

# A Weighting FCM Algorithm for Clusterization of Companies as to their Financial Performances

Francisco Augusto Alcaraz Garcia  
Turku Centre for Computer Science – TUCS  
Åbo Akademi University  
Lemminkäinengatan 14 B, FIN-20520 Åbo, Finland  
E-mail: falcaraz@abo.fi

Adrian Costea  
Turku Centre for Computer Science – TUCS  
Åbo Akademi University  
Lemminkäinengatan 14 B, FIN-20520 Åbo, Finland  
E-mail: acostea@abo.fi

**Abstract**—We apply fuzzy logic to group telecommunication companies into different clusters as to their financial performances. The objective is to build an easy-to-use financial assessment tool that can assist decision makers in their investment planning and be applied regardless of the economic sector to be analyzed. We characterize each cluster in terms of profitability, liquidity, solvency and efficiency. We implement a modified fuzzy C-means (FCM) algorithm and compare the results with those of normal FCM and previously reported SOM clustering. The results show an improvement in pattern allocation with respect to normal FCM and SOM. The interpretation of the clusters is done automatically representing each ratio as a linguistic variable.

## I. INTRODUCTION

The main target for decision makers is to gain more accurate information about their business market as well as their potential investment opportunities. This is done, in most cases, by analyzing the available historical data (qualitative and quantitative) on the market. The process of obtaining more information about a company's competitors and where it is situated against them is called benchmarking [1] [5]. From the investor's point of view, it is important to see the weaknesses and strengths of a business before the decision to invest or not is taken. Managers and potential investors need financial performances in terms of profitability, liquidity, solvency and efficiency of all economic actors (companies) on that business stage. Analysts have to summarize the high dimensional data to make them interpretable. In this process clustering techniques play a central role.

Clustering is "the organization of a collection of patterns – usually represented as a vector of measurements, or a point in a multidimensional space – into clusters based on similarity" [6]. Traditional clustering methods intend to identify patterns in data and create partitions with different structures. These partitions are called clusters and elements within each cluster should share similar characteristics. In principle, every element belongs to only one partition, but there are observations in the data set that are difficult to position. In many cases subjective decisions have to be made in order to allocate these uncertain observations.

In contrast to these methods, fuzzy logic deals with uncertainty that comes from imprecise information and vagueness. The conventional Boolean logic is substituted by degrees or grades of truth, which allows for intermediate values between

true and false. It is common to express the grades of truth by numbers in the closed interval  $[0, 1]$ , and they can be modeled by membership functions. A membership function assigns a degree of truth (membership degree) for every element subject to the use of that function. The membership function defines a set, called fuzzy set, and degrees of 0 and 1 represent non-membership and full membership respectively to that set, while values in between represent intermediate degrees of set membership.

In this framework, fuzzy clustering methods assign different membership degrees to the elements in the data set indicating in which degree the observation belongs to every cluster. The fuzzy logic approach may also deal with multidimensional data and model nonlinear relationships among variables. It has been applied to financial analysis, for example to evaluate early warning indicators of financial crises [11], or to develop fuzzy rules out of a clustering obtained with self organizing map algorithm [4].

One traditional method in fuzzy clustering is the fuzzy C-means clustering method (FCM) [2]. Every observation gets a vector representing its membership degree in every cluster, which indicates that observations may contain, with different strengths, characteristics of more than one cluster. In this situation we usually assign the elements of the data set to the cluster that has the highest membership degree. In spite of the additional information provided by the methodology, there is a problem with the observations that are difficult to position (uncertain observations) when they obtain similar membership values for two or more clusters.

This paper applies a method to allocate the uncertain observations by introducing weights to the FCM algorithm. The weights indicate the level of importance of each attribute in every cluster so that allocation is done depending on the linguistic classification of the partitions. The data set used corresponds to 7 financial ratios of 88 worldwide telecom companies during the period 1995 to 2001 in an annual basis. The results show that the characterization of the clusters by means of linguistic variables gives an easy to understand, yet formal, classification of the partitions. Also, when weights are extracted from these characteristics, the uncertain observations are allocated. The comparison of the results with other methods is discussed.

