

Modelling a Structure of a Fuzzy Data Warehouse

Alain Kuyunsa Mayu¹, Nathanael Kasoro Mulenda², Rostin Mabela Matendo²

¹Faculty of Sciences, Regional Center for Doctoral Education in Mathematics and Computer Sciences, University of Kinshasa, Kinshasa, D. R. Congo

²Faculty of Sciences, Department of Mathematics and Computer Sciences, University of Kinshasa, Kinshasa, D. R. Congo

Email address:

alainkuyunsa2013@gmail.com (A. K. Mayu), kasoro_mulenda@yahoo.fr (N. K. Mulenda), mabelamatendorostin@gmail.com (R. M. Mabela)

To cite this article:

Alain Kuyunsa Mayu, Nathanael Kasoro Mulenda, Rostin Mabela Matendo. Modelling a Structure of a Fuzzy Data Warehouse. *Applied Engineering*. Vol. 1, No. 2, 2017, pp. 48-56. doi: 10.11648/j.ae.20170102.12

Received: April 12, 2017; **Accepted:** April 22, 2017; **Published:** June 28, 2017

Abstract: In this article, we represent the structure of a fuzzy data warehouse. The elements of classification to build the fuzzy data warehouse are presented through the three following tasks: identification of the target-attribute, identification of linguistic terms and definition of membership functions. From these tasks, we present an approach of a fuzzy data warehouse modelling. This allows us to integrate fuzzy logic without affecting the data warehouse base.

Keywords: Target Attribute, Class Membership Attribute, Membership Degree, Membership Degree Attribute, Fuzzy Classification Table, Fuzzy Membership Table

1. Introduction

Ever growing data are generated from all parts by companies. Decision-making has become crucial for managers. The efficiency of this decision-making is based on the provision of relevant information and suitable tools.

DWH systems received at a for the analysis of surrounding operational systems. Owing to the complexity and the number of operating systems, the data received by DWH are heterogeneous. That is why the data must be homogenized in the first instance before being treated by the data warehouse.

The quantity of data which has to be treated in a DWH increases every day and develops into difficult tasks for the administration and the analysis. A part from the high quantity, data coming from operating systems are often incomplete, unclear or uncertain.

This quality problem cannot be completely eliminated in the data pre treatment stage. Consequently, a certain quantity of imprecision affects directly analysis and decision-making which are based on data warehouse information.

The fuzz's theory proposed by Zadeh in [1] can deal with the lack of clarity, uncertainty and imprecision. Contrary to probability systems, it can so be optimized to interpret imprecision in human language and reasoning.

Consequently, the application of fuzzy logic in DWH technologies enhances data analysis and hence leads to a

better decision-making process.

In this paper, a modelling approach of fuzzy data warehouse is presented. It enables the integration of fuzzy logic without affecting the classical data warehouse base.

2. Preliminaries

2.1. Fuzzy Subset

Let E be a classical set used as a frame of reference.

A fuzzy subset of E is the set of couples $(x, \mu(x)) \forall x \in E$ where μ is a mapping from E onto $[0, 1]$.

This definition generalizes that of a classical subset for which the values of μ are taken only in $\{0, 1\}$. It should be noted that μ_A is the characteristic function of the fuzzy subset A .

We write, along with L. A. Zadeh:

$$A = \mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + \dots + \mu_A(x_n)/x_n$$

For E finite, we write $E = \{x_1, x_2, x_3, \dots, x_n\}$ where $A = \int_E \mu(x)/x$ for E infinite.

Often written

$$A = \{(x_1, \mu(x_1)), \dots, (x_n, \mu(x_n))\}$$

2.2. Fuzzy Number

A fuzzy number of average value n , denoted \tilde{n} , a fuzzy subset of IR the characteristic function of which $\mu_{\tilde{n}}$ is weakly increasing on $(-\infty, n]$, is worth 1 at n , and is weakly decreasing on $[n, +\infty)$

A classical real number n will have a characteristic function such that $\mu_n = 0, \forall x \neq n$

A characteristic function (in rectangular-shaped slot) represents the measure of a magnitude in its uncertainty interval.

A characteristic function (bell-shaped curve) models a fuzzy number \tilde{n} worth around n . Number \tilde{n} will be all the more precise if curve $\mu_{\tilde{n}}$ is «sharp».

In all case, $\mu_{\tilde{n}}(x)$ represents the truth value of the proposition « n has value x ».

2.3. Operations on Fuzzy Sets in IR

-The support of a fuzzy set μ is a classical set

$$\text{supp}(\mu) = \{x \in R / \mu(x) > 0\}$$

-The kernel of a fuzzy set μ is a classical set

$$\text{Ker}(\mu) = \{x \in R / \mu(x) = 1\}$$

-A fuzzy set μ is said to be normal if $\text{Ker}(\mu) \neq \emptyset$.

