International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 13, Issue 2, February 2024
DOI: 10.17148/IJARCCE.2024.13201

A SYSTEMATIC REVIEW OF FUZZY DATABASE APPROACHES AND ITS STATE OF THE ART

Chukwu, E. G. [1], Okoronkwo, M. C. [2], Ezema, M. E. [3], Achi, I. K. [4], Nwosu, P. C. [5]

Lecturer II, Federal University of Technology Ikot-Abasi, Akwa-Ibom State, Nigeria[1]

Senior Lecturer, Department of Computer Science, University of Nigeria, Nsukka, Enugu State[2]

Senior Lecturer, Department of Computer Science, University of Nigeria Nsukka, Enugu State[3]

Lecturer II, Department of Computer Science, Rhema University Aba, Abia State[4]

Lecturer II, Department of Computer Science, Rhema University Aba, Abia State[5]

Abstract: This paper addresses two important areas of databases. First, the fuzzy database approach was examined and the current state of fuzzy databases was also discussed in depth. The fuzzy database approach is examined in terms of: the need for fuzzy databases, the techniques used to store and retrieve fuzzy data in fuzzy databases and information retrieval systems, database frameworks for fuzzy databases and the fuzzy database approach for information retrieval systems. The advantages of using object-oriented database frameworks are described. A prototype fuzzy object-oriented database system (FOODS) is developed to demonstrate the feasibility of a fuzzy object-oriented database system. We know that most real-world data is fuzzy, imprecise and incomplete, and that conventional relational databases therefore lack the capacity to integrate and manage it, we then need a thorough knowledge of the state of the art of fuzzy logic to understand it. To managing these imprecise, ambiguous and incomplete data, particularly in multi-criteria decision making, we need fuzzy sets and fuzzy logic to extend the classical relational database model, and they serve as a functional means of support to handling these anomalies and that is achieved by fuzzy database model. This discussion of the state of the art of fuzzy database models is therefore a review and condensation of the various approaches adopted by different authors to integrate precise and incomplete data. The limitations and possible improvements of the models have been taken into account.

Keywords: Object-oriented Database management systems, Fuzzy set theory, Fuzzy Object-Oriented Database Management System, fuzzy Object-Oriented Database System, Fuzzy Logic, Machine learning, Artificial Intelligence.

I. INTRODUCTION

Business systems are designed to be adaptable, intelligent, flexible and efficient in order to respond to the ever-changing business environment. According to (Bellman, R. E., & Zadeh, L. A), decision making in the real world takes place in an environment where the goals, constraints and outcomes are not precisely known. Decision making at the managerial level is usually based on vague or incomplete information. We will consider the various contributions of researchers by reviewing their work and identifying what approaches and recommendations they have formulated for improvement.

According to Phang et al., (1997), advances in computer technology and the realisation that data is an important organisational resource have led to the rapid advancement and development of information systems. Howard L. Resnikoff et al., (1984), then taught us that information systems, such as database management systems (DBMS) and information storage and retrieval systems (ISRS), involve the collection, management, utilisation and dissemination of information. In other words, they are the means by which organisations provide the information they need, Peter et al., (1992). Information systems are superior to traditional filing systems in terms of temporal and spatial efficiency, which significantly improves people's ability to process data.

However, the existing DBMS can only process clear, precise and non-ambiguous data. In other words, these systems are not suitable for fuzzy and ambiguous data. In fact, fuzzy data is constantly present in real life through human thinking and cognitive processes, and we often make decisions based on this data.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13201

It should be noted that information is only useful if it is obtained in a simple and natural way (Maria et al., 1984), and it is clear that with the increasing importance of DBMSs for decision making, the problem of processing fuzzy data that is more compatible with human reasoning is also gaining importance.

The implementation of such a DBMS would improve the human-machine interface and broaden the scope of DBMSs. Attempts to develop DBMSs that can represent and process fuzzy data have recently attracted the attention of researchers. As a result, various models and prototypes have been proposed or implemented.

On the other hand, Otokwala, et al., (2019) stated that a database is an ordered collection of related data elements designed to meet the information needs of an organisation and designed to be shared by multiple users, Amit, et al., (2012). The data type and values of attributes are not always known with sufficient precision, Hudec, et al., (2014). In fact, most data are fuzzy, vague and complex, either due to their nature or due to less than ideal measurement, and the uncertainty resulting from fuzziness is always ambiguous, Ankita, et al., (2012) and Pavese F., (2013). Therefore, the main motivation for using the fuzzy method is the need to solve the fundamental problem of integrating precise and imperfect data into the database, Tuqyah, et al., (2015). By design, relational databases are based on Boolean logic with a bi-stable output of {1 or 0; true or false}. The fuzzy database approach measures information by degree of truth and has become the most practical way to store and manage imprecise data.

The basic components of the fuzzy database model are fuzzy logic and the fuzzy set. While fuzzy logic uses a combination of different mathematical principles to represent knowledge according to a graded degree of membership, the theory of fuzzy sets provides a robust framework for the systematic treatment of fuzzy-based uncertainty, Tuqyah, et al., (2015).

II. OBJECTIVES

- We shall review other works done by different individuals, related to the aim of this study
- We have an In-depth knowledge in fuzzy database approaches through studying the essentials.
- We have an In-depth knowledge of the state of the art in fuzzy database through studying the essentials.
- We then state our observations and spot out the future research areas as captured at the related literatures

III. RELATED WORKS

An overview was given of the approaches to fuzzy databases and the state of the art in models for fuzzy databases, with references to researchers whose work was cited. The problems solved by fuzzy databases and their applications were also discussed. In addition, an overview of Object-Oriented Fuzzy Database Systems (FOODs) was given with details of the various modelling and algebraic expressions that can be performed on these systems. A brief overview of related work by different authors in the areas related to our research was given. The survey focused on fuzzy databases and their advanced models. A presentation was given on the approaches for fuzzy databases such as Information Storage and Retrieval Systems (ISRS), with uncertainty processing, DBMS with uncertainty processing, conventional DBMS with fuzzy queries and fuzzy databases with fuzzy queries.

