

Study of FCM Model for Fault Clustering Using Optimization Method

Yuan Haibin & Xu Jiaojiao

School of Automation Science and Electrical Engineering, Beihang University, Beijing, China

ABSTRACT: In this paper, Fuzzy Cognitive Map (FCM) is employed and discussed to analyze time series that represents fault dynamics. At first, the essential of FCM prototype is shown, and diversified data set that are produced from the prototype to construct candidate FCM model is proposed. The Particle Swarm Optimization (PSO) and Simulated Annealing (SA) learning algorithm is taken as optimization method based on computational intelligence. Secondly, analysis of representation for time series from real world application is carried out with the proposed FCM model to assess the quality of FCM design methodology and algorithmic performance. The dynamics of time series is depicted by fuzzy c-mean clustering, and the concept of the node in the candidate FCM model corresponds to the activation level of each cluster center. In the end, the parameters of node number in the candidate FCM model and steepness of activation function are taken into consideration in order to obtain the better result. The discussion and findings are offered and the results show that the proposed method is desirable in the representation of time series both at numeric value level and linguistic level.

KEYWORD: Fuzzy Cognitive Map (FCM); Prediction; Fault; Clustering; Optimization

1 INTRODUCTION

Application analysis of time series representation is diversified in many fields, such as signal processing and fault management (Sapello et al. 2010), stock market, business and power system (H.K.Yu. 2005; Chow, 2003; Wang et al. 2011). Comprehensive methods and algorithms have been proposed recently in order to exploit the dynamics of time series by representing the features that extracted from it. One of the approaches is knowledge representation regarding the domain of application to clarify the dynamic behavior of the time series. However, this approach lies in the disadvantage of domain knowledge restrictions, subjective inclination and overhead in dealing with large complex system. In this paper, FCM model is proposed to represent the dynamics of the time series using information clustering method and computational intelligence technique in order to develop semi-automated or automated design method to overcome the aforementioned drawbacks from expert based method. FCM was first proposed by Kosko (Kosko. 1986), which can be used to represent the relationships between the elements. The extension of FCM can be used to model complex systems in which a precise mathematical model is difficult to

represent and can also provide an approach to knowledge representation by defining concept according to application requirement. Due to the desirable properties of abstraction, flexibility and adaptability, the design of many different FCM models concerning real world application have been developed, especially in the case of failure analysis and reliability assessment (S.-H. K, 1999). However, little work is carried out to build FCM model for the representation and predication of time series.

In this study, we focus on developing the universal FCM model in an abstract way, and the up-to-date learning method is also employed to train the model so that the scheme of constructed architecture of the model can be suitable for analysis of time series representation. The features that imply the dynamic behavior of the time series is originated from clustered data set, and the treatment of data set with proper FCM model is carried out. In this way, the evolution of the time series is depicted by the state set in the FCM model. It is emphasized that the weights accounts for the cause-effect relationships among the nodes in a FCM model, the weights can be optimized by particle swarm optimization (PSO) method or simulated annealing (SA) method. With the FCM model is proposed, we take a further step toward the performance assessment in order to

capture and emulate the essence of numerical data set produced from real world time series. The main motivations and contributions of the study are as follows:

(1) Comprehensive analysis of FCM model is carried out. Each node within the FCM model is regarded as input and output from current state to the next state respectively. Given the former data value, the successive data points can be reached according to the computation of the developed FCM topology, as a result, the value with regard to the node can be either numerical or linguistic, which can be in the form of specific data or the activation level of fuzzy set.

(2) The formation of candidate FCM model is studied. The determination of proper weight matrix in the model is regarded as optimization objective problem and PSO is employed as learning method. The candidate FCM model is highly abstract in concept and is also easy to construct.

(3) Attempt is made for analysis of time series representation based on proposed method. Amplitude and change of amplitude is clustered to create candidate FCM model. Diversified parameter configuration regarding candidate FCM model is studied as well.

2 VALIDATION OF FCM ARCHITECTURE

2.1 Fundamentals and formation of FCM

The structure of a FCM is based on graph topology and can be conveniently observed visually. Generally speaking, FCM is capable of describing a system by nodes and connections. In a FCM model, the nodes are also referred to as concepts, events, actions, goals or trends of the system, each node in the FCM stands for a state of the system at specific time. Each connection between two nodes also refers to weight, edge, arc or interaction of the nodes. The connection depicts the cause-effect relationship from one node to another. Detailed definition and implementation regarding nodes and connections are determined by different application aspect of system, the nodes and connections can be added or removed on demand as well.

