WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

An Investigative Review of the Cutting- Edge Technique in Fuzzy Database Modeling

Onyima, O. John[1], Okoronkwo C. Matthew[2], Adaora, A. Obayi[3]

[1,2]*Department of Computer Science, University of Nigeria, Nsukka*
[3]*Department of Computer Science, University of Nigeria, Nsukka*
[1]*John.onyima@unn.edu.ng; [2]matthew.okoronkwo@unn.edu.ng; [3]adaora.obayi@unn.edu.ng*

A B S T R A C T

The relational database model has a major limitation in integrating and managing real-world data that are imprecise, vague and imperfect. The concept of fuzzy database system has been introduced to solve this problem. Fuzzy database model incorporate fuzzy data sets and fuzzy logic to enhance the classical relational database models as a well-designed approach for handling fuzziness. This work reviews, the cutting-edge technique in fizzy database model that can handle imperfect and crispy data with highlight on Zvieli and Chen entity relation model, extended relational model of Yacizi and Merdan, Chen and Kerre extended relational model, Chaudhry, Moyne and Rundensteiner model, Ma, Zhang, Ma and Chen fuzzy object oriented database model. Suggestions for further research and possible improvement of these models were discussed.

Keywords: Database, Data, Entity-Relationship, Fuzzy, Imprecision. Uncertainty, Fuzzy Extended Entity-Relationship

1. Introduction

A Database is an organized arrangement mutually dependent data that helps for efficient retrieval, inclusion and deletion of data from the database. This data can be synchronized in the form of tables, views, schemas, reports, etc. A Database Management System (DBMS) is referred to as the technology used in the creation and management of databases. This is a tool used to organize (i.e. create, retrieve, update, and manage) data in a database. Most times, real- world data generated from unconventional means (like radio frequency identification, ecological data sets, remote sensing, satellite, sensors, etc) are imprecise or not defined. The imprecision of data has given rise to the advancement of fuzziness in data manipulations in database management systems. The fundamental rational for applying fuzzy methods in database design is the need to solve the problem of crispy and imperfect data in database [6]. The design of relational database is formulated on Boolean logic with (1 or 0; true or false) output. The fuzzy database approach estimates information on the level of truth and this has become a better and more adapted way of storing and manipulating imprecise data.

In fuzzy database modeling, there are fundamentally two components known as fuzzy logic and fuzzy set. The fuzzy logic rely on the combination of different mathematical assumptions to express knowledge based on a gradual degree of membership, while the fuzzy set theory implement the robust structure for systematically handling of uncertainty based on fuzziness [5,6].

1.1 Definition of Important Terms

The management of uncertainty in database systems is a very important problem [1], as the information is often vague. Motro states that fuzzy information is content-dependent and he classifies it as follows:

- **Uncertainty**: It is impossible to determine whether the information is true or false. For example, “John may be 38 years old.”
- **Imprecision**: The information available is not specific enough. For example, “John may be between 37 and 43 years old,” “John is 34 or 43 years old” (disjunction), “John is not 37 years old” (negative), or even a simple unknown.
- **Vagueness**: The model includes elements (predicates or quantifiers) that are inherently vague, for example, “John is in his early years” or “John is at the end of his youth.” However, after these concepts have been defined, this case would match the previous one (imprecision).
- **Inconsistency**: It contains two or more pieces of information that cannot be true at the same time. For example, “John is 37 and 43 years old, or he is 35 years old”; this is a special case of disjunction
- **Ambiguity**: Some elements of the model lack complete semantics (or a complete meaning). For example, “It is unclear whether the salaries are annual or monthly.”

There exist a meaningful body of research in fuzzy database modelling over the past years and several breakthroughs have been established in this area. Many fuzzy database models (like relational and object-oriented database) have been proposed. In this work, we will be investigating the various approaches in state of the art in fuzzy database modelling.

2. Cutting-edge Technique in Fuzzy Database Model

Fuzzy database can be described as an upgraded relational database that allows fuzzy attribute values and fuzzy truth values; both of these are expressed as fuzzy set. The uncertainty and incomplete data representation that are seen as a draw back in ER model brought about the application of fuzzy sets and fuzzy logic to enhance existing relational database models [8]. The cutting-edge techniques in fuzzy database modeling are the various approaches by various researchers on the extension and implementation of fuzzy database model and techniques. These approaches are discussed below:

2.1 The Zvieli and Chen Approach

Zvieli and Chen considered as very vital imprecision at the modelling level and then introduced an approach to integrate fuzzy logic into the extension of Entity Relational Model. They modified a design methodology for Fuzzy Relational Database that has extensions for indicating the imprecision of data in the Entity-Relationship (ER) data model. They also came up with modalities for the derivation of a Fuzzy Relational Database (FRDB) from this extended ER model [9, 12]. Zvieli and Chen approved fuzzy attributes in entities and relationships and introduced three levels of fuzziness in the ER model. The three levels are: [9]

1. At the very first level, entity set types, relationships and attributes may be fuzzy and hence have a membership degree. For instance, in Figure 1, the fuzzy entity “Company” has membership degree of 0.9; also the relationship “Accepts” has membership degree of 0.7, and the fuzzy attribute “Email Address” membership degree of 0.8.

