Soft Multi-Modal Data Fusion

Sarah Coppock
Lawrence Mazlack
Electrical and Computer Engineering and Computer Science
University of Cincinnati
Cincinnati, OH 45221-0030
{coppocs,mazlack}@uc.edu

Abstract-**Clustering groups items together that are most similar to each other and sets those that are least similar into different clusters. Methods have been developed to cluster records in a data set that are of only qualitative or quantitative data. Data sets exist that contain a mix of qualitative (nominal and ordinal) and quantitative (discrete and continuous) data. Clustering records of mixed kinds of data is a difficult problem. A metric to measure the similarity between records of mixed data types is needed. Once a clustering is found, we do not know how to best evaluate the quality of the clustering when there is a mixture of data varieties.**

1. INTRODUCTION

Grouping together things described by different kinds of data is very important. Records have many diverse types of data. We have a great need to differentially group records containing diverse data. In some sense, this can be considered as "data fusion." Data fusion is a critical necessity for medicine, military sensor integration, and studies of the human population.

Clustering data records offers information. This information includes:

- The discovery of groups of data records that are similar to each other
- The discovery of one or more records that are *outliers*, records which are considered largely dissimilar to the other records; e.g., a fraudulent purchase on a customer's credit card
- A description of similar records, e.g. what type of customer buys a particular product

Note the first two examples rely on the fact that clustering discovers the distribution of the data. The purpose of clustering is up to the user.

Consider a data set containing medical information regarding patients having a particular disease and who were given different treatments. There may exist a combination of different kinds of data: qualitative values such as blood type, and quantitative values such as age and weight. One could use classification on the data set to discover information such as which type of patient responds to a treatment. Clustering is classification's unsupervised counterpart and this offers more potential information. Clustering does not group according to one attribute as classification does. This means that there is potential information offered by clustering that is not offered by classification.

The notion of clustering is closely related to classification and is also used in our own learning of concepts in the real world. Clustering is the grouping of objects into clusters such that the similarity among objects within the same cluster is maximized (*intra-cluster similarity*) and the similarity between objects in different clusters (*inter-cluster similarity*) is minimized [1] [2]. Clustering in data mining and data analysis can discover the general distribution of the data. It allows discovery of similar objects described in the data set. Usually, a good characterization of the resulting clusters is also an objective.

Though the idea of proximity and similarity in clustering rely on the context in which it is used, this experimental research does not seek to include contextual information other than the relative data distributions in the data set. This subjectivity is discussed in the following section. This paper discusses similarity and dissimilarity functions with regard to the goal of grouping together records of different types of data. There are other contexts and goals for which similarity and/or dissimilarity functions are important.

2. CLUSTERING RECORDS

The problem of clustering can be defined as follows: we have a set of attributes, $A=\{A_1, A_2,\ldots,A_k\}$, where each attribute can take on a finite or infinite number of possible values, $Dom(A_i)=\{a_{i1},a_{i2},\ldots,a_{ij}\}$. The domains can be of different data types. *Data type* in the context of this paper refers to the classification of data as either qualitative or quantitative. A data set is composed of records, where each record, r_i is a tuple of values from each attribute's domain. The goal is to cluster the records into groups such that the *intra-cluster similarity*, similarity between records in the same cluster, is maximized and the *inter-cluster similarity*, similarity between clusters, is minimized [1] [2]. Clustering methods that utilize a similarity or distance function assume all domains are of the same kind of data.

2.1 Kinds of Data

Data can be classified by its type and scale, i.e. qualitative or quantitative [1]. Most current clustering algorithms deal with quantitative data. This includes continuous values, such as a person's *height*, and discrete values, such as the *number of cars sold*. Qualitative data on the other hand, is symbols or names with no natural scale between the values. This includes nominal data such as the *color of a car* and ordinal data such as the *doneness of a burger*: *rare*, *medium*, *well*.

A consequence of the lack of a fixed scale for qualitative data is the difficulty in quantitatively measuring the similarity between two qualitative values. It is common to use *simple matching* that assigns a Boolean value of 1 if two values match and 0 if two values don't match (or the converse for measuring dissimilarity). For the similarity of two quantitative values, the normalized magnitude difference between the two values is usually used.

There are different approaches to clustering [1] [2]. Most current research that seeks to cluster records of qualitative data, explicitly or implicitly assume only qualitative data [3] [4] [5] [6] [7].

