MJSAT

Malaysian Journal of Science and Advanced Technology

journal homepage: https://mjsat.com.my/

An Extended Type-2 Fuzzy Relational Database Model for Aggregate and Grouping Operations

Nguyen Hoa*[1, 2]

[1] *Information Technology Faculty, Saigon University, Vietnam.*

[2] *Faculty of Information Technology, Industrial University of Ho Chi Minh City, Vietnam.*

KEYWORDS

Fuzzy number
Type-2 fuzzy set
Type-2 fuzzy relation
Fuzzy aggregate function
ET-2FRDB model

ABSTRACT

This paper introduces an extended type-2 fuzzy relational database model (ET-2FRDB) for aggregate and grouping operations that can represent and query uncertain and imprecise information in the real world applications. In ET-2FRDB, each fuzzy relation is represented by a type-2 fuzzy set whose membership degree of each tuple is a fuzzy number on [0, 1], the fuzzy aggregate functions, the fuzzy relational algebraic, aggregate and grouping operations are defined as extensions of those in the classical relational database model thereby the membership degree of tuples associated by using the minimum and maximum of fuzzy numbers. Also, some properties of the fuzzy relational algebraic, aggregate and grouping operations in ET-2FRDB are formulated and proven.

ARTICLE HISTORY

Received 26 June 2023
Received in revised form 24 July 2023
Accepted 10 August 2023
Available online 22 August 2023

1. INTRODUCTION

As we know, many fuzzy relational database models (FRDB), such as [1], [2], [3], [4] and [5], have been developed to overcome the shortcoming of the classical relational database model (CRDB) for dealing with uncertain and imprecise information of objects in the real world applications. As for CRDB in [6] and [7], each FRDB is defined by a data architecture and a set of the data operations corresponding with this architecture. The set of the data operations in CRDB consists of the algebraic operations and aggregate and grouping operations [7]. The aggregate and grouping operations in a database are extremely important and necessary to perform some queries that the algebraic operations are not able to do. For example, the query "compute the average salary of employees within a department" or "find the maximum treatment cost of all old patients in a hospital" can only be performed by aggregate and grouping operations but not by algebraic operations. However, there has been very little work to date on building the set of aggregate and grouping operations in FRDB models. Most of these models feature only basic algebraic operations and omit aggregate and grouping operations. FRDB models are built by extending CRDB using the fuzzy set theory. However, no model would be so universal that could include all measures and tackle all aspects of uncertainty and imprecision of information in the real world. Thus, new FRDB models still continue to be developed for modeling data objects of the real world.

Some FRDB models, such as [8], [9], [10] and [11], represented a fuzzy relation as a set of tuples whose each attribute may take a fuzzy set or a possibility distribution inferred from a fuzzy set. Fuzzy relational algebraic operations on these models were defined by employing similarity relations on the domain of the attributes and proximity binary relations on fuzzy sets or employing the possibility theory and proximity binary relations on possibility distributions. In such models, only [10] and [11] were concerned with fuzzy aggregate operations but fuzzy grouping operations were not defined.

*Corresponding author:
E-mail address: Nguyen Hoa <nguyenhoa@sgu.edu.vn>.
https://doi.org/10.56532/mjsat.v3i3.169
2785-8901/

Other FRDB models, such as [12], [13], [14], [15], [16] and [17], represented a fuzzy relation as a fuzzy set of tuples whose each attribute only took a single value and the membership degree was a real number in [0, 1] or a fuzzy number on [0, 1]. Fuzzy relational algebraic operations on these models were defined by extending directly classical relational algebraic operations based on fuzzy set operations. A FRDB model is called a type-1 fuzzy relational database model (T-1FRDB), as in [12], [13] and [14], if the membership degree of a relation is a real number in [0, 1], and a type-2 fuzzy relational database model (T-2FRDB), as in [15], [16] and [17], if that is a fuzzy number on [0, 1]. T-2FRDB is better than T-1FRDB in representing imprecise information. In [12], fuzzy aggregate functions were mentioned but not formulared formally. In [13], some forms of fuzzy aggregation queries were presented using SQL but there are no formal definitions for these forms. T-2FRDB models in [15], [16] and [17] were proposed to overcome shortcomings of T-1FRDB models in representing imprecise information, however fuzzy aggregate and grouping operations were not defined therein. Thus, the abilities of expressing and dealing with imprecise information of those T-2FRDB models were limited in the real world applications.

In this paper, we propose an extended type-2 fuzzy relational database model (ET-2FRDB) of CRDB with fuzzy aggregate and grouping operations to overcome the shortcomings of the model in [15] for computing and querying imprecise information in practice. In ET-2FRDB, fuzzy relations are represented as type-2 fuzzy sets, fuzzy relational operations are defined as extensions of those in CRDB [7], T-1FRDB [14] and T-2FRDB [15] thereby the membership degree of tuples associated by using the minimum and maximum of fuzzy numbers in [18]. Some properties of such fuzzy relational operations are also formulated and proven.

Fuzzy sets and fuzzy numbers as the basis of mathematics to develop ET-2FRDB are presented in Section 2. The methodology for building the data model, defining algebraic and aggregate, grouping operations of ET-2FRDB is introduced in Section 3. Section 4 shows out the achieved results and discussion of ET-2FRDB model. Finally, Section 5 concludes the paper and outlines further research directions in the future.

2. FUZZY SETS AND NUMBERS

Fuzzy sets and numbers are used to represent fuzzy relations and execute queries in ET-2FRDB.

2.1 Fuzzy Sets

The fuzzy set is extended from the classical set as in [18], [19] and defined as below.

Definition 1. A *fuzzy set A* on a universal set X is defined by a membership function μ_A from X to the closed interval [0, 1] and denoted by $A = \{x: \mu_A(x) | x \in X\}$. For each $x \in X$, $\mu_A(x)$ or $A(x)$ is the membership degree of x in A.

The *support* of a fuzzy set A on X is the classical set that contains all the elements of X that have nonzero membership degrees in A. The *height* $h(A)$ of a fuzzy set A on X is the largest membership degree obtained by any element in that set. It means $h(A) = \sup_{x\in X} A(x)$. A fuzzy set A is called *normal* if $h(A) = 1$ and *subnormal* if $h(A) < 1$. A fuzzy set A on the real number set $\mathbb{R}$ is called *convex* if for any elements x, y, z in the support of A, the relation $x < y < z$ implies that $\mu_A(y) \geq min(\mu_A(x), \mu_A(z))$.

The standard operations on fuzzy sets are defined by extending the operations on classical sets as follows.

Definition 2. Let A, B be fuzzy sets on a universal set X. The *complement of A, union, intersection and difference of A and B* are fuzzy sets on X and defined by:

1. $\bar{A}(x) = 1 - A(x)$,
2. $(A \cup B)(x) = \max[A(x), B(x)]$,
3. $(A \cap B)(x) = \min[A(x), B(x)]$,
4. $(A\text{-}B)(x) = \min[A(x), 1\text{-}B(x)]$,
   $\forall x \in X$.

