Hierarchical Fuzzy Signature Structure for Complex Structured Data

Kok Wai Wong, Alex Chong, Tamás D. Gedeon, László T. Kóczy and Tibor Vámos

Abstract— **This paper presents an algorithm for constructing a hierarchical fuzzy signature structure by using concepts from interclass separability and hierarchical fuzzy systems. Fuzzy signature is introduced to handle complex structured data and interdependent features problems. Fuzzy signature can also used in cases where data is missing. However, when dealing with a very large data set, it is possible that they hide hierarchical structure that appears in the sub-variable structures. The proposed hierarchical fuzzy signature structure will be used in problems that fall into this category. In the end, reduced hierarchical fuzzy signature structures can assist human experts better by removing unnecessary information when making decisions.**

Index Terms—**Fuzzy signature structure, hierarchical, complex structured data, and large databases.**

I. INTRODUCTION

Fuzzy modelling has become very popular because of the main feature of its ability to assign meaningful linguistic labels to the fuzzy sets [1] in the rule base [2,3]. However, a serious problem is caused by the high computational time and space complexity of rule bases describing systems with multiple inputs with proper accuracy. The complexity allows little general systems application (or real time control application) of classical fuzzy algorithms, where the inputs exceed about 6 to 10. These traditional fuzzy systems deal with very simple structured data, where the number of inputs is well defined, and values for each input occur for most or all data items. This further reduces their general applicability.

Research supported by the Australian Research Council, National Scientific Research Fund OTKA T034233 and T034212, a Main Research Direction Grant 2002 by Széchenyi István University, and the National Research and Development Project Grant NKFP-2/0015/2002.

K. W. Wong is with the School of Computer Engineering, Nanyang Technological University, Singapore 639798 (phone: +65 6790 4327; fax: +65 6792 6559; e-mail: askwwong@ntu.edu.sg).

A. Chong, and T.D. Gedeon are with the School of Information Technology, Murdoch University, Western Australia 6150. (e-mail: {cchong | tgedeon}@murdoch.edu.au).

L.T. Kóczy is with the Department of Telecommunication & Telematics, Budapest University of Technology and Economics, Hungary. He is also with the Institute of Information Technology and Electrical Engineering, Széchenyi István University, Győr, Hungary. He has been a visiting professor to the School of Information Technology, Murdoch University, Western Australia. (e-mail: koczy@ttt.bme.hu or koczy@sze.hu).

T. Vámos is with the Institute of Computing and Automation, Hungarian Academy of Science, Kende u., Budapest, Hungary (e-mail: vamos@sztaki.hu).

Basically, fuzzy rule bases suffer from rule explosion. The number of possible rules necessary is $O(T^k)$ where K is the number of dimensions and T is the number of terms per input. In order to increase the problems solvable by fuzzy rule-base systems, it is essential to reduce T, k, or both. Decreasing T leads to sparse fuzzy systems, i.e. fuzzy rule-bases with "gaps" between the rules [4]. On the other hand, decreasing K leads to hierarchical fuzzy systems [5, 6].

The hierarchical fuzzy system is based on the following idea. Often, the multi-dimensional input space $X = X_1 \times X_2 \times \ldots \times X_k$ can be decomposed into some subspaces, e.g. $Z_0 = X_1 \times X_2 \times \ldots \times X_{k0}$ $(k_0 < k)$, so that in Z_0 a partition $\Pi = \{D_1, D_2, D_3\}$ can be determined. In each D_i, a sub-rule base R_i can be constructed with local validity. The hierarchical rule base structure becomes:

R_0: if Z_0 is D_1 then use R_1
if Z_0 is D_2 then use R_2
……
if Z_0 is D_n then use R_n

R_1: if Z_1 is A_{11} then y is B_{11}
if Z_1 is A_{12} then y is B_{12}
……
if Z_1 is A_{1m1} then y is B_{1m1}

……

R_n: if Z_n is A_{n1} then y is B_{n1}
if Z_n is A_{n2} then y is B_{n1}
……
if Z_n is A_{nmn} then y is B_{nmn}

The fuzzy rules in rule base R_0 are termed meta-rules since the consequences of the rules are pointers to other sub-rule bases instead of fuzzy sets. The complexity of fuzzy systems can be reduced when the suitable Z_0 and Π are found such that in each sub-rule base R_i the input space X_i is a subspace of $X / Z_0 = X_{k0+1} \times X_{k0+2} \times \ldots \times X_k$ [6]. The main difficulty in the automatic construction of such system is largely in finding a suitable subspace Z_0 and Π.