2. Then the second level is associated to the fuzzy occurrences of entities and relationships. Where instances belong to the entity or relationship with diverse membership degrees. For instance, an entity “Young_Employees” must be fuzzy, because its instances, its employees, belong to the entity with diverse membership degrees.

3. Finally the third level has to do with the fuzzy values of attributes of special entities and relationships. For example, attribute “Quality” of a basketball player may be fuzzy (the possibilities may be bad, good, very good, excellent and so on). [11, 12, 14]

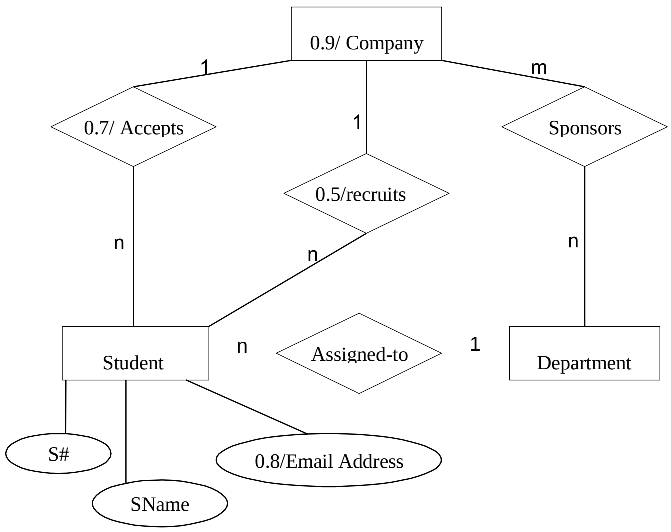

Fig.1. Example with membership degrees to the model in some sets (entities, relationships, or attributes): The first level of the Zvieli and Chen approach. [12]

According to Galindo et al, [12] in investigating the suggestion by Zvieli and Chen, found out that the first level may be helpful, but at the end you have to make decision whether such an entity, relationship, or attribute will or will not come into view in the implementation. The second level is also important, but it is imperative to consider various degree meanings (membership degree, importance degree, fulfilment degree, and so on). The third level is useful and is similar to writing the data type of some attributes, because fuzzy values belong to fuzzy data types.

2.2 Yazici and Merdan Approach

IFO data model is a mathematical data model that uses the basic principles of “semantic” database modeling using a graph-based formalism. Directed graph is used with different types of vertices and edges that represents atomic objects, structured objects, functional fragments and ISA relationships between them [15]. This model was studied by Yazici and Merdan and they adopted the IFO model so as to incorporate imprecise attributes. They then projected an extension of the IFO model to ExIFO (extended IFO) for the processing of imperfect data with exceptional treatment of data where similarity exists in a label [12,15]. The implementation and validation of the representation of a fuzzy conceptual scheme is by looking at a representation of uncertain attributes. Also, they opined three constructors in the conceptual ExIFO model so that the constructors will allow imprecision and uncertainty in database models, based on the IFO conceptual model. Fuzzy values like: true attributes, incomplete-valued attributes, and null- valued attributes were used in their illustration [12]. Firstly, consider a Set of Real numbers, R= {1, 2, 3, 4, 5} and a subset, x= {1, 3, 5}; there will exist a similarity relation between the domain of the real attributes and the subset, x϶R. The second valued-attribute is the incomplete attribute such that the domain is non-specific but will provide a range of numbers (say between 10 and 20) which is a classical incomplete attribute. In the third case, there is availability of true data value but it is not completely precise. An example of this attribute may be, if a certain number exist. Observe that, the major input of this approach is the application of an extended Non-First Normal Form relation (NF2) which is intended at transforming the conceptual design into a logical design [12].