By Janusz, et al., (1989), Human- Consistent' DataBase Querying System Based on Fuzzy Logic with Linguistic Quantifiers was discussed, Fuzzy Querying with SQL: Extensions and Implementation Aspect was mentioned and captured by P. Bosc, et al., (1988), A Document Retrieval System Based on Citations Using Fuzzy Graphs was talked about by Nomoto et al (1990), we talked about Prototype Fuzzy Object- Oriented DataBase System by Phang, (1994), and Proximity Relations in the Fuzzy Relational DataBase Model by Sujeet, et al., (1989). We discussed Fuzzy Data in Traditional Relational Databases by Hudec, (2014), A Study and Comparison of Methods for Fuzzy Data Equivalence was made by Ankita, et al., (2014). State of the Art of Fuzzy Methods for Gene Regulatory Networks Inference was discussed by Tuqyah, et al., (2015), and Kerre, et al., (2001). We also discussed Fuzzy data modeling at a conceptual level: Extending ER/EER concepts. Prof. Rahul, et al, (2014) made a review of Fuzzy Logical Database Models by Anupriya, as well as assessment of benefits and drawbacks of' using fuzzy logic, especially in fire control systems, (N.M. de Reus).

Bookstein also proposed a weighted Boolean retrieval system where the relative importance of each term in a query can be specified in the interval [0, 1]. Note that this model becomes classical when the weights are restricted to {0, 1}.

Nomoto, et al (1990) developed a fuzzy document retrieval system using fuzzy graph theory. Document citations are chosen as the fuzzification criterion. A citation network, i.e. a graph consisting of citations and relations between citations, is created for fuzzy retrieval.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⩰ Peer-reviewed & Refereed journal ⩰ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13201

Miyamoto (1990) developed a fuzzy document retrieval system based on keyword associations. That is, these keywords are related to all the keywords found in the ISRS by fuzzy binary relations. The usefulness of this model has been demonstrated by the implementation of an efficient fuzzy retrieval algorithm in a large bibliographic database.
Ogawa et al (1991) developed a fuzzy document retrieval system based on a keyword connection matrix. This matrix records the similarity between keywords and is used during query processing. The relevance of documents to the query is calculated using fuzzy set theory. The fuzzy approach has shown significant improvements in recall measurement compared to the explicit approach.

Lucarella et al (1991) study a knowledge-based fuzzy information retrieval system. The domain knowledge and inference scheme is based on a fuzzy set framework.

There are still many other authors whose contributions were captured in the work which went a long way, contributing to the development of this work.

IV. IN-DEPTH KNOWLEDGE IN FUZZY DATABASE APPROACHES

4.0 Dealing with Uncertainty in Information Systems.

In order to identify effective practical concepts and evaluate the current state of database management systems (DBMS) in dealing with uncertainty, it is necessary to understand the existence of Information Storage and Retrieval Systems (ISRS) and their ability to deal with uncertainty. Due to the many similarities between ISRS and DBMS, we will focus on ISRS.

4.1 ISRS with Uncertainty Handling

ISRS is a computer system that allows users to obtain information from a collection of documents stored in a database, Abraham, B., (1980). Classical ISRS is built within the framework of Boolean set theory. More recently, approaches based on fuzzy set theory [5] have also been reported.

In general, the use of fuzzy set theory in ISRS provides a more natural query system with efficient algorithms compared to traditional ISRS approaches.

4.2 DBMS with Uncertainty Handling

Applications of fuzzy set theory in DBMSs can be classified into two main classes. Class 1 concerns the study of fuzzy query processing in conventional (non-fuzzy) DBMSs, while class 2 concerns DBMSs that not only have the ability to store and manipulate fuzzy data directly, but also support fuzzy queries.

4.2.1 Conventional DBMS with Fuzzy Queries

The first DBMSs with uncertainty processing were developed in the context of non-fuzzy DBMSs. In general, these systems deal with the construction and evaluation of fuzzy queries with respect to a crisp database and ignore the problem of directly representing fuzzy data in DBMSs. This section gives an overview of the direction and problems that researchers are trying to solve in this area. Chang et al (1978) explore the use of fuzzy queries and propose the concept of the database skeleton, which allows the user to specify the content and meaning of a collection of data. The database skeleton is then used as a semantic base that supports fuzzy queries. The proposed methodology is capable of handling fuzzy queries such as the following:

Query: get the name of the supplier; products are equal to "tv".
This query is fuzzy because it does not provide enough information to the DBMS - no path is given. The translation of the fuzzy query into a full query is formulated as the problem of converting a partially specified query graph into a fully specified graph. A full query for the above query has the following form.

Query: get supply.supplierName; stock.goods equal 'tv'.

In Chang's article, the term "fuzzy" refers to incompletely specified information in the query, such as the absence of a path. It should be noted that the theory of fuzzy sets is not explicitly used in the formulation of the methodology. Therefore, this model is less powerful as it is less suitable for processing fuzzy queries that contain fuzzy data that is vague and ambiguous. Wong (1982) proposes a framework for handling incomplete information in non-fuzzy DBMSs. Sometimes a database cannot provide answers to certain queries due to incomplete or inaccurate information in the DBMS. If the inaccuracy is due to a record error, outdated data, incompatible scaling or measurement error, Wong suggests using a statistical approach to generate approximate but meaningful answers. It should be noted that Wong's

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13201

model is based on probability and statistical theory, not fuzzy set theory. Since fuzzy data is overall possibilistic rather than probabilistic in nature, this model, like Chang's model, fails to correctly process fuzzy queries involving fuzzy data.

Tahani (1977) develops a high-level conceptual framework for processing fuzzy queries in a conventional non-fuzzy database environment. The proposed framework, Fuzzy Retrieval System, uses the associative search scheme approach. In this framework, a fuzzy query is replaced with its associated meaning. Then, the matching operation is applied to compare the fuzzy sentences with the precise data to obtain an answer. Kacprzyk et al (1986 and 1989) present a fuzzy query system called Fquery III. Using Fquery III, data from Dbase III plus (a commercial RDBS based on a non-fuzzy microcomputer) can be analysed using fuzzy queries. Fquery III is based on the framework of fuzzy set theory. Wong et al (1990) develop a fuzzy query language for VAX Rdb/VMS (a conventional non-fuzzy DBMS based on a minicomputer). This fuzzy query system, formulated in the framework of fuzzy set theory, also supports multi-criteria decision making. Bosc et al (1988) discuss the extension of SQL language to process fuzzy queries based on the framework of fuzzy set theory. The extended SQL language has the following format: Choose n/t <attribute> from <relation> where <fuzzy condition>.

Where n and t are the parameters that control the output. Here, the threshold t allows the user to specify the desired threshold for membership. The presence of n can prevent the creation of a solution set by returning only the n-tuples that are relevant to the user. Consequently, processing fuzzy queries tends to become more efficient. In general, in the conventional environment of non-fuzzy DBMSs, fuzzy query processing based on fuzzy set theory is more efficient than fuzzy query processing using an ad hoc approach or probability theory.

