A Flexible Approach to Matching User Preferences with Records in Datasets based on the Conformance Measure and Aggregation Functions

Miljan Vučetić[1] and Miroslav Hudec[2]
[1]Vlatacom Institute of High Technologies, 5 Milutina Milankovića Blvd, Belgrade, Serbia
[2]Faculty of Economic Informatics, University of Economics in Bratislava, Dolnozemská cesta 1, Bratislava, Slovakia

Keywords: Similarity, Conformance Measure, Fuzzy Conjunction, Uni-norms, Geometric Mean, Quantified Fuzzy Aggregation.

Abstract: Matching user preferences with content in datasets is an important task in building robust query engines. However, this is still a challenging task, because the entities' attributes are often expressed by various data types including numerical, categorical, and fuzzy data. Moreover, the user's preferences and data types for particular attributes may not collide, i.e. the user explains his requirements in linguistic term(s), whereas the respective attribute is recorded as a real number and vice versa. Further, the user may provide different relevancies for atomic conditions, where usual one-directional reinforcement aggregation functions, e.g. conjunction, are not suitable. In this paper, we propose a robust framework capable to manage user requirements and match them with records in a dataset. The former is solved by conformance measure, whereas for the latter the suitable aggregation functions have been suggested to cover particular aggregation needs. Finally, we discuss benefits, drawbacks and outline further activities.

1 INTRODUCTION

When searching for suitable entities (customers, products, territorial units, etc.) in a dataset, users may have a variety of requirements in mind (desired values of entities' attributes), which the best matches should meet. Users require that search process provides them with sensible responses to their requests (Snasel et al., 2007).

In a dataset, attributes' values can be stored by a variety of data types and may be heterogeneous, i.e. values of one attribute may be stored for some records as numeric, whereas for others as fuzzy or categorical data. On the other hand, users can explain their expectations linguistically or numerically. Hence, user preferences and datasets are a mixture of data types including numerical, categorical, binary, and fuzzy data. Moreover, users' preferences and the data types for particular attributes may not collide. A user may explain that the desired flat distance to the lake is very short or short, whereas the distance attribute is recorded as a real number greater than 0. In the opposite case, a user may say that he/she expects the distance to the public transport to be within 200 m, but in a dataset the distance is expressed linguistically by one of the following terms: very short walking distance, short walking distance, medium walking distance, long walking distance, beyond walking distance. This makes application of fuzzy queries such as: FQUERY (Kacprzyk and Zadrożny, 1995), FQL (Wang et al., 2007), SQLf (Bosc and Pivert, 1995), GLC (Hudec, 2009), FSQL (Urrutia et al., 2008), PFSQL (Škrbić and Racković, 2009) and their further extensions, hard. Therefore, the promising option is applying conformance measures (Sözat and Yazici, 2001), initially developed for calculating fuzzy functional dependencies (Sachar, 1986). In this paper, the definition of conformance is different from the one presented in the mentioned studies and is in the line with (Vucetic, 2013), which is based on the fuzzy sets and proximity relation.

Further, the overall query condition may consist of higher number of atomic ones (e.g. features of products which should be met). It restricts query answer to few records, but the possibility of the empty answer problem (Bosc et al., 2008) may appear as well. The quantified queries of the

Vučetić, M. and Hudec, M.
A Flexible Approach to Matching User Preferences with Records in Datasets based on the Conformance Measure and Aggregation Functions.
DOI: 10.5220/0006925801680175
In *Proceedings of the 10th International Joint Conference on Computational Intelligence (IJCCI 2018)*, pages 168-175
ISBN: 978-989-758-327-8

