Using Data Mining for Learning and Clustering FCM

Somayeh Alizadeh, Mehdi Ghazanfari, Mohammad Fathian

Abstract—Fuzzy Cognitive Maps (FCMs) have successfully been applied in numerous domains to show relations between essential components. In some FCM, there are more nodes, which related to each other and more nodes means more complex in system behaviors and analysis. In this paper, a novel learning method used to construct FCMs based on historical data and by using data mining and DEMATEL method, a new method defined to reduce nodes number. This method cluster nodes in FCM based on their cause and effect behaviors.

Keywords—Clustering, Data Mining, Fuzzy Cognitive Map (FCM), Learning.

I. INTRODUCTION

IN decision making process, there are some critical components. In most of the time, numbers of these critical components are numerous and they affect each other, so analyzing them is not easy. The efficiency of decision-making depends largely on the ability of decision-makers to analyze the complex cause and effect relationships and take productive actions based on the analysis. In complex systems, different components affect each other, and these cause and effect relations show system behavior. Cause and effect are two different concepts. Causes tell the reason why something happened, whereas effects are the results of that happening. In most of the systems, managers draw a system conceptualization graph to understand all of the system aspect. This diagram shows the cause and effect relations between system components. The information about these relations generated and enriched over time with the experience of managers who are expert in that field. There are two big challenges, at first, if there is no expert to construct the above mental model how this must be drawn and secondly, if there is a way to construct that diagram with more components, how they could be analyzed. Therefore, a new mechanism must be used to bridge these two gaps and constituted with experts in first case and cluster the components into similar categories based on their behaviors for the second one.

. Somayeh Alizadeh is PHD candidate of industrial engineering department of IUST university (corresponding author phone:+982188304468; e-mail: S_alizadeh@ind.iust.ac.ir).

Mehdi Ghazanfari is associate professor of industrial engineering department of IUST university; e-mail: Mehdi@iust.ac.ir).

Mohammad Fathian is assistant professor of industrial engineering department of IUST university, e-mail: Fathian@iust.ac.ir).

This article organized as follows: In this section recalls the basic concepts and definitions of Data mining, Fuzzy Cognitive Map (FCM) learning background, DEMATEL method while in section 3. A new model for learning FCM and clustering it is defined.

II. THEORETICAL BACKGROUND

A. Data mining and clustering

Data mining or knowledge discovery in databases (KDD) is the process of discovering useful knowledge from large amount of data stored in databases, data warehouses, or other information repositories.[1] Data mining is a hybrid disciplinary that integrates technologies of databases, statistics, artificial intelligent. Recently, a number of data mining applications and prototypes have been developed for a variety of domains, including marketing, banking, finance, manufacturing, and health care other types of scientific data. [2, 3] The more common model functions in data mining include Classification, Clustering, Discovering association rules, Summarization, Dependency modeling and Sequence analysis. [3] Soft computing methodologies like fuzzy sets, neural networks, and genetic algorithms are most widely applied in the data mining. Fuzzy sets copes with uncertainty in data mining process. Neural networks are used for classification and rule generation. Genetic algorithms (GAs) simulated Annealing (SA), Tabu search (TS) are involved in various optimization and search processes. [4] All of these functions and methodologies tried to discover knowledge from historical data. This knowledge represented in the form of rules most of the time.

1) Clustering in data mining

The process of grouping a set of objects into classes of similar objects is called clustering. A cluster is a collection of data objects that are similar to one another within the same cluster and are dissimilar to the objects in other clusters. A cluster of data objects can be treated collectively as one group and so may be considered as a form of data compression.

As a branch of statistics, cluster analysis has been extensively studied for many years, focusing mainly on distance-based cluster analysis. Cluster analysis tools based on k-means, k-medoids, and several other methods have also been built into many statistical analysis software packages or systems, such as S-Plus, SPSS, and SAS (Clementine). In machine learning, clustering is an example of unsupervised learning. In general, the major clustering methods can be

classified into some categories: Partitioning methods, Hierarchical methods, Model based method; Grid based method, Clustering high-dimensional data and Constraint-based clustering. Further information could be found in this reference. [3]

B. Fuzzy Cognitive Map (FCM) and causal loop diagram learning

Cognitive maps were initially introduced by Robert Axelrod in 1976 and applied in political science [5]. Also it was used in numerous areas of application such as analysis of electrical circuits [6], medicine [7], supervisory systems [8], organization and strategy planning [9], [10], analysis of business performance indicators [11], software project management [12, 13], Information retrievals [14] modeling of plant control [15], system dynamics and complex systems [16,17,18,19, 20, 21] and modeling virtual world [22].

This model contains components and their corresponding relations, which may be positive, negative, or neutral. A cognitive map is a directed graph that its nodes correspond to relevant concepts and the edges state the relations between every two nodes by a sign. A positive sign implies a positive relation; moreover, any increase in its source value leads to increase in its target value. A negative sign presents negative relation and any increase or decrease in its source value leads to reverse effect to its target value. In a cognitive map if there is no edge between two nodes it means that, there is no relation between them.