2.3 Chen and Kerre Approach

The Chen and Kerre's approach proposed the fuzzy extension that allows the superclass and subclass relationship concepts of the ER model to be extended using fuzzy logic [10,11]. The fundamental plan is to identify a set of members over a universal space such that, if E1 is a superclass of E2 and e ЄE2, then E1(e) ≤ E2(e), such that E1(e) and E2(e) are the membership functions of e to E1 and E2, respectively [13]. Chen and Kerre went further to explain three kinds of constraints with regards to fuzzy relationships. The constraints are:

1. The inheritance constraint. This constraint shows that an instance of a subclass inherits all relationship instances in which it participated as a superclass entity.
2. Total participation constraint. This is when each entity, a*i* >0 in the entity set (E Ǝ □i) occurs in at least one relationship in that relationship set.
3. The cardinality constraints. In an Entity–Relationship (ER) schemas, this constraint shows the dependencies among the entities. A conventional simplified cardinality notation uses 1 for mini and maxi, and a letter (e.g., n) for mini ≥ 0, maxi = N [12,13]. There are kinds of cardinality constraints, and they could be represented thus: 1:1, 1: N, and N: M relationships.

If *E*, *R*, and *A are:* fuzzy entity type, fuzzy interrelation type, and fuzzy attribute set of the fuzzy ER model, and if μ*E*, μ_R, and μ*A* be their membership functions, then Chen and Kerre label types can be represented in the figures below.

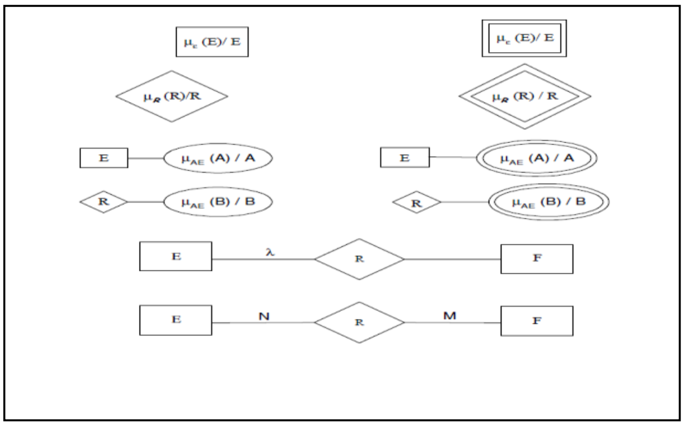

Fig. 2: Showing ER fuzzy notation proposed by Chen [12]

2.4 Chaudhry, Moyne and Rundensteiner Approach

Chaudhry et' al are one of the many authors that anticipated a technique for the extension of the classical relational database. Their technique suggests the extension of the ER model of Zvi- eli and Chen through a series of steps that maps the fuzzy EER model to the fuzzy relational database. They considered two types of imprecision which are: (i) the imprecision in the degree of membership of a tuple in a relation, and (ii) the imprecision in a data value. According these authors, "firstly, present the fuzzy relation construct that expresses the imprecision in the degree of membership of a tuple in a relation, and then the possibilistic relation construct that expresses the imprecision in a data value" [17]. Galindo et' al [12] also defined *n* linguistic labels as *n* fuzzy sets that are over the universe of an attribute with each tuple in the crisp entity transformed up to the level of the *n value* of fuzzy. Each fuzzy tuple (or value) do not store the crisp value but a corresponding linguistic label and a degree of membership to which the corresponding crisp entity belongs in the new entity. The crisp entities and the new fuzzy entity will then be mapped to different tables. Their entity relation model consists of fuzzy relationships as well as relationships with at least one attribute, namely, the membership grade. They recommend FERM, a design methodology for mapping a fuzzy ER data model to a crisp relational database in four steps (formulating a fuzzy ER data model, transforming it to relational tables, normalization, and guaranteeing correct interpretation of the fuzzy relational operators). The authors also presented the application of FERM to build a prototype of a fuzzy database for a discreet control system for a semiconductor manufacturing process. Chaudhry, Moyne, and Rundensteiner extended the model presented in their 1994 paper, with emphasis on their suggestion for the control processes example. In each process, imprecise values are experimental and associated to linguistic labels, and every value involves a process called "DBFuzzifier construct," as can be seen in Figure 3.

