ROBUST ADAPTIVE FUZZY CLUSTERING FOR DATA WITH MISSING VALUES

Yevgeniy Bodyanskiy, Alina Shafronenko

Abstract: *the datasets clustering problem often encountered in many applications connected with Data Mining and Exploratory Data Analysis. Conventional approach to solving these problems requires that each observation may belong to only one cluster, although a more natural situation is when the vector of features with different levels of probabilities or possibilities can belong to several classes.* This situation is subject of consideration of fuzzy cluster analysis, intensively developing today.

In many practical Data Mining tasks, including clustering, data sets may contain gaps, information in which, for whatever reasons, is missing. More effective in this situation are approaches based on the mathematical apparatus of Computational Intelligence and first of all artificial neural networks and different modifications of classical fuzzy c-*means (FCM) method.*

Real data often contain abnormal outliers of different nature too, for example, measurement errors or distributions with "heavy tails". In this situation classic FCM is not effective because the objective function based on the Euclidean metric, only reinforces the impact of outliers. In such conditions it is advisable to use robust objective functions of special form that suppress influence of outliers. For information processing in a sequential mode adaptive procedures for on-line fuzzy clustering have been proposed, which are in fact on-line modifications of FCM, where instead of the Euclidean metric robust objective functions that weaken the influence of outliers were used.

Situation when data set contains missing values and outliers in the fuzzy clustering problem was not analyzed, although such a situation can arise in many practical applications. Therefore the development of twice robust (for missing values and outliers) fuzzy clustering algorithm has theoretical interest and practical sense.

The problem of fuzzy adaptive on-line clustering of data distorted by missing values and outliers sequentially supplied to the processing when the original sample volume and the number of distorted observations are unknown is considered. The probabilistic and possibilistic clustering algorithms for such data, that are based on the strategy of nearest prototype, partial distances and similarity measure of a special kind that weaken or overwhelming outliers are proposed.

Keywords: *Fuzzy clustering, Kohonen self-organizing network, learning rule, incomplete data with gaps and outliers.*

ACM Classification Keywords: *I.2.6 [Artificial Intelligence]: Learning – Connectionism and neural nets; I.2.8 [Artificial Intelligence]: Problem Solving, Control Methods, and Search – Control theory; I.5.1 [Pattern Recognition]: Clustering – Algorithms.*

Introduction

The problem of data sets clustering often occurs in many practical tasks, and for its solution has been successfully used mathematical apparatus of computational intelligence [Rutkowski, 2008] and first of all, artificial neural networks [Marwala, 2009] and soft computing methods [Klawonn, 2006] (in the case of overlapping classes) is usually assumed that original array is specified a priori and processing is made in batch mode. Here as one of the most effective approach based on using FCM [Bezdek, 1981], that is modified for the situation with missing values [Hathaway, 2001] which comes as a result to minimize the objective function with constraints of special form. In [Bodyanskiy, 2012; Bodyanskiy, 2013] adaptive fuzzy clustering procedures have been proposed for processing the data sequences containing an unknown quantity of missing values, realizing the problem in on-line mode and characterized by numerical simplicity. These procedures are in fact hybrid of T. Kohonen neural network [Kohonen, 1995] with the special form of a neighborhood function.

Real data often contain outliers of different nature, for example, measurement errors or distributions with "heavy tails". In this situation classic FCM is not effective because the objective function based on the Euclidean metric, only reinforces the impact of outliers. In such conditions it is advisable to use robust objective functions of special form [Dave, 1997], that suppress influence of outliers. For information processing in a sequential mode, in [Bodyanskiy, 2005; Kokshenev I., 2006] adaptive on-line fuzzy clustering procedures have been proposed, which are in fact on-line modifications of FCM, where instead of the Euclidean metric robust objective functions, weaken the influence of outliers where used.

Situation when data set contains both missing values and outliers in the fuzzy clustering problem are not considered, although such a situation can arise in many practical applications. Therefore the development of twice robust (for missing values and outliers) fuzzy clustering algorithms has theoretical interest and practical sense.