In 1988, Kosko introduced a new extension to cognitive map [23]. It named fuzzy cognitive maps. In a simple fuzzy cognitive map, the relation between two nodes is determined by taking a value in interval [-1, 1]. While -1 corresponds to the strongest negative, +1 corresponds to strongest positive one. The other values express different levels of influence. This model can be presented by a square matrix called *connection matrix*. The value of relation between two nodes is set in their corresponding cell. In the connection matrix, row and column is associated with a source node and a target node, respectively. A simple FCM with five nodes and ten weighted
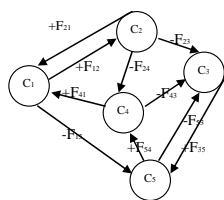

Fig. 1   A simple Fuzzy Cognitive Map (FCM)

arcs is depicted in Fig.1.

FCMs can produced by expert manually or generate by other source of information computationally. Experts developed a FCM or a mental model manually based on their knowledge in related area. In general, the manual procedures for developing FCM have occurred, when at least there is one expert who has expertise in the area under studied.

In some situations, developing a FCM manually becomes very difficult. When the experts were not able to express their expertise or even there is no expert in the related area, therefore there is a gap. For these reasons, the development of computational methods for learning FCM is necessary. For this purpose, not only the edge or casual relations between nodes, but also the strength on each edge must be achieved based on historical data. In this way, the expert knowledge is substituted by the knowledge, which produced from historical data by means and new computational procedures. Many algorithms for learning FCM model structure have been recently proposed.

One way in learning connection matrix of FCM is distance-based algorithm. In 1998, M.Schneider and el. worked on constructing fuzzy cognitive maps automatically. [24] Their method found not only the degree of similarity between any two variables (represented by numerical vectors), but also the relations between variables. D. Kardaras et al. presented similar method in 1998[25] in their paper assumed that every concept in an FCM can be represented by a numerical vector (V), whereas each element (v) of the vector represents a measurement of the concept.

Soft computing like neural network and genetic algorithm help data mining to discover appropriate knowledge in the form of Graph or Fuzzy Cognitive map (FCM) from historical data. Many scientists work on this area and investigated that FCM and its related connection matrix are learned and discovered by historical data. Many researchers worked on these areas by investigating FCM learning methods using historical data.

Kosko proposed a new model by use of simple Differential Hebbian Learning law (DHL) in 1994, but he used this model to learning FCMs without any applications [26]. This learning process modified weights of edges existing in a FCM in order to find the desired connection matrix. In 2002, Vazquez introduced a new extension to DHL algorithm presented by Kosko. He used a new idea to update edge values in a new formula. [27] Another method of learning FCMs based on the first approach (Hebbian algorithm), was introduced by Papageorgiou et al. in 2003. He developed another extension to Hebbian algorithm, called Nonlinear Hebbian Learning (NHL) [28]. Active Hebbian Algorithm (AHL) introduced by Papageorgiu et al. in 2004. In this method, experts not only determined the desired set of concepts, initial structure and the interconnections of the FCM structure, but also identified the sequence of activation concepts [29].

Another category in learning connection matrix of FCM is application of genetic algorithms or evolutionary algorithms. Koulouriotis et al. applied the Genetic Strategy (GS) to learn FCM structure in 2001 [30]. In mentioned model, they focused on the development of an ES-based procedure that determines the values of the cause-effect relationships (causality). Parsopoulos et al. also published other related papers in 2003. They tried to apply Particle Swarm Optimization (PSO) method, which belongs to the class of

Swarm Intelligence algorithms, to learn FCM structure [31, 32]. Khan and Chong worked on learning initial state vector of FCM in 2003. They performed a goal-oriented analysis of FCM and their learning method did not aim to compute the connection matrix, and their model focused on finding initial state vector for FCM [33]. In 2005, Stach et al. applied real-coded genetic algorithm (RCGA) to develop FCM model from a set of historical data in 2005 [34]. In 2005, Parsopoulos et al combined these two categories and published a paper about using evolutionary algorithms to train Fuzzy Cognitive Maps. In their model, they investigated a coupling of differential evolution algorithm and unsupervised Hebbian learning algorithm [35]. Other work to train a FCM was done by Konar in 2005. He worked on reasoning and unsupervised learning in a FCM. In that paper, a new model was introduced for unsupervised learning and reasoning on a special type of cognitive maps realized with Petri nets [36]. In 2007 M.Ghazanfari et al. published a paper about using Simulated Annealing and Genetic algorithm in FCM learning.[37] In that paper, they show that SA algorithm is better than GA in learning FCM with more nodes and introduced a new method to learn connection matrix rapidly. In this research, heuristic algorithms were used to learn FCM matrix. Table 1 shows a comparison between some existing methods:

TABLE I
OVERVIEW OF SOME LEARNING APPROACHES APPLIED TO FCMs

Algorithm	learning goal	Human interaction	type of data	transformation Function	NO of node	learning algorithm
DHL	Connection matrix	No	Single	N/A	N/A	Hebbian
BDA	Connection matrix	No	Single	Binary	5, 7, 9	Modified Hebbian
NHL	Connection matrix	Yes & No	Single	Continuous	5	Modified Hebbian
AHL	Connection matrix	Yes & No	Single	Continuous	8	Modified Hebbian
GS	Connection matrix	No	Multiple	Continuous	7	Genetic
PSO	Connection matrix	No	Multiple	Continuous	5	Swarm
GA	Initial vector	N/A	N/A	Continuous	11	Genetic
RCGA	Connection matrix	No	Single	Continuous	4,6,8,10	Genetic
GA	Connection matrix	No	Single	Continuous	Any Number	GA
SA	Connection matrix	No	Single	Continuous	Any Number	SA

The above table compares the mentioned methods based on several factors, such as learning goal, kind of input historical data, type of transformation function, size of FCM model, type of learning strategy and that whether experts are involved in model or not.