4.2.2 Fuzzy DataBases

The newer DBMSs that deal with fuzziness which are more advanced than the earlier ones. They address the problem of direct representation of fuzzy data in the DBMS and the construction and evaluation of fuzzy queries.

Uncertain information shows up in various ways, typically not in business and administration domains but rather in more empirical domains such as in medicine, criminology, and in all kinds of empirical research. We shall be discussing different types of uncertainty in this study, then present a fuzzy database model for storing and retrieving gradually uncertain information. In various application domains, one encounters uncertain and imprecise information. While the fuzzy set theory of Zadeh, 1965, allows to handle imprecise or vague information, fuzzy logic and possibilistic logic, which are based on the possibility theory of Zadeh, 1979; Dubois and Prade, 1980; Dubois et al., 1994, form the theoretical basis for processing gradually uncertain information.

Classical probability theory, based on the axioms of Kolmogorov, on the other hand, makes very strong assumptions. In many practical cases, it is therefore not applicable to the problem of handling uncertain information. Also, probabilistic methods are computationally much more costly than fuzzy methods because they are non-compositional.

Imprecise information often comes in the form of disjunctive or vague sentences, while uncertain information is often expressed in the form of sentences qualified by a degree of uncertainty like, for instance, 'very likely' or 'with 60% certainty'. Sometimes, such a qualification is called a subjective probability or a degree of belief. A particular form of uncertainty is given by statistical information. Sentences qualified by statistical uncertainty are based on statistical samples measuring relative frequencies. An example of such a sentence is:

80% of all jaundice patients have hepatitis.

Through their empirical grounding, these sentences are more reliable than sentences qualified by a mere subjective degree of belief.

In the literature, there is no clear taxonomy of uncertainty in databases and knowledge bases. An obvious distinction, however, concerns the uncertainty expressed at the level of attribute values, or at the level of the applicability of a predicate to a tuple. While the former is related to the issue of handling vagueness, the latter is related to an account of certainty-qualified sentences. In our formal exposition, we do not treat vagueness expressed by fuzzy-set-valued attributes, but only the more fundamental issue of gradual uncertainty based on fuzzy relations over ordinary ('crisp') attributes.

4.2.2.1 Disjunctive Imprecision

A disjunctive imprecise attribute value specifies a set of possible values without any commitment to (or preference for) a particular one of them. Examples are:

 ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13201

1. 'The culprit escaped with a black BMW or Mercedes-Benz.' This sentence could be formalized as follows:
escape(BMW) V escape(Mercedes-Benz)

2. 'Teleman's concert No. 9 for flute has been played for the first time between 1743 and 1745.' This sentence could be formally expressed by using a set valued attribute like:

premiere(9, {1743, 1744,1745})

or in the form of a disjunction

premiere(9, 1743) V premiere(9, 1744) V premiere(9,1745)

Disjunctive imprecision can be represented in disjunctive databases which are also defined.

4.2.2.2 Contributions of Various Authors on Fuzzy Databases

Buckles et al (1981-1985) propose one of the first versions of the Fuzzy Relational DataBase System (FRDBS) by merging the theory of fuzzy sets and the Relational DataBase System (RDBS). They formulate a robust similarity-based theoretical framework for FRDBS that has the following properties:

1. It admits non-atomic tuple components;
2. It requires a similarity relation for each data domain to preserve the important properties of classical relational databases;
3. It accepts a user-defined acceptance threshold when evaluating the query;
4. It supports only a certain class of fuzzy numbers, but not the Possibility Distribution data type. (For comparison, see the model provided by Umano below).

Shenoi et al (1989-1992) generalize the similarity-based model. They note that preserving the above properties of classical RDBS can also be achieved by restricting the components of fuzzy tuples to non-empty subsets of equivalence classes from partitions of the domain. Since the notion of equivalence classes is more general than the notion of similarity relation, an FRDBS equivalence model has been proposed, which is a generalization of the similarity-based model.

Another different approach in the representation and manipulation of fuzzy data is advanced by Umano (1993) who develops Freedom-0, a FRDBS. Unlike the model by Buckle et al. and Shenoi *et al.* which limits the fuzzy data to specific fuzzy number, Freedom-0 allows for both possibility distribution and Fuzzy number. However, even though Freedom-0 is more powerful in terms of its fuzzy data structure, it lacks the formal DataBase framework which is found in the models of Buckle *et al.* Freedom-0 uses an embedded programming language in FORTRAN for fuzzy data manipulation. Zemankova in (1984) also develops a FRDBS which can handle both fuzzy set and possibility distribution data. RIM, a conventional non-fuzzy RDBS, is chosen as a host in implementing the FRDBS. Vector attribute type supported by RIM is used to represent fuzzy data. Extension is made to the RIM RDBS data structure so that tuple component is not restricted to atomic values. The FRDBS developed can handle fuzzy query.

Generally speaking, the FRDBS's mentioned above demonstrate the following major advantages over the conventional RDBS model, by Maria, et. al, (1984).

1. It allows a more natural way of handling data because fuzzy data are more compatible with human thoughts and cognitions;
2. The use of fuzzy set and possibility distribution theory provides a formal mathematics foundation for the systematic representation and manipulation of crisp and fuzzy data;
3. It provides a Database environment to handle both crisp and fuzzy data.

However, in this model, there is no doubt that RDBS is the key DataBase framework in fuzzy data handling. Like the older generations of DataBase technology, RDBS was developed for conventional data-processing involving crisp and atomic data structure, and is thus not suitable for applications requiring intensive computation and complex data-structure like fuzzy set data, Won Kim, (1990).

4.3 Fuzzy Object- Oriented DataBase System (FOODS)

This is developed by merging the theory of fuzzy set and Object-Oriented DataBase System (OODBS), an idea of Phang and Lee, (1993). FOODS can handle both fuzzy set and possibility distribution data. FOODS is implemented using an object-oriented programming language Smalltalk/V. New object classes were created to represent and manipulate fuzzy

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁚ Peer-reviewed & Refereed journal ⁚ Vol. 13, Issue 2, February 2024
DOI: 10.17148/IJARCCE.2024.13201

data. Extension is made to the FOODS data structure so that DataBase object is not restricted to atomic values. FOODS can also handle fuzzy query, i.e., Fuzzy Object SQL which is an extension of SQL.

Generally speaking, FOODS demonstrates the following major advantages over the conventional fuzzy relational DataBase model:

4.3.1 Compatible Data Structure.

The conventional relational model only supports atomic data types, whereas the data structure of fuzzy data is often not atomic. It is therefore often difficult to represent fuzzy data in a conventional relational model. However, the object-oriented model can easily handle non-atomic data.