Problem statement

Baseline information for solving the tasks of clustering in a batch mode is the sample of observations, formed from N n-dimensional feature vectors $X=\{x_1,x_2,...,x_N\}\subset R^n, x_k\in X, k=1,2,...,N$. The result of clustering is the partition of original data set into m classes $(1<m<N)$ with some level of membership $U_q(k)$ of k-th feature vector to the q-th cluster $(1\le q\le m)$. Incoming data previously are centered and standardized by all features, so that all observations belong to the hypercube $[-1,1]^n$. Therefore, the data for clustering form array $\tilde{X}=\{\tilde{x}_1,...,\tilde{x}_k,...,\tilde{x}_N\}\subset R^n$, $\tilde{x}_k=(\tilde{x}_{k1},...,\tilde{x}_{ki},...,\tilde{x}_{kn})^T$, $-1\le\tilde{x}_{ki}\le 1$, $1<m<N$, $1\le q\le m$, $1\le i\le n$, $1\le k\le N$. Note that traditionally adopted in Kohonen's maps (SOM) data transformation to the form $\|\tilde{x}_k\|=1$ in this case does not make sense, because if x_k contains missing value - calculation rules of such vector is impossible, and if x_k contains outlier in one of the components - $\tilde{x}_k$ will be practically the same as the corresponding unit vector of the feature space. Transformation $-1\le\tilde{x}_{ki}\le 1$ leads to the fact that the non deformed data concentrate in the vicinity of zero, and the data with outliers – near -1 and +1. Furthermore, we introduce additional sub arrays data [Hathaway, 2001]: $\tilde{X}=\{\tilde{x}_1,...,\tilde{x}_k,...,\tilde{x}_N\}\subset R^n$, $\tilde{x}_k=(\tilde{x}_{k1},...,\tilde{x}_{ki},...,\tilde{x}_{kn})^T$, $-1\le\tilde{x}_{ki}\le 1$, $1<m<N$, $1\le q\le m$, $1\le i\le n$, $1\le k\le N$.

We have to develop numerically simple on-line procedure for partitioning in sequential mode to the data processing $\tilde{x}_k$ on m perhaps overlapping classes, while it is not known in advance whether $\tilde{x}_k$ is undistorted or contains missing values and outliers. Furthermore, it is assumed that the amount of information under processing is not known in advance and is increased with time.

Adaptive fuzzy clustering data with missing values based on the nearest prototype strategy

Nearest prototype strategy (NFS), proposed in [Hathaway, 2001], is a modification of FCM-algorithm and leads to the replacement of missing components of the vector observations $\tilde{x}_{ki} \in X_G$ by estimates of the corresponding component prototypes (centroids) of the clusters computed using FCM. Thus for each $\tilde{x}_{ki} \in X_G$ it's possible to find the prototype $w_q = (w_{q1,\ldots,}w_{qi,\ldots,}w_{qn})^T$ nearest to $\tilde{x}_k$ in the sense of the partial distance (PD)

$$D_P^2(\tilde{x}_k, w_q) = \frac{n}{\delta_{k\Sigma}} \sum_{i=1}^{n} (\tilde{x}_{ki} - w_{qi})^2 \delta_{ki} \tag{1}$$

where

$$\delta_{ki} = \begin{cases} 0 \mid \tilde{x}_{ki} \in X_G, \\ 1 \mid \tilde{x}_{ki} \in X_F, \end{cases}$$

$$\delta_{k\Sigma} = \sum_{i=1}^{n} \delta_{ki}$$

$w_q^{(\tau+1)} = \arg\min_q \{D_P^2(\tilde{x}_k, w_1^{(\tau+1)}), \ldots, D_P^2(\tilde{x}_k, w_m^{(\tau+1)})\}$, then instead $\tilde{x}_{ki} \in X_G$ input estimate $\hat{x}_{ki} \in w_{qi}$ used in place of the missing components.