C. DEMATEL method

The Battelle Memorial Institute conducted the DEMATEL method project through its Geneva Research Centre [38]. In recent years, the DEMATEL method has become very popular, because it is especially practical and useful for visualizing the structure of complicated causal relationships with matrixes or digraphs. The matrixes or digraph shows relations between the components of the system. This model distinguishes the causes and effects between components and construct a structure presenting these two groups separately. The DEMATEL method has been successfully applied in many fields [39, 40, 41]. DEMATEL method is designed to find not only direct relations but also indirect relations in a graph. Besides finding the undirected relations, this method defines the cause and effect groups of elements.

A complex system contains a set of elements $C = \{C1, C2, \ldots Cn\}$, which have relations with each other. These relations can be modeled in a graph. The initial direct relation expressed with matrix **M,** connection matrix. The elements of this matrix normalized with Eq. 1:

$$e^{new}_{ij} = \frac{e^{old}_{ij}}{max\left(\sum_{i=1}^{n} e_{ij}\right)} \tag{1}$$

It was proven that the elements of the other matrixes like $(M^2, M^3, \ldots, M^t)$ are normal too. When limit $M^t_{t\to\infty} = 0$ and O is the null matrix and I is the identity matrix, Eq. 2 can be proved:

$$I+ M + M^2 + M^3 + \ldots + M^t = (I - M)^{-1} \tag{2}$$

Because

$S= M + M^2 + M^3 + \ldots + M^t = M \, (I\text{-} M^t)/(I\text{-} M)$ and if limit $M^t_{t\to\infty} = 0$ then

$$M\,(I\text{-}M^t)/(I\text{-}M) = M/(I\text{-}M) = M\,(I-M)^{-1}$$

Therefore $S= M\,(I-M)^{-1}$

Matrix S shows the directed and undirected relations between FCM elements $C=\{C1,C2,\ldots Cn\}$. If only the undirected relations are considered, then the first element of Eq. 2 (M) must be omitted and the equation will be written like:

$$S1 = M^2 + M^3 + \ldots + M^t = M^2\,(I - M)^{-1} \tag{3}$$

This matrix presents the undirected relations or the hidden relations, which had not been considered before. The equivalent FCM with this matrix is considered and the related rules discovered from this new FCM.

In DEMATEL, method cause and effect criteria are recognized and separated from each other. Sum of rows ($\sum_{i=1}^{n} R_i$) and the sum of columns ($\sum_{j=1}^{n} C_j$) of matrix **M** are computed and these two parameters are calculated for causal diagram.

($\sum_{i=1}^{n} R_i + \sum_{j=1}^{n} C_j$ and $\sum_{i=1}^{n} R_i - \sum_{j=1}^{n} C_j$). Max of $\sum_{i=1}^{n} R_i$ is an element which affect on the other elements and is a *dispatcher element* and Max of $\sum_{j=1}^{n} C_j$ is an element which the other elements affect on it and is a *receiver element*. Then a diagram is constructed with the horizontal axis ($\sum_{i=1}^{n} R_i + \sum_{j=1}^{n} C_j$) and the

vertical axis, $\sum_{i=1}^{n} R_i - \sum_{j=1}^{n} C_j$.

The vertical axis separates the elements into cause and effect groups if the value ($\sum_{i=1}^{n} R_i - \sum_{j=1}^{n} C_j$) is positive, it belongs to the cause group, If the value ($\sum_{i=1}^{n} R_i - \sum_{j=1}^{n} C_j$) is negative, it belongs to the effect group. The horizontal axis shows the important of each element.

The steps of DEMATEL method are demonstrated here:

o   Produce *connection matrix* (M) from related FCM

o   Replace all of the elements based on $e^{new}_{ij} = \dfrac{e^{old}_{ij}}{max\left(\sum_{i=1}^{n} e_{ij}\right)}$

o   The directed and undirected relation between all of the elements can be presented by matrix S like bellow: S= $M + M^2 + M^3 + ... + M^t = M (I\text{-}M^t)/(I\text{-}M)$ As limit $\lim_{t\to\infty} M^t = 0$ then $M (I\text{-}M^t) / (I\text{-}M) = M /(I\text{-}M) = M (I-M)^{-1}$

o   The undirected relation between all of the elements can be presented by matrix S1 like bellow: S1= $M^2 + M^3 + ... + M^t = M^2 (I\text{-}M^t)/(I\text{-}M)$ As limit $\lim_{t\to\infty} M^t = 0$ then $M^2 (I\text{-}M^t)/(I\text{-}M) = M^2 /(I\text{-}M) = M^2 (I-M)^{-1}$