-The α -cut or α -level (degree) set of the fuzzy set μ is the classical set

$$\mu^\alpha = \{x \in R, \mu(x) \geq \alpha\}$$

-A set μ is said to be convex if $\mu(x)$ is a quasi-concave function.

-A fuzzy number is a normal and convex set in IR .

2.4. Fuzzy Triangular Numbers (FTNs)

Very often, some data and numbers cannot be specified precisely or exactly because of errors on measurement techniques or instruments used. When we say that someone is 180 cm height, their true height can be written conveniently as a FTN $(180 - \alpha, 180, 180 + \beta)$, where α et β represent the range on the left and the range on the right respectively. In general, a FTN (fuzzy triangular number) « a » can be written as $(a - \alpha, a, a + \beta)$ where α and β represent the left margin error and the right one respectively of a .

Those types of numbers are alternatively represented by (a, α, β) . The mathematical definition of a FTN is given below.

$$\mu_M(x) = \begin{cases} 0 & \text{si } x \leq m - \alpha \\ 1 - \frac{m-x}{\alpha} & \text{si } m - \alpha < x < m \\ 1 & \text{si } x = m \\ 1 - \frac{x-m}{\beta} & \text{si } m \leq x < m + \beta \\ 0 & \text{si } x \geq m + \beta \end{cases}$$

Point m , with membership degree 1, is called *central value* of FTN M , and α and β are respectively the differences on the left and on the right with respect to the central value m .

A FTN M is symmetric if $\alpha = \beta$. Because of the great number of applications of FTNs, several authors have studied the algebraic properties on FTNs.

Let us give some of the most common ones. The definitions of these arithmetic operations are based on those given by Dubois and Prade [2].

Let $M = (m, \alpha, \beta)$ and $N = (n, \gamma, \delta)$ two FTNs:

1. Addition:

$$N + M = (m + n, \alpha + \gamma, \beta + \delta)$$

2. Scalar Multiplication:

If λ is a scalar, we have:

$$\lambda M = \lambda(m, \alpha, \beta) = (\lambda m, \lambda \alpha, \lambda \beta), \quad \text{si } \lambda \geq 0$$

$$\lambda M = \lambda(m, \alpha, \beta) = (\lambda m, -\lambda \alpha, -\lambda \beta), \quad \text{si } \lambda \leq 0$$

In particular, $-M = -(m, \alpha, \beta) = (-m, \beta, \alpha)$

3. Subtraction:

$$M - N = (m, \alpha, \beta) - (n, \gamma, \delta) = (m - n, \alpha + \delta, \beta + \gamma)$$

For 2 NFTs M and N , the three operations, addition, subtraction and scalar multiplication all give FTNs.

4. Multiplication:

It can be shown that the product $M \cdot N$ of 2 FTNs is not necessarily a FTN. However, the MF of the resulting fuzzy number of the product of M and N is bounded by a triangle. A better approximation is the following:

-If $M \geq 0$ and $N \geq 0$ ($M \geq 0$ if $m \geq 0$), then

$$M \cdot N \approx (mn, m\gamma + n\alpha, m\delta + n\beta)$$

-If $M \leq 0$ and $N \geq 0$, then

$$M \cdot N = (m, \alpha, \beta) \cdot (n, \gamma, \delta) \approx (mn, n\alpha - m\delta, n\beta - m\gamma)$$

-If $M \leq 0$ and $N \leq 0$, then

$$M \cdot N = (m, \alpha, \beta) \cdot (n, \gamma, \delta) \approx (mn, n\beta - m\delta, n\alpha - m\gamma)$$

Now we define the quotient of two NFTs using the definitions of multiplication and the inverse of two FTN as follows:

5. Inverse

The inverse of a FTN, $M = (m, \alpha, \beta)$ with $m > 0$ is defined as follows:

$$M^{-1} = (m, \alpha, \beta)^{-1} \approx (m^{-1}, \beta m^{-2}, \alpha m^{-2})$$

It is an approached value of M^{-1} which remains valid only in the neighbourhood of $\frac{1}{m}$.

The inverse of 2 FTNs M and N is thus given by $\frac{M}{N} = M.N^{-1}$.

Given that the inverse and the multiplication are approximations, the quotient will also be an approximation.

The formal definition is given below:

6. Division:

$$\begin{aligned} \frac{M}{N} &= M.N^{-1} \\ &\approx (m, \alpha, \beta) (n^{-1}, \delta n^{-2}, \gamma n^{-2}) \\ &\approx \left(\frac{m}{n}, \frac{m\delta + n\alpha}{n^2}, \frac{m\gamma + n\beta}{n^2} \right) \end{aligned}$$

From the definition of multiplication, we define the power of a FTN as follows:

7. Power

Using the definition of multiplication, it can be shown that:

$$M^n = (m, \alpha, \beta)^n \approx (m^n, -nm^{n-1}\beta, -nm^{n-1}\alpha, nm^{n-1}\beta)$$

3. Fuzzy Data Warehouse

The DWH models enable the construction of data bases dedicated to analysis. That analysis can be carried out at different granularity levels, on large volumes of data represented in an aggregated manner.

