International Research Journal of Engineering Science, Technology and Innovation Vol. 11(1) pp. 1-5, Jan, 2025
Available online http://www.interesjournals.org/IRJESTI

Mini Review

Improving Query Processing Using Fuzzy Logic in Distributed Databases

Waqar Ul Hassan*, Syed Touqir Hussain, Zaheer Abbas, Syed Saqib Hussain, Mohammad Ismail

Department of Computer Science, Karakorum International University, Gilgit, Baltistan

*Corresponding Author's E-mail: gltstar@hotmail.com

Received: 09-May-2024; Manuscript No: irjesti-24-134420; **Editor assigned:** 13-May-2024; Pre-QC No: irjesti-24-134420 (PQ); **Reviewed:** 27-May-2024; QC No: irjesti-24-134420; **Revised:** 08-Jan-2024; Manuscript No: irjesti-24-134420 (R); **Published:** 15-Jan-2024, DOI: 10.14303/2315-5663.2025.99

Abstract

Popular research topics to solve the problem of increasing databases to make them more, user-friendly and to be able to answer vague human queries. To solve this problem many research papers suggested applying fuzziness in data bases. However, this approach is infeasible for real time environment. Still fuzzy databases are not popular due to abundant of pre computation. Having fuzzy databases, it's difficult to run crisp queries on fuzzy databases. This paper proposes as a layer of computation based on fuzzy logic to solve in a real time above traditional query. This fuzzy logic based approach can be used in distributed databases to solve ambiguous queries. The obtained results using the fuzzy logic approach are compared with the traditional approach in terms of accuracy, time taken for each approach and closeness of the results to user's requirements.

Keywords: Terms-Distributed databases, Fuzzy logic, Mumdani's implication, Human queries, Fuzzy databases

INTRODUCTION

Applications of database management systems in many domains like banking, e commerce, web applications and Machine learning (Singhal A, 2001). The Important task of a computer science is to retrieve information from databases, for this purposes many research has been going on since 1950 (Garcia-Molina et al., 2009). Today's computers store huge amount of data and need effectively query processing to obtain desire result in minimum time (Garcia-Molina, 2009).

There are 2 types of query's one is crisp which is known as traditional approach and other is fuzzy, vague or uncertain, based on human language (Singh SS et al., 2011). Crisp approach based on value either 0 or 1 but this approach is limited in this dynamic world applications, now a day we need to run a dynamic query approach to overcome the uncertain query's like "Who is the best student of the college?" This type of vague queries can be solved by using fuzzy logic approach (Bosc P et al., 1994).

This paper presents a fuzzy logic based approach to answer the vague queries from databases and compared the results between traditional and fuzzy approach on user friendly, time taken and answering the vague query's (Ramez Elmasri et al., 2009).

Traditional approach

Steps to approach:

Following steps would involve in traditional approach to solve the above mentioned user specific uncertain query's (Lacroix M et al., 1987).

- User interface and database setup
- User requirement
- Query to search data
- Compile the result

The selection is based over [0,1] as membership grade function. Figure 1 shows the different type of query

processing **(Figure 1)**.

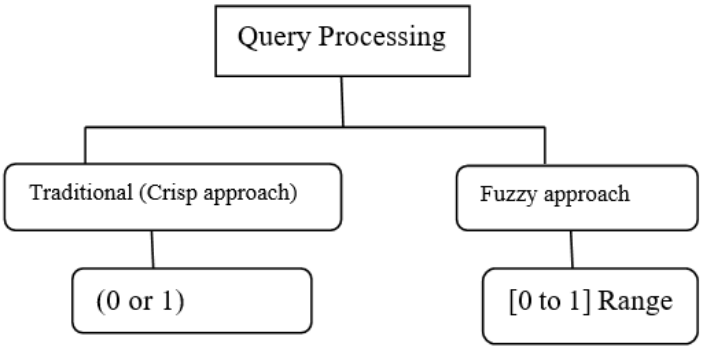

Figure 1. Approaches of query processing.