In [Bodyanskiy, 2013] adaptive fuzzy clustering procedure based on the NPS was introduced:

$$\begin{cases} U_q^{(\tau+1)}(k) = \dfrac{(\|\hat{x}_k^{(\tau)} - w_q^{(\tau)}(k)\|^2)^{\frac{1}{1-\beta}}}{\sum_{l=1}^{m} (\|\hat{x}_k^{(\tau)} - w_l(k)\|^2)^{\frac{1}{1-\beta}}}, \\ \hat{x}_{ki}^{(\tau)} = w_{qi}^{(\tau)}(k), \quad w_q^{(\tau)}(k) = \arg\min_q \{D_P^2(\tilde{x}_k, w_1^{(\tau)}(k)), \ldots, D_P^2(\tilde{x}_k, w_m^{(\tau)}(k))\}, \\ w_q^{(Q)}(k) = w_q^{(0)}(k+1), \\ w_q^{(\tau+1)}(k+1) = w_q^{(\tau)}(k+1) + \eta(k+1)(U_q^{(Q)}(k))^{\beta}(\hat{x}_k^{(\tau)} - w_q^{(\tau)}(k+1)), \end{cases} \tag{2}$$

where $\beta > 1$ - parameter that is called fuzzyfier and defines "vagueness" of boundaries between classes, $\eta(k+1)$ - learning rate parameter, $\tau = 0,1,2,\ldots$ - accelerated computing time between two real-time instance k and $k+1$ occurs Q iteration in accelerated time.

From the last relation (2) it follows that centroids setting made using the Kohonen self-learning rule "Winner Takes More» (WTM) with the neighborhood function $(U_q^{(Q)}(k))^{\beta}$ having the Cauchian form.

The main disadvantage of FCM and other so-called probabilistic fuzzy clustering algorithms associated with a constraint on the levels of membership of each vector-image, which is equal to one, which gives sense of probability and membership, but it is not always correct in terms of the problem being solved. To remove this restriction in [Keller, 2005] possibilistic fuzzy clustering algorithm (PCM) was introduced, and in [Bodyanskiy, 2012; Bodyanskiy, 2013] - its adaptive version for the case of data containing missing values, having the form:

$$\begin{cases} U_q^{(\tau+1)}(k) = \dfrac{1}{1+\left(\dfrac{\left\|\hat{x}_k^{(\tau)} - w_q(k)\right\|^2}{\mu_q^{(\tau)}(k)}\right)^{\frac{1}{\beta-1}}}, \\ \hat{x}_{ki}^{(\tau)} = w_{qi}^{(\tau)}(k), \quad w_q^{(\tau)}(k) = \arg\min_q \{D_P^2(\tilde{x}_k, w_1^{(\tau)}(k)), \ldots, D_P^2(\tilde{x}_k, w_m^{(\tau)}(k))\}, \\ w_q^{(Q)}(k) = w_q^{(0)}(k+1) \\ w_q^{(\tau+1)}(k+1) = w_q^{(\tau)}(k+1) + \eta(k+1)(U_q^{(Q)}(k))^\beta (\hat{x}_{k+1}^{(\tau)} - w_q^{(\tau)}(k+1)), \\ \mu_q^{(\tau+1)} = \dfrac{\sum_{p=1}^{k} (U_q^{(\tau+1)}(p))^\beta \left\|\hat{x}_p^{(\tau+1)} - w_q^{(\tau+1)}(k)\right\|^2}{\sum_{p=1}^{k} (U_q^{(\tau+1)}(p))^\beta}, \end{cases} \tag{3}$$

where the scalar parameter $\mu \geq 0$ determines the distance at which level of membership equals to 0.5, i.e. if $\left\|\tilde{x}_k - w_q\right\|^2 = \mu_q(k)$, then $w_q(k) = 0.5$.

