A MODIFICATION OF TAMURA'S PATTERN CLASSIFICATION METHOD

Viattchenin D.A., Klempach V., Yarkou Y.

Intelligent Information Technologies Department, Belarusian State University of Informatics and Radioelectronics, P. Brovka St. 6, 220013 Minsk, Belarus, e-mail: *viattchenin@mail.ru*

Abstract. A modification of Tamura's method of pattern classification based on fuzzy relation is proposed in the paper. Numerical experiments results are presented Comparative analysis of the original version and the method modification is made. Some conclusions are made too.

Introduction

Since the fundamental Zadeh [8] paper was published, fuzzy sets theory found oneself in many areas, in particular, in the cluster analysis. A fuzzy approach to clustering is very useful because majority of real clusters is fuzzy by their structure. The idea of an application of fuzzy sets approach to clustering problems was outlined by Bellman, Kalaba and Zadeh [1].

Heuristic methods of fuzzy clustering, hierarchical methods of fuzzy clustering and optimization methods of fuzzy clustering were proposed by different researchers. Very interesting and simply classification method was proposed by Tamura, Higuchi and Tanaka [6]. The method was investigated by Dunn [2]. The main purpose of the paper is consideration of a modification of the Tamura's method.

Tamura's pattern classification method and its modification

Let $X = \{x_1, \ldots, x_n\}$ is an initial set of elements and $R : X \times X \to [0,1]$ is some binary fuzzy relation on X and $\mu_R(x_i, x_j), \forall x_i, x_j \in X$ is its membership function.

Definition 1. The fuzzy tolerance is the fuzzy binary intransitive relation, which possesses the symmetric property

$$\mu_R(x_j, x_i) = \mu_R(x_i, x_j), \forall x_i, x_j \in X \tag{1}$$

and the reflexivity property

$$\mu_R(x_i, x_i) = 1, \forall x_i \in X\,. \tag{2}$$

This kind of fuzzy relations is indicated by T symbol.

Definition 2. (max-min)-transitive closure [3] of some fuzzy intransitive relation R on the set X where $card(X) = n$ is a fuzzy relation $\hat{R}$:

$$\hat{R} = R^1 \cup R^2 \cup \ldots \cup R^n, \tag{3}$$

where R^n relations defined as follows:

$$R^1 = R, R^n = R^{n-1} \circ R, n = 2,3,\ldots, \tag{4}$$

and $R \circ R$ is operation of (max-min)-composition of fuzzy relation R on the set X [3]: $\mu_{R \circ R}(x_i, x_k) = \vee_{x_j \in X}(\mu_R(x_i, x_j) \wedge \mu_R(x_j, x_k)), \forall x_i \in X, x_k \in X$.

Let's initial data presented by a matrix $r_{n \times n} = [\mu_T(x_i, x_j)], i = 1,\ldots,n, j = 1,\ldots,n$ of a fuzzy tolerance. Elements of the matrix are coefficients of objects similarity. So, Tamura's classification method can be presented as a next four-step procedure:

1 Construction of the transitive closure matrix $\hat{T}$ of initial fuzzy tolerance T following by (3) formula;
2 Calculating of α-levels values of $\hat{T}$ fuzzy relation, $\alpha \in [0,1]$;
3 Clusters $A^l, l \in [1,n]$ are selected for every value of $\alpha, \alpha \in [0,1]$ in conformity with the rule: if $\mu_{\hat{T}}(x_i, x_j) \geq \alpha$ for some $x_i, x_j \in X$, then the objects $x_i, x_j \in X$ are elements of A^l cluster; so, some partition $P = \{A^1, \ldots, A^c\}, 1 \leq c \leq n$ is correspond to some α-level value, $\alpha \in [0,1]$;
4 A partition $P^* = \{A^1, \ldots, A^c\}$ is selected from the partitions hierarchy on the basis clusters number c parameter, $2 \leq c \leq n$;

Thus, clusters $A^l, l \in [1,c]$ are crisp clusters because every cluster $A^l, l \in [1,n]$ is a α-level of corresponding fuzzy set. So, for monotone increasing sequence of thresholds, $0 \leq \alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_k \leq 1$, one obtains a corresponding α-level hierarchy of clusters. A result of classification is a partition matrix $P_{c \times n} = [\mu_{li}]$ where $\mu_{li} = \begin{cases} 1, x_i \in A^l \\ 0, x_i \notin A^l \end{cases}, l = 1, \ldots, c, i = 1, \ldots, n$.