Moreover, the models of fuzzy data bases are particularly interesting for the representation of imprecise and uncertain data and the inclusion of uncertain quests.

In this section, we present a concept of fuzzy DWH based on the structure of a meta-table.

3.1. The Concept of DWH

Fuzzy concepts can be integrated as the meta table structure without affecting the core of a DWH. Our approach is more flexible, since it enables the integration and definition of the fuzzy concept, without the need for redesigning the DWH core. The use of this DWH storing approach makes it possible to extract and analyse data simultaneously, in a classical for m and a fuzzy manner. The purpose of this section is to represent some concepts of the meta tables, the modelling guidelines and the meta model of the fuzzy DWH approach.

In order to integrate fuzzy concepts in a DWH, we start by identifying and analysing the elements which have to be fuzzily classified in the DWH. Such an element can be a fact

in the fact table or the dimension attribute. An element which has to be fuzzily classified is called the *Target attribute* and the value range of that element instances is called the *Domain of attribute*.

We define below some basic concepts linked to fuzzy DWH, concepts which will be used thereafter.

Domain of attribute: A set of potential values or the range of potential values of a dimension attribute or a fact is called *Domain of attribute or Universe of discourse* of a domain. The *Domain of attribute* of a Domain A is denoted Dom .

Target attribute: A dimension attribute or a fact which is to be fuzzy classified is called a *Target attribute* (TA). Under fuzzy classification, the TA instances are classified over as $et(S)$ represented by a linguistic variable. The linguistic variable consists of a set of non numerical terms called *linguistic terms*, $S = T_1, \dots, T_k$.

The linguistic terms of a linguistic variable are captured in an attribute called *class membership attribute*.

Class membership attribute (CMA): a CMA for a target attribute TA, represented by CMA_{TA} , is an attribute which has a set of linguistic terms T_1, \dots, T_k to which the target attribute may belong. In other words, for all the possible values of a TA (of domain of attribute, Dom_{TA}), there exists a corresponding relation to a CMA value. The values of CMA are the values in the set S .

All values of Dom_{AC} to some fuzzy degree belong to a CMA value. The degree of membership to a CMA value is called membership degree and it defines the relation of an instance TA to a CMA value.

The membership degree (MD $\in [0, 1]$). It is the measure to which the values of a target attribute TA are linked to some linguistic terms T_1, \dots, T_k , respectively with the values of CMA.

The MG is calculated using the membership function.

Membership function (MF): the MF of a CMA class is a function $\mu(TA)$ which is used to calculate the MD of a TA to a CMA $\mu: TA \rightarrow [0, 1]$

The membership degrees generated by the membership functions are captured as membership degree attributes in the fuzzy data warehouse model. A membership degree attribute is defined as follows:

Membership Degree Attribute (MDA): the MDA of a TA is an attribute which has a set of MD of the TA. The value of a MD is calculated by a MF and is represented by $\mu_t(TA) = MG$ where MG is the membership degree of TA for the linguistic term in CMA.

An attribute which has to be fuzzily manipulated is prolonged by two meta tables. The first meta table contains a description of the fuzzy concept and the second meta table contains membership degrees of each instance with regards to the CMA.

The two tables are defined as follows:

Fuzzy Classification Table (FCT): A table which consists of linguistic terms and their unique identifiers is called FCT. There are two attribute tables which consist of an ID attribute and a CMA, where the ID attribute is a unique identifier of the table values. Formally,

$$FCT_{TA} = \{\text{identifier}; CMA_{TA}\}$$

A table which stores the values representing the measure to which a value is linked to a linguistic is called *Fuzzy membership table* (FMT). It is a four-attribute table: the Identifier table attribute, the target attribute identifier TA, the class membership attribute CMA_{TA} in the fuzzy classification table (FCT_{TA}) and the membership degree attribute (MDA) for TA. Formally,

$$FMT_{TA} = \{\text{identifier}; \text{identifier of TA}; \text{Identifier of CMA}_{TA}; MDA_{TA}\}$$

3.2. Model of Fuzzy DWH

The model of fuzzy DWH is a combination of four types of tables. They are dimension tables, fact tables, FMTs and FCTs.

A fuzzy DWH is a set of tables represented the following way

$$FDW = \{\text{Dim}, \text{Fact}, FCT_{TA}, FMT_{TA}\}$$

Where

Dim = {a set of category attributes; level of category attributes}

Fact = {a set of measures}

TA = {TA₁, TA₂, ..., TA_n}, n is the number of FCAs.

It should be noted that the set of TAs is a subset of the dimensions and the set of facts. Formally, TA is a sub set of Dim U Fact (i.e. $\forall TA_i \in \text{Dim U Fact}: 1 \leq i \leq n$).