Algorithms (2), (3) have confirmed working capacity in solving a number of problems [Bodyanskiy, 2013], however, since they are based on the use of Euclidean distance, they do not possess stability to outliers.

Adaptive fuzzy robust data clustering based on the similarity measure

As already mentioned, to solve the problem of fuzzy clustering of data containing outliers the special objective functions of the form [Dave, 1997; Bodyanskiy, 2005; Kokshenev I., 2006] can be used, by some means these anomalies overwhelming, and the problem itself is associated with the minimization of these functions. From a practical point of view it is more convenient to use instead of the objective functions, based on the metrics, the so-called measures of similarity (SM) [Sepkovski, 1974], which are subject to more soft conditions than metrics:

$$\begin{cases} S(\tilde{x}_k, \tilde{x}_p) \geq 0, \\ S(\tilde{x}_k, \tilde{x}_p) = S(\tilde{x}_p, \tilde{x}_k), \\ S(\tilde{x}_k, \tilde{x}_k) = 1 \geq S(\tilde{x}_k, \tilde{x}_p) \end{cases}$$

(no triangle inequality), and clustering problem can be "tied" to maximize these measures.

If the data are transformed so that $-1 \leq \tilde{x}_{ki} \leq 1$ the measure of similarity can be structured so as to suppress unwanted data lying at the edges of interval $[-1,1]$. Figure 1 illustrates the use of similarity measure based on Cauchy function with different parameters width $\sigma^2 < 1$

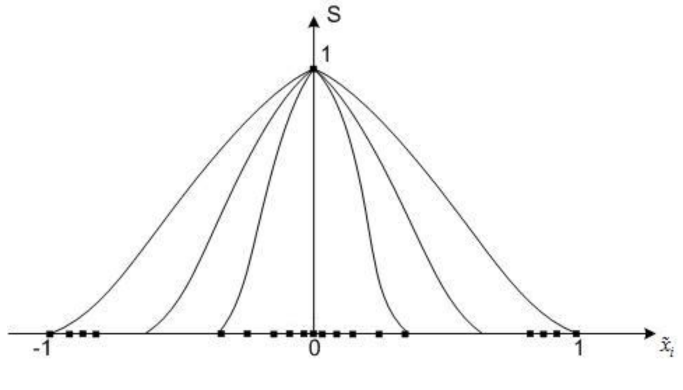


Fig. 1 Similarity measure based on the Cauchy function

By choosing the width parameter σ^2 of functions

$$S(\tilde{x}_k, w_q) = \frac{1}{1+\frac{\|\tilde{x}_k - w_q\|^2}{\sigma^2}} = \frac{\sigma^2}{\sigma^2 + \|\tilde{x}_k - w_q\|^2} = \\ = \frac{\sigma^2}{\sigma^2 + D^2(\tilde{x}_k, w_q)}, \tag{4}$$

is possible to exclude the effect outliers, that in principle cannot be done using the Euclidean metric

$$D^2(\tilde{x}_k, w_q) = \|\tilde{x}_k - w_q\|^2. \tag{5}$$

Further, by introducing the objective function based on similarity measure (4),

$$E_S(U_q(k), w_q) = \sum_{k=1}^{N}\sum_{q=1}^{m} U_q^{\beta}(k) S(\tilde{x}_k, w_q) = \sum_{k=1}^{N}\sum_{q=1}^{m} \frac{U_q^{\beta}(k)\sigma^2}{\sigma^2 + \|\tilde{x}_k - w_q\|^2},$$

probabilistic constraints

$$\sum_{q=1}^{m} U_q(k) = 1,$$

Lagrange function

$$L_S(U_q(k), w_q, \lambda(k)) = \sum_{k=1}^{N}\sum_{q=1}^{m} \frac{U_q^{\beta}(k)\sigma^2}{\sigma^2 + \|\tilde{x}_k - w_q\|^2} + \sum_{k=1}^{N} \lambda(k)\left(\sum_{q=1}^{m} U_q(k) - 1\right) \tag{6}$$