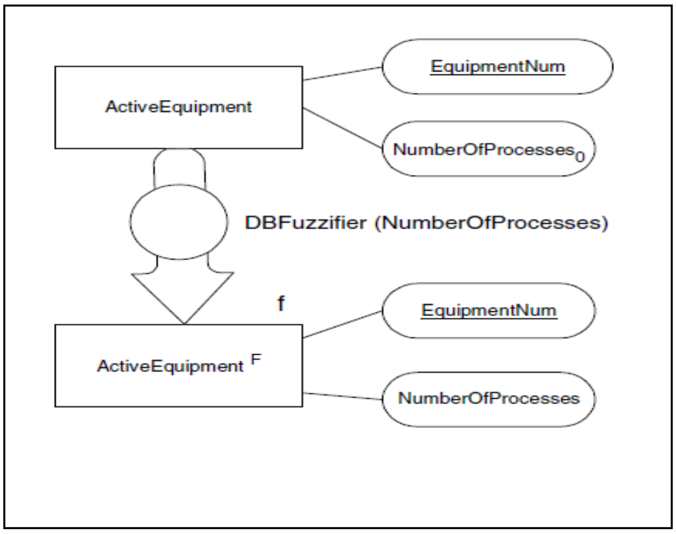

Fig. 3: Showing the model proposed by Chaudhry, Moyne, and Rundensteiner (1994): Example of DBFuzzifier transformation [12,17].

2.5 Proposal of Ma, Zhang, Ma, and Chen Approach

Ma, Zhang, Ma, and Chen did a review on the work of Zvieli and Chen with emphasis on the the three levels which they then included into the Fuzzy Extended Entity-Relationship model (FEER model). This method tries to handle complex objects in the real world at abstract level and correlate their significance to the degree of the various components (attributes, entities, etc.). There were so much constraint was imposed because of their generalization of definitions for specialization, category, and aggregation [12]. Additionally, in 2004, they proposed an extended object-oriented database model to take care of imperfect, imprecise as well as complex objects. The modules of the extended object-oriented database model (EOODBM) that they extended are: objects, classes, objects-classes relationships, sub- class/superclass, and multiple inheritances. Here are **s**ome FEER notations they proposed:

a) fuzzy attributes, entities, and interrelations single-valued attribute type

b) specialization, aggregation, and fuzzy categories

3. Approaches by other Authors

Ruspini in 1986 introduced an extension of the ER model with fuzzy values in the attributes, and a truth value can be connected with each relationship instance. Additionally, some unique relationships such as same-object, subset-of, member-of, and so on are also proposed [2]. In1991 Vandenberghe used Zadeh's extension principle to evaluate the truth value of propositions. For each proposition, a possibility distribution is given on the doubleton true, false of the classical truth values. In this way, the concepts such as entity, relationship, and attribute, as well as subclass, superclass, category, generalization, and specialization, have been extended [3]. The suggestion of Vert, Morris, Stock, and Jankowski is founded on the notation used by Oracle and uses the fuzzy sets theory to treat data sets as a collection of fuzzy objects, applying the result to the area of Geospatial Information Systems (GIS) [4].

More work in fuzzy conceptual data modeling (without using the Entity-Relationship model) is stated in Fujishiro et al., using a graph- oriented schema for modeling a fuzzy database. Fuzziness is taken care of by defining various links between records of the value database (actual data values) and the explanatory database (semantic interpretation of fuzzy attributes, symmetries, and so on) [7]. Extensions that were carried out to allow the modeling of imprecision in semantic data models were also described in Van Gyseghem and de Caluwe discussed two types of imperfect information showing up in database applications: fuzzy information showing information with inherent gradations, for which it is not possible to define sharp or precise borders, and uncertain or imprecise information, representing information that is (temporarily) incomplete as a result of insufficient or more precise knowledge [16]. Handling this type of imperfect information within the formal and crisp environment of a computer, is seen in this paper upon the fuzzy set theory and its related possibility theory, that gives a formal framework to model imperfect information, and upon the object-oriented paradigm, which provides flexible modeling capabilities. The result is the UFO (Uncertainty and fuzziness in object-oriented) database model, a "fuzzy" extension of a full-fledged object-oriented database model.

This research discusses the UFO database model in detail in three steps. Firstly, it shows how fuzzy information is taken care of: Meaningful fuzzifications of numerous object-oriented ideas are introduced so as to store and maintain fuzzy information and to permit a flexible or "soft" modeling of database application. Also, it talks about how uncertainty and imprecision in the information are dealt with: feasible options for the information are stored and maintained by introducing role objects, that are tied like shadows to regular objects in the database and permit the processing of uncertainty and imprecision to the user in an implicit and transparent way, and they also permit the modeling of tentative behavior and of hypothetical information in the database application. Both the static and the dynamic parts of (imperfect) information are developed in the UFO database model, and imperfect information is measured at the data level as well as at the meta level of a database application. The procedure of "extending" an object-oriented database model to the UFO database model, as expressed here, adheres, as closely as possible, to the original principles of the object- oriented paradigm to permit a flexible and transparent, but semantically sound, modeling of imprecise information. The object-oriented database model, from which the extension process starts, keeps to the standard proposal ODMG- 93 to allow for useful implementations of the UFO database model. For the same purpose, the authors' paper also discusses an interface of the UFO database model to an extended relation database model that is competent of handling some imperfect information and for which some prototypes are already available.