For each TA₁, TA₂, ..., TA_n:

$$FCT_{TA_i} = \{\text{identifier}; CMA_{TA_i}\} \text{ où } 1 \leq i \leq n.$$

$$FMT_{TA_i} = \{\text{Identifier}, \text{Identifier of FCT}, \text{Identifier of TA } i, MGA_{TA_i}\} \text{ où } 1 \leq i \leq n.$$

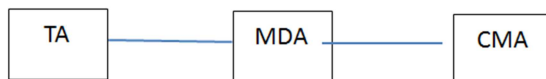
Guide lines for the Modelling of the Fuzzy Data Warehouse

We present in this section a set of guidelines for the design of a fuzzy DWH model and the use of these guide lines for the elaboration of a meta model for the FDW using a real-life case.

i. Distinct Fuzzy Classes / Linguistic Conditions

A set of linguistic terms (also called fuzzy classes) is used to classify instances of a target attribute. In the simplest case, the linguistic terms are distinct, given that there is only one set of non repeated linguistic terms between them. In this case, one instance of a target attribute belongs to only one fuzzy class at a time and the degree of relation is measured by the MF. Formally,

TA—instance (1): Fuzzy classes(1)



Guideline 1. Add a fuzzy classification table (FCT) and a fuzzy membership table (FMT) and a fuzzy membership table (FMT) for each target attribute TA, as indicated below.



Different Membership Degrees for the Same Linguistic Conditions

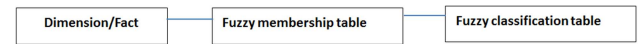
An instance of a TA may belong to a linguistic term but may have different degrees with which it belongs to a linguistic term. It is due to the fact that multiple professional users have different interpretations about one instance of a target attribute i.e. multiple MFs are used for a TA. Formally,

TA—instance (1): Fuzzy classes (1)

But with different membership degree



Guideline2: If an instance of a TA belongs to a fuzzy class, but with multiple degrees of membership, add a FCT and M number of FMTs, As indicated below, where M is the number of distinct membership degrees.



ii. Different Linguistic Conditions for a Target Attribute

An instance of a target attribute may belong to multiple linguistic terms as professional users can have more than ones of classes to which an instance of target attribute may belong i.e. multiple fuzzy classes and multiple MFs. Formally

TA—instance (1): Fuzzy classes (M)

Guideline 3. If an instance of a TA belongs to multiple fuzzy classes, but with the same membership degree, add M number of FCTs and a FMT, where M is the number of distinct linguistic terms.

Guideline 4. If an instance of a TA belongs to more than one fuzzy category with different membership degrees, add an umber M of FCTs and FMTs, as indicated below, where M is the number of distinct linguistic terms, and one FCT is linked to one FMT at the most.

4. Meta Model and Method for Modelling a Fuzzy DWH

4.1. Meta Model

According to Harel et al. Al. [16], a meta model defines the elements of a conceptualization, as well as their relationships. Figure 1 shows the meta model of the proposed fuzzy DWH in which the right side shows the meta model of the classical DWH. The left side shows how the fuzzy concepts are integrated with a classical DWH as a structure of meta tables.