## II. FCM ALGORITHM

The FCM algorithm uses as clustering criterion the minimization of an objective function,  $J_m(U, v)$ , and was developed by Bezdek [2] in 1981. The algorithm partitions a multidimensional data set into a specific number of clusters, giving a membership degree for every observation in every cluster. The objective function to minimize is

$$J_m(U, v) = \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^m (d_{ik})^2 \quad (1)$$

where  $c$  is the number of clusters,  $n$  is the number of observations,  $U \in M_{fc}$  is a fuzzy  $c$ -partition of the data set  $X$ ,  $u_{ik} \in [0, 1]$  is the membership degree of observation  $x_k$  in cluster  $i$ ,

$$d_{ik} = \|x_k - v_i\| = \left[ \sum_{j=1}^p (x_{kj} - v_{ij})^2 \right]^{1/2} \quad (2)$$

is the Euclidean distance between the cluster center  $v_i$  and observation  $x_k$  for  $p$  attributes (financial ratios in our case),  $m \in [1, \infty)$  is the weighting exponent, and the following constraint holds

$$\sum_{i=1}^c u_{ik} = 1. \quad (3)$$

If  $m$  and  $c$  are fixed parameters then, by the Lagrange multipliers,  $J_m(U, v)$  may be globally minimal for  $(U, v)$  only if

$$\forall_{\substack{1 \leq i \leq c \\ 1 \leq k \leq n}} u_{ik} = 1 / \left[ \sum_{j=1}^c \left( \frac{d_{ik}}{d_{jk}} \right)^{2/(m-1)} \right] \quad (4)$$

and

$$\forall_{1 \leq i \leq c} v_i = \left[ \sum_{k=1}^n (u_{ik})^m x_k \right] / \left[ \sum_{k=1}^n (u_{ik})^m \right] \quad (5)$$

When  $m \rightarrow 1$ , the Fuzzy  $c$ -Means converges to the Hard  $c$ -Means (HCM), and when we increase its value the partition becomes fuzzier. When  $m \rightarrow \infty$ , then  $u_{ik} \rightarrow 1/c$  and the centers tend towards the centroid of the data set (the centers tend to be equal). The exponent  $m$  controls the extent of membership sharing between the clusters and there is not theoretical basis for an optimal choice for its value.

The algorithm consists of the following steps:

- Step 1. Fix  $c$ ,  $2 \leq c \leq n$ , and  $m$ ,  $1 \leq m \leq \infty$ . Initialize  $U^{(0)} \in M_{fc}$ . Then, for  $s^{th}$  iteration,  $s = 0, 1, 2, \dots$ :
- Step 2. Calculate the  $c$  cluster centers  $\{v_i^{(s)}\}$  with (5) and  $U^{(s)}$ .
- Step 3. Calculate  $U^{(s+1)}$  using (4) and  $\{v_i^{(s)}\}$ .
- Step 4. Compare  $U^{(s+1)}$  to  $U^{(s)}$ : if  $\|U^{(s+1)} - U^{(s)}\| \leq \varepsilon$  stop; otherwise return to Step 2.

Since the iteration is based on minimizing the objective function, when the minimum amount of improvement between two iterations is less than  $\varepsilon$  the process will stop.

One of the main disadvantages of the FCM is its sensitivity to noise and outliers in data, which may lead to incorrect

values for the clusters' centers. Several robust methods to deal with noise and outliers have been presented in [10]. Here, for simplicity, the outliers and far outliers have been leveled in order to minimize their effect in the FCM and the weighting approach suggested.