III. A NEW MODEL FOR LEARNING FCM AND CLUSTERING ITS NODES INTO SIMILLAR CLUSTERS

If there is no expert to define these relationship, this method can define FCMs based on historical data. This methodology focused on automatic casual loop diagram construction and clustered its nods. Here are steps of this methodology:

o   *Description the problem.*

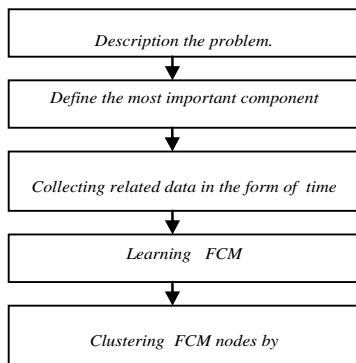o   *Define the most important components and parameters in*

Fig. 2   Methodology steps

that problem

o   *Collecting related data about these parameters during the time and normalizing these data and show them in the form of time series.*

After description the problem and defining the most important parameters, their related data must be collected. As these data show the components behavior during time, they can be shown with the form of time series. But they must be preprocessed and normalized to be prepared for learning. These data are normalized according Eq.4:

$$Normal(c_i) = \frac{c_i - c_{Min}}{c_{Max} - c_{Min}} \qquad (4)$$

In this formula every element (c_i) normalized between [0,1]. c_{Min} is minimum value and c_{Max} is the maximum value.

o   *Learning FCM based on their historical behavior.*

In this step, a solution for automatic construction of Fuzzy Cognitive map is found by using Simulated Annealing, Genetic Algorithm or Tabu search. The focus of this model is to determine cause-effect relationships (causality) and their strength.

As mentioned before, a cause-effect relation is specified by a related Connection matrix. The elements of this matrix are the values of edges in the FCM. The aim of the proposed method is to find these elements. The relations between nodes and edges are calculated as:

$$C_i(t+1) = f\left(\sum_{i=1}^{n} e_{ji} C_j(t)\right) \qquad (5)$$

where e_{ij}'s are the elements of the matrix and f is a transform function which includes recurring relation on t >=0 between C(t + 1) and C(t) that can be presented by a logistic function like:

$$f(x) = \frac{1}{1 + e^{-cx}} \qquad (6)$$

Eq. (5) and Eq. (6) can be expressed by Eq. (7):

$$Output_i(t_{n+1}) = E \times Input_i(t_n) \qquad (7)$$

$Input_i(t_n)$ is input data for node i, $Output_i(t_{n+1})$ is its corresponding output data and E is the Connection matrix of FCM. Eq.(7) implies that corresponding output for every node can be calculated. E (Related Connection Matrix) is a vital factor in Eq.(7) which should be determined in the FCM learning process. The proposed FCM learning method forms structure of a FCM and is able to generate state vector sequences that transform the input vectors into the output vectors. When all real input and output values of a FCM are in hand, the most important step is to find a new solution for the FCM and calculate the estimated output related to this new FCM.

$$Output_i^{estimated}(t_{n+1}) = E^{proposed} \times Input_i(t_n) \qquad (8)$$

According to Eq. (8), $Output_i^{estimated}(t_{n+1})$ is the estimated output and $Input_i(t_n)$ is its corresponding input for the *i*th

node. $E^{proposed}$ is new proposed matrix. The real output is $Output_i^{real}$ (t_{n+1}) and Eq. (9) calculates the difference between real and estimated outputs:

$$Error = Output_i^{estimated} (t_{n+1}) - Output_i^{real} (t_{n+1}) \qquad (9)$$

By using the later two equations, the objective is defined as minimizing the difference between real and estimated outputs. This objective is defined as:

$$Total\_Error = \sum_{n=1}^{K} \sum_{i=1}^{N} Output_i^{estimated} (t_{n+1}) - Output_i^{real} (t_{n+1}) \qquad (10)$$

Where N is the number of nodes and K is the iteration.

$$Input_i(t_n) \rightarrow Output_i(t_{n+1}) \qquad \forall t = 0,..., K-1$$

If $Input_i(t_n)$ defined as an *initial vector*, and $Output_i(t_{n+1})$ as *system response*, *K-1* pairs in the form of {*initial vector*, *system response*} can be generated from the input data.

As mentioned in before, there are many methods to construct FCM matrix automatically, for example, Stach et al. constructed this matrix by a Real Code Genetic Algorithm (RCGA) with simple operators and Ghazanfari et al. constructed this matrix by SA. [34, 37] In this paper, we compared GA, SA and TS to learn FCM and the best FCM is chosen.
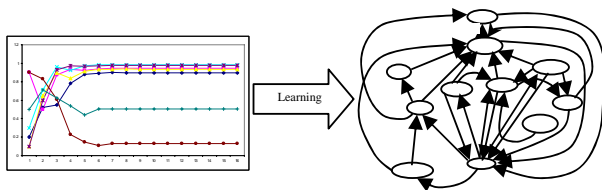

Fig. 3   Learning FCM step

Genetic algorithm is used to find the near optimum solution. Simulated annealing or Tabu search are used not only to find the near optimum solution, but also to escape the local minimum solution and to improve the optimum solutions. In this paper assumed that readers are familiar with GA and SA and TS algorithms. A useful summary about relevant GA and SA and TS can be found in [42, 43, 44].