structure: *most of atomic conditions should be met* (Kacprzyk and Ziółkowski 1986), or relaxing atomic conditions (Bosc et al., 2009) are the possible solutions. The former does not divide atomic conditions into hard (must be at least partially met) and soft (it is nice if they are met as well), i.e. a record is a solution even if it does not meet one of the atomic conditions, regardless the importance of this condition. Hence, the possible solutions are quantifying hard and soft constraints suggested by Kacprzyk and Zadrożny (2013) and Hudec (2017). Relaxing query condition is a complex task of relaxing the most suitable atomic predicates by keeping the semantic meaning as close as possible to the initial query (Bosc et al., 2009). In addition, users can express preferences among atomic conditions by various ways: equal preferences, weights, constraints and wishes, etc. When all atomic conditions should be met at least partially, the often used *and* connective or conjunction expressed by t-norms copes with the non-compensatory effect and downward reinforcement (Beliakov et al., 2007).

This study examines benefits of calculating conformances initially developed in (Vucetic et al., 2013) and recently applied in recommending less-frequently purchased products (Vučetic and Hudec, 2018) to reveal how user's requirements and items (records) in the dataset are conforming with the considered attributes. The second part of this study considers suitable aggregations of atomic conditions in order to cover the most expected preferences among attributes raised by users.

2 CONFORMANCE MEASURE

The fuzzy conformance-based approach is suitable for calculating similarity measures among attributes' values and matching complex user requirements with records in a dataset when mixed-type attributes are considered.

The conformance measure is used to compare expected and existing values of particular attributes. In this sense, the value of conformance in the interval [0, 1] is reasonable for observing how the user's requirements and items in the dataset match. Therefore, amongst many methods, this approach is more natural for comparing given crisp, categorical and fuzzy data that appear in user preferences and attributes' values. Although data may be heterogeneous, we are able to straightforwardly measure the similarity between user requirements and item features by (Vučetić, 2013):

$$C(X_i[t_u,t_j]) = \min(\mu_{tu}(X_i), \mu_{tj}(X_i), s(t_u(X_i), t_j(X_i))) \qquad (1)$$

where C is a fuzzy conformance of attribute X_i defined on the domain D_i between user requirement t_u and record t_j in a dataset, s is a proximity relation and $\mu_{tu}(X_i)$ and $\mu_{tj}(X_i)$ are membership degrees of user preferred value and value in a dataset, respectively.

When we analyse fuzzy data, it is necessary to answer how fuzzy value B belongs to the fuzzy set A (e.g. *price about 1 000* belongs to the fuzzy set *medium price*). This is realized by the possibility measure defined as (Galindo, 2008; Zadeh, 1978):

$$Poss(B, A) = \sup_{x\in X}[t(A(x), B(x))] \qquad (2)$$

where X is a universe of discourse and t is a t-norm. In practice, minimum t-norm is used. This equation is used to get membership degree when fuzzy data appears in user requirements and item features in a dataset.

In order to match user requirements with items in a dataset, the first step is fuzzification of attributes domains and definition of proximity relations. For instance, the attribute *walking distance* is fuzzified into several fuzzy sets, as shown in Figure 1.

The fuzzy conformance relies on proximity relations for each attribute domain. These relations are reflexive and symmetric and do not meet the constraint of max-min transitivity as similarity relation does (Shenoi and Melton, 1999).

Proximity relation is defined on the scalar attribute domain and we integrate it under fuzzified domain for numerical attributes. Specifically, by employing fuzzy sets for domain partitions, it is possible to describe similarities between mixed data types. Algorithms (Vučetić and Hudec, 2018; Tung et al., 2006; De Pessemier et al., 2014) calculate the intensity of compatibility between desired value and values of each record (item) in a dataset.

The distance (to the lake, for example) is in our case fuzzified as very short walking distance, short walking distance, medium walking distance, long walking distance, beyond walking distance as illustrated in Figure 1. In this way we work with numerical data and linguistic terms as is shown in examples below. The same holds for the other attributes. For simplicity reasons, these linguistic terms are mathematically formalized by liner membership functions. The proximity relation among these linguistic terms is shown in Table 1.