## III. WEIGHTING FCM TO ALLOCATE UNCERTAIN OBSERVATIONS

The FCM algorithm gives the membership degree of every observation for every cluster. The usual criterion to assign the data to their clusters is to choose the cluster where the observation has the highest membership value. While that may work for a great number of elements, some other data vectors may be misallocated. This is the case when the two highest membership degrees are very close to each other, for example, one observation with a degree of 0.45 for the first cluster and 0.46 for the third. It is difficult to say in which cluster should we include it and it is possible that, after analyzing the vector components, we realize it does not correspond to the average characteristics of the cluster chosen. We call this data vector as "uncertain" observation. Therefore, it would be useful to introduce in the algorithm some kind of information about the characteristics of every cluster so that the uncertain observations can be better allocated depending on which of these features they fulfil more.

### A. Generation of Linguistic Variables

When we analyze a group of companies by their financial performances, we have to be aware of the economic characteristics of the sector they belong to. Levels of ratios showing theoretical bad performances may indicate, for the specific sector, a good or average situation for a company. Conversely, a good theoretical value for the same indicator may indicate a bad evolution of the enterprise in another sector. Usually, financial analysts use expressions like: "high rate of return", "low solvency ratio", etc. to represent the financial situation of the sector or the company. Expressions like that can be easily modeled with the use of linguistic variables and allow the comparison of different financial ratios in a more understandable way regardless of the sector of activity.

Linguistic variables are quantitative fuzzy variables whose states are fuzzy numbers that represent linguistic terms, such as *very small*, *medium*, and so on [7]. In our study we model the seven financial ratios with the help of seven linguistic variables using five linguistic terms: *very low* (VL), *low* (L), *average* (A), *high* (H), *very high* (VH). To each of the basic linguistic terms we assign one of five fuzzy numbers, whose membership functions are defined on the range of the ratios in the data set. It is common to represent linguistic variables with linguistic terms positioned symmetrically [12]. Since there is no reason to assume that the empirical distributions of the ratios in our data set are symmetric, we applied the normal FCM algorithm to each ratio individually in order to obtain the fuzzy numbers, which appeared not to be symmetric. Therefore, the linguistic terms are defined specifically for the sector into consideration. The value of  $m$  was set to

1.5 because it gave a good graphical representation of the fuzzy numbers, and these were approximated to fuzzy numbers of the trapezoidal form. The graphical representation of the linguistic variable for operating margin is shown in Figure 1 and its trapezoidal approximation in Figure 2.

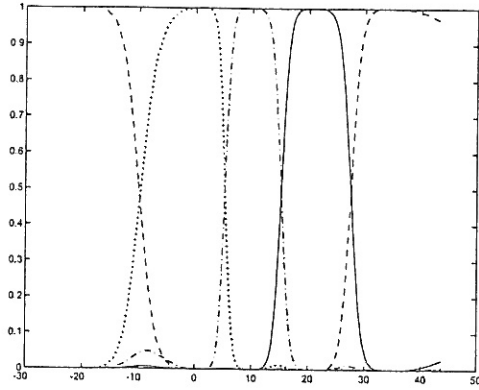


Fig. 1. Linguistic variable representation of operating margin

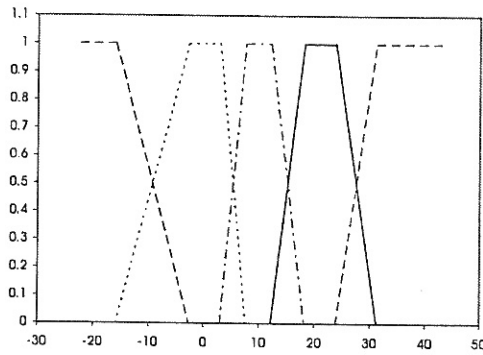


Fig. 2. Trapezoidal approximation for operating margin

Using this approach we can characterize every observation (financial performances of one company in one period), as having *high*, *average*, etc. values in different ratios with respect to the rest of the companies from the same sector. It gives information about the relative situation of the company against its competitors with respect to each individual ratio.