Fuzzy signatures which structure data into vectors of fuzzy values, each of which can be a further vector, are introduced to handle complex structured data [7, 8, 9]. This will widen the application of fuzzy theory to many areas where objects are complex and sometimes interdependent features are to be classified and similarities / dissimilarities evaluated. Often, human experts can and must make decisions based on comparisons of cases with different numbers of data components, with even some components missing. Fuzzy signature is created with this objective in mind. This tree structure is a generalisation of fuzzy sets and vector valued fuzzy sets in a way modelling the human approach to complex problems.

However, when dealing with a very large data set, it is possible that they hide hierarchical structure that appears in the sub-variable structures. This paper is used to address problems having this characteristic, and the possible use of hierarchical structure in creating the fuzzy signature.

II. FUZZY SIGNATURE

The original definition of fuzzy sets was $A: X \to [0,1]$, and was soon extended to *L-fuzzy sets* by Goguen [10],

$$A_S : X \to [a_i]_{i=1}^k,\ a_i = \begin{cases} [0,1] \\ [a_{ij}]_{j=1}^{k_i} \end{cases},\ a_{ij} = \begin{cases} [0,1] \\ [a_{ijl}]_{l=1}^{k_{ij}} \end{cases},$$

$A_L : X \to L$, L being an arbitrary algebraic lattice. A practical special case, *Vector Valued Fuzzy Sets* was introduced by Kóczy [11], where $A_{V,k} : X \to [0,1]^k$, and the range of membership values was the lattice of *k*-dimensional vectors with components in the unit interval. A further generalisation of this concept is the introduction of fuzzy signature and signature sets, where each vector component is possibly another nested vector (right).

Fuzzy signature can be considered as special multi-dimensional fuzzy data. Some of the dimensions are inter-related in the sense that they form sub-group of variables, which jointly determine some feature on a higher level. Let us consider an example. Figure 1 shows a fuzzy signature structure.

The fuzzy signature structure shown in Figure 1 can be represented in vector form as follow:

$$x = \begin{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} \\ \begin{bmatrix} x_{21} \\ \begin{bmatrix} x_{221} \\ x_{222} \\ x_{223} \end{bmatrix} \\ x_{23} \end{bmatrix} \\ \begin{bmatrix} x_{31} \\ x_{32} \end{bmatrix} \end{bmatrix}^T$$

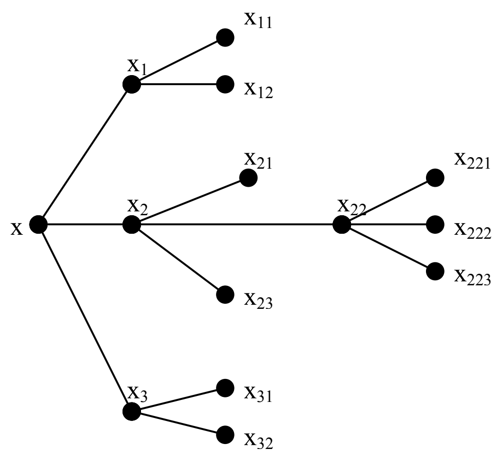

Figure 1: A Fuzzy Signature Structure

Here $[x_{11}\ x_{12}]$ form a sub-group that corresponds to a higher level compound variable of x_1. $[x_{221}\ x_{222}\ x_{223}]$ will then combine together to form x_{22} and $[x_{21}[x_{221}\ x_{222}\ x_{223}]x_{23}]$ is equivalent on higher level with $[x_{21}\ x_{22}\ x_{23}] = x_2$. Finally, the fuzzy signature structure will become $x = [x_1\ x_2\ x_3]$ in the example.