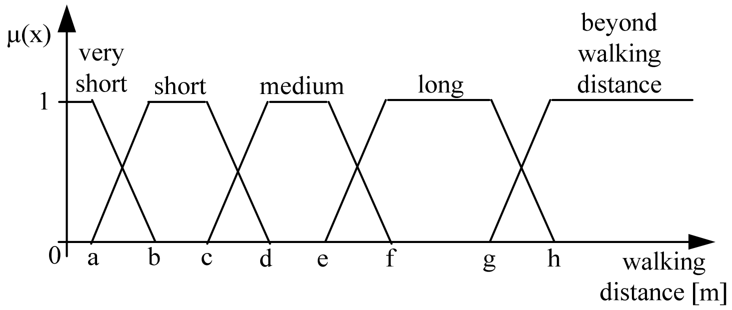

Figure 1: An example of fuzzified attribute *walking distance*.

Table 1: Proximity relation over walking distance domain, where wd stands for walking distance.

S_{wd}	very short wd	short wd	medium wd	long wd	beyond wd
v.sh. wd	1	0.90	0.50	0.10	0
short wd		1	0.80	0.25	0
med. wd			1	0.85	0.45
long wd				1	0.65
beyon wd					1

Let us observe the following examples. The user could start with the preferred walking distance (attribute A_1) $t_u(Walk\_Dist)$ of less than 200 m. Membership degree to the fuzzy set *very short walking distance* is $\mu_{tu}(Walk\_Dist) = 1$ using Eq. (2). For each pair of user requirement and item in the dataset, we use Eq. (1).

In the case of $t_1(Walk\_Dist) = 215$ m, the membership degree to the fuzzy set *very* short walking distance is $\mu_{t1}(Walk\_Dist) = 0.85$ and s(very short, very short) = 1, when parameters $a = 200$ and $b = 300$ in Figure 1:

$C(Walk\_Dist[t_u, t_1]) = \min(\mu_{tu}(Walk\_Dist), \mu_{t1}(Walk\_Dist), s(t_u(Walk\_Dist), t_1(Walk\_Dist)))$
$= \min(1, 0.85, 1) = 0.85$

The conformance of t_u and $t_2(Walk\_Dist)$ = around 670 m (membership degree to the fuzzy set *medium walking distance* is $\mu_{t2}(Walk\_Dist) = 0.70$ using Eq.(2) where $c = 600$, $d = 700$, Figure 1) is given as:

$C(Walk\_Dist[t_u, t_2]) = \min(\mu_{tu}(Walk\_Dist), \mu_{t2}(Walk\_Dist), s(t_u(Walk\_Dist), t_2(Walk\_Dist)))$
$= \min(1, 0.70, 0.50) = 0.50$

The conformance of t_u and $t_3(Walk\_Dist)$, where t_3 contains linguistic term *long wd* is by (1):

$C(Walk\_Dist[t_u, t_3]) = \min(\mu_{tu}(Walk\_Dist), \mu_{t3}(Walk\_Dist), s(t_u(Walk\_Dist), t_3(Walk\_Dist)))$
$= \min(1, 1, 0.10) = 0.10$

It should be noted that conformance may be zero. For example, $C(Walk\_Dist[t_u, t_4])$ between t_u and $t_4(Walk\_Dist) = 2130$ m (membership degree to the fuzzy set *beyond walking distance* is $\mu_{t4}(Walk\_Dist) = 1$,when $h = 2100$ m in Figure 1) for s(very short, beyond) = 0 from Table 1 is calculated as follows:

$C(Walk\_Dist[t_u, t_4]) = \min(\mu_{tu}(Walk\_Dist), \mu_{t3}(Walk\_Dist), s(t_u(Walk\_Dist), t_4(Walk\_Dist)))$
$= \min(1, 1, 0) = 0$

Obviously, the conformance of t_u and t_5, where t_5 contains numerical value of 195 m is 1. These conformances are shown in Table 2, for attribute A_1.