2.2 Fuzzy Numbers

The fuzzy numbers are special fuzzy sets that are used to represent the fuzzy relations in ET-2FRDB. The fuzzy numbers are defined in [18] as follows.

Definition 3. *A fuzzy number A* is a fuzzy set on the real number set $\mathbb{R}$ such that:

1. A is a normal and convex fuzzy set.
2. The support of A is bounded.

Definition 4. Let A and B be two fuzzy numbers. The *minimum and maximum of A and B* are fuzzy numbers, denoted MIN(A, B) and MAX(A, B), that are defined by:

1. $\text{MIN}(A, B)(z) = \sup_{z=\min(x, y)}\min[A(x), B(y)]$,
2. $\text{MAX}(A, B)(z) = \sup_{z=\max(x, y)}\min[A(x), B(y)]$,
   $\forall x, y, z \in \mathbb{R}$.

The fuzzy set defined as above is also called the *type 1 fuzzy set* or the *ordinary fuzzy set*. The type 2 fuzzy set is extended from the ordinary fuzzy set in [18] and [20] as follows.

Definition 5. Let $\wp([0, 1])$ be the set of all ordinary (type-1) fuzzy sets on [0, 1]. A *type-2 fuzzy set A* on a universal set X is defined by a membership function μ_A from X to $\wp([0, 1])$. For each $x \in X$, $\mu_A(x)$ is the membership degree of x for A.

We note that the membership degree of a type-1 fuzzy set is a real number in [0, 1], whereas, that of a type-2 fuzzy set is a fuzzy number or set on [0, 1].

3. METHODOLOGY

ET-2FRDB model including the data model and the set of the fuzzy relational algebraic, aggregate and grouping operations is defined and built by extending the classical relational database model [7] using the fuzzy sets and numbers presented above.

3.1 ET-2FRDB Data Model

As CRDB data model, ET-2FRDB data model is a structure including fundamental components as the schema, relation and database. The ET-2FRDB schema and relation are defined in turn as follows.

Definition 6. An *ET-2FRDB schema* is a pair $R = (\boldsymbol{U}, \mu)$, where

1. $\boldsymbol{U} = \{A_1, A_2, \ldots, A_k\}$ is a set of pairwise different attributes,
2. μ is a mapping from $dom(A_1)\times dom(A_2)\times\ldots\times dom(A_k)$ to $\mathcal{E}([0, 1])$, where $dom(A_i)$ is the domain of the attribute A_i ($i = 1, \ldots, k$) and $\mathcal{E}([0,1])$ is the set of all fuzzy numbers on [0, 1].

As in CRDB, the symbols $R(\boldsymbol{U}, \mu)$ and R can be used to replace $R = (\boldsymbol{U}, \mu)$.

Definition 7. Let $\boldsymbol{U} = \{A_1, A_2, \dots, A_k\}$ be a set of k pairwise different attributes. An ET-2FRDB relation over $R(\boldsymbol{U}, \mu)$ is a finite subset $r = \{t_1, t_2, \dots, t_n\}$ of the set $dom(A_1)\times dom(A_2)\times\dots\times dom(A_k)$ of the domains of the attributes $A_1, A_2, \dots, A_k$ such that each t_i is associated with the fuzzy number $\mu_r(t_i) \in \mathcal{E}([0,1])$ representing the membership degree of t_i in r, for every $i = 1, 2,\dots, n$.

Note that each ET-2FRDB relation r is a type-2 fuzzy set and the fuzzy membership degree $\mu_r(t_i)$ expresses the imprecise degree of the values of attributes $A_1, A_2, \dots, A_k$ of each tuple t_i in r. The ET-2FRDB schema and relation are proper extensions of those in CRDB [7] and T-1FRDB [14], [18] where $\mu_r(t_i)$ was a real number in $\{0, 1\}$ or $[0, 1]$. As in CRDB, each element $t_i = (v_1, v_2, \dots, v_k)$ in an ET-2FRDB relation r is also called a tuple of r and the symbol $t_i.A_j$ or $t_i[A_j]$ is used to denote the value v_j of attribute A_j of tuple t_i. For each set of attributes $H \subseteq \{A_1, A_2, \dots, A_k\}$, the symbol $t[H]$ denotes the rest of t after eliminating the value of attributes not belonging to H. In addition, if we only care about a unique relation over a schema then we can unify its symbol name with its schema's name.

Example 1. An ET-2FRDB fuzzy relation, named DIAGNOSE, over the schema **DIAGNOSE**({D_NAME, P_NAME, P_AGE, P_DISEASE, D_COST}, μ) can be given as Table 1.

Table 1. Relation DIAGNOSE

D_NAME	P_NAME	P_AGE	P_DISEASE	D_COST	μ
Oliver	John	55	Tuberculosis	$30	0.9
Blair	George	46	Cirrhosis	$11	*high*
Anne	Ann	15	Gastritis	$8	*approx_0.7*
Anne	Selena	52	Duodenitis	$8	1.0
Blair	Helen	19	Hepatitis	$10	1.0
Oliver	Paul	60	Lung cancer	$35	{0.8: 0.9, 1:1.0, 0.9:0.9}

In this relation, the attributes P_NAME, P_AGE, P_DISEASE, and D_COST describe the information about the name, age, disease and daily treatment cost of each patient, respectively, the attribute D_NAME expresses the information about the name of the doctors to diagnose and treat for the patients. Fuzzy numbers on [0, 1] such as $\mu(t_2) = high$, $\mu(t_3) = approx\_0.7$ represent the imprecise membership degree of the second and third tuples, where fuzzy numbers *high* and *approx_0.7* may be defined to be *high* = {0.6:0.5, 0.7:0.8, 0.8:0.9, 0.9:1.0, 1:1.0} and *approx_0.7* = {0.6:0.9, 0.7:1.0, 0.8:0.9}, respectively.

Definition 8. An *ET-2FRDB database* over a set of attributes is a set of ET-2FRDB relations corresponding to the set of their ET-2FRDB schemas.

3.2 ET-2FRDB Algebraic Operations

The ET-2FRDB algebraic operations are defined by extending T-1FRDB algebraic operations [14] based on the (type-2) fuzzy set operations using the minimum and maximum of fuzzy numbers [18] to combine the fuzzy membership degree of tuples of relations.

3.2.1 Selection

For defining the selection operation, first, selection conditions are defined as follows.

Definition 9. Let R be an ET-2FRDB schema, X be a set of relational tuple variables and θ be a binary relation from $\{=, \neq, \leq, <, >, \geq\}$, then selection conditions are inductively defined and have one of the following forms:

1. $x.A\ \theta\ c$, where $x \in X$, A is an attribute in R and $c \in dom(A)$.
2. $x.A \Rightarrow f$, where $x \in X$, A is an attribute in R, $\Rightarrow$ is a binary fuzzy relation and $f \in \wp(dom(A))$, where $\wp(dom(A))$ is the set of all fuzzy sets on $dom(A)$.
3. $\neg\ \phi$ if ϕ is a selection condition.
4. $\phi \wedge \psi$ if ϕ and ψ are selection conditions on the same relational tuple variable.
5. $\phi \vee \psi$ if ϕ and ψ are selection conditions on the same relational tuple variable.

