ACL-Scale as a Tool for Preprocessing of Many-Valued Contexts

Tatiana Afanasieva, Nadejda Yarushkina, Gleb Guskov

Ulyanovsk State Technical University, Information System, Ulyanovsk, Russia
tv.afanasjeva@gmail.com, jng@ulstu.ru,guskovgleb@gmail.com

Abstract. One of the formal technique in Data mining is Formal Concept Analysis (FCA). During preprocessing of a many-valued context many applications of FCA require the partitioning of numerical data attributes into some smaller intervals. Designation of such numerical intervals with linguistic terms without domain experts will help researchers to understand attributes and their dependencies better. To solve this task we propose the notion of a special ACL-scale, which can be considered as a linguistic variable with ordered linguistic terms, modeled by fuzzy sets. The notion of ACL-scale, algorithms of its creation and application are presented. The example how many-valued context can be transformed into formal context using ACL-scale is shown in the paper. The main contribution is a new uniform tool for preprocessing of numerical attributes of given tables which simplify their transformation into a formal context with linguistic attributes.

Keywords: data mining, data preprocessing, ACL-scale, formal context, linguistic values

1 Introduction

One of the formal techniques in Data Mining and Knowledge Discovery in Databases (DM&KDD) process for extraction and representation of useful information, of objects (attributes) and of data dependencies is the Formal Concept Analysis (FCA) [1,2]. The first steps in applying of FCA is data preprocessing, where a many-valued context has to be transformed into a formal context to represent a data table with values of suitable granularity. When the input values are numerical, they have to be partitioned into numerical intervals. There are three main approaches to do this transformation, based on scaling theory. The conceptual scaling approach is well established and it uses conceptual scales [3,4] to derive a formal context. Logical scaling was introduced in [5] as a method using some expert knowledge to transform given data into the data from which conceptual hierarches can be explored. The fuzzy scaling approach beeing considered for example in [6,7,8] applies the notion of a linguistic variable [9]. The latter adds information to the structure of a formal context and can give linguistic description of numerical values of attributes and their dependences. The comparison of conceptual and fuzzy scaling theories for FCA was considered in [6]. The different approaches to embed fuzzy logic into FCA and application in KDD are given in [10]. The authors described the most important theories connected with fuzzy attributes, fuzzy concepts and fuzzy concept lattice.

The main problems in applying fuzzy scaling theory to FCA were discussed in [11] and some solutions were presented. One of the problems the author mentioned was the problem of using and interpreting the membership functions in FCA, so the short alternative conceptual description of fuzziness without using membership functions was given in [11].

In this paper we propose the approach for transforming numerical attributes of a many-valued context into linguistic variables. This transformation is considered as preprocessing based on the fuzzy scaling theory, where the membership functions are used to derive linguistic values of the partitions of numerical attributes only. The advantage of this approach is a linguistic granulation of numerical attributes in a many-valued context. This linguistic granulation can be useful in segmentation of objects with similar features. Mining the dependencies among several objects expressed in linguistic terms is another application of that linguistic granulation. To solve this task we propose the notion of a special Absolute & Comparative scale (ACL-scale). Using ACL-scale the partitions of numerical data and their linguistic descriptions can be derived. Therefore, the formal context can be presented in a traditional form, and well-known algorithms for FCA can be applied without computing of membership functions.

2 Problem Definition

Here we recall the definition of many-valued context [12] in respect to attributes *m* having numerical values *w*.

Definition 1. A many-valued context $K = (G, M, W, J)$ is a set of objects G, a set of attributes M, a set of possible values W, and a ternary relation $J \subseteq G \times M \times W$, with

$$(g, m, w) \in J, (g, m, v) \in J \Rightarrow w = v,$$

where $(g, m, w) \in J$ indicates that object g has the attribute *m* with value *w*. In this case, we also write $m(g) = w$, regarding the attribute *m* as a partial function from G to W.

Definition 2. A formal context is a triple $C = \langle G, Y, I \rangle$ where G is a set of objects, Y is a set of attributes and $I \subseteq G \times Y$ is a binary relation between G and Y. For $\langle g, y \rangle \in I$ it is said "The object g has the attribute y".

The task is to transform given many-valued context into a formal context. We denote this transformation as $K \Rightarrow C$.