The aim of the experiments is to assess quality of the proposed methods for learning FCMs. In general, the goal of learning FCM is to find FCM connection matrix that generates the same state vector sequence as the input data for a given initial state vector. Now, an important parameter is considered. This variable measures similarity between the input data, and data generated by simulating the candidate FCM with the same initial state vector as the input data. The criterion is defined as a normalized average error between corresponding concept values at each iteration between the two state vector sequences:

$$error = (\frac{1}{(K-1) \times N}) \sum_{n=1}^{K} \sum_{i=1}^{N} \left(Output_i^{estimated} (t_{n+1}) - Output_i^{real} (t_{n+1}) \right)^2 \qquad (11)$$

As mentioned before, in above formula $Output_i^{real} (t_{n+1})$ is the value of a node at iteration t in the input, $Output_i^{estimated} (t_{n+1})$ is the estimated value of that node at iteration t from simulation of the candidate FCM, K is the input data length and N is the number of nodes. All experiments were performed with logistic transformation function, this function is parameterized as:

$$F(x) = \frac{1}{1 + e^{-0.2x}} \qquad (12)$$

In our experiment, for comparing the metaheuristic results, some test problems were solved by using GA, SA and TS on a PC Pentium IV, 1.6 GHz. The metaheuristic algorithms were developed using Visual Basic 6. Three algorithms, GA, SA and TS ran with different Node numbers: 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15 and for every run, the Error and time consuming saved. Each considered FCM, in terms of the number of nodes, was simulated 10 times with the three algorithms, which totaled in 12*10*3 experiments. The obtained results are shown in table 2.

Considering the results of Tables (2), the presented metaheuristic algorithms are able to find and report the near-

TABLE II
EXPERIMENT RESULTS WITH DIFFERENT FITNESS FUNCTIONS AND TIMES

Node NO	Genetic algorithm		Simulated Annealing		Tabu Search	
	Fitness Function	Time	Fitness Function	Time	Fitness Function	Time
	(Error)	sec	(Error)	sec	(Error)	sec
3	3.49E-06	50	4.50E-05	0.7187	7.17E-07	2.046
4	4.40E-06	50	3.15E-05	0.9843	6.20E-07	4.2343
5	5.88E-06	50	3.44E-05	1.2343	3.49E-07	8.21
6	7.80E-06	50	5.37E-05	1.4843	4.73E-07	14.1093
7	2.33E-05	50	4.82E-05	1.8125	8.83E-07	22.79
8	3.52E-05	50	7.53E-05	2.1875	5.59E-07	35.54
9	4.02E-05	50	9.20E-05	2.5625	7.24E-07	50
10	3.90E-05	50	9.84E-05	2.8906	1.29E-06	50
11	1.52E-04	50	8.31E-05	3.3281	3.46E-06	50
12	1.56E-04	50	8.17E-05	3.8125	3.72E-06	50
13	1.91E-04	50	1.07E-04	4.3593	1.41E-05	50
14	3.06E-04	50	1.31E-04	4.9531	7.31E-05	50
15	4.85E-04	50	1.61E-04	5.5781	8.63E-05	50

optimal and promising solutions in a reasonable computational time. This indicates the success of the proposed algorithms. In general, we can conclude the following results:

All heuristic algorithms found the near-optimal solutions and the results of these experiments show that these algorithms gradually converge into a high-quality candidate

FCM. Three examples of FCM learning experiments based on GA, SA and TS are plotted in Fig.4-1, 4-2 and 4-3.
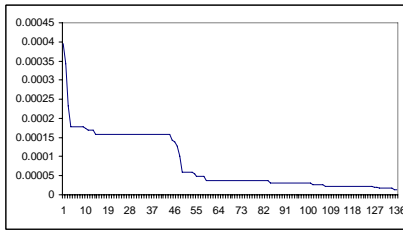

Fig. (4-1) An example of GA fitness Function which show that error coverage to near zero
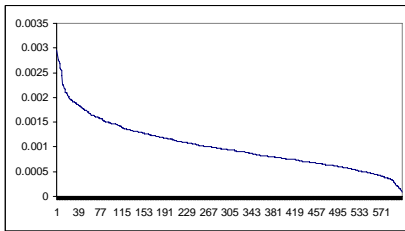

Fig. (4-2) An example of SA fitness Function which show that error coverage to near zero
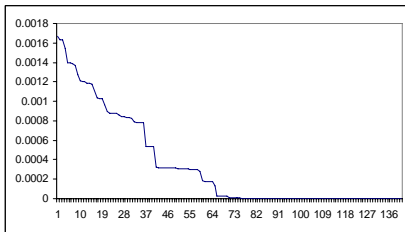

Fig. (4-3) An example of TS fitness Function which show that error coverage to near zero

o *Clustering FCM nodes by DEMATEL and Clustering algorithm*

Now there is a connection matrix which is corresponding to causal effect diagram and was learned by historical data. In this step, nodes which have similar behaviour categorized in one cluster. This clustering is useful for FCM with more nodes. In those case, all of the nodes which have simmilar effects on other nodes are in the same category and it is better to cluster them in one cluster. For this reason we have to define their cause and effect behavior based on causal patterns and categorize them in some clusters.