Example 2. Consider the ET-2FRDB schema **DIAGNOSE** in Example 1, the selection of "all patients who are *young* and get hepatitis" can be expressed by the selection condition x.P_AGE $\Rightarrow$ *young* $\wedge$ x.P_DISEASE =*hepatitis*, where *young* is the fuzzy set that represent imprecise ages of patients given as in Fig.1.

The interpretation of a selection condition as the semantics of it is defined as below.

Definition 10. Let R be an ET-2FRDB schema, r be a relation over R, x be a tuple variable and t be a tuple in r. The *interpretation of selection conditions* with respect to R, r and t, denoted by $\mathfrak{I}_{R,r,t}$, is the partial mapping from the set of all selection conditions to $\mathcal{E}([0,1])$ that is inductively defined as follows:

1. $\mathfrak{I}_{R,r,t}(x.A\ \theta\ c) = \text{MIN}(\mu_r(t), t.A\ \theta\ c)$.
2. $\mathfrak{I}_{R,r,t}(x.A \Rightarrow f) = \text{MIN}(\mu_r(t), \mu_\varphi(t))$, where $\varphi = x.A \Rightarrow f$.
3. $\mathfrak{I}_{R,r,t}(\neg\phi) = 1 - \mathfrak{I}_{R,r,t}(\phi)$.
4. $\mathfrak{I}_{R,r,t}(\phi \wedge \psi) = \text{MIN}(\mathfrak{I}_{R,r,t}(\phi), \mathfrak{I}_{R,r,t}(\psi))$.
5. $\mathfrak{I}_{R,r,t}(\phi \vee \psi) = \text{MAX}(\mathfrak{I}_{R,r,t}(\phi), \mathfrak{I}_{R,r,t}(\psi))$.

Note that $\varphi = x.A \Rightarrow f$ is the fuzzy set whose elements are tuples in r. For each $t \in r$, $\mu_\varphi(t) = f(t.A)$. Intuitively, $\mathfrak{I}_{R,r,t}(x.A\ \theta\ c)$ and $\mathfrak{I}_{R,r,t}(x.A \Rightarrow f)$ respectively are the satisfied degrees of the conditions $t.A\ \theta\ c$ and $t.A \Rightarrow f$ for the tuple t in r.

Now, the selection on a fuzzy relation under a selection condition is defined as follows.

Definition 11. Let R be an ET-2FRDB schema, r be a relation over R and ϕ be a selection condition. The *selection* on r with respect to ϕ, denoted by $\sigma_\phi(r)$, is the fuzzy relation r^* over R defined by $r^* = \{t \in r \mid \mu_{r^*}(t) = \mathfrak{I}_{R,r,t}(\phi) \neq 0\}$.

Example 3. Let r be the relation DIAGNOSE in Example 1 and *middle_aged* be the fuzzy set whose membership function and graph as in Fig.1 below.

$$middle\_aged(x) = \begin{cases} \frac{x-20}{15}, \forall x \in [20,35) \\ 1, \quad \forall x \in [35,45] \\ \frac{60-x}{15}, \forall x \in (45,60] \\ 0, \quad \forall x \notin [20,60]. \end{cases}$$

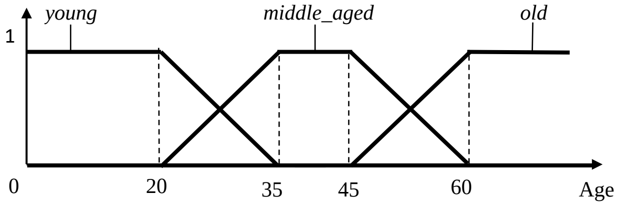

Fig.1. Fuzzy Set Values for Imprecise Ages of Patients

Then the query "find all patients who are middle-aged and have cirrhosis" can be done by the selection operation σ_ϕ(DIAGNOSE), where ϕ = x.P_AGE ⇛ *middle_aged* ∧ x.P_DISEASE = *cirrhosis*.

The selection σ_ϕ(DIAGNOSE) is implemented by checking the satisfaction of all tuples in DIAGNOSE for the selection condition ϕ. Using Definition 10 and the membership function of the fuzzy set *middle_aged* we have the interpretation of ϕ for t_2 being:

$$\mathfrak{T}_{R,r,t_2}(\phi) = \text{MIN}(\mathfrak{T}_{R,r,t_2}(x.\text{P\_AGE} \Rrightarrow middle\_aged), \mathfrak{T}_{R,r,t_2}(x.\text{P\_DISEASE} = cirrhosis))$$
$$= \text{MIN}(\text{MIN}(\mu_r(t_2), middle\_aged(46)), \text{MIN}(\mu_r(t_2), t_2.\text{P\_DISEASE} = cirrhosis))$$
$$= \text{MIN}(\text{MIN}(high, 0.93), high)$$
$$= \{0.6{:}0.5, 0.7{:}0.8, 0.8{:}0.9, 0.9{:}1.0, 0.93{:}1.0\}.$$

For the other tuples, one has

$$\mathfrak{T}_{R,r,t_i}(x.\text{P\_DISEASE} = cirrhosis) = 0, \forall i \neq 2.$$

Thus, the result of the query is the relation σ_ϕ(DIAGNOSE) that consists of one tuple as Table 2.

Table 2. Relation σ_ϕ(DIAGNOSE)

D_NAME	P_NAME	P_AGE	P_DISEASE	D_COST	μ
Blair	George	46	Cirrhosis	$11	{0.6:0.5, 0.7:0.8, 0.8:0.9, 0.9:1.0, 0.93:1.0}

3.2.2 Projection

The projection operation of an ET-2FRDB relation is extended from the projection operation of CRDB [7] and T-1FRDB relations [14] with the fuzzy membership degree of tuples and is defined as follows.

Definition 12. Let $R(\boldsymbol{U}, \mu)$ be an ET-2FRDB schema, r be a relation over R and $\boldsymbol{H} = \{A_1, A_2, \ldots, A_m\}$ be a subset of $\boldsymbol{U}$. The *projection* of r on $\boldsymbol{H}$, denoted by $\Pi_{\boldsymbol{H}}(r)$, is the fuzzy relation r^* over the ET-2FRDB schema R^*, determined by:

1. $R^* = (\boldsymbol{H}, \mu^*)$, where μ^* is the mapping from $dom(A_1)\times dom(A_2)\times\ldots\times dom(A_m)$ to $\mathcal{E}([0, 1])$.
2. $r^* = \{t^* = t[\boldsymbol{H}] \mid t \in r$ and $\mu^*_{r^*}(t^*) = \text{MAX}_{t\in r}\{\mu_r(t) \mid t^* = t[\boldsymbol{H}]\}\}$.

Example 4. The projection of the relation DIAGNOSE in Table 1 on $\boldsymbol{H}$ ={P_NAME, P_DISEASE, D_COST} is the relation $\Pi_{\boldsymbol{H}}$(DIAGNOSE) as in Table 3.