(here $\lambda(k)$ - indefinite Lagrange multipliers) and solving the system of Karush-Kuhn-Tucker equations, we get the solution

$$\begin{cases} U_q(k) = \dfrac{(S(\tilde{x}_k, w_q))^{\frac{1}{1-\beta}}}{\sum_{l=1}^{m} (S(\tilde{x}_k, w_l))^{\frac{1}{1-\beta}}}, \\ \lambda(k) = -\left(\sum_{l=1}^{m} (\beta S(\tilde{x}_k, w_l))^{\frac{1}{1-\beta}}\right)^{1-\beta}, \\ \nabla_{w_q} L_S(U_q(k), w_q, \lambda(k)) = \sum_{k=1}^{N} U_q^{\beta}(k) \dfrac{\tilde{x}_k - w_q}{(\sigma^2 + \|\tilde{x}_k - w_q\|^2)^2} = \vec{0}. \end{cases} \tag{7}$$

The last equation (7) has no analytic solution, so to find a saddle point of the Lagrangian (6) we can use the procedure of Arrow-Hurwitz-Uzawa, as a result of which we obtain the algorithm

$$\begin{cases} U_q(k+1) = \dfrac{(S(\tilde{x}_{k+1}, w_q))^{\frac{1}{1-\beta}}}{\sum_{l=1}^{m}(S(\tilde{x}_{k+1}, w_l))^{\frac{1}{1-\beta}}}, \\ w_q(k+1) = w_q(k) + \eta(k+1)U_q^{\beta}(k+1)\dfrac{\tilde{x}_{k+1} - w_q}{(\sigma^2 + \|\tilde{x}_{k+1} - w_q\|^2)^2} = w_q(k) + \eta(k+1)\varphi_q(k+1)(\tilde{x}_{k+1} - w_q) \end{cases} \quad (8)$$

where

$$\varphi_q(k+1) = \frac{\tilde{x}_{k+1} - w_q}{(\sigma^2 + \|\tilde{x}_{k+1} - w_q\|^2)^2}$$

neighbourhood robust functions of WTM-self-learning rule.

Assuming the fuzzifier value $\beta = 2$ we get a robust variant of FCM:

$$\begin{cases} U_q(k+1) = \dfrac{(S(\tilde{x}_{k+1}, w_q))}{\sum_{l=1}^{m}(S(\tilde{x}_{k+1}, w_l))}, \\ w_q(k+1) = w_q(k) + \eta(k+1)\dfrac{U_q^2(k+1)}{(\sigma^2 + \|\tilde{x}_{k+1} - w_q\|^2)^2}. \end{cases}$$

Further, using the concept of accelerated time, it's possible to introduce robust adaptive probabilistic fuzzy clustering procedure in the form

$$\begin{cases} U_q^{(\tau+1)}(k) = \dfrac{(S(\tilde{x}_k, w_q^{(\tau)}(k)))^{\frac{1}{1-\beta}}}{\sum_{l=1}^{m}(S(\tilde{x}_k, w_l^{(\tau)}))^{\frac{1}{1-\beta}}}, \\ w_q^{(Q)}(k) = w_q^{(0)}(k+1), \\ w_q^{(\tau+1)}(k+1) = w_q^{(\tau)}(k+1) + \eta(k+1)\dfrac{(U_q^{(Q)}(k))^{\beta}}{(\sigma^2 + \|\tilde{x}_{k+1} - w_q^{(\tau)}(k+1)\|^2)^2}(\tilde{x}_{k+1} - w_q^{(\tau)}(k+1)), \end{cases} \quad (9)$$

with the decision of each membership $\tilde{x}_k$ to a specific cluster takes on the maximum value of similarity measure.