DEMATEL method calculates new properties for every node. These properties show whether they are a *dispatcher element* or a *receiver element* and their importance. As mentioned before, every element normalized based on $max\left(\sum_{i=1}^{n} e_{ij}\right)$ and this new calculated connection matrix used to produce a new matrix, which shows not only direct but also

indirect relations $S = M \ (I - M)^{-1}$. After that, Sum of rows ($\sum_{i=1}^{n} R_i$) and the sum of columns ($\sum_{j=1}^{n} C_j$) of S matrix are computed and these two parameters are calculated for causal diagram. Then a diagram is constructed with the horizontal axis ($\sum_{i=1}^{n} R_i + \sum_{j=1}^{n} C_j$) and the vertical axis, ($\sum_{i=1}^{n} R_i - \sum_{j=1}^{n} C_j$).
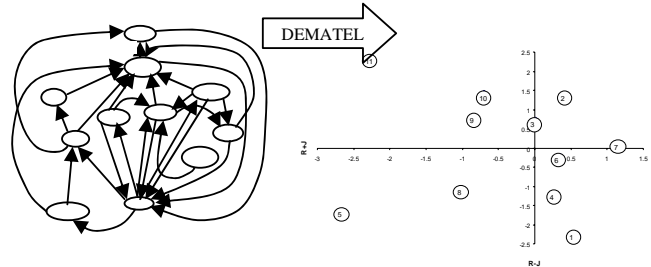

Fig. 5 DEMATEL method

Above figure shows this step and draw all graph elements based on their properties in a diagram. As mentioned before, horizontal axis shows the importance and the vertical axis shows whether they are *dispatcher* or *reciever*. Now there are two properties ($\sum_{i=1}^{n} R_i + \sum_{j=1}^{n} C_j$) and, ($\sum_{i=1}^{n} R_i - \sum_{j=1}^{n} C_j$) that all nodes can be clusterd based on them. As mentioned ibefore, there is some clustering methods. Here we used K-Means method in Clementine software. In this step number of clusters is considered 5. There is two properties for these nodes ($\sum_{i=1}^{n} R_i + \sum_{j=1}^{n} C_j$) and ($\sum_{i=1}^{n} R_i - \sum_{j=1}^{n} C_j$). At first, all nodes are clustered based on $\sum_{i=1}^{n} R_i - \sum_{j=1}^{n} C_j$ property. Nodes which are in one cluster have similar behavior and are different with the others in other clusters. In FCM with more nodes these clusters are very helpful. Because center of these clusters are representative of all nodes in that clusters and analyzing of this FCM become easier than the one with more nodes. The related graph is constructed in the last step with 11 nodes. Here are these new clusters:
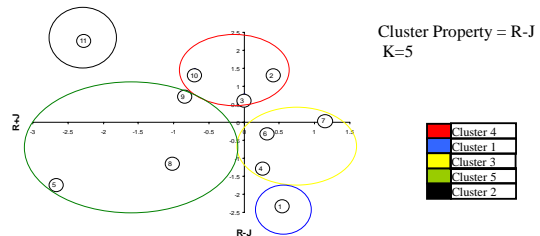

Fig. 6 clustering nodes with one property

In this way, nodes with number 10, 2, 3 are in cluster 4 , nodes with number 9,8,5 are in cluster 5, nodes with number 4,6,7 are in cluster 3, node 1 in cluster 1 and node 11 in

cluster2. Center of these clusters is shown in Table 3.

TABLE III
CLUSTER CENTER FOR FIG 6

NO	Center
1	-2.263
2	2.239
3	0.177
4	-1.328
5	1.159

As mentioned before, horizontal axis shows the importance and the vertical axis shows whether they are *dispatcher* or *receiver*. Now there are two properties ($\sum_{i=1}^{n} R_i + \sum_{j=1}^{n} C_j$) and,

($\sum_{i=1}^{n} R_i - \sum_{j=1}^{n} C_j$) that all nodes can be clustered based on them. For example in above figures all nodes can be clustered based on their behavior into 5 clusters.

If the importance of *dispatcher* or *receiver* nodes were important, so the other property must be considered in clustering is $\sum_{i=1}^{n} R_i + \sum_{j=1}^{n} C_j$. Now in new clustering not only their behaviors but also their importance considered. Here are these new clusters and their centers:

Fig. 7 clustering nodes Two property

TABLE IV
CLUSTER CENTER

NO	R-J	R+J
1	-1.73	0.44
2	2.24	-2.24
3	-1.65	-2.66
4	0.67	0.07
5	-1.14	-1.02

In these two above methods, experts defined number of clusters. Some ways choose the best cluster numbers. These numbers is the optimum numbers for clusters that all similar elements are together and all different elements are separated in the optimum way. In Clementine software there is a clustering method, which clusters nodes based on their properties with the optimum solution. It is Two Step clustering. Here are the clusters which produced by this new clustering method

TABLE V
CLUSTER CENTER

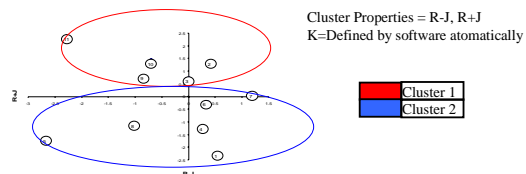| NO	R-J	R+J
1	1.5	-0.5
2	-0.95	-0.44

Fig. 8 clustering nodes with two properties,
without Number of cluster

Based on nodes clustering in a FCM, all similar nodes come into a cluster and center of these clusters are representative of all that similar nodes.