Example

We present an example to elaborate the issues in traditional ways (Motro A, 1988). A database of mobiles with following 4 attributes–brand, year, model and price (Ichikawa T et al., 1986). Suppose a customer wants to find the best mobile according to his/her priorities (Chang CL, 1982). Budget and needs. But the traditional way can search the database using a single attributed at a time (Ganesh M, 2006). So it is time costly for multiple attributes. To overcome these issues a fuzzy approach frame work is needed to provide more satisfactory results (Baghel A Set al., 2013).

Fuzzy logic

The fuzzy logic theory was introduced in 1965 by Lofi Zadeh (Bezdek V, 2011). Fuzzy logic is the form of multi valued logic in which value of variables may be real number between [0 to 1]. Fuzzy logic contains the fuzzy rules in the form of IF (antecedent) THEN (consequent) format. The fuzzy expert system consists of fuzzy based rules and membership functions to derive conclusion on user input. The process involves the following 4 major steps. Figure 2 defines the process involves in an expert system (Raipurkar A et al., 2013).

Fuzzification: Fuzzification is the process of converting crisp input **(Figure 2)**.

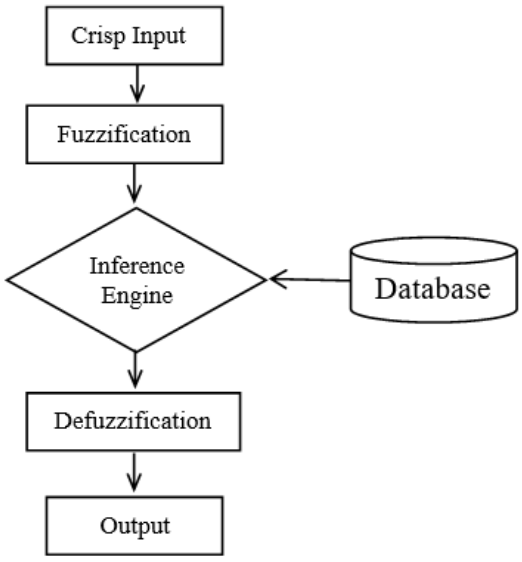

Figure 2. Steps perform in fuzzy expert system.

Value to a fuzzy value and determine the degree which that belong to a fuzzy set.

Inference engine: Provided input value is applied on antecedent part of rule and matched consequent part of rule is being fired.

Aggregation: Aggregation is the process of combination of all outputs in a single fuzzy set.

Defuzzification: Transferring fuzzy inference results into a crisp output.

Proposed method

It is difficult for the user to find specific output because some queries of users contain some amount of fuzziness

and vagueness. Below are the steps involved in proposed method.

Dataset filtering: This step involves to set the requirements on range based or equality based query. The result is returned on the bases of relational algebra query.

Fuzzification: The membership attribute is calculated for each attribute. Figure 3 shows the membership grade for each attribute. Now a rule is fired from database using fuzzified value.

Inference: The fired rule is calculated through Yagars intersection formula

Yager's intersection formula

Truth value = $1 - \min (1, (p1*(1-a1)^w + p2*(1-a2)^w + \ldots + pn*(1-an)^w)^{1/w})$

Where, p1, p2, ..pn are the priorities of attributes a1, a2, .. an respectively and w is the Norm chosen.

Our contribution in this paper depends on the modification of the general Yager's intersection formula to the one mentioned above which includes the priorities of each of the attributes.

Using Mamdani's implication we find the suitable region for each rule.

Aggregation: Using the weighted centre of gravity method and the area of the region obtained from each rule we obtain the final score value of each record.

Aggregated value = $\sum_{i=1}^{n} ci * Ai*Wi / \sum_{i=1}^{n} Ai*Wi$

Deffuzification: The final result obtained by using center of gravity gives us the final score, hence we have been able to handle the vague queries fired on the database. Figure 3 shows the output.

2 types of query processing- central and distributed querying. In the centralized system all the data is stored on a single database which increases the load, high communication costs and system failure risks. In the big data, this approach is very limited, so our paper proposes the use of distributed query processing system that aims to regulate operating costs and response time. Local processing phase, reduction phase and final processing phase over each node above which exists the fuzzy based computation layer processing the input data and obtaining precise results for the ambiguous query **(Figure 3)**.