A weight exists between two nodes can be indicated as positive causality, negative causality and neutral causality. In FCM, the weight takes the value in the interval $[-1, 1]$, which describes the strength of the corresponding causality. In Figure 1, the weight W_{ij} indicates the cause-effect relationship from source node i to node j holds some conditions. The condition of $W_{ij} \in [-1, 0)$ means the increase (decrease) in the value of node i can lead to the increase (decrease) in the value of node j , the condition of $W_{ij} \in (0, 1)$ means the increase (decrease) in the value of node i can lead to the decrease (increase) in the value of node j , whereas

$W_{ij}=0$ means the change in the value of node i does not effect the value of node j .

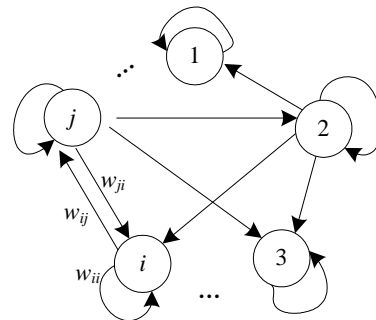


Figure 1. Topology of a FCM model

In general, the inference process in a FCM model is regarded as an iterative process, and feedback can be presented during the process as well. Once the initial state is determined in the FCM, the next state can be calculated and updated by applying the scalar product and activation function to each node synchronously. The values of each time state for all nodes are combined in order to form a discrete time system that can indicate cause-effect propagation in the value from the initial state to the final state. The iteration will continue until one of the following requirements on convergence is satisfied: equilibrium, limited cycle and chaotic behavior. The inference mechanism and state update of FCM can be described in the following formula:

$$v_i(t+1) = f\left(\sum_{j=1}^n v_j(t) \times w_{ij}(t)\right) \quad (1)$$

where $i, j=1,2,\dots,n; t=1,2,\dots,k; t$ denotes the iteration step, $v_i(t)$ denotes the state value of node i , f denotes threshold of activation function. Activation level of the nodes will change over iteration, which describes the dynamic behavior and characteristics of FCM. By comparison with neural network and other expert system, FCM can provide more flexible interaction and distinct meanings for the problem of real world application, and the weights and nodes can be configured and visually observed transparently (Tsadiras. 2008).

2.2 Design of candidate FCM model

The main objective for the development of candidate FCM model is to emulate a specific series of data set that are clustered from time series. To begin with, a FCM prototype with fully connected weights and self-feedback of each node is taken into consideration. In order to make it possible for further evaluation of both linguistic and numerical purpose, the feature that can be used to describe the trend of the time series is taken into account and will be presented to describe the behavior of the time series, which means the data values of the time series is not represented directly. Suppose input time series that

is used for fault clustering $x(t)$, the amplitude and change of amplitude is then combined together to show the behavior. More specifically, we can use amplitude and change of amplitude to capture the dynamics of the time series, which is in the form of two dimensional data. The dynamics of time series can be described by information granules in the form of amplitude and change of amplitude $(x(t), \Delta x(t))$ using clustering method. Each data value in time series is given a numerical value of activation level with regard to the information granules. Therefore, for a given time series, the behavior of each data value centers around information granules, which are also regarded as the features of time series, and each data value has a degree of belonging to cluster center. The best way to form the information granules is to cluster time series. Fuzzy c-mean clustering is a useful way and can be used to divide data into different category. The cluster center will constitute the certain level of granules that imply the behavior of the original time series. In this way, the features with regard to information granules are extracted. The information granules are used. Each cluster is defined as a node in the FCM model, the concept of the node is described as the degree of belonging to the cluster.