Similarly, we calculate conformances for the other attributes. For instance, attribute A_2 is energy consumption expressed by linguistic terms. The user may express preferred value as a subset {very low, low}, whereas stored data may be expressed by one term when the observation is clear, or by two terms when expert has doubts between, e.g. *low* and *medium*.

The conformance on binary data usually gets value 0 or 1, when the proximity between Yes and No is 0. Theoretically, the proximity can be greater than 0, when these two opposite cases are not fully exclusive for users. For instance, in Table 2 attribute A_4 expresses presence of the elevator in the block of flats. The rest of attributes may be any attribute, e.g. size of flat, storey and aggregated opinion about location on social networks.

Our notion of fuzzy conformance is related to the calculated degree of similarity between user requirements and items in a dataset per particular attribute.

Table 2: Fuzzy conformances of attributes A_1 to A_7 between user preferences expressed by vector of ideal values t_u and records t_1 to t_5.

record	$C(A_1[t_u, t_j])$	$C(A_2[t_u, t_j])$	$C(A_3[t_u, t_j])$	$C(A_4[t_u, t_j])$	$C(A_5[t_u, t_j])$	$C(A_6[t_u, t_j])$	$C(A_7[t_u, t_j])$
t_1	0.85	0.85	0.85	1.00	0.85	0.85	0.85
t_2	0.50	0.25	0.26	1.00	0.29	0.24	0.27
t_3	0.10	0.65	0.46	1.00	0.41	0.88	0.44
t_4	0.00	0.95	0.88	1.00	1.00	0.90	0.85
t_5	1.00	0.25	0.65	0.00	0.25	0.00	0.35

Table 3: Aggregation by t-norms, uni-norm and geometric mean.

record	$C(A_1)$	$C(A_2)$	$C(A_3)$	$C(A_4)$	$C(A_5)$	$C(A_6)$	$C(A_7)$	min t-norm (4)	product t-norm (5)	uninorm (6)	geometric mean (7)
t_1	0.85	0.85	0.85	1.00	0.85	0.85	0.85	0.85	0.3771	1.00	0.8699
t_2	0.50	0.25	0.26	1.00	0.29	0.24	0.27	0.24	0.0006	1.00	0.3474
t_3	0.10	0.65	0.46	1.00	0.41	0.88	0.44	0.10	0.0047	1.00	0.4656
t_4	0.00	0.95	0.88	1.00	1.00	0.90	0.85	0.00	0.0000	0.00	0.0000
t_5	1.00	0.25	0.65	0.00	0.25	0.00	0.35	0.00	0.0000	0.00	0.0000

In the next step, each fuzzy conformance is combined with the aggregation operator to meet user preferences in accordance with his expectations regarding all of the attributes.

The simplest case for finding the best matching record is when a record/item is dominant by all atomic conditions, or is equal to all but one atomic condition and is better than the last one, i.e.

$$t_j \succ t_k \Leftrightarrow V_1(t_j) \geq V_1(t_k) \wedge \cdots \wedge V_{m-1}(t_j) \geq V_{m-1}(t_k) \\ \wedge V_m(t_j) > V_m(t_k) \;\; j,\, k = 1 \ldots n,\, j \neq k \quad (4)$$

where for clarity conformances are expressed as $C(A_i[t_u, t_j]) = V_i(t_j)$.

However, in reality, a record can be more suitable by one and less suitable by another atomic condition or conformance. This case is illustrated in Table 2 for conformance of seven attributes between user preferences t_u and records t_1 to t_5 in a dataset.

The next section is focused on the aggregation of conformances in order to cope with different characters of user preferences.

3 AGGREGATION OF ATOMIC CONDITIONS EXPRESSED BY CONFORMANCE MEASURES

This section examines several most expected cases of aggregation of conformances among attributes covering different kinds of preferences which might be raised by users.