#### B. Calculation of Weights for the FCM

Once we have the linguistic variables for all financial ratios in our data set, we can obtain an importance coefficient (weight) for every ratio in every cluster and introduce it in the clustering algorithm. The objective is to better allocate uncertain observations taking into consideration the linguistic characterizations of the ratios from the certain observations in every cluster.

In order to separate between certain and uncertain observations the FCM algorithm was applied to the initial data set using  $m = 1.5$  and  $c = 7$ . Other clustering methods like SOM [8] showed the appropriateness of seven clusters for the given data set, therefore seven clusters were chosen to make comparisons possible.

We considered as uncertain observations those for which the difference between the two maximum membership degrees was less than twice the equal membership level for every cluster:  $2 * 1/c$ , which seems a reasonable assumption in our data set to clearly define linguistic structures in the clusters. By removing the uncertain observations from the clusters we can represent in a better way the true properties of the clusters and, therefore, obtain clearer classification rules.

Once we have the clusters with the certain observations we can apply the linguistic variables obtained in the previous section to determine the linguistic characterizations. In every cluster and for every ratio we can obtain how many times every linguistic term appears and also the percentage with respect to the total number of observations in the cluster. Clearly, a ratio will be important for the cluster if it has a high percentage of occurrences concentrated in few linguistic terms. In the contrary, if one ratio has a number of occurrences evenly distributed among the linguistic terms, it will not be a good definer of the cluster. As a measure of how evenly or unevenly the percentages of the occurrences are distributed we use the standardized variation coefficient ( $SVC_{ij}$ ). Let us denote with  $perc_{ij}$  the vector of percentages of ratio  $j$  in cluster  $i$ . One element of this vector,  $perc_{ij}(k)$ , will denote the percentage of occurrences of linguistic term (LT)  $k$  for ratio  $j$  in cluster  $i$ .

$$perc_{ij}(k) = \frac{\text{nr of occurrences of LT } k \text{ for ratio } j \text{ in cluster } i}{\text{nr of certain observations in cluster } i} \quad (6)$$

where  $k = 1(VL), 2(L), 3(A), 4(H), 5(VH)$ .

The variation coefficients and the standardized variation coefficients are

$$VC_{ij} = \frac{\text{standard deviation}(perc_{ij})}{\text{mean}(perc_{ij})} \quad (7)$$

and

$$SVC_{ij} = \frac{VC_{ij}}{\sum_{j=1}^p VC_{ij}} \quad (8)$$

A high variation coefficient of the percentages indicates that the ratio clearly defines the cluster. After we split the data in certain and uncertain observations, we calculate the weights ( $SVC_{ij}$ ) using only the certain information. These weights remain constant throughout the iterations of the algorithm. In every iteration, after allocating new uncertain observations, we obtain new clusters' centers and new membership degree values for those observation that remain uncertain.

#### C. Modified FCM

The previous weights are introduced in the Euclidean distance term of the FCM algorithm in the following form:

$$d_{ik} = \left[ \sum_{j=1}^p (x_{kj} - v_{ij})^2 SVC_{ij} \right]^{1/2} \quad (9)$$

where  $SVC_{ij}$  is the standardized variation coefficient of cluster  $i$  for the ratio  $j$ , and it fulfils the constraint (10) since

they are standardized before introducing them in the objective function.

$$\sum_{j=1}^p SVC_{ij} = 1 \quad (10)$$

At each iteration  $s$  we should find the membership degrees that minimize the following objective function:

$$J_m(U, v) = \sum_{k \in I} \sum_{i=1}^c (u_{ik}^{(s)})^m (d_{ik}^{(s)})^2 (1 - u_{ik}^{(s-1)}) \quad (11)$$

where  $I$  is the set of certain observations in iteration  $s$  and  $u_{ik}^{(s-1)}$  is the membership degrees of the certain observations for cluster  $i$  corresponding to the previous iteration. This term is introduced to avoid that lower membership degrees from the uncertain observations become more important in the new allocation. A higher previous membership degree value  $u_{ik}^{(s-1)}$  should lead to a lower recalculated distance from that uncertain observation to the center of that cluster. Therefore,  $1 - u_{ik}^{(s-1)}$  is used when calculating the new distances.