4.3.2 Better Budget Management.

The Fuzzy DataBase query process is computationally intensive, and fuzzification, defuzzification and operations on fuzzy data are difficult to manage within the RDBS framework. The application programmer must manage these calculations explicitly; this is not the case in the object-oriented fuzzy database model. Here, the calculations are called up automatically through message passing, which is a powerful feature of an object-oriented system.

4.3.3 Better Modelling Capability.

The relational model does not support semantic concepts such as aggregation and generalization relationships. Application programmers have to deal explicitly with these semantic concepts in their programs. Unlike its relational counterpart, the object-oriented model allows for complex data structures and also offers greater modelling power, Won Kim (1990).

V. IN-DEPTH KNOWLEDGE OF THE STATE OF THE ART IN FUZZY DATABASE

5.1 Contributions from Researchers

According to Parsons in (2005), real-world applications, data and information are often imprecise, uncertain and vague, and many sources contributed to this. Thus, we are increasingly confronted with large amounts of data generated by both traditional and non-traditional means (e.g. sensors, cameras, genomes, biological and geographical systems, etc.). Since this data is imprecise, imperfect and uncertain, it poses a considerable problem in terms of inclusion, representation and processing in the traditional relational database, Ma, et al., (2010). The introduction of fuzzy logic by Zadeh has contributed to the extension and integration of fuzzy data by different data models, Adnan, et al., (2000).

Okoronkwo. et al., (2020) appraises state of the arts in fuzzy database model that can handle uncertainty data with emphasis on the Zvieli and Chen entity relation model, Relational model of Chaudhry, Moyne and Rundensteiner, extended relational model of Yacizi and Merdan, extended entity relational model of Chen and Kerre, fuzzy object oriented database model of Ma, Zhang, Ma and Chen.

Challa, et al., (2023) look at, examine, and build a conceptual framework for future research on the present scientific discoveries on the use or progress of machine learning using fuzzy logic in a variety of disciplines such as Computer science, Engineering, Mathematics/Statistics, Medical, Finance, and Agriculture fields. Moreover, the study contrasts fuzzy querying with conventional data models. The survey study suggests prospective topics for further research in fuzzy data processing and provides a broad overview of the approaches for fuzzy predictive modeling and retrieval.

5.2 State Of The Art of Fuzzy Database Models

Uncertainty and incomplete data representation, seen as drawbacks of the ER model, have made it necessary to use fuzzy sets and fuzzy logic to extend existing relational database models, B.P. Buckles and F.E. Petry (2011). The state of the art of the fuzzy database model consists of different authors' approaches to the extension and implementation of the fuzzy database model and techniques. These approaches are as follows:

5.2.1 Chen and Kerre's Approach

The approach of Chen and Kerre introduced fuzzy outer tension, which extends the concept of superclass and subclass relations in the ER model using fuzzy logic, by Chen, G.Q. and Kerre, E.E. in (1998 and 2002). The basic idea is that if E1 is a superclass of E2 with $e \in E2$, then $E1(e) \leq E2(e)$, where E1 (e) and E2 (e) are membership functions of e for E1 and E2 respectively, Mira Balaban et al (2002). Chen and Kerre introduced also three types of constraints regarding fuzzy relations, these constraints are:

- Inheritance constraints. This constraint proposes that instances of a subclass inherit instances of all relations in

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁞ Peer-reviewed & Refereed journal ⁞ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13201

which it has participated as an entity of a superclass.

- Total participation constraint. Defined when all entities in the set of entities (E ∃ □i), ai >0, appear in at least one relation in this set of relations.
- Cardinality constraint, a constraint specifying the dependencies between entities in an Entity-Relationship (ER) diagram. In the simplified traditional notation of cardinality, 1 is used for mini and maxi, and a letter (e.g. n) is used for mini≥0 and maxi=N, Mira, B., Galindo, J., et al (2002). There are several types of cardinality constraints, which can be expressed as 1:1, 1: N or N: M relationships according to Adegoke, et al. (2020), If E, R and A are the fuzzy entity types, fuzzy interrelationship types and fuzzy attribute sets of the fuzzy ER model, and μE, μR and μA are their membership functions, the Chen label type and Kerre label type can be represented as shown in the diagram below.