Our research seeks to group records of both quantitative and qualitative data such as in TABLE1. *Color* and *edible* are qualitative and *weight* and *height* are quantitative. One possible clustering could be:

- $\{r_1,r_2\},\{r_3\},\{r_4,r_5\}$ if the physical attributes (*height* and *weight*) are equally or more important than the *edible* attribute or
- $\{r_1,r_2,r_3\},\{r_4,r_5\}$ if the *edible* attribute dominates, or is most important, and *height* and *weight* are irrelevant.

Our research does not seek to limit any results to only one specific clustering if two or more are found to be reasonable.

There are many possible partitions of a data set. When grouping together multi-modal records in an unsupervised manner, we do not know for what purpose the partition will be used. That is, in the above example, we do not know whether edibility is more important than the physical features or vice versa.

TABLE 1

	color	**weight**	**height**	**edible**
r_1	*orange*	*.001*	*12”*	*no*
r_2	*Red*	*.002*	*12”*	*no*
r_3	*orange*	*10.2*	*12”*	*no*
r_4	*green*	*9.8*	*11”*	*yes*
r_5	*orange*	*.98*	*4”*	*yes*

2.1 Approaches to Clustering

There are more than one approach to clustering. Two of the more commonly found approaches are hierarchical clustering and partitioning. While hierarchical clustering is more dependent on similarity between two single records (or a record and a cluster representation), the partitioning approach relies on a similarity function over all the records, i.e. a function of the intra- and inter-cluster similarities.

There are two types of hierarchical approaches to clustering: *agglomerative* and *divisive* [2] [8]. Agglomerative begins with all objects in their own cluster and combines clusters for which the similarity is the greatest. This is done repeatedly until all objects are in the same cluster. Divisive begins with all objects in the same cluster and works in the reverse direction, until all records are in their own cluster.

As these approaches are based on a similarity metric, an appropriate similarity metric must be defined for measuring the similarity between records (or cluster representation) for any combination of data. Such metrics already exist for records that contain only qualitative or quantitative data; e.g. Minkowski metric for quantitative records [8] or Simple Matching Coefficient for qualitative records [9]. In deciding to merge clusters, a representation for clusters may be used. How best to represent clusters with multiple records which contain both qualitative and quantitative data is also an open problem.

2.1.2 Clustering by Partitioning

Another approach to clustering is to use an initial, possibly arbitrary, partition of the records and to refine this initial partition [2] [8]. The refinement of the clustering is achieved by redistributing records to other clusters according to some similarity criterion. In this approach, it is common to use a representation for each cluster, e.g. a mean, in order to decide how to redistribute the records. Alternatively, the algorithm can use an overall evaluation function of the goodness of the clustering to search for an optimal clustering.

Usually, this approach requires the number of clusters to be known a priori—which can become a difficulty if the decided number is not appropriate to the data distribution. Advances have been made in autonomously recognizing cluster counts and cluster centers [10]. Recognizing clusters is a problem that exists independent of the kinds of data. It is more difficult for mixed data. The difficulty is in evaluating the goodness of the clustering and finding a suitable cluster representation is not completely solved for records composed of different kinds of data.

Clustering qualitative and quantitative data is a difficult problem. When all attributes are of the same kind then the inter- and intra-cluster similarity can be defined according to a similarity measure between records. Similarity metrics are defined for records of one data type. When attributes are of different data types, we do not have a sufficiently defined similarity metric to use in measuring the similarity between two records. This makes it difficult to generalize clustering algorithms to cluster records of mixed data.

3. SIMILARITY METRICS

To be able to use developed clustering algorithms, it would be useful to have a similarity metric that is useful on any mix of data types. Most current similarity metrics use some combination of the similarity between individual attribute values to derive the overall similarity between records [19].

For $r_1=\{a_1,b_1,c_1\}$ and r2=$\{a_2,b_2,c_2\}$, the similarity would be defined as $sim(a_1,a_2) \oplus sim(b_1,b_2) \oplus sim(c_1,c_2)$ where $\oplus$ indicates some combination operator. Typically, the combination operator is the addition operator [9] [11] [21].

Although the measure itself can be either quantitative or qualitative, most current metrics derive a quantitative measure. It is assumed that finding a quantitative measure of similarity would be best for the purpose of clustering records.

Similarity between values is according to the type of data the values are.