IV. CONCLUSION

In this study, a comprehensive learning method has been proposed to construct FCMs based on historical data. Some metaheuristics (GA, SA, and TS) have been used to extract FCMs in the proposed method and their results have been compared. When FCM is learned, some nodes with their relation is costructed. The FCM with more nodes could not be analysed easily. In this paper, clustering method is used to reduce the number of FCM nodes to analyse easily. One of interesting and open issues is using the other heuristic methods like Ant Colony for learning FCMs and comparing them with the others. Another interesting direction concerns the use of the learned FCM is ranking components in complex systems.

REFERENCES

[1] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, Eds., Advances in Knowledge Discovery and Data Mining. Menlo Park, CA: AAAI/MIT Press, 1996.
[2] S. K. Pal and S. Mitra, Neuro-Fuzzy Pattern Recognition: Methods in Soft Computing. New York: Wiley, 1999.
[3] Jiawei Han and Micheline Kamber, Data Mining : concepts and techniques, Morgan kaufman publishers, 2006
[4] Hussein Aly Abbass, Ruhul Amin, Sarker, Charles S. Newton. Data mining: a heuristic approach, 2002, Idea Group Publishing.
[5] R.Axelrod, Structure of Decision: The Cognitive Maps of Political Elites, Princeton University Press, Princeton, NJ, 1976.
[6] M.A. Styblinski, B.D. Meyer, Signal flow graphs versus fuzzy cognitive maps in application to qualitative circuit analysis, Internet. J. Man Mach. Studies 35 (1991) 175–186.
[7] V.C. Georgopoulos, G.A. Malandraki, C.D. Stylios, A fuzzy cognitive map approach to differential diagnosis of specific language impairment, J. Artif. Intel Med. 29 (3) (2003) 261–278.
[8] C.D. Stylios, P.P. Groumpos, Fuzzy cognitive map in modeling supervisory control systems, J. Intel. & Fuzzy Systems 8 (2) (2000) 83–98.
[9] M. G. Bougon, "Congregate Cognitive Maps: a Unified Dynamic Theory of Organization and Strategy," Journal of Management Studies, 29:369-389, (1992)
[10] K.C. Lee,W.J. Lee, O.B. Kwon, J.H. Han, P.I.Yu, Strategic planning simulation based on fuzzy cognitive map knowledge and differential game, Simulation 71 (5) (1998) 316–327.
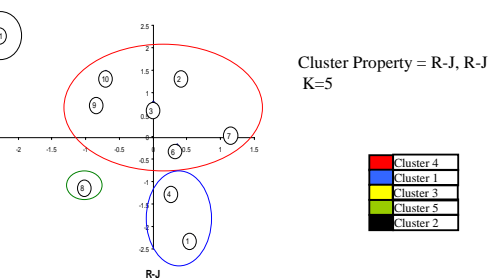
[11] D. Kardaras, G. Mentzas, Using fuzzy cognitive maps to model and analyze business performance assessment, in: J. Chen, A. Mital (Eds.), Advances in Industrial Engineering Applications and Practice II, 1997, pp. 63–68.

[12] W. Stach, L. Kurgan, Modeling software development project using fuzzy cognitive maps, Proc. 4th ASERCWorkshop on Quantitative and Soft Software Engineering (QSSE'04), 2004, pp. 55–60.

[13] W. Stach, L. Kurgan,W. Pedrycz, M. Reformat, Parallel fuzzy cognitive maps as a tool for modeling software development project, Proc. 2004 North American Fuzzy Information Processing Society Conf. (NAFIPS'04), Banff, AB, 2004, pp. 28–33.

[14] A. R. Montazemi, D. W. Conrath, "The Use of Cognitive Mapping for Information Requirements Analysis," MIS Quarterly, 10:44-55, (1986)

[15] K. Gotoh, J. Murakami, T.Yamaguchi, Y.Yamanaka, Application of fuzzy cognitive maps to supporting for plant control, Proc. SICE Joint Symp. 15th Systems Symp. and Tenth Knowledge Engineering Symp., 1989, pp. 99–104.

[16] Carvalho, J.P., Tomé, J.A.,"Rule Based Fuzzy Cognitive Maps and Fuzzy Cognitive Maps - A Comparative Study", Proceedings of the 18th International Conference of the North American Fuzzy Information Processing Society, NAFIPS99, New York

[17] Carvalho, J.P., Tomé, J.A.,"Rule Based Fuzzy Cognitive Maps- Fuzzy Causal Relations", Computational Intelligence for Modeling, Control and Automation, Edited by M. Mohammadian, 1999

[18] Carvalho, J.P., Tomé, J.A., "Fuzzy Mechanisms for Causal Relations", Proceedings of the Eighth International Fuzzy Systems Association World Congress, IFSA'99, Taiwan

[19] Carvalho, J.P., Tomé, J.A.,"Rule Based Fuzzy Cognitive Maps – Qualitative Systems Dynamics", Proceedings of the 19th International Conference of the North American Fuzzy Information Processing Society, NAFIPS2000, Atlanta

[20] D.E. Koulouriotis, I.E. Diakoulakis, D.M. Emiris, E.N. Antonidakis, I.A. Kaliakatsos, Efficiently modeling and controlling complex dynamic systems using evolutionary fuzzy cognitive maps (Invited Paper), Internet. J. Comput. Cognition 1 (2) (2003) 41–65.