Each value $y \in Y$ is a linguistic value (some linguistic description of a numerical value w), derived by scaling. This means that for each attribute $m \in M$ on the set of its possible numerical values W a special scale has to be defined and then applied to transform a given numerical value w into a linguistic value y. Therefore we consider a task of a scale construction for each attribute $m \in M$ on the set of its possible numerical values W. The main demands for this scale construction are simple adaptation to a set of numerical values W and minimizing of an expert participation. To solve this task the scale must be formed in automatic way using uniform quantity of parameters and of operations. Beside that the scale must be considered as a linguistic variable to associate its linguistic terms to the scaling values.

So, the problem is to denote the notion of a special scale, which satisfies the mentioned above demands, and algorithms of its construction and its application. Application of this special scale will allow to decrease preprocessing time of a transformation of a given many-valued context into a formal context using uniform formal tool.

3 Notion of an ACL-scale

In this section we propose a special scale, named an ACL-scale (Absolute & Comparative scale) to do the transformation of given many-valued context into a formal context.

Let $\{x_i \in W, W \subseteq \mathbb{R}, i = 1,2, \ldots, n\}$ be the set of possible ordered values of a numerical attribute *m* in respect to definition 1.

We assume that the binary relation $x \leq y$ is defined possessing the following properties:

- reflexivity: $x \leq x, \forall x \in W$.
- transitivity: if $x \leq y$ and $y \leq z$, then $x \leq z, \forall x, y, z \in W$.
- anti-symmetry: if $x \leq y$ and $y \leq x$, then $x = y, \forall x, y \in W$.

Let suppose several partially ordered intervals of equal length cover a set W and they are used for building a linguistic variable $\tilde{X}$ with fuzzy terms $\tilde{x}_k = \{x_i, \mu_{\tilde{x}_k}(x_i)\}$, $x_i \in W$, $\tilde{x}_k \in \tilde{X}$, $i = 1,2, \ldots, n, k = 1,2, \ldots, r, r < n$. Here $\mu_{\tilde{x}_k}(x_i), i = 1,2, \ldots, n$ denotes the membership function of a fuzzy term with a linguistic value $\tilde{x}_k$. Therefore it can be said that a set of linguistic values covers a set W. Each linguistic value $\tilde{x}_k \in \tilde{X}$ can be considered as an ordered gradation of a scale and as linguistic estimation of every numerical value with some truth value.

Definition 3. ACL-scale for an attribute *m* with possible numerical values from the set *W* is an algebraic system

$$ACL = \{H, \Psi, \Omega\},$$

where the set $H = \{W, \tilde{X}\}$ denotes possible numerical values and possible fuzzy terms for an attribute m; $\Psi = \{nmin, nmax, r, MF\}$ is a set of parameters of an ACL-scale; $\Omega = \{Fuzzy, DeFuzzy\}$ is a set of operations, defined on a set H.

Below the components Ψ and Ω of an ACL-scale will be considered in details.

3.1 Parameters of an ACL-scale

Parameterization of an ACL-scale is useful as a tool for domain specific adaptation.
To adopt an ACL-scale to real values of a set *W* we consider two alternatives. The first one corresponds to the case when experts evaluate quantaty, parameters and shape of membership functions of linguistic variables $\tilde{X}$. Unfortunately this case is difficult to realize in practice. In the second alternative the goal is to minimize the work of expert and some algorithm is used to adopt an ACL-scale to real values of a set *W*. We apply the second alternative and consider four parameters of an ACL-scale adaptation:

$$\Psi \ = \{nmin, nmax, r, MF\}, \tag{1}$$

where $nmin \ = \ inf(W), nmax \ = \ sup(W)$; MF is the uniform shape of the membership functions of fuzzy terms (for example in a triangular form) [13]; r is the quantity of fuzzy terms, *r*+1 is the quantity of numerical intervals of equal length *d*, used for membership functions construction:

$$[x - d, x] \subset \ W, \qquad d = \frac{nmax - nmin}{r + 1}. \tag{2}$$

Notice, that these intervals are the result of partitioning of the set *W* and any numerical value $w \ \in \ [x - d, x]$ is considered according to an ACL-scale as identical, with the same linguistic value, but having different truth degree. According to (2) the length of numerical intervals *d* depends on quantity of fuzzy terms.
In this case researcher must define the shape and the quantity of fuzzy terms r . Parameter *r* determines a quantity of numerical intervals and their length *d*. It means that parameter *r* determines a level of linguistic granulation: smaller value of parameter *r* corresponds to larger linguistic granulation and vice versa. Therefore the quantity of fuzzy terms *r* depends on research goals and required level of granulation. Taking into account human perception the recommendation for choosing the value of parameter *r* are: $3 < r < 10$.
The example of ACL-scale for a numerical attribute *m* with possible values defined in $W = [-26, 66]$ is shown on the Figure 1. Here partitioning into six ordered intervals was done, on which five triangular fuzzy terms (*r*=5) were constructed with linguistic values $\tilde{X} = \{A_{0-1}, A_0, A_1, A_2, A_{2+1}\}$.

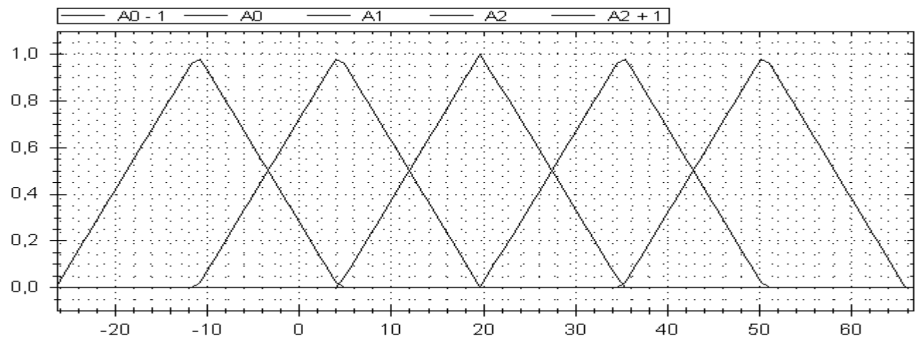

Fig. 1. Example of an ACL-scale

We assume that the following is fulfilled for an ACL-scale:

1. The numerical values *w* of attributes *m* corresponding to real or ideal objects are estimated.
2. Numerical and linguistic estimates are various, but they are equally essential aspects at the different levels of granularity.
3. Linguistic values of numerical attributes can be estimated by expert or a modeling estimation procedure.

The usage of parameters of an ACL-scale for linguistic description of numerical attributes allows to determine linguistic values practically in an automatic way, better understood by researchers.

3.2 The operations of an ACL-scale

The set of operations, defined on a set *H*, can be based on fuzzified/defuzzified functions. The operation $Fuzzy$ for linguistic description of each numerical value is defined as the following function:

$$\tilde{x}_i = \tilde{x}_s, if\ \tilde{x}_s(x_i) \geq \tilde{x}_j(x_i), s \in \{1, 2, \ldots, r\}, \forall j = 1, 2, \ldots, \mathrm{r}. \quad (3)$$

In respect to (3) for every $x_i \in W$ there will be only one linguistic value $\tilde{x}_s \in \tilde{X}$ with the maximum value among all of membership functions, *s* – is a number of that membership function.

We denote the operation $deFuzzy$ for numerical estimation of linguistic value as function $x_i' = DeFuzzy(\tilde{x}_i), x_i \in W, \tilde{x}_i \in \tilde{X}$, for example, as centroid of area:

$$x_t' = \frac{\int_{nmin}^{nmax} x \cdot \tilde{x}(x)dx}{\int_{nmin}^{nmax} \tilde{x}(x)dx}.$$

It is obvious, that $DeFuzzy$ function calculates approximate value with some error of estimation, and the latter can be computed in different ways, for example in a form:

$$Er_{x_i} = |x_i' - x_i|,$$

where the approximate value is $x_i' = DeFuzzy(\tilde{x}_i)$; x_i is the actual numerical value of some attribute.

The usage of uniform scaling by an ACL-scale will allow to transform given many-valued context into a formal context in automatic way and to explore the concepts having linguistic values which are better understood by researchers.

4 Transformation of numerical values into linguistic ones using an ACL-scale

The transformation of a numerical value $x_i \in W, i = 1,2, \dots, n$ into a linguistic value $\tilde{x}_i \in \tilde{X}$ with an ACL-scale means, that it is possible to define several fuzzy terms $\tilde{x}_j(x_i), j = 1, 2, \dots, \mathrm{r}$ with different truth degree for $\forall x_i$.