Similarly, it's possible to synthesize a robust adaptive algorithm for possibilistic [Klawonn, 1998] fuzzy clustering using criterion

$$E_S(U_q(k), w_q, \mu_q) = \sum_{k=1}^{N}\sum_{q=1}^{m}U_q^{\beta}(k)S(\tilde{x}_k, w_q) + \sum_{q=1}^{m}\mu_q\sum_{k=1}^{N}(1 - U_q(k))^{\beta}.$$

Solving the problem of optimization, we obtain the solution:

$$\begin{cases} U_q(k+1) = \left(1 + \left(\dfrac{S(\tilde{x}_{k+1}, w_q(k))}{\mu_q(k)}\right)\right)^{-1}, \\ w_q(k+1) = w_q(k) + \eta(k+1) U_q^{\beta}(k+1) \dfrac{\tilde{x}_{k+1} - w_q(k)}{(\sigma^2 + \|\tilde{x}_{k+1} - w_q(k)\|^2)^2}, \\ \mu_q(k+1) = \dfrac{\sum_{p=1}^{k+1} U_q^{\beta}(p) S(\tilde{x}_p, w_q(k+1))}{\sum_{p=1}^{k+1} U_q^{\beta}(p)}, \end{cases} \tag{10}$$

receiving at $\beta = 2$ the form

$$\begin{cases} U_q(k+1) = \dfrac{1}{1 + \dfrac{S(\tilde{x}_{k+1}, w_q(k))}{\mu_q(k)}}, \\ w_q(k+1) = w_q(k) + \eta(k+1) \dfrac{U_q^2(k+1)}{(\sigma^2 + \|\tilde{x}_{k+1} - w_q(k)\|^2)^2} (\tilde{x}_{k+1} - w_q(k)), \\ \mu_q(k+1) = \dfrac{\sum_{p=1}^{k+1} U_q^2(p) S(\tilde{x}_p, w_q(k+1))}{\sum_{p=1}^{k+1} U_q^2(p)}. \end{cases}$$

And, finally, introducing the accelerated time we obtain the procedure

$$\begin{cases} U_q^{(\tau+1)}(k) = \dfrac{1}{1 + \left(\dfrac{S(\tilde{x}_k, w_q^{(\tau)}(k))}{\mu_q^{(\tau)}(k)}\right)^{\frac{1}{\beta-1}}}, \\ w_q^{(Q)}(k) = w_q^{(0)}(k+1), \\ w_q^{(\tau+1)}(k+1) = w_q^{(\tau)}(k+1) + \eta(k+1) \dfrac{(U_q^{(Q)}(k))^{\beta}}{(\sigma^2 + \|\tilde{x}_{k+1} - w_q^{(\tau)}(k+1)\|^2)^2} (\tilde{x}_{k+1} - w_q^{(\tau)}(k+1)), \\ \mu_q^{(\tau+1)}(k) = \dfrac{\sum_{p=1}^{k} (U_q^{(\tau+1)}(p))^{\beta} S(\tilde{x}_p, w_q^{(\tau+1)}(k))}{\sum_{p=1}^{k} (U_q^{(\tau+1)}(p))^{\beta}}. \end{cases} \tag{11}$$

Adaptive fuzzy robust data clustering with missing values

For solving the problem of robust data clustering with missing values let's introduce the partial similarity measure (PCM), which is a hybrid of a partial distance (PD) (1) and similarity measure (SM) (4). It is easily to see that such PSM has the form

$$S_P(\tilde{x}_k, w_q) = \frac{\sigma^2}{\sigma^2 + D_P^2(\tilde{x}_k, w_q)}, \tag{12}$$

that allows to obtain the desired properties of algorithms based on procedures described above.