2.3 Learning method and FCM training process

Training process is important in the design of FCM model, and the learning method plays vital rule during training process. Evolutionary strategy, particle swarm optimization and simulated annealing are dominate learning method. Having the data set available, the supervised learning method is taken into account. We are interested in the learning method of particle swarm optimization method for it is promising in tackling optimization problems. Apart from that, the activation function is crucial in manipulating the value of state vector. The proper selection of the activation function is dependant on some factors, for example, the nature of the problem, the required representation capabilities of the problem and the level of inference required by the case. Here, sigmoid function is generally used for it is suitable for qualitative and quantitative problems. The determination of the weights is made according to the objective function described as follows:

$$F_{obj} = \frac{1}{n(k-1)} \sum_{i=1}^k \sum_{j=1}^n \left\| f \left(\sum_{j=1}^n v_i(t) \times \bar{w}_{ij}(t) \right) - v_i(t+1) \right\|^2 \quad (2)$$

The objective function is calculated for each iteration during optimization process, and the weights will be updated at the next iteration according to corresponding learning method. With particle swarm optimization is introduced (Clerc, 2002), the swarm consisting of m particles is created initially, with each particle is the vector of $n \times n$ dimensions, and the dimension of each particle

is corresponding to the number of cells in weight matrix W .

3 APPLICATION AND ANALYSIS

3.1 Data set and pre-processing

In this paper, a time series is introduced to demonstrate the proposed method. For simplicity, the time series under analysis comes from collection of electricity demand in New South Wales, Australia January, 2010. The data set is shown in Figure 2.

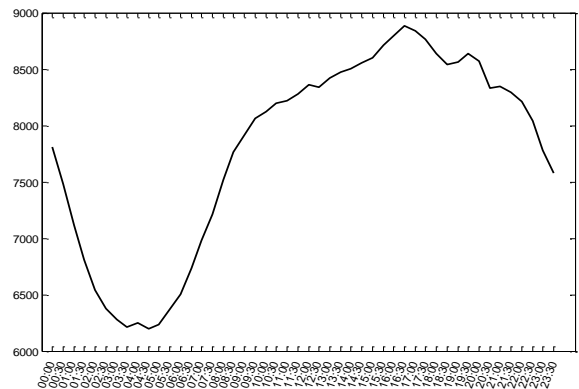


Figure 2. Introduced time series

The values are recorded at an interval of 30 minutes. There are 48 data points altogether for one day of 24 hours. The procedure for analysis of representation begins with pre-processing of data set and setting the number of linguistic terms. Figure 3 shows the data set for one day, and the number of linguistic terms that is used to describe amplitude and change of amplitude is set to 2. In this way, the cluster center and its activation level matrix is obtained by using fuzzy c-mean clustering based on section 2. Simulation is performed regarding the process of FCM construction. In order to observe the influence of FCM parameters, the steepness of activation function is set to 1 and 6 separately for the same data points, the quality of representation from candidate FCM model is validated and assessed as well.

3.2 Result and analysis

In this section, the real world time series is taken to explore the proposed method, parameters of configuration are taken into account to observe the performance and quality of candidate FCM model as well. Figure 3 shows the activation level regarding data set, the number of linguistic terms is 2 labels for each of the two signals, i.e., amplitude and change of amplitude. The number of cluster is therefore set to 4, which is the same as the number of nodes in the candidate FCM model, after training process is completed, the candidate FCM model will be used subsequently to imitate the dynamics of the

activation level as shown in Figure 3 over all data points.

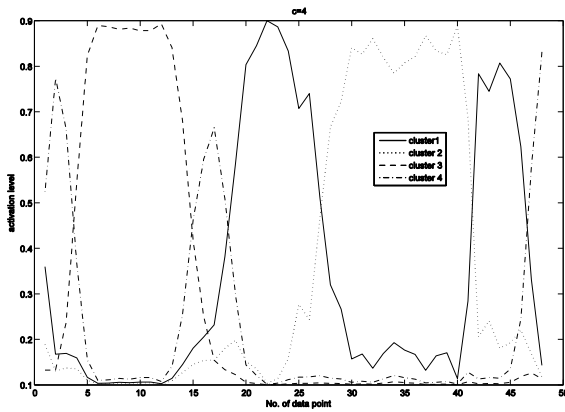


Figure 3. Cluster for data set from time series

In this respect, the parameters regarding candidate FCM model is initialized, the node number is set to 4, and the data points that are used to train the candidate FCM model is 48, the steepness of activation function is set to 1 and 6 respectively. The SA algorithm is also taken into consideration for comparison purpose. The red line denotes the original cluster, the blue line denotes the representation with PSO method, and the black line denotes the representation with SA method. Figure 4 and Figure 5 show the result.