4. Similarity Between Records Of Mixed Data Types

Data sets with a mixture of types of data are common. Applying an existing algorithm that assumes only one type of data would not be meaningful. In order to cluster these records using an existing clustering algorithm, a meaningful way of measuring the similarity between records must be developed. Therefore, measuring the similarity or distance between two records requires that the metric must be able to handle a mixture of types of data. There are at least two possible ways of developing a useful metric:

- a metric developed for one type of data, e.g. Euclidean distance for quantitative data, can be extended by some form of mapping to include both types or
- a metric that usefully combines two or more metrics, some qualitative and some quantitative, can be developed.

Both of these approaches to solving the problem of clustering mixed data have difficulties. It is possible that when developing a similarity metric for mixed data types, the utility of the metric is compromised.

4.1 Extending Quantitative Metrics

The most straightforward way to extend metrics developed for quantitative data is to use some form of mapping from the qualitative data to quantitative values. The difficulty in doing this is discovery of a useful mapping. This is due to the lack of scale between qualitative values. With nominal data, the lack of order along with lack of scale causes difficulty.

Unfortunately, it does not make sense to map these values into a form appropriate for the typical distance measures. Rather, it is difficult to discover a meaningful mapping even if we have contextual information regarding the data.

Let one arbitrary mapping (π_1) be:

fruit={orange: 1, apple: 2},
color={red: 1, orange: 2, green: 3}

and another arbitrary mapping (π_2) be:

fruit={orange: 2, apple: 1},
color={red: 0, orange: 1, green: 6}

for the nominal values in TABLE 2.

TABLE 2

	fruit	color	bag size
r_1	apple	red	5
r_2	orange	orange	3
r_3	apple	green	5

Euclidean distance is defined as $d(X,Y)= (\Sigma ([x_i-y_i])^2)^{1/2}$, where X and Y are the records being compared, x_i and y_i are the i^th attribute values of X and Y, and $d(X,Y)$ is the distance between records X and Y. Using this distance function, $d(r_1,r_2) > d(r_1,r_3)$ with π_1, but $d(r_1,r_2) < d(r_1,r_3)$ with π_2.

The difficulty is that the metric is defined for quantitative values, but we are imposing an artificial ordering and a scale. Therefore, we can select an arbitrary mapping, but we can't be sure about the usefulness of the resulting measure. If we put records in order of similarity from an arbitrary single record, i.e. $r_1<r_2<\ldots<r_n$ where $r_i<r_j<r_k$ indicates that r_j is more similar to r_i than r_k, then we can change the similarity order according to the mapping used. This could then affect the resulting clustering.

4.2 Extending Qualitative Metrics

Extending qualitative metrics to deal with quantitative data is also difficult. Most metrics useful for records of qualitative data use a form of value matching to decide the measure [9]. Usually, this measure is some proportion of the total number of attribute values. For example, the simple matching coefficient [9] is defined as the number of attribute-value pairs that the two records have in common divided by the total number of attributes in the records.

Extending a metric such as this to quantitative data results in a loss of information; this would not be desirable. For example, if using simple matching, as defined in section 3.1, between the quantitative values, 3.4, 3.5, and 4.2, the similarity measure will be the same between each pair. Information such as the fact that 4.2 is more dissimilar to 3.4 than it is to 3.5 is lost.

Even if matching were to be used on quantitative ranges or intervals such as [3.0,3.5], it is possible that information would still be lost. For example, if two different values lay in the same interval, then the information that they are dissimilar will be lost when measuring similarity. In addition, discovering the optimal intervals in order to make this useful would be difficult and computationally expensive.

4.3 Combining Existing Metrics

Since extending existing metrics for one data type has difficulties, the next possible solution to determine similarity between records of mixed data type is to look at combining existing metrics. For example, extending the k-means algorithm to handle both qualitative and quantitative data has been attempted [15] [18] [12]. This particular approach is further discussed in section 4.3.1.

Since extending one metric to handle multiple data types is difficult, the other alternative is to combine two types of metrics, one measure for the qualitative attributes and one for the quantitative attributes. But, combining in a simple manner causes the utility of the measure to be compromised.

4.3.1 Similarity Metric Utility

Although the following discussion is about distance measures, the same argument can be made for any pair of similarity measures. This is because distance is a *mathematically metric* measure of dissimilarity and dissimilarity is the complement of similarity. By *mathematically metric*, it is meant that the measure meets the following criteria:

- $\text{measure}(x,y) \geq 0$

- measure(x,x) = 0
- measure(x,y) = measure(y,x)
- measure(x,y) + measure(y,z) ≤ measure(x,z)

This definition will be used in later sections.