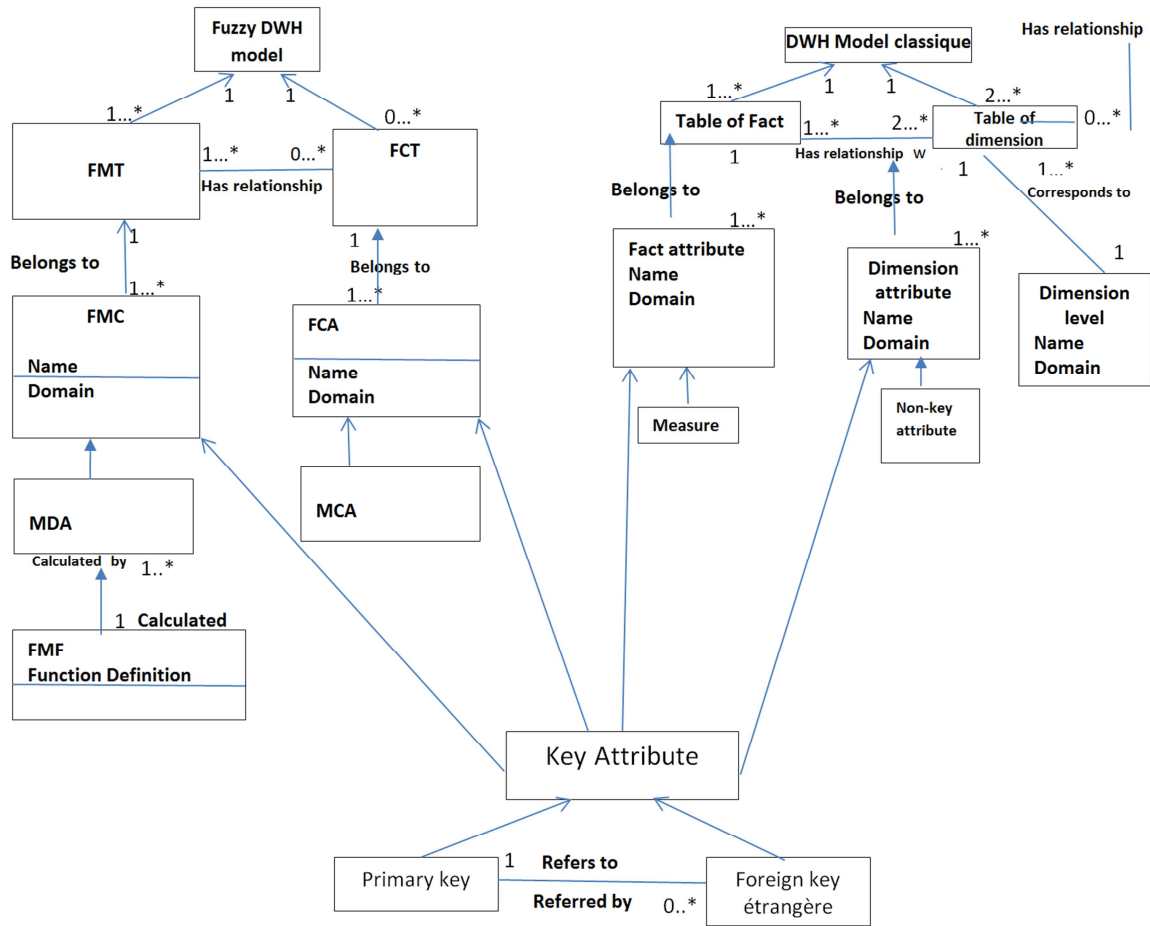


Figure 1. Meta model of fuzzy DWH.

The model class of a DWH in the meta model refers to a DWH schema which is made up of one or more fact tables, and two or more dimension tables. A fact table is located at the centre of a DWH model and it essentially captures the commercial measures of the process (Kimball [17]). The relationship with dimension tables is realized with the help of fact attributes. A fact attribute could be a measure (also called fact) or a key attribute (primary or foreign key). A measure (a sub-class of the fact attribute) captures critical values of a business process i.e. where a set of key attributes are used to capture the relationship with dimension tables.

In a classical DWH, two or more dimension tables surround a fact table. A fact table can also be linked to one or more other dimension tables to form hierarchies. In this case, each dimension table is at a different level of hierarchy (in order to comply with the snow flake schema). The level of hierarchy is referred to by the class of the dimension level in the meta model. A dimension table contains some dimension attributes which represent the category attributes or key attributes of a dimension table. The key attributes capture the relationship between dimension tables, respectively between fact and dimension tables. Moreover, other non-key attributes characterize the category attributes of a dimension table [19].

The model class of fuzzy DWH considered in the meta model refers to the fuzzy concept integrated with in a DWH. For each target attribute identified a model of fuzzy DWH

can be added. Thus, a classical model of DWH can have more than one fuzzy DWH model. Fuzzy concepts may exist without a linguistic variable. These fuzzy concepts are represented by fuzzy DWH models without a FCT.

Each FCT has a relation to one or more FMTs. Therefore, the fuzzy DWH model is made up of one or more FMTs and zero or more FCTs.

AFMT is built of FMAs. AFMA might be a key attribute to denote primary key or foreign keys. A second type of FMA is the MDA.

The instances of the MDA are calculated by the FMFs of the fuzzy DWH model.

The FCT contains the fuzzy CMAs which can be, similarly to FMAs, key attributes. Furthermore, the fuzzy class membership attributes can be a class membership attribute that describes the linguistic term of the fuzzy concept.

4.2. A Method of Modelling a Fuzzy Data Warehouse

In order to create a fuzzy DWH, a method is shown that guide the translation of a warehouse of a crisp DWH into a fuzzy DWH. The input of the method is a classical DWH and the output is a fuzzy DWH. The process is broken down into two phases: in the first phase, the first phase elements of classification are defined and in the second phase, the fuzzy DWH is built. Figure 2 shows the tasks and order in which they are carried out.