However, crisp clusters cannot present real structure of initial set because typicalness degree is exists for some element $x_i \in X$ which is element of some cluster $A^l, l \in [1,c]$. Some value $\alpha, \alpha \in [0,1]$ corresponds to some partition $P = \{A^1, \ldots, A^c\}, 1 \leq c \leq n$. Thus, a value of α can be considered as a similarity threshold. Further, if some element $\tau^l = x_l \in X$ is a centre of A^l cluster, then membership function $\mu_T(x_i, \tau^l) = \mu_{li}$ is similarity degree of some point $x_i \in A^l$ with centre τ^l of A^l cluster. So, the membership function can be considered as the degree of association of an element $x_i \in X$ with the cluster A^l.

Thus, a modification of the Tamura's classification procedure can be presented as follows:

1 Construction of the transitive closure matrix $\hat{T}$ of initial fuzzy tolerance T following by (3) formula;
2 Calculating of α-levels values of $\hat{T}$ fuzzy relation, $\alpha \in [0,1]$;
3 Clusters $A^l, l \in [1,n]$ are selected for every value of $\alpha, \alpha \in [0,1]$ in conformity with the rule: if $\mu_{\hat{T}}(x_l, x_i) \geq \alpha$ for some $x_l, x_i \in X$, then the objects $x_l, x_i \in X$ are elements of A^l cluster where $A^l = \{(x_i, \mu_{li}) \mid \mu_{li} \geq \alpha, x_i \in X, l = 1, \ldots, c\}$;
4 A family of fuzzy clusters $\{A^1, \ldots, A^c\}$ is selected from the fuzzy clusters families hierarchy on the basis clusters number c parameter, $2 \leq c \leq n$;

Degrees of association of an element $x_i \in X$ with the cluster A^l are defined as $\mu_{li} = \begin{cases} \mu_{\hat{T}}(x_l, x_i), x_i \in A^l \\ 0, x_i \notin A^l \end{cases}$. So, every element of the family $\{A^1, \ldots, A^c\}$ is α-level fuzzy set in the Radecki [4] sense.

Numerical experiments results

Experimental results of the Tamura's method modification application to Tamura's real data of portraits similarity in comparison with results of original version of Tamura's classification procedure application are demonstrated. For this purpose, Tamura's data of portraits similarity is used. Let $X = \{x_1, \ldots, x_{16}\}$ is a set of portraits. Figure 1 presents the objects real assignment.