Huang [14] and [15] extends the distance-based method k-means algorithm to handle categorical data. By using an integer value, 1 or 0, to indicate non-matching and matching respectively, a categorical attribute is incorporated into the distance metric. In [14], similarity is computed as the sum of square differences for the numerical attributes simply added to a weighted summation of matches for the categorical attributes. In other words, the similarity is a sum of two metrics, one for quantitative and one for qualitative. This is like adding apples to oranges. This is because quantitative values contribute their magnitude whereas the qualitative attributes have no magnitude to contribute (simply an integer value between 0 and the number of qualitative attributes). The magnitudes of the quantitative attributes therefore contribute to the measure differently, a point previously made by [19].

TABLE 3

	A_1	A_2	A_3
r_1	1	a	2
r_2	1	b	2
r_3	4	a	2
r_4	2	a	1
r_5	1	b	4
r_6	3	a	3

For TABLE 3, we have $d(r_1,r_2)=1$ and $d(r_1,r_3)=9$ where $d(r_i,r_j)$ is the distance (dissimilarity) for objects r_i and r_j as defined by [14]. It is important to note that the suggested weight was approximately 1.27 and would make little, if any, difference in the following discussion. The question raised is: should the magnitude associated with a numerical attribute give considerably more (or less) weight to the (dis)similarity measure? Note that r_1 has two values in common with both r_2 and r_3. In this case, it makes more sense to find a quantitative metric consistent across all of the attributes being considered in the measure. This follows from the idea of standardizing the data before the computation of similarity [1].

Li [16] developed a clustering method using the Goodall similarity metric [19]. This metric measures the amount of weight that a categorical value contributes to the overall similarity measure. For example, if two records have the same value for an attribute k, then the similarity isn't necessarily 1 while most previous metrics allow only 0 or 1, non-match or match respectively. The given value for a match is therefore a real number between 0 and 1. The metric allocates a value proportional to the frequency as compared with other values. This weight allows for a more consistent contribution between the two types of attributes.

For the same three items we have $d(r_1,r_2) \approx 0.31$ and $d(r_1,r_3) \approx 0.82$ using Goodall's function ($sim(r_1,r_2) \approx 0.69$ and $sim(r_1,r_3) \approx 0.18$). This metric takes the distribution of values into account along with the magnitude of the quantitative values. The chi-squared ($\chi 2$) statistic is used for computing the measure between records. Because this statistic is used, there is the assumption of independence among the attributes, which cannot be guaranteed. It also may not always be the best idea to have the probability distribution influence the measure of similarity. In other words, if two records have the same qualitative value for an attribute and the value happens to be less common, should it contribute more to the measure than if it were common? A possible solution would be to discover a meaningful and feasible weighting on the attributes if one exists.

The above briefly discusses two different distance measures. Because of the different possible values for the two measures, it does not make sense to compare the two directly. The measure in [14] has no upper bound (it lies on the range of positive reals) where the measure in [16] is on the real line interval 0 to 1. In fact, one measure is mathematically metric while the other measure is. This means that we cannot say whether the measures are comparable with regard to the similarity of the records. It in fact begs the question of how to decide the usefulness of the measures.

One approach that has been proposed by [13] uses a metric similar to [14], but differs only in the metric for measuring the distance between the quantitative attributes. [13] uses the Mahalanobis metric. This metric for quantitative data solves the difficulty of the covariance of the data. The clustering uses this metric with rough sets to find an optimal clustering. The use of rough sets seems to be a useful approach to clustering records of mixed data types, but it is unclear whether this method works for mixed data.

In developing a suitable metric, the question of whether the metric should be mathematically metric arises. It would be natural to say that it should; yet, the answer to this question may be context-dependent. For example, the Goodall measure [19] was developed with biological taxonomy in mind where the commonalities of a characteristic influence how similar two objects with the characteristic are. This may or may not be desirable in some applications. In addition, the importance of one attribute may outweigh another.