Table 3. Relation $\Pi_{\boldsymbol{H}}$(DIAGNOSE)

P_NAME	P_DISEASE	D_COST	μ*
John	Tuberculosis	$30	0.9
George	Cirrhosis	$11	*high*
Ann	Gastritis	$8	*approx_0.7*
Selena	Duodenitis	$8	1.0
Helen	Hepatitis	$10	1.0
Paul	Lung cancer	$35	{0.8: 0.9, 1:1.0, 0.9:0.9}

3.2.3 Cartesian Product

For defining Cartesian product of ET-2FRDB relations, as in CRDB, we assume that the set of attributes of ET-2FRDB schemas is disjoint and every k-tuple $t = (v_1, v_2,\ldots, v_k)$ is an unordered list. The Cartesian product of two ET-2FRDB relations is extended with the fuzzy membership degree of tuples from that of two T-1FRDB relations [18] as follows.

Definition 13. Let $\boldsymbol{U_1}$, $\boldsymbol{U_2}$ be two sets of attributes that have not any common element, $R_1(\boldsymbol{U_1}, \mu_1)$, $R_2(\boldsymbol{U_2}, \mu_2)$ be two ET-2FRDB schemas, r_1, r_2 be two relations over R_1 and R_2, respectively. The *Cartesian product* of r_1 and r_2, denoted by $r_1\times r_2$, is the fuzzy relation r over the ET-2FRDB schema R, determined by:

1. $R = (\boldsymbol{U}, \mu)$, where $\boldsymbol{U} = \boldsymbol{U_1} \cup \boldsymbol{U_2}$, μ is the mapping from $D_1\times D_2\times\ldots\times D_{k+m}$ to $\mathcal{E}([0, 1])$, $k = |\boldsymbol{U_1}|$, $m = |\boldsymbol{U_2}|$, D_i is $dom(A_i)$, $A_i \in \boldsymbol{U_1} \cup \boldsymbol{U_2}$.
2. $r = \{t = (v_1, v_2, \ldots, v_k, v_{k+1}, v_{k+2},\ldots, v_{k+m}) \mid t_1 = (v_1, v_2, \ldots, v_k)$, $t_2 = (v_{k+1}, v_{k+2},\ldots, v_{k+m})$, $t_1 \in r_1$, $t_2 \in r_2$ and $\mu_r(t) = \text{MIN}(\mu_{1r_1}(t_1), \mu_{2r_2}(t_2))\}$.

3.2.4 Join

The join of ET-2FRDB relations is extended from the natural join of T-1FRDB relations in [14] and [18] with the fuzzy membership degree of tuples as below.

Definition 14. Let $\boldsymbol{U_1}$ and $\boldsymbol{U_2}$ be two sets of attributes such that if they have the same name attributes, respectively in those two sets then such attributes have the same value domain. Let $R_1(\boldsymbol{U_1}, \mu_1)$ and $R_2(\boldsymbol{U_2}, \mu_2)$ be two ET-2FRDB schemas, r_1, r_2 be two relations over R_1 and R_2, respectively and $\{A_h, \ldots, A_k\} = \boldsymbol{U_1} \cap \boldsymbol{U_2}$. The *natural join* of r_1 and r_2, denoted by $r_1 \bowtie r_2$, is the fuzzy relation r over the ET-2FRDB schema R, determined by:

1. $R = (\boldsymbol{U}, \mu)$, where $\boldsymbol{U} = \boldsymbol{U_1} \cup \boldsymbol{U_2}$, μ is the mapping from $D_1\times D_2\times\ldots\times D_n$ to $\mathcal{E}([0, 1])$, $n = |\boldsymbol{U}|$, D_i is $dom(A_i)$, $A_i \in \boldsymbol{U_1} \cup \boldsymbol{U_2}$.
2. $r = \{t = (v_1, \ldots, v_j, v_h, \ldots, v_k, v_m, \ldots, v_n) \mid t_1 = (v_1, \ldots, v_j, v_h, \ldots, v_k)$, $t_2 = (v_h, \ldots, v_k, v_m, \ldots, v_n)$, $t_1 \in r_1$, $t_2 \in r_2$ such that $v_h = t_1[A_h] = t_2[A_h]$, …, $v_k = t_1[A_k] = t_2[A_k]$ and $\mu_r(t) = \text{MIN}(\mu_{1r_1}(t_1), \mu_{2r_2}(t_2))\}$.

Example 5. Let $\boldsymbol{U_1}$ = {P_ID, P_DISEASE}, $\boldsymbol{U_2}$ = {P_NAME, P_DISEASE} be two sets of attributes and PATIENT$_1$, PATIENT$_2$ be two fuzzy relations over two schemas $R_1 = (\boldsymbol{U_1}, \mu_1)$ and $R_2 = (\boldsymbol{U_2}, \mu_2)$ respectively as Tables 4 and 5. The result of the join of PATIENT$_1$ and PATIENT$_2$ is the relation PATIENT$_1 \bowtie$ PATIENT$_2$ over the schema $R = (\boldsymbol{U_1} \cup \boldsymbol{U_2}, \mu)$ computed as in Table 6.

Table 4. Relation PATIENT$_1$

P_ID	P_DISEASE	μ_1
P004	Cirrhosis	1.0
P005	Lung cancer	*high*
P006	Hepatitis	{1: 1, 0.9: 0.8, 0.8: 0.3}

Table 5. Relation PATIENT$_2$

P_NAME	P_DISEASE	μ_2
Blair	Cirrhosis	0.9
Oliver	Hepatitis	{0.6: 0.3, 0.5: 1, 0.4: 0.4}

Table 6. Relation PATIENT$_1 \bowtie$ PATIENT$_2$

P_ID	P_NAME	P_DISEASE	μ
P004	Blair	Cirrhosis	0.9
P006	Oliver	Hepatitis	{0.6: 0.3, 0.5: 1, 0.4: 04}

Here, note that MIN({1: 1, 0.9: 0.8, 0.8: 0.3}, {0.6: 0.3, 0.5: 1, 0.4: 0.4}) = {0.6: 0.3, 0.5: 1, 0.4: 0.4}.

3.2.5 Intersection, Union and Difference

The intersection, union and difference operations of ET-2FRDB relations are straight extended from those of CRDB [7] and T-1FRDB [18] using MAX, MIN of fuzzy numbers as follows.

Definition 15. Let r_1 and r_2 be two relations over the same ET-2FRDB schema $R(\boldsymbol{U}, \mu)$, then the *intersection*, *union*, *difference of* r_1 and r_2, denoted by $r_1 \cap r_2$, $r_1 \cup r_2$, $r_1 - r_2$, respectively is the fuzzy relation r over R defined by:

1. $r_1 \cap r_2 = \{t \mid \mu_{r_1 \cap r_2}(t) = \text{MIN}(\mu_{r_1}(t), \mu_{r_2}(t))\}$,
2. $r_1 \cup r_2 = \{t \mid \mu_{r_1 \cup r_2}(t) = \text{MAX}(\mu_{r_1}(t), \mu_{r_2}(t))\}$,
3. $r_1 - r_2 = \{t \mid \mu_{r_1 \cap \neg r_2}(t) = \text{MIN}(\mu_{r_1}(t), 1\text{-}\mu_{r_2}(t))\}$.