So, on the basis of the procedures (2) and (9) we can introduce the robust adaptive probabilistic fuzzy clustering algorithm for data with missing values:

$$\begin{cases} U_q^{(\tau+1)}(k) = \dfrac{(S_P(\hat{x}_k^{(\tau)}, w_q^{(\tau)}(k)))^{\frac{1}{\beta-1}}}{\sum_{l=1}^{m}(S_P(\hat{x}_k^{(\tau)}, w_l^{(\tau)}))^{\frac{1}{\beta-1}}}, \\ \hat{x}_{ki}^{(\tau)} = w_{qi}^{(\tau)},\ w_q^{(\tau)}(k) = \arg\max_q\{S_P(\tilde{x}_k^{(\tau)}, w_1^{(\tau)}(k)), \dots, S_P(\tilde{x}_k^{(\tau)}, w_m^{(\tau)}(k))\}, \\ w_q^{(Q)}(k) = w_q^{(0)}(k+1), \\ w_q^{(\tau+1)}(k+1) = w_q^{(\tau)}(k+1) + \eta(k+1)\dfrac{(U_q^{(Q)}(k))^\beta}{(\sigma^2 + \|\hat{x}_{k+1}^{(\tau)} - w_q^{(\tau)}(k+1)\|^2)^2}(\hat{x}_{k+1}^{(\tau)} - w_q^{(\tau)}(k+1)), \end{cases} \tag{13}$$

based on procedures (3) and (11), also we can write the robust adaptive algorithm for possibilistic fuzzy clustering of data with missing values:

$$\begin{cases} U_q^{(\tau+1)}(k) = \dfrac{1}{1 + \left(\dfrac{S^{-1}(\hat{x}_k, w_q^{(\tau)}(k))}{\mu_q^{(\tau)}(k)}\right)^{\frac{1}{\beta-1}}}, \\ \hat{x}_{ki}^{(\tau)} = w_{qi}^{(\tau)},\ w_q^{(\tau)}(k) = \arg\max_q\{S_P(\tilde{x}_k^{(\tau)}, w_1^{(\tau)}(k)), \dots, S_P(\tilde{x}_k^{(\tau)}, w_m^{(\tau)}(k))\} \\ w_q^{(Q)}(k) = w_q^{(0)}(k+1), \\ w_q^{(\tau+1)}(k+1) = w_q^{(\tau)}(k+1) + \eta(k+1)\dfrac{(U_q^{(Q)}(k))^\beta}{(\sigma^2 + \|\hat{x}_{k+1} - w_q^{(\tau)}(k+1)\|^2)^2}(\hat{x}_{k+1}^{(\tau)} - w_q^{(\tau)}(k+1)), \\ \mu_q^{(\tau+1)}(k) = \dfrac{\sum_{p=1}^{k}(U_q^{(\tau+1)}(p))^\beta S_P^{-1}(\hat{x}_p, w_q^{(\tau+1)}(k))}{\sum_{p=1}^{k}(U_q^{(\tau)}(p))^\beta}. \end{cases} \tag{14}$$

Thus, the use of partial similarity measure based on partial distance (1), allows us to solve the problem of fuzzy clustering of data containing both missing values and outliers.

Conclusion

The problem of robust adaptive fuzzy clustering algorithms is considered, allowing in on-line mode to process distorted data containing both outliers and missing values is considered. The basis of the proposed algorithms is using of classical procedures as fuzzy c-means of J. Bezdek, T. Kohonen self-learning, as well as specially introduced similarity measure allowing to process distorted information. The algorithms are simple in numerical implementation, being essentially gradient optimization procedures for objective functions of special form.

Bibliography

[Rutkowski, 2008] L.Rutkowski. Computational Intelligence. Methods and Techniques. Berlin-Heidelberg: Springer-Verlag, 2008.

[Marwala, 2009] T Marwala. Computational Intelligence for Missing Data Imputation, Estimation, and Management: Knowledge Optimization Techniques. Hershey-New York: Information Science Reference, 2009.

[Hathaway, 2001] R.J. Hathaway, J.C Bezdek. Fuzzy c-means clustering of incomplete data. IEEE Trans. on Systems, Man, and Cybernetics, 31, №5, 2001, P. 735-744.

