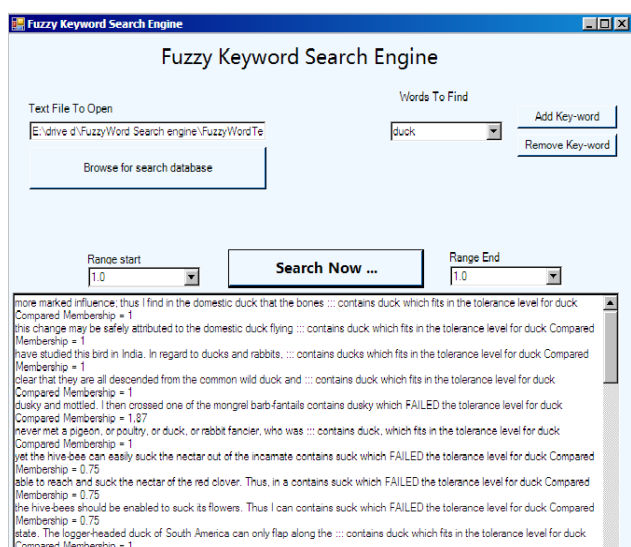


Fig. 2 Fuzzy Search Engine

And add a word which we want search in a select document and get the search result.

4.1 Browse for search database

In that if user click on that button OpenFileDialog will open a window of file selection dlg.Filter will filer a file and at same time add a index at the indexer to search a file.

4.2 Add Word to Find

In this section we have add word which we have to search on clicking a add button the word get added which we want search in a given document and if we want to remove that word we have remove keyword button.

4.3 Range Value

This value are used to fixed a lower value and upper vale of a fuzzy interval and the fuzzy interval varies depend on a situation and a condition. In this search engine we have taken a difference of 12 in start range and stop range. If range is from 1 to 1 then it searches only for the exact word. If we increase the range of the probability of finding an approximate word goes increase but the accuracy of word decrease. E.g. suppose the start range is 1.87 and the lowest range is 0.12 then obviously [1.87 - 0.12] ranges is larger than the [1 - 1] range. In a 1.87 to 0.12 they show the result even if 1or 2 alphabet is also matches with fail tolerance and low membership value.

4.4 Search Function

In a search function the added word get compare with the document which we have selected line by line. First calculated a fuzzy number of enter value and then compare it with calculated membership see if it is in the range witch we have provided if it matches then show the result with its membership value.

4.5 Fuzzy logic function

In a fuzzy logic we calculate a membership value for different aspect depending upon provided maximum and minimum value and number witch is given. In a printout function we provided an output on a screen with its tolerance level and a its membership value. In a fuzzy logic function checks as if we enter a word or a number.

4.5.1 We have use a different class to handle if we enter a word then we have call a function overloading to handle a different scenario. Set the maximum and minimum value of range, for membership comparison set some value for the minimum letter which should have to present for comparison, set value for maximum allowable incorrect word and allow for the plurals word also.

If a search word exactly matches with the document word then membership set to 1 or set the membership by a comparison function.

4.5.2 If we enter a number then we have develop a different class to handle it. We have call a function

overloading to handle a different scenario and in that there are different variable we have set like minimum, maximum, membership variable to store a number etc. calculate a membership for different scenario.

5. Flow diagram of fuzzy search engine

In Fig 3 first we select a file from which we search a word and then add a word in indexer and start searching that word in a selected file if the exact word is found then then return with a membership value 1 or if the exact word is not found then set a fuzzy logic calculate a membership value and display a result on screen.