We note that the mathematical base for Definition 15 is Definition 2 of the intersection, union and difference operations on fuzzy sets. Moreover, because the relations r_1 and r_2 are the type-2 fuzzy sets, so instead of using *min* and *max* of real numbers, we employ MIN and MAX of fuzzy numbers.

Example 6. Given two fuzzy relations DIAGNOSE_1 and DIAGNOSE_2 over the schema **DIAGNOSE**({P_ID, D_ID, P_AGE, P_DISEASE, DATE}, μ) as in Tables 7 and 8, then the intersection of DIAGNOSE_1 and DIAGNOSE_2 is the relation $\text{DIAGNOSE}_1 \cap \text{DIAGNOSE}_2$ over **DIAGNOSE** computed as in Table 9.

Table 7. Relation DIAGNOSE_1

P_ID	D_ID	P_AGE	P_DISEASE	DATE	μ
P215	D093	60	Tuberculosis	18/3/2023	1.0
P234	D102	41	Hepatitis	20/3/2023	*high*

Table 8. Relation DIAGNOSE_2

P_ID	D_ID	P_AGE	P_DISEASE	DATE	μ
P383	D102	68	Lung cancer	20/3/2023	0.9
P234	D102	41	Hepatitis	20/3/2023	1.0
P242	D025	15	Gastritis	18/3/2023	*approx_0.7*

Table 9. Relation $\text{DIAGNOSE}_1 \cap \text{DIAGNOSE}_2$

P_ID	D_ID	P_AGE	P_DISEASE	DATE	μ
P234	D102	41	Hepatitis	20/3/2023	*High*

3.3 ET-2FRDB Aggregate and Grouping Operations

For defining aggregate and grouping operations in ET-2FRDB, first, fuzzy aggregate functions in ET-2FRDB are defined as extensions of aggregate functions in CRDB [7] as below.

Definition 16. Let r be a fuzzy relation over an ET-2FRDB schema $R(\boldsymbol{U}, \mu)$, A be an attribute in $\boldsymbol{U}$. Then

1. The *fuzzy aggregate function max* for A on r, denoted by $max_r(A)$, is the operation that computes a pair (v, χ) determined by:
   - $v = \max(\{t[A] \mid t \in r\})$, where $dom(A)$ is an ordered set,
   - $\chi = \text{MIN}(\{\mu_r(t) \mid t[A] = v\})$.
2. The *fuzzy aggregate function min* for A on r, denoted by $min_r(A)$, is the operation that computes a pair (v, χ) determined by:
   - $v = \min(\{t[A] \mid t \in r\})$, where $dom(A)$ is an ordered set,
   - $\chi = \text{MIN}(\{\mu_r(t) \mid t[A] = v\})$.
3. The *fuzzy aggregate function count* for A on r, denoted by $count_r(A)$, is the operation that computes a pair (v, χ) determined by:
   - $v = |\{t[A] \mid t \in r\}|$,
   - $\chi = \text{MIN}_{t \in r}(\{\delta \mid \delta = \text{MAX}_{u \in r}(\{\mu_r(u) | u[A] = t[A]\})\})$.
4. The *fuzzy aggregate function sum* for A on r, denoted by $sum_r(A)$, is the operation that computes a pair (v, χ) determined by:
   - $v = \sum_{t \in r} t[A]$, where the operation of the sum exists on $dom(A)$,
   - $\chi = \text{MIN}_{t \in r}(\{\delta \mid \delta = \text{MAX}_{u \in r}(\{\mu_r(u) | u[A] = t[A]\})\})$.
5. The *fuzzy aggregate function avg* for A on r, denoted by $avg_r(A)$, is the operation that computes a pair (v, χ) determined by:
   - $v = (\sum_{t \in r} t[A])/|r|$, where $r \neq \varnothing$ and the operations of the sum and average exist on *dom(A)*.
   - $\chi = \text{MIN}_{t \in r}(\{\delta \mid \delta = \text{MAX}_{u \in r}(\{\mu_r(u) | u[A] = t[A]\})\})$.

Note that the value χ in the pair (v, χ) represents the imprecise degree of the value v computed in a fuzzy aggregate function. In addition, the symbol $f(A)$ can be used to replace $f_r(A)$ if we only care about a relation r, the expressions $f(A).v$ and $f(A).\chi$ are used to denote the values v and χ in (v, χ). We use the function *count*(*) to count all tuples in a relation.

Now, the ET-2FRDB aggregate and grouping operations are defined as extensions of those in CRDB [7] under fuzzy aggregate functions as follows.

Definition 17. Let r be a fuzzy relation over an ET-2FRDB schema $R(\boldsymbol{U}, \mu)$ and $A_1, A_2, ..., A_k$ be the attributes in $\boldsymbol{U}$ such that the aggregate functions $f_1(A_1), f_2(A_2), ..., f_k(A_k)$ are valid on r and $A_{1f_1}, A_{2f_2}, ..., A_{kf_k}$ be the derived attributes from $f_1(A_1), f_2(A_2), ..., f_k(A_k)$ for taking the values $(v_1, \chi_1), (v_2, \chi_2), ..., (v_k, \chi_k)$ of $f_1(A_1), f_2(A_2), ..., f_k(A_k)$, respectively on r. The *fuzzy aggregate operation* on r with respect to $f_1(A_1), f_2(A_2), ..., f_k(A_k)$, denoted by $\Im_{f_1(A_1), f_2(A_2), ..., f_k(A_k)}(r)$, is the fuzzy relation s over the ET-2FRDB schema $\Im(\{A_{1f_1}, A_{2f_2}, ..., A_{kf_k}\}, \mu_{f_1 f_2, ..., f_{k_s}})$ defined by $s = \{t \mid t[A_{1f_1}] = f_1(A_1).v_1, t[A_{2f_2}] = f_2(A_2).v_2, ..., t[A_{kf_k}] = f_k(A_k).v_k, \mu_{f_1 f_2, ..., f_{k_s}}(t) = \text{MIN}(f_1(A_1).\chi_1, f_2(A_2).\chi_2, ..., f_k(A_k).\chi_k)\}$.

Example 7. Considering the relation DIAGNOSE in Table 1, the query "find the minimum daily treatment cost of patients" can be done by the aggregate operation $\Im_{min(\text{D\_COST})}$(DIAGNOSE) with the result being the relation over the schema $\Im(\{\text{MIN\_DCOST}\}, \mu_{min})$ as Table 10.