3.1 Conjunction of Equally Important Atomic Conditions Expressed by Conformance

The simplest case is when all conditions are equally important and should be at least partially met. This naturally leads to the aggregation by conjunction, expressed through t-norms. On the other hand, t-norms lack compensation effect, i.e. minimum t-norm (Beliakov et al., 2007) adjusted for conformances (1) for record t_j:

$$t_{min\_tj} = \min_{i=1,\dots,n} C(A_i[t_u, t_j]) \quad (4)$$

(where n is the number of atomic conditions), or have property of downward reinforcement, i.e. product t-norm (Beliakov et al., 2007), also adjusted to conformances (1):

$$t_{prod\_tj} = \prod_{i=1}^{n} C(A_i[t_u, t_j]) \quad (5)$$

More precisely, except the minimum t-norm all other t-norms have the property of downward reinforcement.

This problem is illustrated in Table 3 on the data from Table 2. When six attributes are conforming with value of 0.85 each (record t_1), and one is conforming with value of 1 (neutral element) the overall similarity to the user requirements is 0.3771 calculated by product t-norm (5) (downward reinforcement).

T-norms map result into the unit interval, i.e. $[0, 1]^n \rightarrow [0, 1]$, where 1 is the ideal case. It might lead user to conclude that the record t_1 is not very similar to the desired one by (5); that is, it is far from the ideal value. The solution based on the minimum t-norm (4) reveals the problem of non-compensatory effect, ranking t_2 higher than t_3, even though t_3 is significantly more suitable in all but one conformance and worse in attribute A_1, i.e. values higher than the minimum are not considered.

The disjunction is not the option, because it is not restrictive (value 1 is annihilator), and t-conorms, which model disjunction, also have one-directional, in this case the upward reinforcement property (Beliakov et al., 2007).

Therefore, an alternative may be uni-norms. They meet the property of full reinforcement (Beliakov et al., 2007) punishing low values (as conjunction does) and emphasizing high values (as disjunction does), in our case values of conformances. The 3-∏ function (Yager and Rybalov, 1996) is adjusted to calculate conformance (1) for record *tj* as:

$$u_{3P\_tj} = \frac{\prod_{i=1}^{n} C(A_i[t_u, t_j])}{\prod_{i=1}^{n} C(A_i[t_u, t_j]) + \prod_{i=1}^{n}(1 - C(A_i[t_u, t_j]))} \quad (6)$$

The product in numerator (6) ensures that only the records (items) that at least partially meet all conditions are considered, i.e. value 0 is annihilator. The consequence of being mixed aggregation functions is that value 1 is also annihilator. The uni-norm has desired behaviour when matching degrees of conformances are in the open interval (0, 1). Applying (6) on data in Table 3, has shown that $t_1 - t_4$ fully meet the condition whereas t_4 and t_5 are fully rejected. Record t_4 has conformance equal to 0 for attribute A_1 and therefore is excluded by both: t-norms and uni-norms.

Another options are averaging aggregation functions, but only the borderline case with conjunction functions, to meet the requirement that all atomic conformances should be at least partially met, is suitable. Thus, the solution is geometric mean:

$$av_{geom\_tj} = \sqrt[1/n]{\prod_{i=1}^{n} C(A_i[t_u, t_j])} \quad (7)$$

Applying (7) on data in Table 3 (last column), has shown that t_1 is emphasized, but not as by uni-norm (6), t_3 got better evaluation than t_2 as is expected due to better behaviour in majority of conformances. Records t_4 and t_5 have got conformances equal to 0 for one or more attributes and therefore are excluded by all functions: t-norms, uni-norms and geometric mean.

Although, t-norms are widely used in computing matching degrees for conjunction, the benefit of geometric mean and in the restricted cases of uni-norms should not be neglected, especially when a high number of atomic conformances is considered. In the case of a small number of atomic conditions, t-norms are suitable.