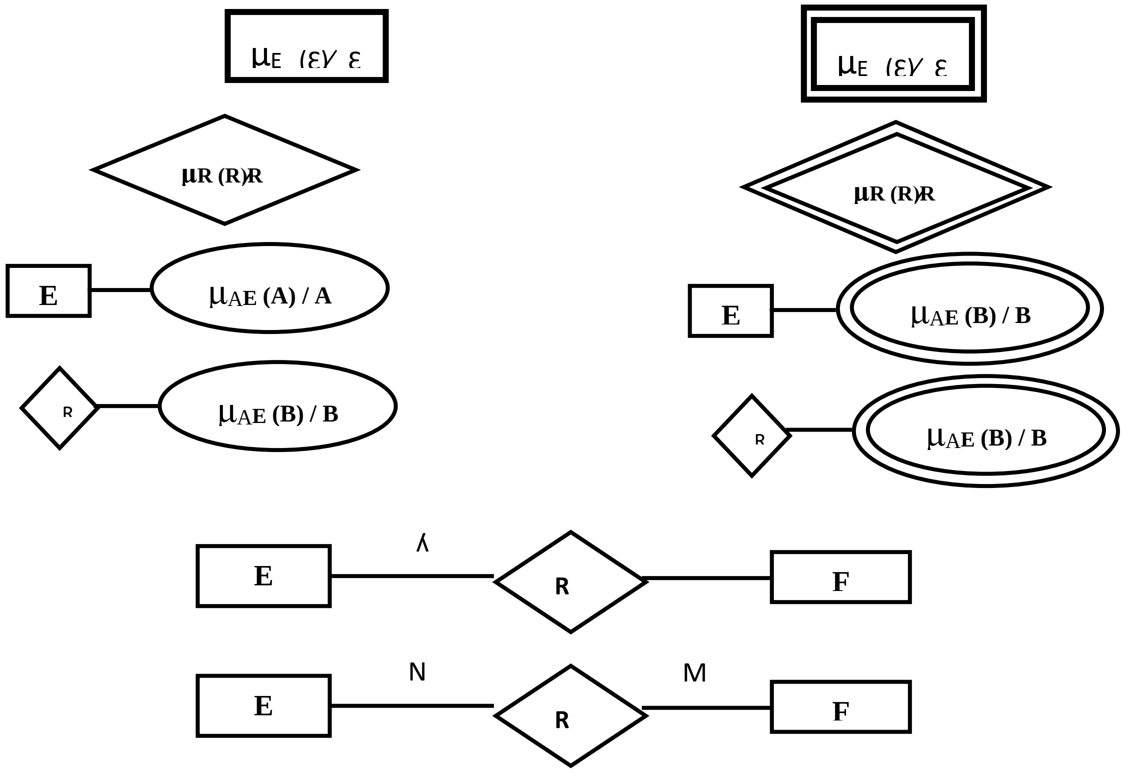

Figure 1 showing ER fuzzy notation proposed by Chen

5.2.2 The Zvieli and Chen Approach

Zvieli and Chen proposed the first approach to incorporating fuzzy logic into ER model extensions, as modelling imprecision is of paramount importance. They adopted the FRDB design methodology and provided the ER data model with an extension to represent data imprecision. They also proposed a series of steps to derive a fuzzy relational database (FRDB) from this extended ER model, Zvieli A. and Chen P. (2008). Zvieli and Chen introduced three fuzzy degrees in the ER model by allowing fuzzy attributes for entities and relations. The three levels are as follows:

1. At the first level, entity set types, relationships and attributes are fuzzy and have degrees of membership. For example, in Figure 2, the fuzzy entity "Company" has a degree of membership of 0.9, the relation "Accept" has a degree of membership of 0.7 and the fuzzy attribute "E-mail address" has a degree of membership of 0.8. The degree of membership is 0.8.

2. The second level concerns the fuzzy occurrence of entities and relations. Here, instances are longer than entities or relationships with different degrees of membership. For example, the entity "Young Employees" must be fuzzy because its position, employees, belongs to entities with different degrees of membership.

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2024
DOI: 10.17148/IJARCCE.2024.13201

3. The third level concerns the fuzzy values of the attributes of entities or special relationships. For example, the "Quality" attribute of a basketball player could be fuzzy (possibilities include bad, good, very good, etc.). Kerre E. E. et al, Galindo, J. et al, and Mira Balaban et al, from (2001-2006).

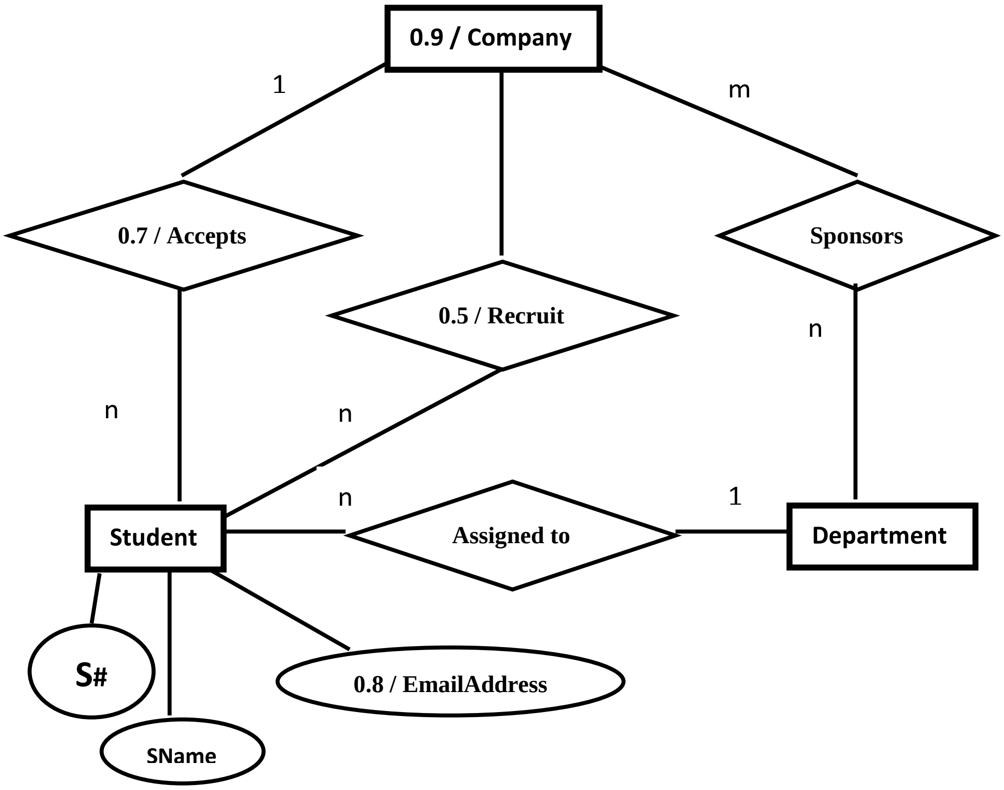

Figure 2, showing membership degrees to the model in some sets (entities, relationships, or attributes): The first level of the Zvieli and Chen approach. [12]

Galindo et al (2006) analyzed Zvili and Chen's proposal and said that the first level can be useful, but it is necessary to decide at the implementation stage whether to show these entities, relationships or attributes. The second level is useful, but it is important to consider other levels of importance (membership, importance, degree of realization, etc.). The third level is useful only to the extent that it is similar to the data type of some attributes.

5.2.3 Yazic and Merdan's Approach

The IFO data model is a mathematically defined data model that combines the basic principles of 'semantic' database modelling with a graph-based formalism. The model uses a directed graph with different types of vertices and edges to represent atomic objects, structured objects, functional fragments, and ISA relationships between them, Adnan Yazicit and Ali INAR (2000). Yazici and Merdan studied this model and adopted the IFO model to incorporate imprecise attributes. They then proposed to extend the IFO model to ExIFO to handle incomplete data through special handling of data where similarities exist in the labels, J. Galindo et al (2006) and Adnan Y., et al (2000). The implementation and validation of the fuzzy conceptual schema representation is done by examining the uncertain attribute representation. They also proposed three constructors in the ExIFO conceptual model based on the IFO conceptual model to allow constructors in the database model to tolerate imprecision and uncertainty. They use fuzzy values such as true attributes, attributes with incomplete values, and attributes with null values, Galindo J., et al (2006). In the first case, for example, if we consider a set of real numbers R= {1, 2, 3, 4, 5} and a subset x= {1, 3, 5}, there is a similarity relation between the domain of the real number attribute and the subset $x \epsilon R$. The second value is an incomplete property, which is a classical incomplete property whose domain is not specific, but only gives a range of numbers (e.g., between 10 and 20).