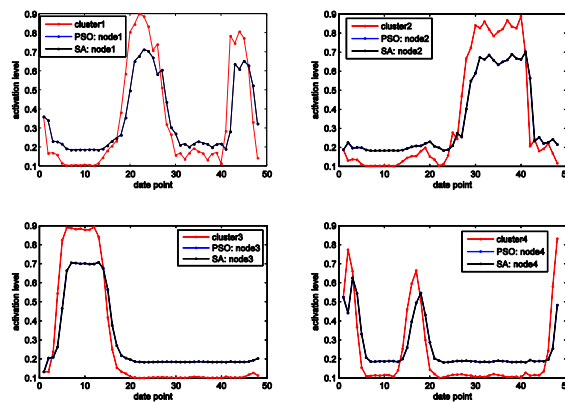


Figure 4. Representation of activation level for cluster ($\sigma=1$)

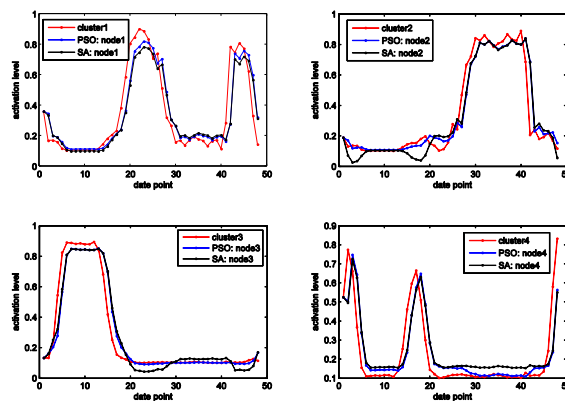


Figure 5. Representation of activation level for cluster ($\sigma=6$)

It can be seen in Figure 4 that the steepness has significant effect on the result of data representation,

the performance is not desirable for each cluster in the case of $\sigma=1$, although the trend of decrease and increase is depicted well, the quality is not good. In Figure 5, σ is set to 6, we can see that the quality of representation is improved. It is therefore concluded that the steepness of activation function is more sensitive in the design process of candidate FCM model. The result also shows that the data representation is advisable provided that the parameters are taken into consideration and the optimization performance is obtained.

4 CONCLUSIONS

In this paper, a design method based on FCM model is proposed for the representation of time series analysis. The up-to-data learning method is used to optimize the weight matrix, the steepness of activation function that affect FCM model is studied. The feature of the time series can be depicted by adjusting the parameters. It can be used for trend prediction and analysis for the application of real world time series.

5 ACKNOWLEDGMENTS

This research is supported by National Natural Science Foundation of China under Grant No. 61273165 and SRF for ROCS, SEM.

REFERENCES

- [1] Sapello, A. Sethi, M. Nodine, R. Chadha. 2010, Application of Time Series Analysis to Fault Management in MANETs. International Conference on Network and Service Management, (2010), October 25-29; Niagara Falls, Canada
- [2] Tsadiras. 2008, Comparing the inference capabilities of binary, trivalent and sigmoid fuzzy cognitive maps, Information Sciences. 178: 3880-3894
- [3] Kosko. 1986, Fuzzy cognitive maps, International Journal of Man-Machine Studies :24- 65
- [4] G. C. Chow. 2003, A Time-Series Analysis of the Shanghai and New York Stock Price Indices, General Economics and Teaching, EconWPA.
- [5] H.K.Yu. 2005, Statistical Mechanics and its Applications, Physica A. 346: 3-4
- [6] J. Wang, D. Chi, J. Wu, H.-y. Lu. 2011, Chaotic time series method combined with particle swarm optimization and trend adjustment for electricity demand forecasting, Expert System with Applications, 38 :8419-8429
- [7] M. Clerc, J. Kennedy. 2002, The particle swarm - explosion, stability, and convergence in a multidimensional complex space, IEEE Trans. Evolutionary Computation. 6: 58-73
- [8] S.-H. K. Kee-sang Lee. 1999, , International Journal of Man-Machine Studies 24Fault diagnostic system based on fuzzy time cognitive map, Transaction on Control, Automation and System Engineering. 1: 62-68