The relationship between higher and lower level is govern by a set of fuzzy aggregations. The results of the parent signature at each level are computed from their branches with appropriate aggregation of their child signature. Let a_1 be the aggregation associating x_{11} and x_{12} used to derive x_1, thus $x_1 = x_{11}\ a_1\ x_{12}$. By referring to Figure 1, the aggregations for the whole signature structure would be a_1, a_2, a_{22}, and a_3. The aggregations a_1, a_2, a_{22}, and a_3 are not necessarily identical or different. The simplest case for a_{22} might be the *min* operation, the most well known t-norm. Let all aggregations be *min* except a_{22} be the averaging aggregation. We will show the operation based on the following fuzzy signature values for the structure in the example.

$$x = \begin{bmatrix} \begin{bmatrix} 0.3 \\ 0.4 \end{bmatrix} \\ \begin{bmatrix} 0.2 \\ \begin{bmatrix} 0.6 \\ 0.8 \\ 0.1 \end{bmatrix} \\ 0.9 \end{bmatrix} \\ \begin{bmatrix} 0.1 \\ 0.7 \end{bmatrix} \end{bmatrix}^T$$

After the aggregation operation is perform to the lowest branch of the structure, it will be described on higher level as:

$$x = \begin{bmatrix} 0.3 \\ \begin{bmatrix} 0.2 \\ 0.5 \\ 0.9 \end{bmatrix} \\ 0.1 \end{bmatrix}^T$$

Finally, the fuzzy signature structure will be:

$$x = \begin{bmatrix} 0.3 \\ 0.2 \\ 0.1 \end{bmatrix}^T$$

Each of these signatures contains information relevant to the particular data point x_0; by going higher in the signature structure, less information will be kept. In some operations it is necessary to reduce and aggregate information to become compatible with information obtained from another source (some detail variables missing or simply being locally omitted). Such is when interpolation within a fuzzy signature rule base is done, where the fuzzy signature flanking an observation are not exactly of the same structure. In this case the maximal common sub-tree must be determined and all signature must be reduced to that level in order to be able to interpolate between the corresponding branches or roots in some cases.

III. HIERARCHICAL DATA STRUCTURE

When dealing with a very large data set, it is possible that they hide hierarchical structure that appears in the sub-variable structures. In this case, the normal fuzzy signature structure as described in the previous section will be huge. In this case, it is necessary to identify the hierarchical structure hidden in the large data set. This section discusses the method used to generate hierarchical fuzzy system from data with hierarchical nature. In the next section, the method is extended to generate hierarchical fuzzy signature.

The main requirement of a reasonable $\prod$ is that each of its elements D_i can be modelled by a rule base with local validity. In this case, it is reasonable to expect D_i to contain homogeneous data. The problem of finding $\prod$ can thus be reduced to finding homogeneous structures within the data. This can be achieved by clustering algorithms.

The subspace Z_0 is used by meta-rules to select the most appropriate sub-rule base to infer the output for a given observation (i.e. system input). In general, the more separable the elements in $\prod$, the easier the sub-rule base selection becomes. Therefore, the proper subspace can be determined by means of some criterion that ranks the importance of different subspaces (combination of variables) based on their capability in separating the components D_i.

Unfortunately, the problems of finding Π and Z_0 are not independent of each other. Consider the following helicopter control fuzzy rules extracted from the hierarchical fuzzy system in [5]:

If *distance* (from obstacle) is *small* then *Hover*
Hover: if (helicopter) *body* rolls *right* then
move *lateral* stick *leftward*
if (helicopter) *body* pitches *forward* then
move *longitudinal* stick *backward*

Geometrically, an example of the data corresponding to the fuzzy rules is illustrated in Figure 2. From the figure, three clusters can be clearly observed in x_1. However, from x_2 and x_3, the data points are evenly distributed across the entire domain. Therefore, finding a 'good' subspace Z_0 where the clusters are clearly separable is crucial for the success of the clustering algorithm.

On one hand, we need to rely on the feature selection technique to find a 'good' subspace Z_0 for the clustering algorithm to be effective. However, most of the feature selection criteria assume the existence of the clusters beforehand.