The Lagrange function to minimize the objective function (11)

$$J_{m,\lambda}(U, v) = \sum_{k \in I} \sum_{i=1}^c (u_{ik}^{(s)})^m (1 - u_{ik}^{(s-1)}) + \sum_{j=1}^p (x_{kj} - v_{ij}^{(s)})^2 SVC_{ij} - \sum_{k \in I} \lambda_k \left( \sum_{i=1}^c u_{ik}^{(s)} - 1 \right) \quad (12)$$

leads to the partial derivatives

$$\frac{\partial J_{m,\lambda}(U, v)}{\partial u_{ik}^{(s)}} = m(u_{ik}^{(s)})^{m-1} (d_{ik}^{(s)})^2 (1 - u_{ik}^{(s-1)}) - \lambda_k \stackrel{!}{=} 0 \quad (13)$$

and

$$\frac{\partial J_{m,\lambda}(U, v)}{\partial \lambda_k} = \sum_{i=1}^c u_{ik}^{(s)} - 1 \stackrel{!}{=} 0 \quad (14)$$

We obtain from (13)

$$u_{ik}^{(s)} = \left[ \frac{\lambda_k}{m(d_{ik}^{(s)})^2 (1 - u_{ik}^{(s-1)})} \right]^{1/(m-1)} \quad (15)$$

and with (14) leads to

$$\left( \frac{\lambda_k}{m} \right)^{1/(m-1)} = 1 / \sum_{i=1}^c \left( \frac{1}{(d_{ik}^{(s)})^2 (1 - u_{ik}^{(s-1)})} \right)^{1/(m-1)} \quad (16)$$

that together with (15) gives the expression for the membership degrees

$$u_{ik}^{(s)} = 1 / \sum_{r=1}^c \left( \frac{(d_{ik}^{(s)})^2 (1 - u_{ik}^{(s-1)})}{(d_{rk}^{(s)})^2 (1 - u_{rk}^{(s-1)})} \right)^{1/(m-1)} \quad (17)$$

The necessary condition for the cluster centers is

$$\frac{\partial J_{m,\lambda}(U, v)}{\partial v_{ij}^{(s)}} = -2 \sum_{k=1}^n (u_{ik}^{(s)})^m (1 - u_{ik}^{(s-1)}) (x_{kj} - v_{ij}^{(s)}) SVC_{ij} \stackrel{!}{=} 0 \quad (18)$$

giving

$$\sum_{k \in I} (u_{ik}^{(s)})^m (1 - u_{ik}^{(s-1)}) x_{kj} = v_{ij}^{(s)} \sum_{k \in I} (u_{ik}^{(s)})^m (1 - u_{ik}^{(s-1)}) \quad (19)$$

and the expression for the cluster centers is

$$v_{ij}^{(s)} = \frac{\sum_{k \in I} (u_{ik}^{(s)})^m (1 - u_{ik}^{(s-1)}) x_{kj}}{\sum_{k \in I} (u_{ik}^{(s)})^m (1 - u_{ik}^{(s-1)})} \quad (20)$$

We use equations (20) and (17) to update the centers and membership degrees in our algorithm. We propose the following algorithm:

- Step 1. Fix  $c$  and  $m$ . Initialize  $U = U^{(1)}$ . Apply normal FCM (see Section II) to all dataset and determine the certain ( $I$ ) and uncertain ( $I'$ ) sets of observations. Determine  $SVC_{ij}$  based on the certain observations. We will denote the final  $U$  obtained at this step with  $U^{(l)}$ . Next (steps 2-5 iteratively), allocate the uncertain observations into the certain clusters. Every iteration  $s \in \mathbb{N}$  allocating the uncertain elements consists of following steps:
- Step 2. In the iteration  $s$ , calculate the centers of the clusters using equation (20) with the membership degrees  $u_{ik}^{(s)}$  and  $u_{ik}^{(s-1)}$  corresponding to the certain observations of the current and previous iterations, respectively. When  $s = 1$ ,  $u_{ik}^{(1)} = U^{(l)}$  and  $u_{ik}^{(0)} = 0$ ,  $\forall i = \overline{1, c}$ ,  $\forall k = \overline{1, n}$ .
- Step 3. Calculate  $u_{ik}^{(s+1)}$  of the uncertain observations using equation (17) with the centers obtained in Step 2, and the previous degrees  $u_{ik}^{(s)}$ ,  $k \in I'$  where  $I'$  is the set of uncertain data.
- Step 4. Identify the new certain observations from  $I'$  (based on  $u_{ik}^{(s+1)}$  from the previous step) and attempt to allocate them in the corresponding clusters. Update  $I$  with the new certain observations from  $I'$ . The remaining uncertain observations will become  $I'$  in the next iteration.
- Step 5. If at least one uncertain observation was allocated go to Step 2. If not, exit.

#### IV. IMPLEMENTATION

We have applied the normal FCM and the modified version presented in Subsection III-C to our dataset trying to find clusters of financial performance. The dataset consists of 630 observations of 88 companies from five different regions (Asia, Canada, Continental Europe, Northern Europe, and USA) during the period 1995 to 2001. Every observation contains seven financial ratios of a company for a year calculated from



companies' annual reports, using the Internet as the primary source. The ratios used were: *operating margin*, *return on equity*, and *return on total assets* (profitability ratios); *current ratio* (liquidity ratio); *equity to capital*, and *interest coverage* (solvency ratios); and *receivables turnover* (efficiency ratio). The ratios were chosen from Lehtinen's [9] comparison of financial ratios' reliability and validity in international comparisons. We have used  $m = 1.5$  in the implementation of the algorithm as we have done in the calculation of the linguistic variables, and  $c = 7$  to make the results comparable with SOM algorithm from our previous work [3]. We have characterized each cluster by using the linguistic variables of the certain observations obtained in Step 1 of the algorithm (Table I).

TABLE I  
CHARACTERIZATION OF CLUSTERS

	OM	ROTA	ROE	Current	E to C	IC	Rec. T.	Order
Cluster 1	VL	VL	VL&L	-	A&H	VL&L	-	Bad
Cluster 2	A	A	A&H	-	A	A	-	Average
Cluster 3	VL&L	VL	VL	-	VL&L	L	-	Worst
Cluster 4	H	H	VH	VL	A	A	A	Good
Cluster 5	A	A	A&H	H	H	VH	-	Good
Cluster 6	L	L	A	L	L	L	-	Bad
Cluster 7	VH	VH	H	VH	VH	VH	-	Best

We considered that one linguistic term characterizes one cluster if it represents more than 40 % out of total number of observations for that cluster. For example, for Cluster 1, and ratio ROE, we have two linguistic terms that have more than 40 % of the occurrences (VL&L). When all linguistic terms for one cluster and one ratio are under 40 % we consider that the ratio is not a good definer for that specific cluster. It seems that Receivables Turnover does not have any discriminatory power among data except for one cluster. By simply comparing the clusters we can easily label them as being good, bad, worst, etc. depending on their linguistic terms, as it is shown in Table I.