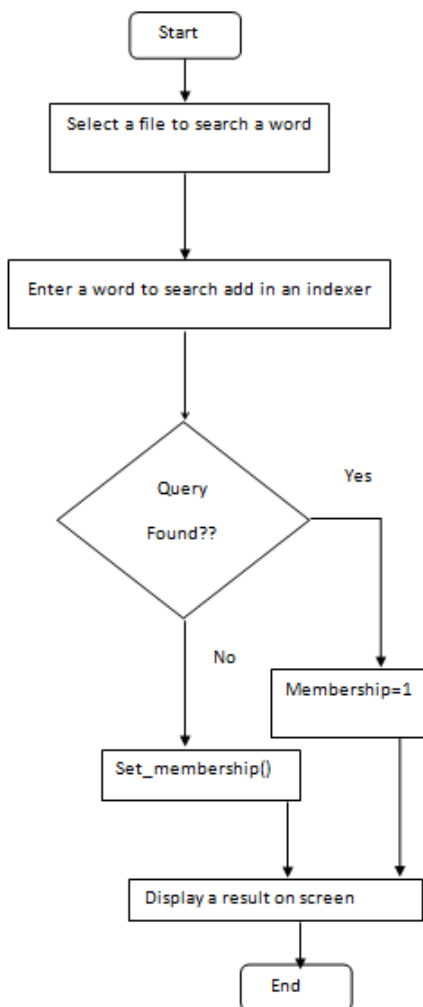


Fig. 3 Flow chart of fuzzy search engine

Conclusion and Future Work

In this paper we develop a fuzzy search engine and Fuzzy search engine uses an approximate word finding method in which if word contain some incorrect alphabet then also it search its relative word. Fuzzy search engine handle a singular and plural word concept and search more than one word in a document we have implemented a fuzzy search engine on a local pc because on internet there is more data and if we implemented a fuzzy on such huge data then search will become very slow.

In Fuzzy search engine it not handles a question based answer search mean if we enter when we enter where is India? Search engine will not give me a direct answer it just provide me a document related to it.

References

Lien-Fu, Lai Chao-Chin, Wu Pei-Ying Lin, Liang-Tsung Huang (June 27-30, 2011), Developing a Fuzzy Search Engine Based on Fuzzy Ontology and Semantic Search, *2011 IEEE International Conference on Fuzzy Systems Taipei, Taiwan*, ISSN:1098-7584. DOI:10.1109/FUZZY.2011.6007378.

Madhu Kumari, Kamal K. Bharadwaj (2010), Fuzzy Logic Based Effective Bidding Range Computation and Bidder's Behavior Estimation in Keyword Auctions, *2010 IEEE 2nd International Advance Computing Conference-IACC*, DOI: 10.1007/978-3-642-22555-0_7

Elmira Hajimani, Eslam Nazemi (2010), Improving the Performance of Proposed Multi-Agent Domain Specific Search Engine Using Query Refinement Component, *2010 5th International Symposium on Telecommunications (IST'2010)*, DOI: 10.1109/ISTEL.2010.5734092

Tomohiro TAKAGI, Masanori TAJIMA (2001), Query Expansion Using Conceptual Fuzzy Sets For Search Engine, *Fuzzy Systems, 2001. Fuzzy Systems, The 10th IEEE International Conference, Vol 3*, ISSN:1303-1308, DOI:10.1109/FUZZ.2001.1008898

Fang Zheng Li, DaYong Luo, Dong Xie (2009), Fuzzy Search on Non-numeric Attributes of Keyword Query over Relational Databases, *Computer Science & Education, 2009. ICCSE'09. 4th International Conference on*, ISBN- 978-1- 4244-3520-3. DOI:10.1109/I CCSE.2009. 5228162.

Marlene Goncalves, Graciela Perera and Stephen Rodabaugh (2010), Detecting Illegal File Sharing in Peer-to-Peer Networks using Fuzzy Queries, *2010 IEEE International Conference*, 978-1-4244-8126-2/10/2010 IEEE, ISSN :1098-7584. DOI: 10.1109/FUZZY. 2010.5583952.

Gloria Bordogna, Gabriella Pasi (2013), A Fuzzy Approach to the Conceptual Identification of Ememes on the Blogosphere, *2013 IEEE International Conference*, ISSN: 1098-7584, DOI:10.1109/FUZZ-IEEE.2013.6622390.