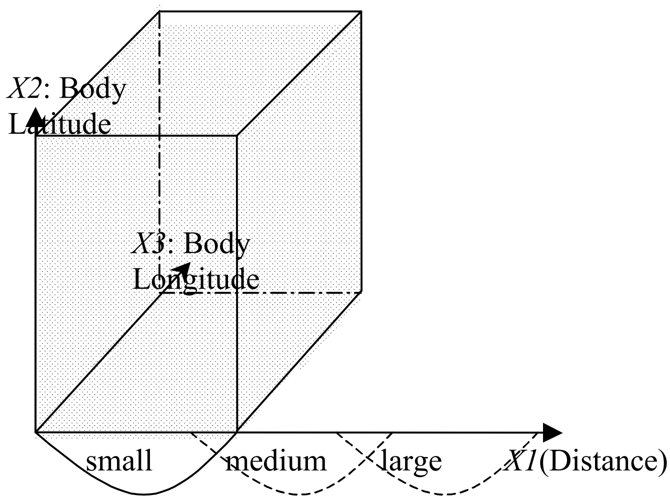

Figure 2: Geometric representation data in a hierarchical data structure

One way to tackle this problem is to adopt a projection-based approach. The data is first projected to each individual dimension and fuzzy clustering is performed on the individual dimensions. The number of clusters *C* for clustering can be chosen arbitrarily large. With the set of one-dimensional fuzzy clusters obtained, the separability (importance) of each input dimension can be determined separately. With the ranking, we can then select the *n* most important features to form the *n*-dimensional subspace Z_0.

In this case here, a hierarchical fuzzy system can be obtained using the following steps:

1. Perform fuzzy c-means clustering on the data along the subspace Z_0. The optimal number of clusters, C, within the set of data is determined by means of the FS index [12].

2. The previous step results in a fuzzy partition $\prod = \{D_1,...,D_C\}$. For each component in the partition, a meta rule is formed as:

   If Z_0 is D_i then use R_i

3. From the fuzzy partition Π, a crisp partition of the data points is constructed. That is, for each fuzzy cluster D_i, the corresponding crisp cluster of points is determined as

   $$P_i = \{p \mid \mu_i(p) > \mu_i(p) > \mu_j(p) \forall j \neq i\}.$$

4. Construct the sub-rule bases. For each crisp partition P_i, apply a feature extraction algorithm to eliminate unimportant features [13]. The essence of this feature selection is to look for cylindricity in the input variables. The remaining features (know as true inputs) are then used by a fuzzy rule extraction algorithm to create the fuzzy rule base R_i. Here, the use of the projection-based fuzzy modeling approach could be used[14].

IV. HIERARCHICAL FUZZY SIGNATURE STRUCTURE

In the previous section, the concept of finding interclass separability and the algorithm of constructing the hierarchical fuzzy rules have been discussed. In this section, we will take a look at how those concepts can be used to construct the hierarchical fuzzy signature structure. Normally in a large database, it consists of different hierarchical structure and sub-variables for different purposes. It is therefore necessary to reduce the structure and remove any redundant information when presenting the final fuzzy signature structure to human. The main objective of this research is to deal with large data set, hopefully to generate a reduced fuzzy signature structure that omits unnecessary information. Each item of the data has their own fuzzy signature structure, i.e. a set of treelike structured features.

The propose algorithm for constructing the hierarchical fuzzy signature structure technique is as follows:

1. Rank the importance of input variables, using the interclass separability criterion described in the previous section.

2. For each cluster C_i, eliminate the variable that cylindricity is found.

3. If there is more level in that variable, move one level down and perform sub cluster identification, i.e. repeat steps 1 and 2.
4. Step 3 is repeated until no more variable can be removed.

5. The reduced fuzzy signature structure is considered a hierarchical fuzzy signature structure.

6. Repeat step 2 for the next identified cluster C_i, until all clusters have been processed.

By using the example in Figure 1. Let us assume that a large database is given: $S = \{x_i\}$. A hierarchical fuzzy signature structure based on S can be constructed as follow:

1. Reduce the data to the highest level (to vectorial fuzzy values), which represents the least details model of the data set, i.e.

   $$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}^T$$

2. By applying the interclass separability technique, where potential cylindric clusters C_i can be determined, i.e. where data points are evenly distributed along the dimension.

3. For each cluster C_i, identification of the cylindricity is performed. Suppose, C_1 and C_2 were found in x_1, where C_1 is cylindric in x_2 and C_2 is cylindric in x_3.

4. First, investigate C_1, and within C_1, x_3 can be moved one level down to x_{31} and x_{32}.

5. The interclass separability is performed at this local area to see whether a local reduction can be done.

6. Similarly for C_2, the model is split into x_{21}, x_{22}, and x_{23} by going down one level.

7. Again a local interclass separability is performed. If x_{23} was found most separating and it gives C_{21} and C_{22}, further in C_{21}, x_{21} is eliminated, while in C_{22}, x_{22} is eliminated.