Gower [21] developed a metric similar to Goodall's in the field of taxonomy. His metric is defined as:

$$sim(X,Y)=\Sigma\,(wt_{xy(i)}*sim_{xy(i)}) \,/\, \Sigma\, wt_{xy(k)}$$

where $wt_{xy(j)}$ is the weight apportioned to the j^{th} attribute. This weight is usually 0 or 1, and allows for the handling of missing attribute values of one or both records being compared. $sim_{xy(k)}$ is the similarity measure between the k^{th} attribute values for records X and Y. For qualitative data, $sim_{xy(k)}$ is either 1 or 0, whether the values match or not, respectively. For quantitative values, $sim_{xy(k)}$ is the proportion of difference in magnitude and the range for the k^{th} attribute. This metric may be more useful. The portion contributed by the quantitative attributes to the measure relies on the range of values, where as the portion contributed by the qualitative attributes is fixed for any qualitative attribute. It is unclear whether the difference between the two measures combined compromises the utility of the metric. It would seem that this is essentially the same difficulty: combining measures that lie in different spaces. As of yet, this metric has not been used in clustering data sets of mixed type.

6. Rough Sets

Rough set theory is a mathematical tool introduced by [22] and is used to deal with uncertainty and vagueness. What follows is a general introduction. For a more detailed, mathematical introduction, see [23]. The basic idea behind rough sets is the notion of *indiscernibility* and *equivalence*. For a set X, an item is either in the set, not in the set or it's membership in the set is unknown. That is, if we are given TABLE 4 and we are interested in the set for swim:yes ({r2,r5}). This set is defined by:

- *Lower approximation*: the records that are definitely in the set {r5}
- *Boundary region*: those records that whose membership cannot be determined {r2,r4} and
- The *upper approximation:* those records definitely in the set and those in the boundary region {r2,r5,r4}

In this case, the set {r2,r5} is termed *rough*, the boundary region is not empty [22]. Another way to view the concept of rough is visually, as in Fig. 4. The concept, *A* is represented by the shaded irregular shape. The area in between the two rectangles, the *upper* and *lower* approximations, is the boundary region.

We can extend the above concepts to grouping records together in different ways. The most straightforward way is to consider the concept defined by the rough set as the cluster or group to be approximated, as in [24]. Rough set theory can be a particularly helpful when using a scalar distance or similarity metric for mixed data; this is a result of the uncertainty that arises from the artificial impositions (scale and/or order) or combining established metrics of different scales.

TABLE 4

	sky	temp	swim
r1	*rainy*	*warm*	*no*
r2	*sunny*	*hot*	*yes*
r3	*rainy*	*cold*	*no*
r4	*sunny*	*hot*	*no*
r5	*sunny*	*warm*	*yes*

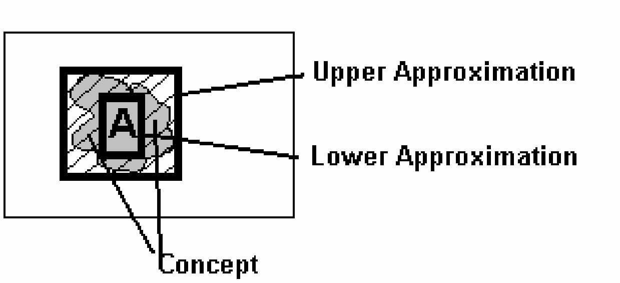

Figure 1 Visual representation of a rough set.

Pawlak [23] defines associated quantitative terms which describe degrees of approximation for a concept. Similar terms are used to modify the equivalence relations in [13]. One of the hardships with combining existing similarity functions is that the multiple functions have different domains and ranges while maintaining ordering properties. It is this difficulty of fusing and the fact that rough sets can be used in approximating concepts that leads to the idea that rough sets would be a useful tool in fusing the information.

7. Epilogue

Clustering groups items together that are most similar to each other and sets those that are least similar into different clusters. Methods have been developed to cluster records in a data set that are of only qualitative or quantitative data. Data sets exist that contain a mix of qualitative (nominal and ordinal) and quantitative (discrete and continuous) data. Clustering records of mixed kinds of data is a difficult problem. A metric to measure the similarity between records of mixed data types is needed. Once a clustering is found, we do not know how to best evaluate the quality of the clustering when there is a mixture of data varieties. The proposed research is to find a useful way of clustering data sets containing both quantitative and qualitative data. This could be a developed metric general enough to be used in existing algorithms or the development of a new approach incorporating soft computing, specifically rough sets.