Let $W \subseteq \mathbb{R}$ be a set of possible numerical values of an attribute.

First of all, it is required to construct an ACL-scale on the set W, containing the ordered fuzzy terms with linguistic values $\tilde{x}_k \in \tilde{X}$, $k = 1,2, \dots, r$.

Below we propose the Algorithm 1 for an ACL-scale creation by the determining its parameters on the set of possible numerical values *W* of a many-valued context.

Algorithm 1.

Step 1. Define the parameter *r* (the number of fuzzy terms) of ACL-scale.

Step 2. Compute the parameter *nmin* as the minimum value on a set of *W*.

Step 3. Compute the parameter *nmax* as the maximum value on a set of *W*.

Step 4. Order the possible values on W. Partition the ordered set of possible values $W \subseteq \mathbb{R}$, into *r*+1 intervals in respect to (2).

Step 5. Define the shape of the membership functions *MF* of fuzzy terms. Determine the linguistic values of fuzzy terms $\tilde{x}_k \in \tilde{X}$, $k = 1,2, \dots, r$.

To output the linguistic values for the numerical values of the set *W*, using an ACL-scale, Algorithm 2 is proposed.

Algorithm 2.

For each numerical value $x_i \in W, i = 1,2, \dots, n$ do the following:

Step 1. Using operation $Fuzzy$ (3) and well-known notion of fuzzy terms of chosen shape (for details you can see [13]) compute the values of their membership functions $\tilde{x}_k = \{x_i, \mu_{\tilde{x}_k}(x_i)\}, \tilde{x}_k \in \tilde{X}$, $k = 1,2, \dots, r$.

Step 2. Determine the fuzzy term $\tilde{x}_s(x_i)$ with the maximum value of membership function according to (3).

Step 3. Assign the output linguistic value as $\tilde{x}_i = \tilde{x}_s$ for input x_i Here *s* is the number of linguistic value on the set $\tilde{X}$, corresponding to an ACL-scale for the set of numerical values W.

5 Example

To illustrate how the ACL-scale can be applied to transform a many-valued context into a formal context we use the input data, which characterize hardware by two at-

tributes x_{cpu} ="Load of the central processor - CPU" and x_{ram} ="Load of the memory - RAM" (see Table 1).
We created one ACL-scale using the Algorithm 1 for both attributes, as their numerical values are contained in the same set of possible numerical values [0,100] presented in percentage. For this domain we defined $nmin = 0\%, nmax = 100\%$. Then seven fuzzy terms ($r = 7$) with linguistic values "very low", "low", "below an average", "average", "above an average", "high", "very high" were defined.

Table 1. Input many-valued data

id_obiect	x_{cpu}, %	x_{ram}, %
1	84,31	82,94
2	50,67	58,93
3	66,89	68,18
4	97,06	77,56
5	92,04	33,58
6	97,33	93,42
7	97,44	94,78
8	88,30	80,05
9	66,64	48,49

The shape of membership function was chosen as triangular with parameters shown in Table 2 (a - left, c - right, b - middle of numerical interval on which membership function is build).

Table 2. The parameters of membership functions of fuzzy terms in the form of triangular fuzzy number for attributes of hardware

	Linguistic values	The parameters of membership functions		
		a	***b***	***c***
x_{ram} x_{cpu}	very low	0	0	16,5
	low	0	16,5	33
	below an average	16,5	33	50
	average	33	50	66,5
	above an average	50	66,5	83
	high	66,5	83	100
	very high	83	100	100

After an ACL-scale has been created, it was used to output the linguistic value for every numerical value of the hardware attributes, applying the Algorithm 2. Table 3 illustrates the results of transformation of input data (see Table 1) into linguistic values.

Table 3. The results of linguistic estimation of the numerical values of the hardware attributes, using ACL-scale

id_obiect	**linguistic values x_{cpu}**	**linguistic values x_{ram}**
1	high	high
2	average	above an average
3	above an average	above an average
4	very high	high
5	high	below an average
6	very high	very high
7	very high	very high
8	high	high
9	above an average	average

Table 4 presents the formal context with linguistic values of hardware numerical attributes (here vl =”very low”, lo= “low”, ba = “below an average”, av = “average”, aa=”above an average”, hi=”high”, vh =”very high” for short).