3.2 Quantified Condition of Atomic Conformances

In the aggregations by t-norms, uni-norms and geometric mean the record is excluded when all but one condition are met. It especially holds when the user provides higher number of atomic conditions. However, not all of them must always be met. In this case, we can consider quantified query condition: *most of atomic conditions should be met* (Kacprzyk and Ziółkowski, 1986) or, in our case, most of conformances should be greater than 0. For this purpose we adjusted equation from fuzzy quantified queries (Hudec, 2017) to conformances in the following way:

$$v(t_j) = \mu_Q\left(\frac{1}{n}\sum_{i=1}^{n} C(A_i[t_u, t_j])\right) \quad (8)$$

where v is the validity or matching degree for item t_j to quantified condition, n is the number of conformances and μ_Q is the function of relative quantifier *most of* in the sense of Zadeh (1983) which can be re-formalized as:

$$\mu_Q(y) = \begin{cases} 0 & y \le 0.5 \\ \frac{y - 0.5}{0.4} & y \in (0.5,\ 0.9) \\ 1 & y \ge 0.9 \end{cases} \quad (9)$$

Obviously, the ideal record is one with conformance values equal to 1 for all attributes, regardless the applied aggregation.

Regarding Tables 2 and 3, the best match is t_1 with validity of 0.929, followed by t_4 with validity 0.743. Record t_3 has low validity, more precisely 0.157, and the validities of records t_2 and t_5 are zero. Although, record t_2 met all atomic preferences, these low values are reflected in the proportion. On the other hand record t_4 failed to meet one conformance, but significantly met other ones. This aggregation is suitable when all conditions are considered as soft ones, i.e. it is not imperative that a particular atomic condition should be met, but majority of them.

We should be careful, because this approach is not suitable when several conformances should be imperatively greater than zero. For instance, when one of the considered attributes is price, and the user cannot afford the product that is beyond budget even if all other features are excellently met. The next subsection examines this case.

3.3 Merging Quantified Query Conditions with Conjunction and Other Aggregations

We should be careful with quantified conditions because some of the atomic conditions may be hard constraints like price. If price is beyond the limit, it is irrelevant whether other conditions are met. Such conditions we call hard ones, which we should manage in quantified queries separately. The suitable solution is aggregating hard conditions with the soft ones, managed by quantified condition (8), by conjunction:

$$v(t_j) = (\wedge_{i=1}^{p} C(A_i[t_u, t_j])) \wedge \mu_Q(\frac{1}{q}\sum_{l=1}^{q} C(A_l[t_u, t_j])) \qquad (10)$$

where p is the number of hard conditions and q is the number of soft conditions.

In Section 3.2, the second option is record t_4 form Table 3. However, if the conformance of attribute A_1 is a hard condition, e.g. instead walking distance it represents price, this record is irrelevant and therefore the aggregated value should be 0. The aggregation by (10) provides the expected results shown in Table 4. The results differ in comparison to Table 3 and Section 3.2 because the nature of preferences is changed.

Table 4: Aggregation of hard conditions and quantified condition by (10).

record	hard condition	quantified condition	solution by min t-norm in (10)
t_1	0.85	0.875	0.85
t_2	0.50	0.00	0.00
t_3	0.10	0.350	0.10
t_4	0.00	1.000	0.00
t_5	1.00	0.000	0.00

For conjunction in (10), we can use any t-norm, but we should be aware of the strengths and weaknesses discussed in Section 3.1. We can also apply uni-norm (6) or geometric mean (7) in (10) instead of t-norms.

3.4 Discussion

The inspiration for this work were problems with buying flats, where higher number of attributes is considered. Further, collected data may be mixed data types, i.e. numerical, categorical or fuzzy for the same attribute. In addition, user may explain large scale of preferences among attributes. Moreover, we cannot fully rely on recommender systems for less-frequently bought products, because the history of similar customers is weak (Vučetić and Hudec, 2018).