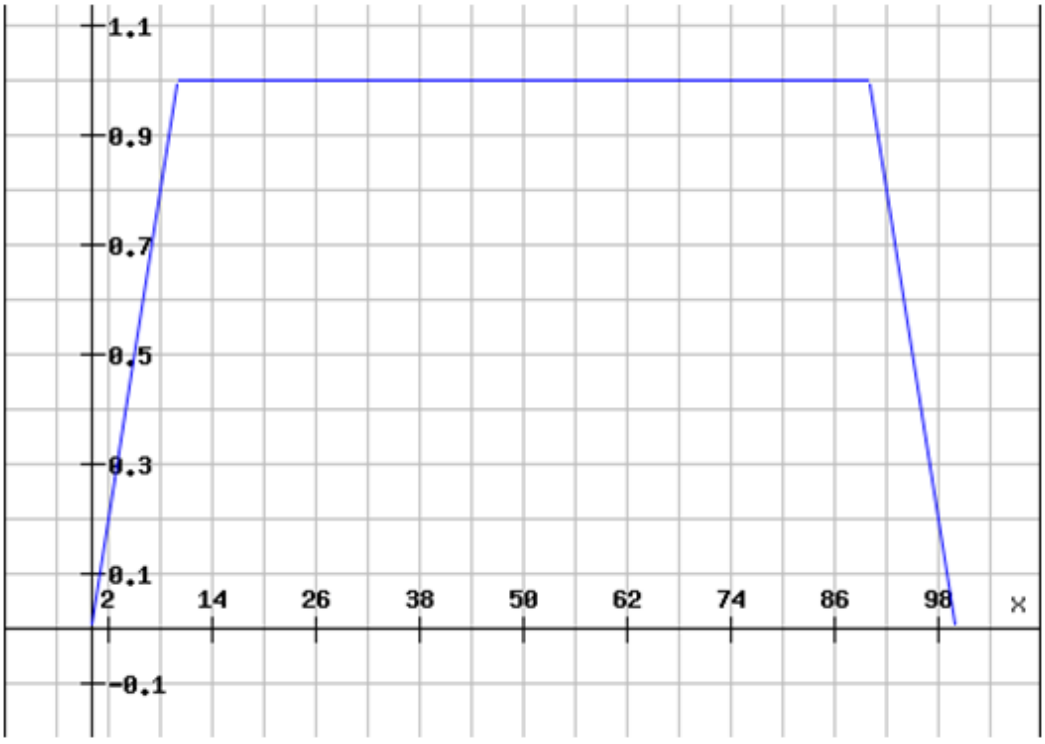

Figure 3. Graph representation of deffuzification

Distributed environment

Figure 4 shows the proposed model for query processing in distributed environment **(Figure 4)**.

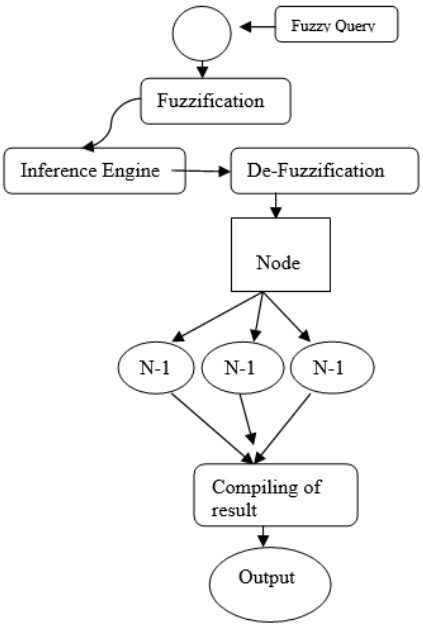

Figure 4. The proposed model for query processing in distributed environment.

RESULTS

Today's users use the Internet to find the information related to their requirements. So the above fuzzy based approach is applied on traditional web based query searching. Consider a dataset of second hand mobiles. The objective is to find the best mobile according to the user's preferences. Table 1 shows a subset of original dataset used for analysis. Consider that the crisp filtering conditions on the dataset and the preferences for each attribute are provided by the user. This section shows the traditional approach of solving this problem and then compares the results with the proposed approach.