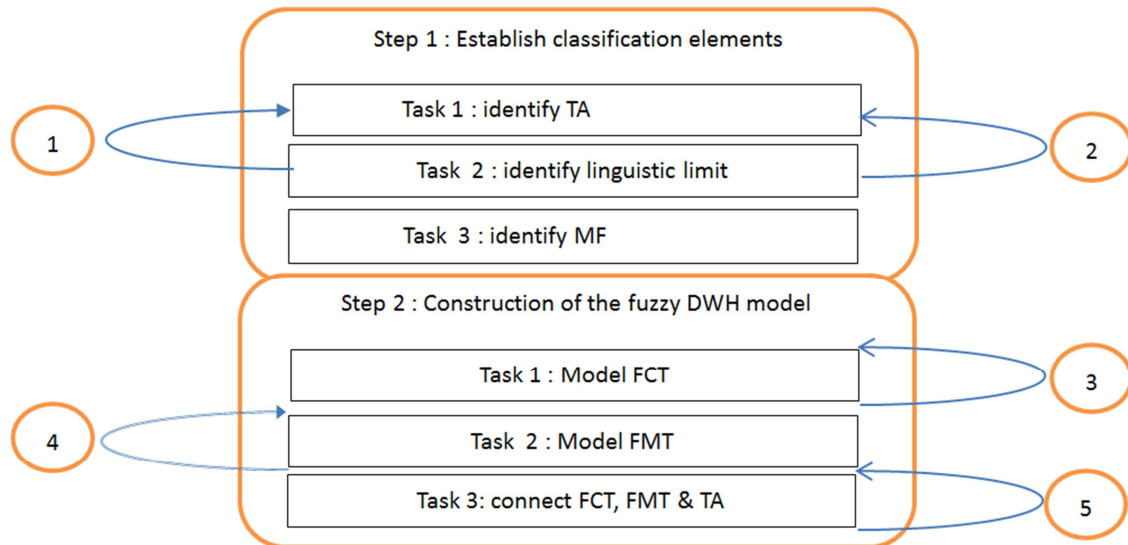


Figure 2. Graphical overview of the modelling method of a fuzzy DWH.

To illustrate the different steps of the method, the following example is used:

Example 1. ADWH contains the dimension product. With only one hierarchy, the unique level dimension containing a table product. Each product has the following attributes: id_product, name, date_expiration and price. From the unit price attribute, we can calculate the product price using the unit price multiplied by the quantity. Fig. 3 shows the dimension product.

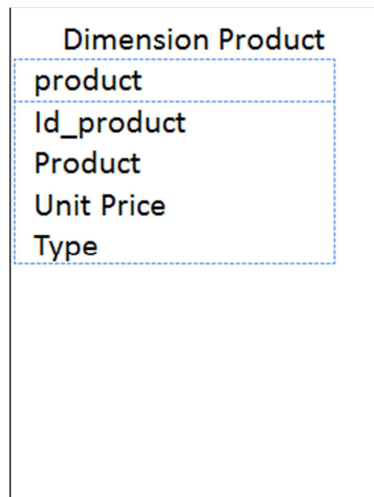


Figure 3. Dimension Product.

The purpose of this stage is to define the elements of classification which are used in the second stage to build the fuzzy DWH model. It involves three tasks: identify the target attribute, identify the linguistic terms and define the membership functions. Here are the details:

First task: This task consists in identifying what has to be classified i.e. the TA which contains the values destined to be classified fuzzily.

This will be done in a way that takes into account the input

of the end user. In the simplest, a TA is identified. For Example 1, consider product price as a TA.

Second task: It consists in determining how the values of the identified TA should be classified i.e. identifying these linguistic terms which are used to classify the instances of a TA. Repeat this task for all TAs. It is showed by Loop-through 1 in Figure 2. There are two possibilities:

Case 1—Distinct linguistic conditions: it is the simplest case in which the linguistic terms are distinct i.e. there is only one set of linguistic terms. Formally,

TA—instance (1): Fuzzy classes (1)

For the example of the product price, let us consider the set of linguistic terms for {price_high; price_average; prix_low}.

Case 2 —Different linguistic terms for a TA: it is a case where there are more than ones of linguistic terms to classify the TA instances. In this case, instances of the TA belong to more than one linguistic term, as identified by professional users. Formally,

TA—instance (1): fuzzy classes (M)

For the example on the product prices, consider that the following sets of linguistic terms are identified. These sets are {price_low; price_average; price_high } and { cheaper; cheap; expensive}. The linguistic terms might already exist in a classical DWH modelling form of instances of a dimension category. In that case, these terms can be used for classifying the cases of TA.

Third task: it consists in defining a MF (denoted μ) for each linguistic term. It is done in such way that the values can be determined over a scale of 0 to 1. Repeat the task for each identified linguistic term. It is showed by Loop-through 2 in Figure 2. It could be the case only for different users i.e. a TA belongs to the same set of linguistic terms with different membership degrees. The case is as follows:

The case is as follows:

Case3— Different MDs for the same linguistic terms: it is a case in which an instance of a TA belongs to a linguistic term with different membership degrees. It can be due to the fact

that multiple business users have different interpretations of a single instance of a TA i.e. the multiple MFs are used for a TA. Formally,

TA—instance (1): fuzzy classes (1)
But with different membership degrees.