8. For C_{22}, it can be kept as such, however for C_{21}, it can be split further more to x_{221}, x_{222}, and x_{223}.

9. Until this point without going further into x_{221}, x_{222}, and x_{223}, we already know x_{22} and x_3 can be eliminated. Thus, the reduced signature structure for C_2 at this point is:

$$x = \begin{bmatrix} x_1 \\ \begin{bmatrix} x_{221} \\ x_{222} \\ x_{223} \end{bmatrix} \\ x_{23} \end{bmatrix}^T$$

If further processing is continued for the above case, the interclass separability check can be performed on x_{221}, x_{222}, and x_{223} and may be further sub-clusters C_{22i} can be obtained. Thus possible further reduction of the three may be performed locally. The steps are repeated until all lower levels have been visited to obtain a full hierarchical fuzzy signature structure.

V. Conclusions

We have described a technique for fuzzy signature structure for dealing with large database. The hierarchical fuzzy signature structure presented can be constructed by feature selection and interclass separability to arrive at a reduced fuzzy signature structure which best describes the complex and interdependent features. This hierarchical fuzzy signature structure can also handle missing data and remove any redundant information to solve a problem. This will greatly assist human expert in understanding the problem and making decision. The future work for this hierarchical fuzzy signature structure will be to examine the use of fuzzy rules interpolation concept to reduce rules in cases where cylindricity cannot be found.

References

[1] L.A. Zadeh, "Fuzzy Algorithm", *Information and Control*, vol. 12, 1968, pp. 94-102.

[2] M. Sugeno, and T. Takagi, "A New Approach to Design of Fuzzy Controller," *Advances in Fuzzy Sets, Possibility Theory and Applications*, 1983, pp. 325-334.

[3] H.T. Nguyen, and M. Sugeno, (Eds) *Fuzzy Systems: Modeling and Control*, The Handbook of Fuzzy Sets Series, Kluwer Academic Publishers, 1998.

[4] L.T. Kóczy, and K. Hirota, "Approximate reasoning by linear rule interpolation and general approximation," *Int. J. Approx. Reason*, Vol. 9, 1993, pp.197-225.

[5] M. Sugeno, T. Murofushi, J. Nishino, and H. Miwa, "Helicopter flight control based on fuzzy logic," *Proceedings of Fuzzy Engineering toward Human Friendly Systems '91*, 1991, pp. 1120-1124.

[6] L.T. Kóczy, and K. Hirota, "Approximative inference in hierarchical structured rule bases," *Proceedings of Fifth IFSA World Congress,* 1993.

[7] L.T. Kóczy, T. Vámos, T., and Gy. Biró, "Fuzzy Signatures," *Proceedings of EUROFUSE-SIC '99*, 1999, pp. 210-217.

[8] T. Vámos, L.T. Kóczy, Gy. Biró, "Fuzzy Signatures in Data Mining," *Proceedings of IFSA World Congress and 20th NAFIPS International Conference, 2001*, pp. 2842-2846.

[9] T.D. Gedeon, L.T. Kóczy, K.W. Wong, and P. Liu, "Effective Fuzzy Systems for Complex Structured Data," *Proceedings of IASTED International Conference Control and Applications (CA 2001)*, 2001, pp. 184-187.

[10] J.A. Goguen, "L-fuzzy Sets," *J. Math. Anal. Appl.,* vol. 18, 1967, pp. 145-174.

[11] L.T. Kóczy, "Vectorial I-fuzzy Sets," in: M. M. Gupta & E. Sanchez (eds.): *Approximate Reasoning in Decision Analysis*, North Holland, Amsterdam, 1982, pp. 151-156.

[12] Y. Fukuyama, and M. Sugeno, "A new method of choosing the number of clusters for fuzzy c-means method," in *Proceedings of the 5th Fuzzy Systems Symposium*, 1989, pp. 247-250.

[13] A. Chong, T.D. Gedeon, and L.T. Kóczy, "Feature selection and subspace clustering for hierarchical fuzzy rule extraction," CIMSA. 2003, submitted, Paris.

[14] A. Chong, T.D. Gedeon, and L.T. Kóczy, "Projection Based Method for Sparse Fuzzy System Generation," in *Proceedings of 2nd WSEAS Int. Conf. on Scientific Computation and Soft Computing*, 2002, pp. 321-325.