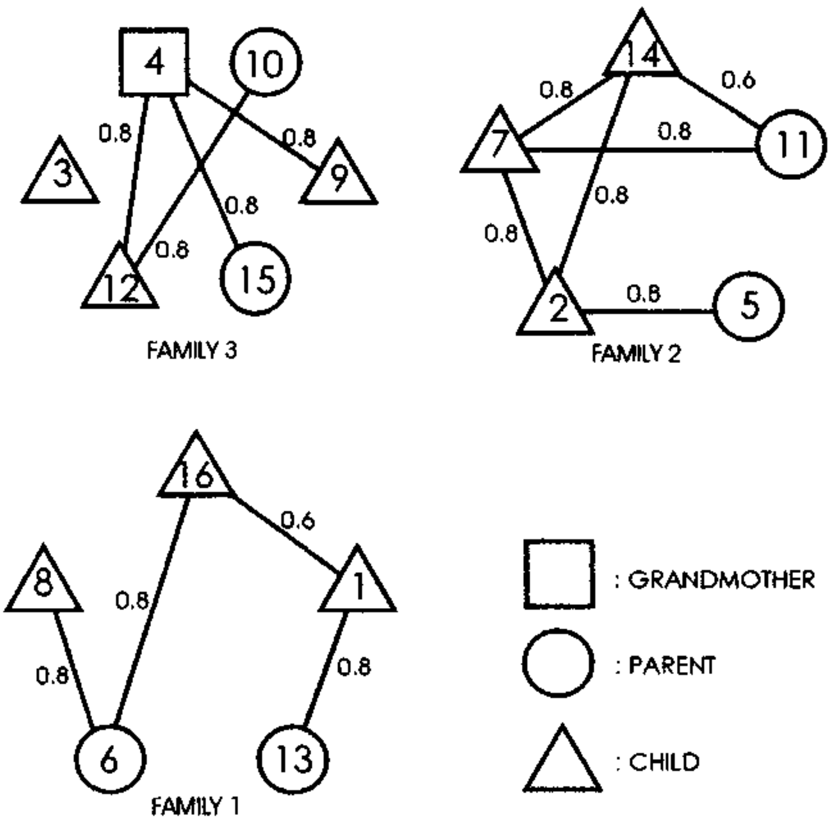

Fig. 1 The objects assignment

Results of the Tamura's method original version application for three classes to the data are illustrated by Table 1:

Class	Object															
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
1	1	0	0	1	0	1	0	1	1	1	0	1	1	0	1	1
2	0	1	0	0	1	0	1	0	0	0	1	0	0	1	0	0
3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 1 Matrix of the crisp portrait assignment

The crisp partition received for similarity threshold value $\alpha = 0.5$. Results of the Tamura's method modification application for three classes to the data can be illustrated by Table 2:

Class	Object															
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
1	1.0	0.0	0.0	0.5	0.0	0.6	0.0	0.6	0.5	0.5	0.0	0.5	0.8	0.0	0.5	0.6
2	0.0	1.0	0.0	0.0	0.8	0.0	0.8	0.0	0.0	0.0	0.8	0.0	0.0	0.8	0.0	0.0
3	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 2 Matrix of the fuzzy portrait assignment

These results can be presented by a linear diagram which is depicted on a Figure 2.