Below, we discuss the examples of the third task for each case:

Example of Task 3, Case 1—Distinct linguistic terms. For the product price example, a MF is defined for each linguistic term. The MFs μ_{low} , $\mu_{average}$, μ_{high} become:

$$\mu_{low}(product_price) = \begin{cases} \text{if } product_price \leq 500, MD_PriceGroup = 1 \\ \text{if } product_price \geq 3000, MD_PriceGroup = 0 \\ \text{else} \quad ,DA_PriceGroup = \frac{3000-product_price}{3000-500} \end{cases}$$

$$\mu_{average}(product_price) = \begin{cases} \text{if } product_price \leq 500, DA_PriceGroup = 0 \\ \text{if } product_price \geq 5000, DA_PriceGroup = 0 \\ \text{if } 2000 \leq product_price \leq 3000, DA_PriceGroup = 1 \\ \text{if } 500 < product_price < 2000, DA_PriceGroup = \frac{product_price-500}{2000-500} \\ \text{else} \quad ,DA_PriceGroup = \frac{5000-product_price}{5000-3000} \end{cases}$$

$$\mu_{high}(product_price) = \begin{cases} \text{if } product_price \leq 3000, MD_PriceGroup = 0 \\ \text{if } product_price \geq 5000, MD_PriceGroup = 1 \\ \text{else} \quad ,DA_PriceGroup = \frac{product_price-3000}{5000-3000} \end{cases}$$

Example of Task 3 for case 2 – Different linguistic terms for a TA. For case 2 of the example of the product price ,a MF is defined for each linguistic term i.e. μ_{low} , $\mu_{average}$, μ_{high} , $\mu_{cheaper}$, μ_{adult} , μ_{old} . The MFs μ_{low} , $\mu_{average}$, μ_{high} are the same as above and $\mu_{cheaper}$, μ_{cheap} , μ_{cher} become:

$$\mu_{low}(product_price) = \begin{cases} \text{if } product_price \leq 500, MD_PriceGroup = 1 \\ \text{if } product_price \geq 3000, MD_PriceGroup = 0 \\ \text{else} \quad ,MD_PriceGroup = \frac{3000-product_price}{3000-500} \end{cases}$$

$$\mu_{average}(product_price) = \begin{cases} \text{if } product_price \leq 500, MD_PriceGroup = 0 \\ \text{if } product_price \geq 5000, MD_PriceGroup = 0 \\ \text{if } 2000 \leq product_price \leq 3000, MD_PriceGroup = 1 \\ \text{if } 500 < product_price < 2000, MD_PriceGroup = \frac{product_price-500}{2000-500} \\ \text{else} \quad ,MD_PriceGroup = \frac{5000-product_price}{5000-3000} \end{cases}$$

$$\mu_{high}(product_price) = \begin{cases} \text{if } product_price \leq 3000, MD_PriceGroup = 0 \\ \text{if } product_price \geq 5000, MD_PriceGroup = 1 \\ \text{else} \quad ,MD_PriceGroup = \frac{product_price-3000}{5000-3000} \end{cases}$$

Example of Task 3, Case 3– Different membership degrees for the same linguistic terms. For Case 3 of the product price example, a MF for each linguistic term becomes:

$$\mu_{low}(product_price) = \begin{cases} \text{if } product_price \leq 500, MD_PriceGroup = 1 \\ \text{if } product_price \geq 3000, MD_PriceGroup = 0 \\ \text{else} \quad , MD_PriceGroup = \frac{3000-product_price}{3000-500} \end{cases}$$

$$\mu_{average}(product_price) = \begin{cases} \text{if } product_price \leq 500, MD_PriceGroup = 0 \\ \text{if } product_price \geq 5000, MD_PriceGroup = 0 \\ \text{if } 2000 \leq product_price \leq 3000, MD_PriceGroup = 1 \\ \text{if } 500 < product_price < 2000, MD_PriceGroup = \frac{product_price-500}{2000-500} \\ \text{else} \quad , MD_PriceGroup = \frac{5000-product_price}{5000-3000} \end{cases}$$

$$\mu_{high}(product_price) = \begin{cases} \text{if } product_price \leq 3000, MD_PriceGroup = 0 \\ \text{if } product_price \geq 5000, MD_PriceGroup = 1 \\ \text{else} \quad , DA_PriceGroup = \frac{product_price-3000}{5000-3000} \end{cases}$$

In order to connect the product price with the same linguistic terms using another MD, we define another MF for each linguistic term i.e. μ_{low1} , $\mu_{average1}$, μ_{high1} . The definitions of the MF become:

$$\mu_{low1}(product_price) = \begin{cases} \text{if } product_price \leq 1750, MD_PriceGroup = 1 \\ \text{if } product_price \geq 3500, MD_PriceGroup = 0 \\ \text{else} \quad , DA_PriceGroup = \frac{3500-product_price}{3500-1750} \end{cases}$$

$$\mu_{average1}(product_price) = \begin{cases} \text{if } product_price \leq 3000, MD_PriceGroup = 0 \\ \text{if } product_price \geq 7000, MD_PriceGroup = 0 \\ \text{if } 3500 \leq product_price \leq 4000, MD_PriceGroup = 1 \\ \text{if } 3000 < product_price < 3500, MD_PriceGroup = \frac{product_price-3000}{3500-3000} \\ \text{else} \quad , MD_PriceGroup = \frac{7000-product_price}{7000-4000} \end{cases}$$

$$\mu_{high1}(product_price) = \begin{cases} \text{if } product_price \leq 4000, MD_PriceGroup = 0 \\ \text{if } product_price \geq 7000, MD_PriceGroup = 1 \\ \text{else} \quad , MD_PriceGroup = \frac{product_price-4000}{7000-4000} \end{cases}$$

5. Conclusion

DWH systems are used for the analysis of company performance. A potential of a classical DWH is that the numerical values of a DWH can be difficult to interpret by professional users, or can be interpreted in a wrong way.