Table 10. Relation $\Im_{\boldsymbol{min}(\text{D\_COST})}$(DIAGNOSE)

MIN_DCOST	μ_{min}
$8	*approx_0.7*

Definition 18. Let r be a fuzzy relation over an ET-2FRDB schema $R(\boldsymbol{U}, \mu)$, $H_1, H_2, ..., H_m$ and $A_1, A_2, ..., A_k$ be the attributes in $\boldsymbol{U}$ such that r is partitioned into n subsets $r_1, r_2, ..., r_n$ by the equation binary relations on $dom(H_1), dom(H_2), ..., dom(H_m)$ and the aggregate functions $f_1(A_1), f_2(A_2), ..., f_k(A_k)$ are valid on r and $A_{1f_1}, A_{2f_2}, ..., A_{kf_k}$ be the derived attributes from $f_1(A_1), f_2(A_2), ...,$

$f_k(A_k)$ for taking the values (v_1, χ_1), (v_2, χ_2), ..., (v_k, χ_k) of $f_1(A_1)$, $f_2(A_2)$, ..., $f_k(A_k)$, respectively on r_1, r_2, ..., r_n. The *fuzzy grouping operation* on r with respect to H_1, H_2, ..., H_m and $f_1(A_1)$, $f_2(A_2)$, ..., $f_k(A_k)$, denoted by ${}_{H_1, H_2, ..., H_m}\mathcal{G}_{f_1(A_1), f_2(A_2), ..., f_k(A_k)}(r)$, is the fuzzy relation g over the ET-2FRDB schema $\mathcal{G}(\{H_1, H_2, ..., H_m, A_{1f_1}, A_{2f_2},..., A_{kf_k}\}, \mu_{f_1 f_2, ..., f_{kg}})$ defined by $g = \{t_i \mid t_i[H_1] = t[H_1], t_i[H_2] = t[H_2], ..., t_i[H_m] = t[H_m], t \in r_i, t_i[A_{1f_1}] = f_{r_i}(A_1).v_1, t_i[A_{2f_2}] = f_{r_i}(A_2).v_2,..., t_i[A_{kf_k}] = f_{r_i}(A_k).v_k, \mu_{f_1 f_2, ..., f_{kg}}(t_i) = \text{MIN}(f_{1r_i}(A_1).\chi_1, f_{2r_i}(A_2).\chi_2,...., f_{kr_i}(A_k).\chi_k), i = 1, ..., n\}$.

Example 8. Considering the relation DIAGNOSE in Table 1, the query "find the sum and the daily average treatment cost of patients under the diagnostician" can be done by the grouping operation ${}_{\text{D\_NAME}}\mathcal{G}_{count(*), avg(\text{D\_COST})}(\text{DIAGNOSE})$ with the result being the relation over the schema $\mathcal{G}(\{\text{D\_NAME, P\_COUNT, AVG\_DCOST}\}, \mu_{count,avg})$ as Table 11.

Table 11. Relation ${}_{\text{D\_NAME}}\mathcal{G}_{count(*), avg\,(\text{D\_COST})}(\text{DIAGNOSE})$

D_NAME	P_COUNT	AVG_DCOST	$\mu_{count,avg}$
Oliver	2	\$32.5	{0.8: 0.9, 0.9:1.0}
Blair	2	\$10.5	*high*
Anne	2	\$8	1.0

3.4 Property of ET-2FRDB Operations

The properties of the algebraic, aggregate and grouping operations in ET-2FRDB are extended from those in CRDB as equations below.

Proposition 1. Let R be an ET-2FRDB schema, r be a relation over R, ϕ_1 and ϕ_2 be two selection conditions. Then

$$\sigma_{\phi_1}(\sigma_{\phi_2}(r)) = \sigma_{\phi_2}(\sigma_{\phi_1}(r)) \quad (1)$$

Proof. Let $s = \sigma_{\phi_2}(r)$, we have

$$\begin{aligned}\sigma_{\phi_1}(\sigma_{\phi_2}(r)) &= \{t\in s \mid \mathfrak{T}_{R,s,t}(\phi_1) \neq 0\} \text{ (Definition 11)}\\ &= \{t\in r \mid \mathfrak{T}_{R,r,t}(\phi_2) \neq 0 \wedge \mathfrak{T}_{R,s,t}(\phi_1) \neq 0\}\\ &= \{t\in r \mid \mathfrak{T}_{R,r,t}(\phi_2) \neq 0 \wedge \mathfrak{T}_{R,r,t}(\phi_1) \neq 0\} \\ &\qquad \text{(Because } s \subseteq r)\\ &= \{t\in r \mid \text{MIN}(\mathfrak{T}_{R,r,t}(\phi_2), \mathfrak{T}_{R,r,t}(\phi_1)) \neq 0\}\\ &\qquad \text{(Definition 10)}\\ &= \{t\in r \mid \mathfrak{T}_{R,r,t}(\phi_2\wedge\phi_1) \neq 0\}\\ &= \sigma_{\phi_1\wedge\phi_2}(r).\end{aligned}$$

Since $\phi_1 \wedge \phi_2 \Leftrightarrow \phi_2 \wedge \phi_1$ (the logical conjunction of selection conditions is commutative) hence $\sigma_{\phi_1\wedge\phi_2}(r) = \sigma_{\phi_2\wedge\phi_1}(r)$. Therefore, it results in $\sigma_{\phi_1}(\sigma_{\phi_2}(r)) = \sigma_{\phi_2}(\sigma_{\phi_1}(r))$ □

Proposition 2. Let R be an ET-2FRDB schema, r be a relation over R, $\boldsymbol{A}$ and $\boldsymbol{B}$ be two subsets of attributes of R, $\boldsymbol{A} \subseteq \boldsymbol{B}$. Then

$$\Pi_{\boldsymbol{A}}(\Pi_{\boldsymbol{B}}(r)) = \Pi_{\boldsymbol{A}}(r) \quad (2)$$

Proof. Because $\boldsymbol{A} \subseteq \boldsymbol{B}$, so $\boldsymbol{A}\cap\boldsymbol{B} = \boldsymbol{A}$. From Definition 12, it is easy to see that the sides of (2) are the relations over the same schema with the set of attributes $\boldsymbol{A}\cap\boldsymbol{B} = \boldsymbol{A}$. By the property of the projection of the classical relations, Definition 7 and Definition 12, it follows that two classical sets of tuples which are collected respectively from two relations $\Pi_{\boldsymbol{A}}(\Pi_{\boldsymbol{B}}(r))$ and $\Pi_{\boldsymbol{A}}(r)$ are the same. Moreover, by Definition 12, the operation MAX in two sides of (2) is executed on the same value set of the membership degrees of tuples of r. From that $\Pi_{\boldsymbol{A}}(\Pi_{\boldsymbol{B}}(r)) = \Pi_{\boldsymbol{A}\cap\boldsymbol{B}}(r) = \Pi_{\boldsymbol{A}}(r)$. Thus, the equation (2) is proven □

Proposition 3. Let R_1, R_2 and R_3 be ET-2FRDB schemas such that if they have the same name attributes then such attributes have the same value domain, r_1, r_2 and r_3 be relations over R_1, R_2 and R_3 respectively. Then

$$r_1 \bowtie r_2 = r_2 \bowtie r_1 \quad (3)$$

$$(r_1 \bowtie r_2) \bowtie r_3 = r_1 \bowtie (r_2 \bowtie r_3) \quad (4)$$

The equations (3) and (4) say that the join operation of ET-2FRDB relations is commutative and associative.