In traditional approach one attribute is processed at a onetime and ranking the results. Table 2, 3 shows the results of sorting the mobiles according to least price and manufacture year. Table 4 shows the aggregate results **(Tables 1,2)**.

Table 1. Shows data of used mobiles.

ID	Make	Year	Price	Used since
1	Samsung	2009	10000	2009
2	Black berry	2008	12000	2008
3	Nokia	2005	10000	2005
4	Samsung	2001	8000	2002

Table 2. Shows sorted dataset on Price.

ID	Make	Year	Price	used since	Price rank
2	Black berry	2008	12000	2008	1
1	Samsung	2009	10000	2009	2
3	Nokia	2005	10000	2005	3
4	Samsung	2001	8000	2001	4

Fuzzy logic approach

The fuzzy logic approach takes into account all the attribute preferences of the user at the same time in order to generate the score.

In order to rank all the mobiles according to the preferences of user. The final score is obtained by using the fuzzy rule base and giving different priorities to each of the attributes and aggregating the centre of gravities of each of the rules fired using the Yager's intersection and the Mamdani's implication formulas described above. Using the priorities for User Id 1(equal for all), the following results were obtained using the fuzzy-based approach **(Table 3)**.

Table 3. Shows the aggregate ranking on fuzzy approach.

ID	Make	Year	Price	used since	Price rank	Year rank	Used rank
2	Black berry	2008	12000	2008	1	2	2
1	Samsung	2009	10000	2009	2	1	1
3	Nokia	2005	10000	2005	3	3	3
4	Samsung	2001	8000	2001	4	4	4

The proposed fuzzy based approach not only runs all the individual attributes but also the aggregate rank. Hence, fuzzy logic approach better accommodates the user specification and provide better results for vague queries.

CONCULUSIONS

In the traditional approach, data need to be ordered based on all attributes one by one and requires multiple aggregations and hence a lot of processing is required as compared to the fuzzy based approach.

REFERENCES

1. Singhal A (2001). Modern information retrieval: A brief overview. Data Eng Bull. 24(4):35-43.
2. Garcia-Molina, Ullman, Widom (2009). The Worlds of Database Systems. 2nd Edition. Pearson Education Limited. Harlow, England.
3. Garcia-Molina (2009). Database systems: The complete book. 2nd Edition. Pearson Prentice Hall. Upper Saddle River, New Jersey, United States of America.
4. Singh SS, Sayal R, Rao V (2011). Analysis and usage of fuzzy logic for optimized evaluation of database queries. Int J Comput Appl. 16(3).
5. Bosc P, Pivert O (1994). Fuzzy queries and relational databases. Proc 1994 ACM Symp Appl Comput. 170-174.
6. Ramez Elmasri, Shamkant B (2009). Fundamentals of Database System. 5th Edition. Pearson Education. Second Impression. Massachusetts, USA.
7. Lacroix M, Lavency P (1987). Preferences: Putting more knowledge into queries. Proc 13th Conf VLDB. Brighton, United Kingdom. 217-225.
8. Motro A (1988). Vague: A user interface to relational databases that permits vague queries. ACM Trans Inf Syst. 6(3):187-214.
9. Ichikawa T, Hirakawa M (1986). ARES: A relational database with the capability of performing flexible interpretation of queries. IEEE Trans Softw Eng. 12(5):624-634.
10. Chang CL (1982). Decision support in an imperfect world. IBM San Jose. New York, USA.
11. Ganesh M (2006). Introduction to Fuzzy sets and Fuzzy Logic. PHI Learning Private Limited. Rajasthan, India.
12. Baghel A, Sharma T (2013). Survey on fuzzy expert system. Int J Emerg Technol Adv Eng. 3(12):230-233.
13. Bezdek V (2011). Possible use of fuzzy logic in database. J Syst Integr. 2(2):31.
14. Raipurkar A, Bamnote GR (2013). Query processing in distributed database through data distribution. Int J Adv Res Comput Commun Eng. 2(2):1134-1139.