Aggregation operators should be able to cover variety of preferences among atomic conditions, or in our case conformances. The conformance measure reveals how user requirements and items (records) in the dataset are conforming to the considered attributes.

When small number of atomic conditions is included and all should be at least partially met, the options are t-norms, which formalize conjunction in the fuzzy environment.

On the other hand, when higher number of conformances is included, where all of them should be at least partially, the best matches emphasized (upward reinforcement) and the weak matches punished (downward reinforcement), the solution seems to be reached by uni-norm which have property of full reinforcement, e.g. 3-Π function (6). The product in nominator eliminates items which fail to meet at least one conformance. But, when only one conformance is fully met, item ideally meets requirements regardless other conformances.

The aggregation function, which meets the following requirements: 0 as annihilator, compensation effect without downward reinforcement and value 1 is the neutral element not the annihilator, is geometric mean. This function is the borderline case between conjunctive and averaging aggregation functions.

Further, when a user provides a larger number of atomic conditions, where not all of them must be met, the aggregations by t-norms, uni-norms and geometric mean are not suitable. The solution is quantified aggregation of the structure *most of conformances should be (significantly) met*.

Finally, when several conformances must be met and at least majority of others, the solution is aggregation between hard conditions (conformances which must be at least partially met) and soft conditions (it is beneficial if majority of these conformances are met, i.e. quantified aggregation) by t-norms or geometric mean.

An illustrative example was used to demonstrate various options of conformances among mixed data types and aggregations. Anyway, this approach is a universal framework for working with the real-life data.

4 CONCLUDING REMARKS

In queries, users may be interested in higher number of atomic conditions expressed through preferred values of respective attributes. Fuzzy conformance has been proven to be a very useful approach to measure how user preferences conform to the values stored in datasets. Our work addresses the problem of matching data that contain numerical, categorical, binary and fuzzy data in attributes. The goal is building a framework that automatically handles these mixed data types and different characterization of user preferences. Fuzzy conformance is also the object of intense research activities in other fields such as discovering fuzzy functional dependencies, product recommendation techniques, data fusion in fuzzy relations etc.

Users may also express different natures of preferences among attributes in queries. Although t-norms are widely used in computing matching degrees of atomic conditions, the benefit of geometric mean and possibilities of uni-norms should not be neglected when higher number of atomic conformances is considered due to non-compensatory effect or downward reinforcement property of t-norms. The geometric mean is a suitable solution, because the product of atomic conformances ensures that only the records that at least partially meet all conditions are considered.

Further, higher number of atomic condition may lead to the problem known as *empty answer problem*. The suggested solution is a quantified condition of the structure *most of atomic conformances should be met*. But, when several atomic conditions are hard, (e.g. if price is beyond the budget limits, record is irrelevant regardless it met other requirements ideally), the solution is connective expressed by t-norms, uni-norms or geometric mean between hard conditions and soft conditions in a quantified query.

This study may help software developers to include further flexibility into the data retrieval tasks for data users, when the users consider higher number of atomic features, mixed data types and large scale of possible aggregations among atomic conformances. The overall matching degree in the unit interval clearly indicates how far the considered records to the ideal one are.

REFERENCES

Beliakov, G., Pradera. A., Calvo. T. 2007. *Aggregation Functions: A Guide for Practitioners*. Springer-Verlag, Berlin Heidelberg.

Bosc, P., Pivert, O. 1995. SQLf: a relational database language for fuzzy querying. *IEEE Transactions on Fuzzy Systems*, 3:1–17.

Bosc, P., Hadjali, A., Pivert, O. 2008. Empty versus overabundant answers to flexible relational queries. *Fuzzy Sets and Systems*, 159:1450–1467.

Bosc, P., Brando, C., Hadjali, A., Jaudoin, H., Pivert, O. 2009. Semantic proximity between queries and the empty answer problem. In *Joint 2009 IFSA-EUSFLAT Conference*, Lisbon, EUSFLAT.