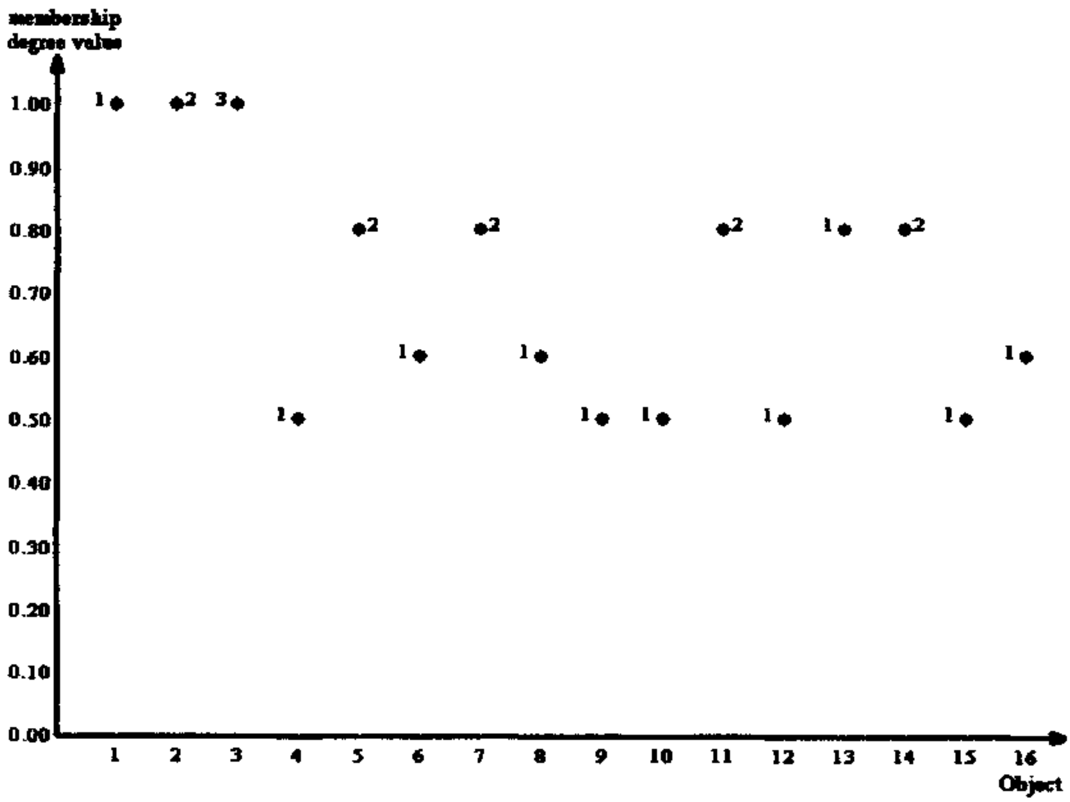

Fig. 2 The linear diagram of objects assignment

Summary

In both cases, the second class is corresponds to second family. However, first family and third family are originates one class and third object is originates separate class. In general, these results are corresponds to real objects assignment approximately. An interpretation of membership function as a degree of object association to class is main property of the Tamura's classification method modification.

Notable, that the Tamura's classification method modification has some peculiarities. Firstly, the objects assignment is not satisfied Ruspini's fuzzy partition condition [5], because a condition $0 < \sum_{l=1}^{c} \mu_{li} \leq 1, i = 1,\ldots,n, l = 1,\ldots,c$ is met. Secondly, if the fuzzy tolerance is a powerful fuzzy tolerance in the Viattchenin's sense [7], then the modification can be applied to problems classification solving fruitfully.

Thus, the modification is of Tamura's classification method can be used for preliminary clusters centers determination and preliminary consideration of the objects assignment. This property of the method modification can be used for initial fuzzy partition constructing for some optimization fuzzy clustering method, because fuzzy clustering optimization methods convergence speed is depends on initial fuzzy partition. So, the modification is useful techniques for preliminary data analysis and initial fuzzy partition constructing for further some optimization procedure application.

References

[1] R. Bellman, R. Kalaba, L.A. Zadeh "Abstraction and Pattern Classification", *Journal of Mathematical Analysis and Applications*, 13, (1966), 1-7.

[2] J.C. Dunn "A Graph Theoretic Analysis of Pattern Classification via Tamura's Fuzzy Relation", *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-4, (1974), 310-313.

[3] A. Kaufmann "Introduction to the Theory of Fuzzy Subsets", New York: Academic Press, Vol. 1, (1975).

[4] T. Radecki "Level fuzzy sets", *Journal of Cybernetics*, 7, (1977), 189-198.

[5] E.H. Ruspini "A New Approach to Clustering", *Information and Control*, 15, (1969), 22-32.

[6] S. Tamura, S. Higuchi, K. Tanaka "Pattern Classification Based on Fuzzy Relations", *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-1, (1971), 61-66.

[7] D.A. Viattchenin "Some Remarks To Concept of Fuzzy Similarity Relation For Fuzzy Cluster Analysis". In: *Pattern Recognition and Information Processing: Proceedings of Fourth International Conference (20-22 May 1997 Minsk, Republic of Belarus)* /Ed. by V. Krasnoproshin et al., Szczecin: Wydawnictwo Uczelniane Politechniki Szczecinskiej, Vol.1, (1997), 35-38.

[8] L.A. Zadeh "Fuzzy Sets", *Information and Control*, 8, (1965), 338-353.