Proof. Clearly, $r_1 \bowtie r_2$ and $r_2 \bowtie r_1$ are two relations over the same schema. By the property of the join of the classical relations, Definition 7 and Definition 14, it follows that two classical sets of tuples which are collected respectively from two relations $r_1 \bowtie r_2$ and $r_2 \bowtie r_1$ are the same. In addition, the operation MIN of two fuzzy numbers has commutativity. From that leading the join of tuples has commutativity. So, by Definition 14 we have $r_1 \bowtie r_2 = r_2 \bowtie r_1$.

By Definition 14, clearly $(r_1 \bowtie r_2) \bowtie r_3$ and $r_1 \bowtie (r_2 \bowtie r_3)$ are two relations over the same schema. By the property of the join of the classical relations, Definition 7 and Definition 14, it follows that two classical sets of tuples which are collected respectively from two relations $(r_1 \bowtie r_2) \bowtie r_3$ and $r_1 \bowtie (r_2 \bowtie r_3)$ are the same. Let A be a common attribute in $\boldsymbol{U_1}$, $\boldsymbol{U_2}$ and $\boldsymbol{U_3}$ of R_1, R_2 và R_3, because the operation MIN of two fuzzy numbers and the identical operation of attribute values have associativity, from that the join of tuples has associativity. Thus, by Definition 14, it results in $(r_1 \bowtie r_2) \bowtie r_3 = r_1 \bowtie (r_2 \bowtie r_3)$ □

Because the Cartesian product (Definition 13) is a particular case of the join, it yields the straight corollary of Proposition 3 below.

Corollary 1. Let R_1, R_2 and R_3 be ET-2FRDB schemas such that each pair of them has not any common attribute, r_1, r_2 and r_3 be relations over R_1, R_2 and R_3, respectively. Then

$$r_1 \times r_2 = r_2 \times r_1 \quad (5)$$

$$(r_1 \times r_2) \times r_3 = r_1 \times (r_2 \times r_3) \quad (6)$$

Proposition 4. Let r_1, r_2 and r_3 be relations over the same ET-2FRDB schema. Then

$$r_1 \cap r_2 = r_2 \cap r_1 \quad (7)$$

$$(r_1 \cap r_2) \cap r_3 = r_1 \cap (r_2 \cap r_3) \quad (8)$$

$$r_1 \cup r_2 = r_2 \cup r_1 \quad (9)$$

$$(r_1 \cup r_2) \cup r_3 = r_1 \cup (r_2 \cup r_3) \quad (10)$$

Equations of (7), (8), (9) and (10) say that the intersection and union of ET-2FRDB relations are commutative and associative.

Proof. Because the intersection and union operations of sets and the MIN and MAX operations of fuzzy numbers have commutativity and associativity. So by Definition 15, it follows that the equations (7), (8), (9) and (10) □

Proposition 5. Let R be an ET-2FRDB schema, r be a relation over R and A_1, A_2, ..., A_k be attributes of R such that the aggregate functions $f_1(A_1)$, $f_2(A_2)$, ..., $f_k(A_k)$ are valid on r. Then

$$\mathfrak{I}_{f_1(A_1), f_2(A_2), ..., f_k(A_k)}(r) = \mathfrak{I}_{f_i(A_i), f_j(A_j), ..., f_m(A_m)}(r) \quad (11)$$

where the sequence $f_i(A_i)$, $f_j(A_j)$, ..., $f_m(A_m)$ is a permutation of the sequence $f_1(A_1), f_2(A_2), ..., f_k(A_k)$.

This property says that the ET-2FRDB aggregate operation does not depend on the implemented order of the aggregate functions.

Proof. It is easy to see that two sides of (11) are the relations over the same schema (Definition 17). Because $f_i(A_i)$, $f_j(A_j)$, ..., $f_m(A_m)$, $i, j, m \in \{1,..., k\}$ compute at two sides of (11) on the same relation r, hence return the same result at both sides. By Definition 4, MIN, MAX of fuzzy numbers have commutativity, consequently, the membership degree of tuples of the relations at two sides of (11) is the same. Thus, the equation (11) is proven □

The ET-2FRDB grouping operation does not depend on the implemented order of the aggregate functions as formulated by the following proposition.

Proposition 6. Let R be an ET-2FRDB schema, r be a relation over R and $H_1, H_2, ..., H_m, A_1, A_2, ..., A_k$ be attributes of R such that the aggregate functions $f_1(A_1), f_2(A_2), ..., f_k(A_k)$ are valid on r. Then

$$_{H_1,H_2, ..., H_m}\mathcal{G}_{f_1(A_1),f_2(A_2), ..., f_k(A_k)}(r) = {}_{H_1,H_2, ..., H_m}\mathcal{G}_{f_i(A_i),f_j(A_j), ..., f_m(A_m)}(r) \quad (12)$$

where the sequence $f_i(A_i), f_j(A_j), ..., f_m(A_m)$ is a permutation of the sequence $f_1(A_1), f_2(A_2), ..., f_k(A_k)$.

Proof. Clearly, the relations at two sides of (12) have the same attribute set, thus they are the relations over the same schema (Definition 18). Because the order and number of grouping attributes $H_1, H_2, ..., H_m$ at two sides of (12) is the same, so the number of partitioned groups by r and the attributes of each group at two sides of (12) also are respectively the same. Moreover, the grouping operation computes on each group of a relation similarly to the aggregate operation does on that relation, by Proposition 5, we have the equation (12) □

4. RESULT AND DISCUSSION

We are easy to see that ET-2FRDB is an extension of CRDB and T-1FRDB with type-2 fuzzy sets for aggregate and grouping operations. ET-2FRDB is capable of expressing imprecise information better than T-1FRDB (the conventional fuzzy relational database model). In addition, ET-2FRDB also has the capability of manipulating data effectively.

4.1 Extension of ET-2FRDB in representing and handling data

As presented above, there are two classes of the conventional fuzzy relational database model (FRDB). The first one represents a fuzzy relation as a set of tuples whose each attribute may take a fuzzy set or type 1 fuzzy set. The membership degree of a tuple in a relation is a number in the set {0, 1}. The fuzzy relational algebraic, aggregate and grouping operations are defined by employing similarity (fuzzy) relations on the domain of the attributes and proximity binary relations on fuzzy sets.

The second one represents a fuzzy relation as a type 1 fuzzy set of tuples whose each attribute only takes a single value. The membership degree of a tuple in a relation is a real number in the interval [0, 1]. The fuzzy relational algebraic, aggregate and grouping operations are defined by extending directly CRDB operations based on fuzzy set operations.

Both of two classes of the conventional fuzzy relational database model only use type 1 fuzzy sets and are called the type 1 fuzzy relational database model (T-1FRDB). The extension of T-1FRDB with type 2 fuzzy sets for expressing fuzzy relations where the membership degree of a tuple is a fuzzy number on [0, 1] to create the type 2 fuzzy relational database models (T-2FRDB). To our knowledge, except ET-2FRDB, no T-2FRDB molel is built for fuzzy aggregate and grouping operations. In other words, ET-2FRDB is an extension of T-2FRDB models with fuzzy aggregate and grouping operations. Fig.2 illustrates the extension of ET-2FRDB in comparison with the conventional fuzzy relational database model.