This is a classic incomplete attribute. The third case is when the actual value of the data is available but not explicitly precise. Examples of this type of attribute may be, whether a certain number exist. Note, the main contribution of this approach is the use of an extended Non-First Normal Form relation (NF2) which is aimed at transforming the conceptual design into a logical design, Galindo J., et al (2006).

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13201

5.2.4 Approach of Chaudhry, Moyne and Rundensteiner

Chaudhry et al. are one of many authors who have proposed a method for extending the classical relational database. Their method proposes the extension of Zvieli and Chen's ER model through a sequence of steps that map the fuzzy ER model to the fuzzy relational database. The two types of imprecision they consider are: (i) imprecision in the degree to which a tuple belongs to a relation and (ii) imprecision in a data value. According to them, "we must first introduce the concept of fuzzy relation, which expresses imprecision in the degree to which a tuple belongs to a relation, and then the concept of possibilistic relation, which expresses imprecision in a data value", Nauman A. Chaudhry (1994). Galindo et al (2006) defined n linguistic labels as n fuzzy sets over the universe of an attribute, where each tuple of the raw entity is transformed to the level of the n value of fuzziness. Each fuzzy tuple (or value) stores not the raw value, but a corresponding linguistic label and a degree of membership of the corresponding raw entity to the new entity. The simple entities and the new fuzzy entity are then mapped into separate tables. The design sequence for the extension of FRDB is as follows: see Figure 6.

Step1: Construction of an extended fuzzy ER data model.
Step2: Conversion of ER model into relational tables.
Step3: Normalization of the relationships.
Step4: Ensuring correct interpretation of fuzzy relational operators.

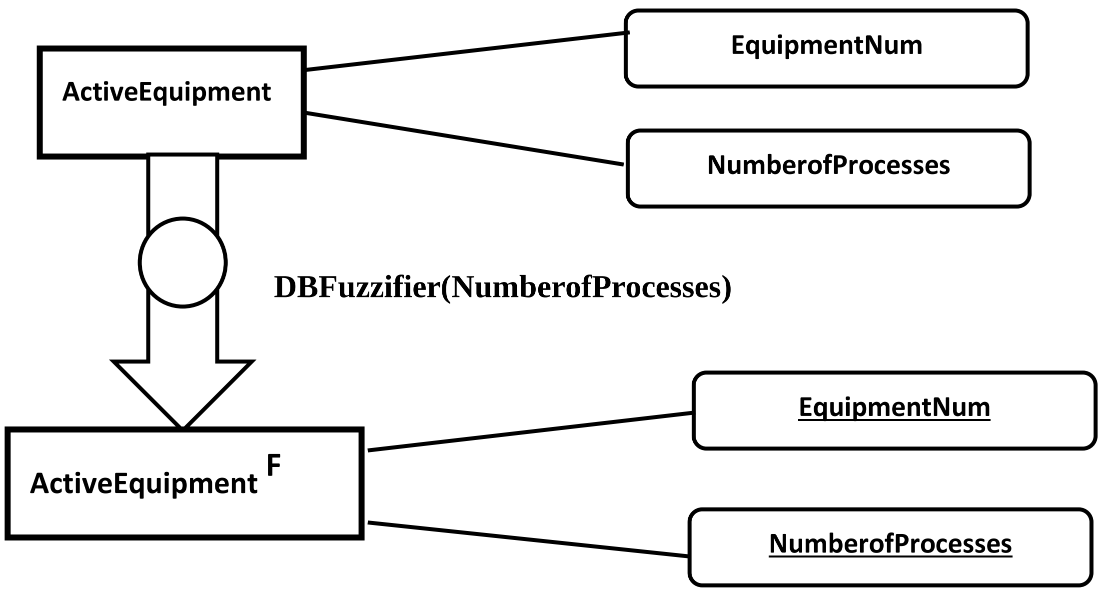

Figure 3. Showing the model proposed by Chaudhry, Moyne, and Rundensteiner (1994): Example of DBFuzzifier transformation

5.2.5 Buckles and Petry model

The Buckles and Petry model is the first model to use similarity relations in the relational model. It provides a structure for representing imprecise information in the form of a relational database. This structure differs from ordinary relational databases in two ways: (1) the components of the tuples do not necessarily have to be unique values and (2) a similarity relation is required for each set of domains in the database. In this model, it is assumed that a fuzzy relation is defined as a subset of the following Cartesian product: P(D1) × ... × P(Dm), where P(Di) represents the elements of the domain (Di), including all subsets that could be considered as part of the domain Di, Peter Chen (1976). The three types of data that this proposal admits are: a finite set of scalars, a finite set of numbers and a set of fuzzy numbers Nauman A. et al (1994).

5.2.6 Approach of Ma, Zhang, Ma and Chen

Ma, Zhang, Ma and Chen studied the work of Zvieli and Chen, especially the three levels, which they then incorporated into the Fuzzy Extended Entity-Relationship model (FEER model). This approach attempts to deal with complex objects in the real world at the conceptual level and attribute their meaning to the degree of each of their components (attributes, entities, etc.). However, just as many constraints have been imposed by the generalization of the definitions of

Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13201

specialization, category and aggregation, Galindo J. et al (2006). In addition, an extended object-oriented database model was introduced in 2014 to deal with imperfect, imprecise and complex objects.
The EOODBM modules they extended are: Objects, Classes, Object-Class Relationships, Subclass/Superclass and Multiple Inheritance. Here are some of the FEER notations they propose:
a) Fuzzy attributes, entities and interrelationships monovalent attribute type
b) Specialization, aggregation and fuzzy categories

5.2.7 Possibility Models.

Possibility theory is based on the idea that linguistic variables are linked to fuzzy sets. In this way, it is possible to evaluate the possibility that the variable X belongs to the set Y, as well as the degree to which the element X belongs to Y, Prof. Rahul Rishi, et al (2013). Some examples of possibilistic models are discussed below:

5.2.7.1 The Prade-Testemale model.

This is an FRDB model that allows incomplete or uncertain data to be integrated into the theory of possibilities and the relations corresponding to the knowledge base. After integration, the uncertain data is stored in the form of tables, regardless of the fact that there may be differences in the type of values in the columns. For example, an attribute A, having a domain X with e as a special element, denotes a scenario where A is not applied to y, Anupriya, Prof. Rahul Rishi, et al (2013). Thus, the values of A for an object y can be represented by a possibility distribution $\pi A(y)$ over $X \cup \{e\}$ such that the DP, $\pi A(y)$ is an application that runs from $X \cup \{e\}$ to the interval [0, 1], Anupriya, Prof. Rahul Rishi, et al (2013).