[Klawonn, 2006] F. Klawonn. Reducing the Number of Parameters of a Fuzzy System Using Scaling Functions. Soft Computing 10, 2006, P 749-756

[Bezdek, 1981] J.C. Bezdek. Pattern Recognition with Fuzzy Objective Function Algorithms. – N.Y.: Plenum, 1981.

[Bodyanskiy, 2012] BodyanskiyYe., Shafronenko A., Volkova V. Adaptive clustering of incomplete data using neuro-fuzzy Kohonen network. In “Artificial Intelligence Methods and Techniques for Business and Engineering Applications” – Rzeszow-Sofia: ITHEA, 2012. – P. 287-296.

[Bodyanskiy, 2013] BodyanskiyYe., Shafronenko A., Volkova V. Adaptive fuzzy probabilistic clusrering of incomplete data. Int.J. Information Models and Analysis. – 2013. – 2. - №2. – P. 112-117.

[Bodyanskiy, 2013] BodyanskiyYe., Shafronenko A., Volkova V. Neuro fuzzy Kohonen network for incomplete data clustering using optimal completion strategy// Proc. East West Fuzzy Coll., 20th Zittau Fuzzy Coll. – Zittau / Goerlitz: HS, 2013. – P. 214 – 223.

[Kohonen, 1995] T. Kohonen. Self-Organizing Maps. Berlin: Springer-Verlag, 1995.

[Dave, 1997] Dave R.N., Krishnapuram R. Robust clustering methods: A unified view// IEEE Trans. on Fuzzy Systems. – 1997. – 5. - №2. – P.270-293.

[Bodyanskiy, 2005] BodyanskiyYe., Gorshkov Ye., Kokshenev I., Kolodyazhniy V. Robust recursive fuzzy clustering algorithms//Proc. East West Fuzzy Coll. – Zittau/Goerlitz: HS, 2005. – P. 301-308.

[Bodyanskiy, 2005] BodyanskiyYe. Computational intelligence techniques for data analisis. – Lectures Notes on Informatics Vol. P. – 72. – Bonn: GI, 2005. – P. 15-36.

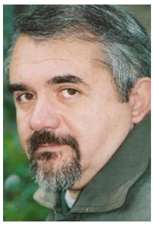
[Kokshenev I., 2006] Kokshenev I., BodyanskiyYe., Gorshkov Ye., Kolodyazhniy V. Outlier resistant recursive fuzzy clustering algorithm / In “Computational Intelligence: Theory and Application”. – Ed. by B.Reusch-Advances in Soft Computing, Vol. 38. – Berlin-Heidelberg: Springer-Verlag, 2006. – P. 647-652.

[Keller, 2005] L. Keller, R. Krishnapuram, N.R. Pal. Fuzzy Models and Algorithms for Pattern Recognition and Image Processing. – N.Y.:Springer Science + Business Media, Inc., 2005.

[Sepkovski, 1974] Sepkovski J.J. Quantified coefficients of association and measurement of similarity // J. Int. Assoc. Math. – 1974. – 6. – №2. – P. 135-152.

[Klawonn, 1998] F. Klawonn, A. Keller. Fuzzy Clustering with Evolutionary Algorithms. Intelligent Systems 13, 1998, P. 975-991.

Authors' Information



Yevgeniy Bodyanskiy – Professor, Dr. – Ing. habil., Scientific Head of Control Systems Research Laboratory, Kharkiv National University of Radio Electronics, 14 Lenin Ave., Office 511, 61166 Kharkiv, Ukraine; e-mail: bodya@kture.kharkov.ua

Major Fields of Scientific Research: Artificial neural networks, Fuzzy systems, Hybrid systems of computational intelligence

Alina Shafronenko –Ph.D. student in Artificial Intelligence dept., Kharkiv National University of Radio Electronics, 14 Lenin Ave., Office 517, 61166 Kharkiv, Ukraine;

e-mail: alinashafronenko@gmail.com

Major Fields of Scientific Research: neural networks, neural network processing of data with missing values, fuzzy clustering, clustering of data