[21] C.D. Stylios, P.P. Groumpos, Modeling complex systems using fuzzy cognitive maps, IEEE Trans. Systems Man, Cybern. Part A: Systems Humans 34 (1) (2004).

[22] D.E. Koulouriotis, I.E. Diakoulakis, D.M. Emiris, Anamorphous of fuzzy cognitive maps for operation in ambiguous and multi-stimulus real world environments, 10th IEEE Internet. Conf. on Fuzzy Systems, 2001, pp. 1156–1159.

[23] B. Kosko, Hidden patterns in combined and adaptive knowledge networks, Internet. J. Approx Reason. 2 (1988) 377–393.

[24] M. Schneider, E. Shnaider, A. Kandel, G. Chew, Automatic construction of FCMs, Fuzzy Sets and Systems 93 (2) (1998) 161–172.

[25] D. Kardaras, B. Karakostas "The use of fuzzy cognitive maps to simulate the information systems strategic planning process". Information and Software Technology 41 (1999) 197–210

[26] J.A. Dickerson, B. Kosko, Fuzzy virtual worlds, Artif. Intel. Expert 7 (1994) 25–31.

[27] A. Vazquez, A balanced differential learning algorithm in fuzzy cognitive maps, Technical Report, Departament de Llenguatges I Sistemes Informatics, Universitat Politecnica de Catalunya (UPC), 2002.

[28] E. Papageorgiou, C.D. Stylios, P.P. Groumpos, Fuzzy cognitive map learning based on nonlinear Hebbian rule, Australian Conf. on Artificial Intelligence, 2003, pp. 256–268.

[29] E. Papageorgiou, C.D. Stylios, P.P. Groumpos, Active Hebbian learning algorithm to train fuzzy cognitive maps, Internat. J. Approx. Reason. 37 (3) (2004) 219–249.

[30] D.E. Koulouriotis, I.E. Diakoulakis, D.M. Emiris, Learning fuzzy cognitive maps using evolution strategies: a novel schema for modeling and simulating high-level behavior, IEEE Congr. On Evolutionary Computation (CEC2001), 2001, pp. 364–371.

[31] K.E. Parsopoulos, E.I. Papageorgiou, P.P. Groumpos, M.N. Vrahatis, A first study off uzzy cognitive maps learning using particle swarm optimization, Proc. IEEE 2003 Congr. on Evolutionary Computation, 2003, pp. 1440–1447.

[32] E.I. Papageorgiou, K.E. Parsopoulos, C.D. Stylios, P.P. Groumpos, M.N. Vrahatis, Fuzzy cognitive maps learning using particle swarm optimization, J. Intel. Inform. Systems, in press. 2003

[33] M. Khan, A. Chong, Fuzzy cognitive map analysis with genetic algorithm, Proc. 1st Indian Internat. Conf. on Artificial Intelligence (IICAI-03), 2003

[34] W. Stach, Lukasz Kurgan,Witold Pedrycz, Marek Reformat Genetic learning of fuzzy cognitive maps, Fuzzy Sets and Systems 153 (2005) 371–401

[35] E Papageorgiou,C.Stylios P. Groumpos, Unsupervised learning techniques for fine-tuning fuzzy cognitive map causal links, Int. J. Human-Computer Studies 64 (2006) 727–743

[36] Amit Konar, Uday K. Chakraborty , Reasoning and unsupervised learning in a fuzzy cognitive map , Information Sciences 170 (2005) 419–441

[37] M.Ghazanfari, S.Alizadeh, M.Fathian, D.E.Koulouriotis, Comparing Simulated Annealing and Genetic Algorithm in Learning FCM, Applied Mathematics and Computation (2007), doi: 10.1016/ j.amc.2007.02.144

[38] Gabus, A. and Fontela, E. Perceptions of the World Problematique: Communication Procedure, Communicating With Those Bearing Collective Responsibility (DEMATEL Report No.1). Battelle Geneva Research Centre, Geneva, Switzerland. (1973)

[39] Yamazaki, M., Ishibe, K. and Yamashita, S. An analysis of obstructive factors to welfare service using DEMATEL method. Reports of the Faculty of Engineering, Yamanashi University, pp. 48 25-30. (1997).

[40] S.M. Seyed-Hosseini, N. Safaei, M.J. Asgharpour , Reprioritization of failures in a system failure mode and effects analysis by decision making trial and evaluation laboratory technique, Reliability Engineering and System Safety (2005) 1–10

[41] Goodman, R. (1988). Introduction to Stochastic Models. Benjamin/Cummings Publishing Company Inc., California, U.S.A.

[42] Marc Pirlot, General local search methods, European journal of operational research 92, 1996 , 493-511

[43] Duc Truong Pham and Dervis Karaboga, Intelligent Algorithms, tabu search, simulated annealing and neural networks, Springer, New York, 1998

[44] van Laarhoven, P. and Aarts, E. (1987): Simulated Annealing: Theory and Applications. Dordrect: Reidel.