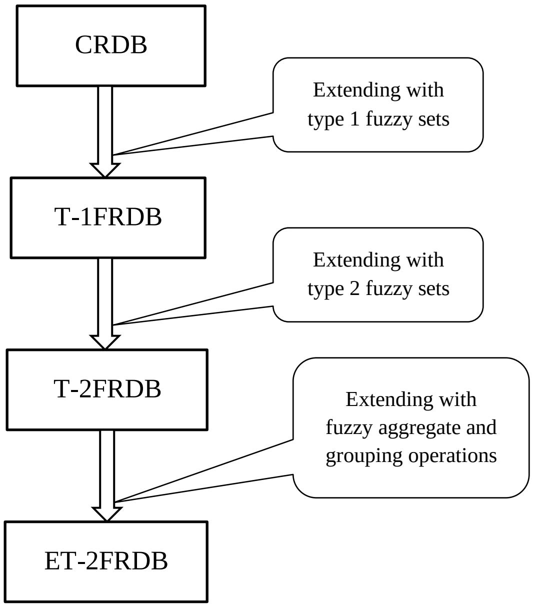

Fig.2. Extension of ET-2FRDB

4.2 Efficiency of ET-2FRDB in computing and querying data

By definitions of fuzzy relational algebraic and aggregate, grouping operations of ET-2FRDB, we can see that the computing complexity of these operations is a polynomial under the size of relations. Indeed, regarding the selection operation, since the computation time of the membership degree of fuzzy sets, MAX and MIN of the membership degrees of tuples is bounded above by some constant (Definition 1 and 4), then the cost for the interpretation of a selection condition for a tuple (Definition 10) also is some constant (i.e. O(1)). It results in the computing time complexity of the selection operation on an ET-2FRDB relation (Definition 11) having n tuples is $O(n)$. Similarly, the computing time complexity of Cartesian product, join, intersection, union and difference operations on two ET-2FRDB relations having n and m tuples is $O(nm)$.

Because the computation time of MAX and MIN of the membership degrees of tuples is bounded above by a constant, we can see that the computing complexity of the projection and aggregate operations on an ET-2FRDB relation (Definition 12, 17) having n tuples is $O(n)$. For the grouping operation, the cost for the grouping time on an ET-2FRDB relation having n tuples is $O(n)$. We have at most n groups and each group has at most m tuples ($m \leq n$), then the computing complexity of the grouping operation is $O(n^2m)$. Thus, we can say that the performance of ET-2FRDB model in computing, manipulating and querying imprecise information is good and can apply it in practice.

5. CONCLUSION

In this paper, we have introduced an extended type-2 fuzzy relational database model (ET-2FRDB) with aggregate and grouping operations. In this extended model, the membership degrees of tuples in a relation are represented by the fuzzy numbers on [0, 1], each relation is a type-2 fuzzy set. The data model and fuzzy relational algebraic and aggregate, grouping operations in ET-2FRDB have been defined formally and coherently. The computation and combination of the membership degrees of tuples in the algebraic and aggregate, grouping operations on ET-2FRDB relations are implemented by the operations MAX and MIN of fuzzy numbers. ET-2FRDB is able to represent and execute the soft queries associated with fuzzy sets as well as the aggregate and grouping computations on fuzzy relations for dealing with imprecise information in real databases.

In the next step, we will build a management system for ET-2FRDB with the familiar querying and manipulating language like SQL that allows expressing and handling imprecise information in the real world.

REFERENCES

[1] B.P. Buckles and F. Petry, "A fuzzy representation of data for relational databases," Journal of Fuzzy sets and Systems, vol.7, pp. 213-226, 1982.

[2] E. Doumard, O. Pivert, G. Smits and V. Thion, "Processing fuzzy relational queries using fuzzy views," Proc. of the 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), New Orleans, Louisiana, USA, 2019, pp. 332-337.

[3] G. Nowakowski, "Fuzzy queries on relational databases," Proc. of the 2018 International Interdisciplinary PhD Workshop, Swinoujscie, Poland, 2018, pp.293-299.

[4] A.A. Sabour, A.M. Gadallah and H.A. Hefny, "Flexible querying of relational databases: fuzzy set based approach," Advanced Machine Learning Technologies and Applications Communications in Computer and Information Science, vol.488, pp.446-455, 2014.

[5] M. Umano and S. Fukami, "Fuzzy relational algebra for possibility-distribution-relational model of fuzzy data," International Journal of Intelligent Systems, vol.3, pp.7-28, 1994.

[6] E.F. Codd, "A relational model of data for large shared data banks," Communications of the ACM, vol.13, no.6, pp.377-387, 1970.

[7] A. Silberschatz, H.F. Korth and S. Sudarshan, Database System Concepts, Seventh Edition, McGraw-Hill, 2019.

[8] H. Prade and C. Testemale, "Generalizing database relational algebra for the treatment of incomplete/uncertain information and vague queries," International Journal of Information Sciences, vol.34, no.2, pp.115-143, 1984.

[9] X. Tang and G. Chen, "Equivalence and transformation of extended algebraic operators in fuzzy relational databases," Journal of Fuzzy Sets and Systems, vol.157, no.12, pp.1581-1596, 2006.

[10] O. Pivert and H. Prade, "Dealing with aggregate queries in an uncertain database model based on possibilistic certainty," Proc. of the 15th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Montpellier, France, 2014, pp.150-159.

[11] E.A. Rundensteiner and L. Bic, "Aggregates in possibilistic databases," Proc. of the Fifteenth International Conference on Very Large Data Bases, Amsterdam, Netherlands, 1989, pp.287- 295.

[12] A. Aguilera, R.M. Toledo, A. Subero, M. Monger and P. Gupta, "On an extension of fuzzy aggregate functions for databases," Journal of Information Systems&Operations Management, vol.7, no.1, pp.1-8, 2013.

[13] X.L. Ye, H. Wang and Y.F. Chen, "Fuzzy aggregation queries in relational databases," Applied Mechanics and Materials, vol.40, pp.195-200, 2010.

[14] F.E. Petry, Fuzzy Databases: Principles and Applications, Kluwer Academic Publishers, 1996.

[15] H. Nguyen, "A type-2 fuzzy relational database model," Journal of Information & Communication Technology: Research and Development on Information & Communication Technology, vol.E–3, no.14, pp.19-26, 2017.

[16] H. Nguyen, T.U.N. Nguyen and N.D Le, "Fuzzy relational database model and management system for imprecise information," Journal of Computer Science and Cybernetics, vol.37, no.2, pp.145-162, 2021.

[17] P.C. Saxena, and D.K. Tayal, "Normalization in type-2 fuzzy relational data model based on fuzzy functional dependency using fuzzy functions," International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol.20, no.1, pp.99-138, 2012.

[18] G.J. Klir and B. Yuan, Fuzzy Sets and Fuzzy Logic–Theory and Applications, Prentice Hall PTR, 1994.

[19] L.A. Zadeh, "Fuzzy sets," Information and Control, vol.8, pp.338-353, 1965.

[20] H. Uesu, "Triangular shaped type-2 fuzzy number and UESU product," ICIC Express Letters, vol.16, no.8, pp. 869-876, 2022.