5.2.7.2 The Zemankova-Kaendel model.

This model dates from 1984 and 1985 and is based on three databases: a value database, an explanatory database and a set of translation rules. The data in the value database is ordered in a similar way to possibilistic models, while the explanatory databases contain fuzzy subsets and fuzzy relations. The set of translation rules corresponds to the different processing measures for adjectives and modifiers, Anupriya, Prof. Rahul Rishi, et al (2013). The possibility measure, PA(S) is used to find the compatibility of the fuzzy subset, S of the condition, with an attribute A value for each tuple in the relation is given as $PA(S) = \sup_{x \in X} \{\mu F(x) \cdot \pi A(x)\}$.

5.2.7.3 The GEFRED Model.

According to Prof. Rahul Rishi, et al (2013), the GEFRED (Generalized Fuzzy Relational Database) model was proposed in 1994 by Medina-Pons-Vila. When it was developed, the fuzzy domain was considered in the context of the possibilistic model. GEFRED is designed to contain unknown, undefined and null values, which means it can handle different types of data. It also redefines relational algebra operators such as union, intersection, difference, Cartesian product, projection, selection, join and division in generalized fuzzy relational algebra.

VI. OBSERVATIONS AND FUTURE RESEARCH

The approaches by the authors on the use of fuzzy model to integrate imprecise and imperfect data into the database have aligned with the benefits of the fuzzy model which are based on a generality of function estimators: clarity, modularity, ability to be explained, easy handling of uncertainty, and parallel processing of rules. Tuqyah, et al., (2015) there are however, some very important drawbacks which portends major limitations to the fuzzy model that the authors did not factor in.

The drawbacks are: the high computational costs, severe computing power restrictions, comprehensibility, and optimization. Reus de N. M. (1994) Future research should therefore incorporate the computational complexities and the severe computing restriction. Typical references are the limitations inherent in bioinformatics settings such that the computational complexities have created hurdles for crispy data defuzzification. Tuqyah, et al., (2015)

VII. CONCLUSION

This article gives an overview of the different models of fuzzy database systems with regard to the representation of fuzzy data and the database framework. It seems that among these models, the object-oriented fuzzy database model offers many advantages. We believe that this approach to fuzzy database design is feasible, versatile and has many advantages. To demonstrate the feasibility of the proposal, a prototype of an object-oriented fuzzy database, FOODS, has been created. An overview of fuzzy sets and fuzzy logic, which have become useful tools for accurately modelling and integrating uncertain, imprecise and imperfect real-world data, is also provided. The summary of the different approaches to augment traditional, conventional databases in this research paper should provide an up-to-date overview of the state of the art in fuzzy database modelling, limitations and future areas of research.

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁓ Peer-reviewed & Refereed journal ⁓ Vol. 13, Issue 2, February 2024
DOI: 10.17148/IJARCCE.2024.13201

REFERENCES

[1] Abraham Bookstein. "Fuzzy Requests: An Approach to weighted Boolean Searches", *Journal of the American Society for Information Science*, July 1980.
[2] Adegoke, O. Folake et al, International Journal of Computer Science and Mobile Computing, Vol.9 Issue.6, June-2020, pg. 120-131
[3] Adnan YAZICI and Ali ~INAR. Conceptual Modeling for the Design of Fuzzy Object Oriented Databases. Springer-Verlag Berlin Heidelberg 2000 https://link.springer.com/chapter/10.1007/978-3-7908-1865-9_2
[4] Amit Garg & Dr. Rahul Rishi. *Global Journal of Computer Science and Technology Volume 12 Issue 6 Version 1.0 March 2012*
[5] Ankita Srivastava*and Awadhesh Kumar Sharma. *A Study and Comparison of Methods for Fuzzy Data Equiva- lence Volume 3, No. 3, May-June 2012 International Journalof Advanced Research in Computer Science ISSN No. 0976-5697.*
[6] Anupriya, Prof. Rahul Rishi, Mira Balaban et tal. Re- view of Fuzzy Logical Database Models. *IOSR Journal of Computer Engineering (IOSRJCE) ISSN: 2278-0661, ISBN: 2278-8727 Volume 8, Issue 4 (Jan. - Feb. 2013), PP 24-30.* www.iosrjournals.org www.iosrjournals.org
[7] Billy P. Buckles. "Fuzzy DataBases and Their Application" in [29], pp. 361-371.

[8] Billy P. Buckles and F. E. Petry. "A Fuzzy Representation of Data for Relational DataBases", *Fuzzy Sets and Systems*, Vol. 7, 1982, pp. 213-226.
[9] Billy P. Buckles and F. E. Petry. "Information-Theoretical Characterisation of Fuzzy RelationalDataBases", *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC- 13, No. 1,January/February 1983.
[10] Billy P. Buckles and F. E. Petry. "Uncertainty Models in Information and DataBase Systems", *Journal of Information Science*, Vol. 11, 1985 pp. 77-87.
[12] Bellman, R. E., & Zadeh, L. A., Decision making in a fuzzy environment in Management Sciences, 17(4), pp. 141-175. [Online] Available: http://ayouty.com/___Gam_Shorouq_final/5info_sys/FUZZY_DECISION.pdf
[13] Bosc P. et al. "Fuzzy Querying with SQL: Extensions and Implementation Aspect", *Fuzzy Sets and Systems*, Vol. 28, 1988, pp. 333-349.
[14] Buckles B. P. and Petry F.E., "A fuzzy representation of data for relational databases" *Fuzzy Sets and Systems*, 7, pp. 213–226, 2011.
[15] Chen G.Q. and Kerre E.E., "Extending ER/EER conceptstowards fuzzy conceptual data modeling" *Proc. of the IEEE International Conference on Fuzzy Systems*, 1998, pp. 1350-1325.
[16] De Reus N. M., Assessment of benefits and drawbacksof' using fuzzy logic. NWO especially in fire control systems II *https://apps.dtic.mil/dtic/tr/fulltext/u2/a285428.pdf*
[17] Eugene Wong. "A Statistical Approach toIncomplete Information in DataBase Systems", *ACM Transactions on DataBase Systems*, Vol. 7, No. 3, September 1982, pp. 470-488.
[18] Gupta M. M. and Sanchez E., *Fuzzy Information and Decision Process*, North- Holland 1982.
[19] Howard L. Resnikoff et al. "Information Science" in*Identifying Research Areas in the ComputingIndustry to 1995*, Robert F. Cotellessa. Noyes Publications, 1984.
[20] Hudec M., *Fuzzy Data in Traditional Relational Databases, IEEE conference of Neurel, Belgrade (2014), 195-199*
[21] Janusz Kacprzyk et al. "FQUERY lll+: A 'Human- Consistent' DataBase Querying System Based on Fuzzy Logic with Linguistic Quantifiers", *Information Systems*, Vol. 14, No. 6, 1989.
[22] Janusz Kacprzyk and A. Ziolkowski. "DataBase Queries with Fuzzy Linguistic Quantifiers", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-16, No. 3, May/June 1986.
[23] Galindo J., Urrutia A., Piattini M., (2006), *Fuzzy Databases: Modeling Design and Implementation*, Hershey, USA: IDEA Group,
[24] Kerre E. E. and Chen G.Q., "Fuzzy data modeling at a conceptual level: Extending ER/EER concepts" *Knowledge management in fuzzy databases*, 2001, pp. 3-11
[25] Lucarella D. and Morara R., "FIRST: Fuzzy Information Retrieval System", *Journal of Information Science*, Vol. 17, 1991, pp. 81-91.
[26] Maria Zemankova-Leech and A. Kandel. Fuzzy*Relational DataBases - a Key to Expert Systems.* Verlag TUV Rheinland 1984.
[27] Ma+ Z. M. and Li Yan. A Literature Overview of FuzzyConceptual Data Modeling Journal *of Information Science and Engineering 26, 427-441 (2010) 427*
[28] Mira Balaban et tal, *Enforcing Cardinality Constraints in the ER Model with Integrity Methods1*. Ben-Gurion University of the Negev, Israel Journal of Database Man- agement, 11(4), 2002, pp. 14-23
[29] Miyamoto S., "Information Retrieval Based on Fuzzy Citations", *Fuzzy Sets and Systems*, Vol. 38, 1990, pp. 191-205.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 13, Issue 2, February 2024
DOI: 10.17148/IJARCCE.2024.13201