De Pessemier, T., Dooms, S., Martens, L. 2014. Comparison of group recommendation algorithms. *Multimedia Tools and Applications*, 72(3):2497-2541.

Galindo, J. 2008. Introduction and Trends to Fuzzy Logic and Fuzzy Databases. In *Handbook of Research on Fuzzy Information Processing in Databases*, pages. 1-33. Information Science Reference, Hershey.

Hudec, M. 2017. Constraints and wishes in quantified queries merged by asymmetric conjunction. In *First International Conference Fuzzy Management Methods (ICFMsquare 2016)*, Fribourg, Springer.

Hudec, M. 2009. An approach to fuzzy database querying, analysis and realisation. *Computer Sciences and Information Systems*, 6:127–140.

Kacprzyk, J., Zadrożny, S. 1995. FQUERY for Access: fuzzy querying for windows-based DBMS. In *Fuzziness in Database Management Systems*, pages. 415–433. Physica-Verlag, Heidelberg.

Kacprzyk, J., Ziółkowski, A. 1986. Database queries with fuzzy linguistic quantifiers. *IEEE Transactions on Systems Man and Cybernetics*, SMC-16:474–479.

Kacprzyk, J., Zadrożny, S. 2013. Compound bipolar queries: combining bipolar queries and queries with fuzzy linguistic quantifiers. In *8th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2013)*, Milano, EUSFLAT.

Sachar, H. 1986. *Theoretical aspects of design of and retrieval from similarity-based relational database systems*. Ph.D. Dissertation, University of Texas at Arlington, Arlington.

Shenoi, S., Melton, A. 1999. Proximity relations in the fuzzy relational database model. *Fuzzy Sets and Systems*, 100:51-62.

Snasel, V., Kromer, P., Musilek, P., Nyongesa, H. O., Husek, D. 2007. Fuzzy Modeling of User Needs for Improvement of Web Search Queries. In *Annual Meeting of the North American Fuzzy Information*

Processing Society (NAFIPS 2007), San Diego.
Sözat, M., Yazici, A. 2001. A complete axiomatization for fuzzy functional and multivalued dependencies in fuzzy database relations. *Fuzzy Sets and Systems*, 117(2):161-181.
Škrbić, S., Racković, M. 2009. PFSQL: a fuzzy SQL language with priorities. In *4th International Conference on Engineering Technologies*, Novi Sad.
Tung A.K.H., Zhang, R., Koudas, N., Ooi, B.C. 2006. Similarity search: a matching based approach. In *the 32nd international conference on Very large data bases*, Seoul.
Urrutia, A., Tineo, L., Gonzales, C. 2008. FSQL and SQLf: Towards a standard in fuzzy databases. In *Handbook of Research on Fuzzy Information Processing in Databases*, pages. 270–298. Information Science Reference, Hershey.Vučetić, M. (2013). *Analysis of functional dependencies in relational databases using fuzzy logic*. PhD thesis, Faculty of Organizational Sciences, Belgrade.
Vučetić M., Hudec M. 2018. A fuzzy query engine for suggesting the products based on conformance and asymmetric conjunction. *Expert Systems with Applications*, 101:143-158.
Vucetic, M., Hudec, M., Vujošević, M. 2013. A new method for computing fuzzy functional dependencies in relational database systems. *Expert Systems with Applications*, 40:2738–2745.
Wang, T-C., Lee, H-D., Chen, C-M. 2007. Intelligent queries based on fuzzy set theory and SQL. In *39th Joint Conference on Information Science*, Salt Lake City.
Yager, R., Rybalov, A. 1996. Uninorm aggregation operators. *Fuzzy Sets and Systems*, 80:111–120.
Zadeh, L.A. 1978. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3-28.
Zadeh, L.A. 1983. A computational approach to fuzzy quantifiers in natural languages. *Computers & Mathematics with Applications*, 9:149–184.