After Step 1 of the algorithm we obtained 110 uncertain observations, while the remaining 520 certain observations were distributed among different clusters. Our algorithm allocated all uncertain observations except two. We will treat these two observations separately. A total of 19 observations were clustered differently by our algorithm compared to normal FCM. We characterized each one of these observations using our linguistic variables (see Table II). Column X of Table II shows how many ratios of each observation are characterized by the same linguistic term as the characterization of the cluster (shown in Table I) given by the normal FCM, while column Y has the same meaning but for the cluster given by the modified FCM. If we consider that a method clusters better if it gives a higher number of coincidences in the linguistic terms, 9 out of 19 observations (77, 115, 158, 257, 273, 274, 301, 436, 539) were better clustered by our algorithm compared with 6 (42, 213, 233, 265, 391, 619) clustered

better by normal FCM. 4 observations (221, 443, 490, and 614) have an equal number of linguistic term coincidences with the clusters. From this point of view, our implementation overcame, overall, normal FCM.

Last column in Table II shows how SOM clustered these uncertain observations in our previous work. We can also see that our method (modified FCM) has an overall better clustering performance than SOM as well. SOM clustered the nine observations for which modified FCM is better than normal FCM in: a) the same clusters as normal FCM for observations (77, 115, 273, 274, 301), b) the same cluster as modified FCM for (158, 436, 539), and c) a different cluster for observation 257. This means that in 8 out of 9 cases modified FCM outperformed SOM or they performed similarly. For those cases when normal FCM outperformed modified FCM, only in one case (265) SOM outperforms modified FCM by clustering this observation in the same cluster as normal FCM.

Observations 321 and 442 were not allocated by our algorithm because their two highest membership degree values are too close to each other. Observation 321 has a membership degree of 34,87 % for Cluster 4 and 31,76 % for Cluster 5, while observation 442 has 34,9 % for Cluster 2 and 26,77 % for Cluster 1.

Observation 321 corresponds to IDT Company for the year 2001, which experienced rather strange financial results: VL operating margin, VH return on total assets, and VH return on equity (all being profitability ratios). It is difficult, therefore, to make an assessment regarding its profitability performance. Subjectively, Clusters 4 and 5 being labeled as good clusters, we can consider IDT financial performance in 2001 as being "good".

Observation 442 corresponds to the average of US companies in the year 1999. Normal FCM clustered this observation in a good cluster, while our approach was more pessimistic by placing the observation in an average cluster (Cluster 2). This observation shows a pattern similar to IDT in 2001 (observation 321) in the sense that has opposite values for different profitability ratios. Moreover, being an average of US telecommunication companies, we would place it (as our modified FCM shows) in an average cluster rather than in a good one as normal FCM recommends.

## V. CONCLUSIONS

We have implemented a modified version of the traditional fuzzy C-mean algorithm by introducing some weights measures which better characterize each cluster and each ratio.

Firstly, we have built the clusters using certain information (observations with high differences between the highest two membership degree values). The weights were calculated using seven linguistic variables (one for each ratio) using five linguistic terms: *very low* (VL), *low* (L), *average* (A), *high* (H), *very high* (VH). The remaining uncertain observations were reallocated in the certain clusters by using these weights to calculate new distances between the uncertain observations and the new centers of the certain clusters.