In regard to knowledge discovery and to clustering in particular, it would be useful to have a fundamental similarity measure for data records. Unfortunately, few exist that account meaningfully for any combination of types of data. The more meaningful metrics known are restrictive to a particular area or science. How to combine difference in magnitude and simple matching so that it is general enough for mining is a question that is yet to be reasonably answered. In clustering records of mixed data types, questions exist that need to be answered.

References

[1] B. Everitt. *Cluster Analysis*, 3rd ed. Hodder & Stoughton, London, 1993.

[2] Jain and R. Dubes. *Algorithms for Clustering Data*. Prentice Hall, New Jersey, 1988.

[3] V. Ganti, J. Gehrke, and R. Ramakrishnan. “CACTUS: Clustering Categorical Data Using Summaries” in *Knowledge Discovery and Data Mining*, pp. 73-83, 1999.

[4] D. Gibson, J. Kleinberg, and P. Raghavan. “Clustering Categorical Data: an Approach Based on Dynamical Systems” in *Proceedings of the 24th VLDB Conference*, Vol. 8(3/4), pp. 222-236, 2000.

[5] S. Guha, R. Rastogi, and K. Shim, “ROCK: A Robust Clustering Algorithm for Categorical Attributes”, in *Information Systems*, Vol. 25(5), pp. 345-366, 2000.

[6] K. Wang, C. Xu, and B. Liu. “Clustering Records Using Large Items”, in CIKM 1999. pp. 483-490.

[7] Y. Zhang, A. Wai-chee Fu, C. Cai, and P. Heng. "Clustering Categorical Data", in *16th International Conference on Data Engineering,* 2000.

[8] J. Han and M. Kamber. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers, San Fransisco, 2001.

[9] P. Sneath and R. Sokal. *Numerical Taxonomy*. Freeman and Company, San Fransisco, 1973.

[10] Y. Zhu. *Unsupervised Database Discovery Based on Artificial Intelligence Techniques*, Masters Thesis, June 2002.

[11] K. Gowda and E. Diday. "Symbolic Clustering Using a New Dissimilarity Measure" in *Pattern Recognition*, Vol. 24(6), pp. 567-578, 1991.

[12] S. Gupta, K. Rao, and V. Bhatnagar, "K-Means Clustering Algorithm For Categorical Attributes", in *Data Warehousing and Knowledge Discovery*, pp. 203-208, 1999.

[13] S. Hirano, T. Okuzaki, Y. Hata, S. Tsumoto, and K. Tsumoto. "A Rough Set-Based Clustering Method with Modification of Equivalence Relations", in *PAKDD 2001*, D. Cheung et al eds., pp. 513-518, 2001.

[14] Z. Huang. "Clustering Large Data Sets With Mixed Numeric and Categorical Values", in *Proceedings of 1st Pacific-Asia Conference on Knowledge Discovery & Data Mining*, 1997.

[15] Z. Huang and M. Ng. "A Fuzzy k-Modes Algorithm for Clustering Categorical Data", in *IEEE Records on Fuzzy Systems*, Vol 7(4), pp. 446-452, August 1999.

[16] C. Li and G. Biswas. "Conceptual Clustering With Numeric-and-Nominal Mixed Data-A New Similarity Based System", in *IEEE Transcript on KCE* 1998.

[17] E. Han, G. Karypis, V. Kumar, and B. Mobasher. "Clustering Based on Association Rule Hypergraphs", in *Proceedings of SIGMOD '97 Workshop on Research Issues in Data Mining and Knowledge Discovery (DMKD'97)*, May 1997.

[18] H. Ralambondrainy. "A Conceptual Version of the K-Means Algorithm" in Pattern Recognition Letters, Vol. 16, pp. 1147-1157, 1995.

[19] D. Goodall. "A New Similarity Index Based On Probability" in Biometrics, Vol. 22(4), pp. 882-907, 1966.

[20] B. Everitt and S. Rabe-Hesketh, *The Analysis of Proximity Data*, Wiley, New York, 1997.

[21] J. Gower. "A General Coefficient of Similarity and Some of Its Properties" in Biometrics, Vol. 27(4), pp. 857-871, 1971.

[22] Pawlak, Z. "Rough Sets", in *International Journal of Computer and Informational Sciences*, Vol. 11, no. 5, pp. 341-356, 1982.

[23] Pawlak, Z. *Rough Sets: Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publishers, Dordrecht, 1991.

[24] Pawlak, Z. "Rough Classification" in International Journal of Man-Machine Studies, Vol. 20(5), pp. 469-483, 1984.