Table 4. The formal context for a many-valued data derived by ACL-scale

id_obiect	x_{cpu}							x_{ram}						
	vl	lo	ba	av	aa	hi	vh	vl	lo	ba	av	aa	hi	vh
1						x							x	
2				x								x		
3					x							x		
4							x						x	
5						x				x				
6							x							x
7							x							x
8						x							x	
9					x						x			

The results in Table 4 show the transformation of the numerical attributes of a many-valued context (see Table 1) into linguistic variables for more understandable description of these attributes, which can be used for mining dependencies or for clustering. For further analysis the additional characteristics of a linguistic value of attributes are useful: the truth degree and the membership function.

6 Conclusion

During the past years preprocessing became an important step of data mining. For better understanding and analyzing numerical data, it is useful to have their linguistic description. To derive the latter description the transformation tecniques based on scaling are used usually.

In this paper the notion of an ACL-scale as the tool for transformation a many-valued context with a numerical attributes into a formal context with linguistic attributes is proposed. The algorithm of an ACL-scale creation by adaptation of its parameters on a set of numerical values is described. Application of an ACL-scale provides the linguistic granulation which can be useful in segmentation and investigation of objects with similar features. Mining the dependencies among attributes and among several objects expressed in linguistic terms is another application of that linguistic granulation. In these tasks time reduction on preprocessing stage will be obtained due to usage of the proposed uniform scaling algorithm for different numerical attributes.

The given example shows applicability and suitability of an ACL-scale for the preprocessing of a many-valued context with numerical attributes and deriving formal context with linguistic values.

7 Acknowledgements

The authors acknowledge that this paper was partially supported by the project no. 2014/232 of the Ministry of Education and Science of Russian Federation "Development of New Approach to the Intellectual Analysis of Information Resources" and by the project no. 16-07-00535 "Development and research of data mining algorithms for organizational and technical systems based on fuzzy models" of the Russian Foundation of Basic Research.

References

1. Poelmans, J., Ignatov, D.I., Kuznetsov, S.O., Dedene G.: Formal Concept Analysis in knowledge processing: A survey on models and techniques. Expert Systems with Applications, 40(16), 6601–6623 (2013).
2. Kumar, Ch. A.: Knowledge discovery in data using Formal Concept Analysis and Random Projections. In Int. J. Appl. Math. Comput. Sci., 21, 4, 745–756 (2011).
3. Ganter, B., Wille, R.: Conceptual Scaling. In: F.Roberts (ed.): Applications of combinatorics and graph theory to the biological and social sciences,139-167. Springer-Verlag, New York, (1989).
4. Ganter, B., Wille, R.: Formal Concept Analysis. In Mathematical Foundations. Springer Verlag, Berlin (1999).

5. Prediger S.: Logical scaling in formal conceptual analysis. In D. Lukose et.al. (eds.): Conceptual Structures: Fulfilling Peirce's Dream. Proceedings of the ICCS'97, LNAI 1257, Springer, Berlin, 332-341 (1997).
6. Wolff, K. E. : Concepts in fuzzy scaling theory: order and granularity. In Fuzzy Sets and Systems, 132(1):63-75 (2002).
7. Belohlavek, R., Vychodil, V.: What is a fuzzy concept lattice? In Proc. of 3rd International Conf. on Concept Lattice and their Applications (CLA-2005):34-45 (2005).
8. Yan, W., Baoxiang, C.: Fuzzy Many-Valued Context Analysis Based on Formal Description. In Proc. of 8th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (2007).
9. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning. Memorandum ERL-M 411 Berkeley, October (1973).
10. Poelmans, J., Ignatov, D.I., Kuznetsov, S.O., Dedene G.: Fuzzy and rough formal concept analysis: a survey. In: International Journal of General Systems, 43, 2, 105-134 (2014).
11. Wolff, K. E.: Position Paper: Pragmatics in Fuzzy Theory. In Proc.of 13th International Conf. RSFDGrC, LNAI 6743:135-138 (2011).
12. Gugisch R.: Many-valued Context Analysis using Descriptions. In ICCS 2001, LNAI 2120, 157-168 (2001).
13. Zimmermann, H.J.: Fuzzy Set Theory and its Applications (Third edition). Boston/Dordrecht/London: Kluwer Academic Publishers (1996).