4. Observations and Future Research

The various approaches by the authors and researchers on the use of fuzzy model to incorporate imprecise and imperfect data into the database have agreed with the benefits of the fuzzy model which are founded on a generality of function estimators: clarity, modularity, ability to be explained, easy handling of uncertainty, and parallel processing of rules [6]. We will not fail to mention some very imperative drawbacks which shows major limitations to the fuzzy model that were not factored in by the authors. The drawbacks are: the very high computational costs, severe computing power restrictions, comprehensibility, and optimization [18]. Future research should therefore integrate the computational complexities and the severe computing constraint. Classic references are the limitations intrinsic in bioinformatics settings such that the computational complexities have created hurdles for crispy data defuzzification [6].

5. Conclusion

Fuzzy set and fuzzy logic have developed into functional tools for accurate modeling and integration of real-world uncertain, imprecise and imperfect data. Critically reviewing the various approaches of the extension of the traditional classical databases in this research work was with a view to putting forward an up-to-date cutting-edge technique in fuzzy database modeling, the limitations and the future research areas.

References

[1] Motro, A., Management of uncertainty in database system. In W. Kim (Ed.), Modern database system: The object model, interoperability and beyond. Addison-Wesley.

[2] Ruspini, E. Imprecision and uncertainty in the entity-relationship model. In H. Prade & C. V. Negoita, Fuzzy logic in knowledge engineering (pp. 18-22). Verlag TUV Rheinland.

[3] Vandenberghe, R. M. (1991). An extended entity-relationship model for fuzzy databases based on fuzzy truth values. In Proceedings of the Fourth International Fuzzy Systems Association World Congress, Brussels (pp. 280-283).

[4] Vert, G., Morris, A., Stock, M., & Jankowski, P. (2000). Extending entity-relationship modelling notation to manage fuzzy datasets. In Proceedings of the Eighth International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Spain (pp. 1131-1138).

[5] Miroslav Hudec. Collecting, storing and managing fuzzy data in statistical relational databases

[6] Tuqyah Abdullah Al Qazlan, Aboubekeur Hamdi- Cherif, and Chafia Kara-Mohamed. State of the Art of Fuzzy Methods for Gene Regulatory Networks Infer- ence Scientific World Journal, Volume 2015 (2015), Article

ID,11 pages http://dx.doi.org/10.1155/2015/148010

[7] Fujishiro, I. et al. (1991). The design of a graph-oriented schema for the management of individualized fuzzy data. Japanese Journal of Fuzzy Theory and Systems, 3(1), 1-14.

[8] B.P. Buckles and F.E. Petry, "A fuzzy representation of data for relational databases" Fuzzy Sets and Systems, 7, pp. 213–226, 2011.

[9] A. Zvieli and P. Chen, "ER modeling and fuzzy data- bases" Proc. of the Second International Conference on Data Engineering, 2008, pp. 320-327.

[10] G.Q. Chen and E.E. Kerre "Extending ER/EER concepts towards fuzzy conceptual data modeling" Proc. of the IEEE International Conference on Fuzzy Systems, 1998, pp. 1350-1325.

[11] E.E. Kerre and G.Q. Chen "Fuzzy data modeling at a conceptual level: Extending ER/EER concepts" Knowledge management in fuzzy databases, 2001, pp. 3-11

[12] J. Galindo, A. Urrutia, M. Piattini, Fuzzy Databases: Modeling Design and Implementation, Hershey, USA: IDEA Group, 2006

[13] Mira Balaban et tal, Enforcing Cardinality Constraints in the ER Model with Integrity Methods1. Ben-Gurion Uni- versity of the Negev, Israel Journal of Database Man- agement, 11(4), 2002, pp. 14-23

[14] Z. M. Ma+ And Li Yan. A Literature Overview of Fuzzy Conceptual Data Modeling Journal of Information Science and Engineering 26, 427-441(2010) 274

[15] 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

[16] Van Gyseghem, N., de Caluwe, R., & Vandenberghe, R. (1993, March). UFO: Uncertainty and fuzziness in an object-oriented model. In Proceedings of the Second IEEE International Conference Fuzzy Systems, San Francisco (pp. 773-778).

[17] 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/1994/6610/00/0033

[18] N.M. de Reus. Assessment of benefits and drawbacks of' using fuzzy logic. NWO especially in fire control systems II https://apps.dtic.mil/dtic/tr/fulltext/u2/a285428.pdf