[30] Motohide Umano. "Freedom-0: A Fuzzy DataBase System" in [29], pp. 339-347.
[31] Ms. Poonam, Banasthali Vidyapith, Rajasthan. Fuzzy Database Data Models and Its Framework. *International Journal of Computer Science Trends and Technology – Vol- ume 3 Issue 2, Mar-Apr 2015.*
[32] Nomoto K. et al. "A Document Retrieval System Based on Citations Using Fuzzy Graphs", *Fuzzy Sets and Systems*, Vol. 38, 1990, pp. 207-222.
[33] Nauman A. Chaudhry, James R. Moyne, Elke A. Run- densteiner. A Design Methodology for Databases with Uncertain Data. Https://www.computer.org/csdl/proceedings/ssdm/1994/6610/00/00336964.pdf
[34] Otokwala, U. J., and Ofut, O. T., Systematic Review of the State of the Art inFuzzy Database Model, *International Journal of Scientific & Engineering Research*, ISSN 2229-5518Volume 10, Issue 7, July-2019
[35] Parsons S., "Current approaches to handling imperfect information in data and knowledge bases," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 8, 2005, pp. 353-372
[36] Pavese F., Why should correction values be better known than the measured true value? *Journal of Physics*: Conference series 459, (2013).
[37] Peter Gray, K. G. Kulkarni and N. W. Paton. *Object-Oriented DataBases - A Semantic Data Model Approach*. Prentice-Hall, 1992.
[38] Phang K. K. and Lee P. A. "The Application of Fuzzy Mathematics in DataBase Systems", *International Seminar on Mathematical Science,Kuching, Sarawak*, 1993.
[39] Phang K. K. and Mashkuri Hj. Yaacob. "Fuzzy Object-Oriented DataBase System", *IASTAD, Washington D.C., June 1996.*
[40] Phang K. K. *FOODS - a Prototype Fuzzy Object- Oriented DataBase System*, M. Comp. Sc. Thesis, University of Malaya, 1994
[41] Phang K. K. and Mashkuri Hj. Yaacob, Development of Fuzzy Database Systems, *Malaysian Journal of Computer Science*, Vol. 10 No. 1, June 1997, pp. 42-46
[42] Shi-Kuo Chang and Wu-Haung Cheng. "DataBase Skeleton and Its Application to Logical DataBase Synthesis", *IEEE Transaction on Software Engineering*, Vol. SE-4, No. 1, January 1978.
[43] Sujeet Shenoi and Austin Melton. "Proximity Relations in the Fuzzy Relational DataBase Model", *Fuzzy Sets and Systems*, Vol. 31, 1989, pp. 285-296.
[44] Sujeet Shenoi and Austin Melton. "AnEquivalence Classes Model of Fuzzy Relational DataBases", *Fuzzy Sets and Systems*, Vol. 38, 1990,pp. 153-170.
[45] Sujeet Shenoi and Austin Melton. "Functional Dependencies and Normal Forms in the Fuzzy Relational DataBase Model", *Information Science*, Vol. 60, 1992, pp. 1-28.
[46] Tuqyah Abdullah Al Qazlan, Aboubekeur Hamdi-Cherif, and Chafia Kara-Mohamed. State of the Art ofFuzzy Methods for Gene *Regulatory Networks Infer- ence Scientific World Journal, Volume 2015 (2015), ArticleID, 11 pages* http://dx.doi.org/10.1155/2015/148010 https://books.google.com.ng/books?isbn=3319159704
[47] The Entity-Relationship and Enhanced Entity- Relationship Model. *http://www.peterlo.com/Teaching/IT354/L02.pdf*
[48] Valiollah Tahani. "A Conceptual Framework for Fuzzy Query Processing - A Step toward very Intelligent DataBase Systems", *InformationProcessing & Management*, Vol. 13, 1977.
[49] Won Kim. *Introduction to Object-Oriented DataBases*, MIT, 1990.
[50] Wong M. H. and Leung K. S., "A Fuzzy DataBase-Query Language", *Information Systems*, Vol. 13, No. 5, 1990.
[51] Yahushi Ogawa, T. Morita and K. Kobayashi. "A Fuzzy Document Retrieval System using the keyword connection matrix and a learning method",*Fuzzy Sets and Systems*, Vol. 39, 1991, pp. 163-179.
[52] Zadeh L. A., "Fuzzy Set as a Basis for a Theory ofPossibility", *Fuzzy Set and System*, Vol. 1, 1978, pp. 3-28.
[53] Zvieli A. and Chen P., "ER modeling and fuzzy data- bases" *Proc. of the Second International Conference on Data Engineering*, 2008, pp. 320-327.