TABLE II  
UNCERTAIN OBSERVATIONS

Obs	OM	ROTA	ROE	Current	E to C	IC	Rec. T.	Normal FCM	X	Modif FCM	Y	SOM
42	A	H	H	A	H	VH	L	5 - G	5	2 - A	4	A
77	A	L	VL	A	A	A	L	6 - B	2	3 - W	3	B
115	H	A	A	L	L	L	VH	2 - A	4	6 - B	5	A
158	L	L	L	VL	VH	L	H	5 - G	1	1 - B	4	B
213	A	L	A	VL	VH	L	A	5 - G	3	7 - Be	2	A
221	L	L	L	L	A	L	L	6 - B	5	1 - B	5	B
233	A	A	A	H	VH	VH	L	5 - G	6	7 - Be	3	A
257	L	L	L	VH	A	L	A	6 - B	4	1 - B	5	W
265	H	H	H	VL	H	A	H	2 - A	4	5 - G	3	A
273	H	H	H	L	A	A	L	4 - G	4	2 - A	5	G
274	H	H	H	L	A	A	L	4 - G	4	2 - A	5	G
301	H	H	VH	VH	H	H	L	2 - A	2	4 - G	3	A
321 <sup>1</sup>	VL	VH	VH	VH	H	VH	VL	5 - G	3	4 or 5 - G	-	A
391	A	A	A	VH	VH	H	A	5 - G	4	7 - Be	3	A
436	A	VH	H	H	H	L	VH	7 - Be	3	5 - G	5	G
442 <sup>1</sup>	A	H	VL	H	A	VH	H	5 - G	4	2 or 1 - A	-	A
443	L	A	L	H	A	VH	A	5 - G	4	2 - A	4	A
490	VL	VL	VL	VL	L	VL	A	3 - W	6	1 - B	6	W
539	L	L	A	A	A	A	A	6 - B	4	2 - A	5	A
614	L	L	L	VH	L	VL	L	3 - W	4	6 - B	4	W
619	L	L	A	A	A	L	A	6 - B	5	2 - A	4	A

<sup>1</sup>Unallocated uncertain observations      W – worst, B – bad, A – average, G – good, Be – best

We have compared the results of this approach with normal FCM and SOM using a dataset of 88 worldwide telecommunication companies. Our version outperformed both normal FCM and SOM clustering techniques finding better clusters for the uncertain observations. Also, compared with the other two methods, the use of linguistic variables gave our method a better explanatory power of each cluster. We can now, automatically, characterize each cluster and, also, find those observations that need to be treated carefully due to their specifics.

#### REFERENCES

- [1] T. Bendell, L. Boulter, P. Goodstadt, *Benchmarking for Competitive Advantage*, Pitman Publishing, London, 1998.
- [2] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.
- [3] A. Costea, T. Eklund, J. Karlsson, "A framework for predictive data mining in the telecommunication sector", *Proceedings of the IADIS International Conference - WWW/Internet*, Published by IADIS Press, Lisbon, Portugal, November 2002.
- [4] M. Drobits, W. Winiwarter, U. Bodenhofer, "Interpretation of Self-Organizing Maps with Fuzzy Rules", *Proceedings of the ICTA 2000 – The Twelfth IEEE International Conference on Tools with Artificial Intelligence*, Vancouver, 2000.
- [5] J. Ellis, D. Williams, *Comparative Financial Analysis in Corporate Strategy and Financial Analysis: Managerial, Accounting and Stock Market Perspectives*, Financial Times/Pitman Publishing, London, pp. 203–247.
- [6] A. K. Jain, M. N. Murty, P. J. Flynn, "Data Clustering: A Review", *ACM Computing Survey*, vol. 31, no. 3, September 1999.
- [7] G. J. Klir, B. Yuan, *Fuzzy Sets and Fuzzy Logic. Theory and Applications*, Prentice Hall PTR, Upper Saddle River, New Jersey, 1995.
- [8] T. Kohonen, *Self-Organizing Maps*, 2nd edition. Springer-Verlag, Heidelberg, 1997.
- [9] J. Lehtinen, *Financial Ratios in an International Comparison. Validity and Reliability*, Acta Wasaensia, no. 49, Vasa, 1996.
- [10] J. Leski, "Towards a robust fuzzy clustering", *Fuzzy Sets and Systems* 137, pp. 215–233, 2003.
- [11] C. K. Lindholm, S. Liu, "Fuzzy Clustering Analysis of the Early Warning Signs of Financial Crisis", *Proceedings of the FIP2003, an International Conference on Fuzzy Information Processing: Theory and Applications*, Beijing, March 1–4, 2003.
- [12] T. Lindström, "A fuzzy design of the willingness to invest in Sweden", *Journal of Economic Behavior and Organization*, vol. 36, pp. 1–17, 1998.