For the precise comprehension of numerical values, professional users require an interpretation in terms which

are significant but non numerical.

However, if classification between the linguistic terms is strong, the true values cannot be measured. The solution is the use of a fuzzy – founded representation. The fuzzy representation allows the integration of fuzzy concepts with in the dimensions and the facts, while preserving the clear data structures and that leads to a fuzzy modelling.

In our study, we have applied fuzzy concepts at the level

of a dimension hierarchy. The FMT provides the relation to the dimension table in which the dimensional attribute resides. An example of this representation is carried out, modelled when applied in a DWH of a product price hierarchy of the dimension product.

General Abbreviations

CMA	Class Membership Attribute
DWH	Data Warehouse
FCT	Fuzzy Classification Table
FMA	Fuzzy Membership Attribute
FMT	Fuzzy Membership Table
FTN	Fuzzy Triangular Number
MD	Membership degree
MDA	Membership Degree Attribute
MF	Membership function
MG	Membership Grade
MGA	Membership Grade Attribute
TA	Target attribute

References

- [1] Zadeh, L. A. (1973). The Concept of a Linguistic Variable and its Application to Approximate Reasoning-1, *Informations Sciences* Vol.8/3, p.199-249.
- [2] Dubois, D., Prade, H.(1985). Théories des possibilités, Applications à la représentation des connaissances en informatique, Masson, 2^e édition. Academic Press, NewYork.
- [3] Zadeh, L. A. (1978). Fuzzy Sets as a Basis for a Theory of Possibility, *Fuzzy Sets and Systems* Vol.1, 3-28.
- [4] Buckley, J. J. (1989). Solving Possibilistic Linear Programming Problems, *Fuzzy Sets and Systems* Vol. 31, 329-341.
- [5] Kaufmann (1977). Introduction à la théorie des sous-ensembles flous, A l'usage des Ingénieurs, (Fuzzy Sets Theory), Tome 1, Éléments théoriques de base, Masson, Paris.
- [6] Klir, G. J., Folger, T. A. (1988). Fuzzy Sets, Uncertainty, and Information, Prentice-Hall, Englewood Cliffs, NJ.
- [7] Buckley, J. J. (1988a). Possibility and Necessity in Optimization, *Fuzzy Sets and Systems* Vol.25, 1-13.
- [8] Pedersen, T. B., Jensen, C. S., and Dyreson C. E.. Supporting Imprecision in Multidimensional Data bases Using Granularities. Eleventh International Conference on Scientific and Statistical Database Management, 1999.
- [9] Inmon W., Building the data warehouse, John Wiley & Sons, 1996.
- [10] Kimball R., The Data warehouse Toolkit, John Wiley & Sons, 1996.
- [11] Jarke M., Lenzerini M., Vassiliou Y., Vassiliadis P., Fundamentals of Data Warehouses, Springer-Verlag, 1998.
- [12] Sapir, L., Shmilovici A., and Rokach, L.. A Methodology for the Design of a Fuzzy Data Warehouse. In Intelligent Systems, 2008. IS'08. 4th International IEEE Conference, volume 1, 2008.
- [13] Inmon W., "The operational Data Store", White Paper, www.billinmon.com/library/whiteprs/earlywp/ttds.pdf, 2000.
- [14] Codd E., Codd S., Salley C., Providing OLAP(On-Line Analytical Processing) to User-Analysts: An IT Mandate, Report, Arbor Software White Paper, 1993.
- [15] L. A. Zadeh. The Concept of a Linguistic Variable and its Application of Approximate Reasoning-PartI. *Information Science*, (8): 199-249, 1975.
- [16] D. Hareland B. Rumpe. Meaningful Modeling: What's the Semantics of "Semantics"? Computer, 37(10): 64-72, October 2004.
- [17] Ralph Kimball and Joe Caserta. The Data Warehouse ET L Tool kit. Wiley Publishing, Inc., 2004.
- [18] K. V. N. N. Pavan Kumar, P. Radha Krishna, and Supriya Kumar The Fuzzy OLAP Cube for Qualitative Analysis. In Intelligent Sensing and Information Processing, pages 290-295, 2005.
- [19] Heiko Schepperle, Andreas Merkel, and Alexander Haag. Erhalt von Imperfektion in einem Data Warehouse. Internationales Symposium: Data-Warehouse-Systeme und Knowledge - Discovery, 2004.