Annals of Nuclear Energy 76 (2015) 63–74

Contents lists available at ScienceDirect

Annals of Nuclear Energy

journal homepage: www.elsevier.com/locate/anucene

Performance analysis of fuzzy rule based classification system for transient identification in nuclear power plant

Manas Ranjan Prusty *, T. Jayanthi, Jaideep Chakraborty, H. Seetha, K. Velusamy

Indira Gandhi Centre for Atomic Research, Kalpakkam 603102, India

ARTICLE INFO

Article history:
Received 9 June 2014
Received in revised form 15 September 2014
Accepted 18 September 2014
Available online 10 October 2014

Keywords:
Fuzzy modeling
Pattern classification
Interpretability and accuracy of fuzzy systems
Feature selection
Fuzzy rule based system

ABSTRACT

Even though fuzzy rule based classification system (FRBCS) has been useful in event identification, it has led to strong clash in terms of better interpretability along with adequate percentage of accuracy. Basically for classification in nuclear power plant (NPP) which receives data within a cycle time of few milliseconds, either the accuracy or the interpretability of the FRBCS would get jeopardized. Online event identification of any abnormality or transient using FRBCS becomes really critical for the plant which has such a short cycle time. For such cases, the output from a FRBCS may not be conducive to classify the event every cycle. Thus, it is necessary to monitor the output of a classification system for certain amount of cycles till the static nature is attained. This gives a high level of confidence on the classifier output to be accurate. A FRBCS can produce this level of confidence by choosing the best input features with high interpretability and acceptable accuracy. The best feature selection out of a lot of input variables and preparing the rule base is again a very critical and challenging task in FRBCS. There is always a dilemma on judiciously choosing the number of input features for the FRBCS to achieve an optimized interpretable and accurate fuzzy system. It is always advisable to select least number of features with proper output error margin for a FRBCS. On adding extra features along with some rules as input to the system, certainly increases the accuracy of the system but degrades its interpretability. The other thing which is of equal importance is to keep an agile notice on the output which has to be within the output error margin. The output error margin denotes the response time and static nature of the classifier output. Here, the classification could be confirmed only after observing the output for a subsequent number of cycles. Both the number of inputs and the output error margin needs to be properly balanced in order to get an efficient online FRBCS for such kind of low cycle time systems. This paper concentrates on the performance analysis of some FRBCS with different number of input features and their behavior towards different datasets from Prototype Fast Breeder Reactor (PFBR) which is a NPP. Based on the PFBR Operator Training Simulator (OTS) data, some FRBCS are developed to identify some of the transients or abnormalities occurring in PFBR. The comparison is done based on the number of input features, response time and static nature of the output in each of the produced systems.

1. Introduction

The beauty of fuzzy modeling lies in its interpretability and accuracy similar to a human behavior during a particular situation (Lughofer, 2013). This depends on the representation of the fuzzy system during the learning procedure called as the knowledge base. Interpretability of a fuzzy system involves the skill of the specific end-user who interprets its linguistic description with the aim of conceiving the significance of the system behavior (Alonso et al., 2011). Accuracy of a fuzzy system means the ability of the system to produce an output which follows the actual output in real world scenario. There are several input features on which the interpretability and accuracy depends in a fuzzy rule based systems (FRBS). These input features are the model structure, number of input variables, number of rules, number of linguistic terms, shape of the membership functions, etc (Mansoori et al., 2007; Gactoa et al., 2011). The choice of perfect membership functions with proper centre and width also plays its part in producing an acceptable result. The number of input variables and the fuzzy rules decides the interpretability and accuracy of the fuzzy system. An increase in the number of input variables may increase the accuracy of the fuzzy system jeopardizing the interpretability. There is always a trade-off between interpretability and accuracy of the fuzzy systems (Gactoa et al., 2011; Yu and Xiao, 2009; Mikut et al., 2005; Mencar et al., 2011; Casillas et al., 2007; González et al., 2007). A

* Corresponding author. Tel.: +91 4427480500x21153; fax: +91 4427480210.
E-mail address: manas.iter144@gmail.com (M.R. Prusty).

http://dx.doi.org/10.1016/j.anucene.2014.09.039
0306-4549/

fuzzy system with user controllable trade-off between accuracy and interpretability is also a solution to find a well acceptable system based on the requirement and satisfaction level (Yu and Xiao, 2009). For non linear systems, whose dependent parameters are allowed to change throughout the course of simulation run such as temperature based thermal conductivity, modeling with proper accuracy and interpretability is a challenging task. For designing different fuzzy systems for such non linear modeling, there is a method which takes care of the interpretability issues (Cpałka et al., 2014).

Classification is a data mining function which maps a given data to a collection of items along with their classes. The goal of classification algorithm is to predict the target classes of a given data. Fuzzy class depicts the real world scenario in a better way using linguistic variables, overlapping classes and approximate reasoning. A specific category of FRBS which deals with classification is called as fuzzy rule base classification system (FRBCS). There are many instances using FRBCS for classification effectively (Samantaray, 2013; Fernández et al., 2008). There are some modifications also done to the FRBCS in order to increase the modularity of the present system (Sanz et al., 2010; Jahromi and Taheri, 2008). The FRBCS is widely recognized for its robustness to imperfect data and interpretability (Śaez et al., 2011). But the generation of the fuzzy rules in such systems with huge data set becomes really cumbersome. In this case, generation of the fuzzy rules automatically based on the training data set is an appreciable effort (Nakashima et al., 2007). In order to get a smooth classification boundary, an adaptive method to construct a FRBCS is certainly useful for high performance pattern classification (Nozaki et al., 1996). The importance of selection of the rule weights in the fuzzy rules and its effect is something really needs to be looked upon in order to generate a highly accurate classification system (Ishibuchi and Nakashima, 2001). A multi class classification is also a field which FRBCS can deal with efficiently using pair wise learning and preference relations (Fernández et al., 2010). The implementation of genetic algorithm into FRBCS is also a novel way of decreasing the computational time (Fazzolaria et al., 2013b).

In nuclear power plant (NPP) like Prototype Fast Breeder Reactor (PFBR), sensors at the plant site transmits data every 200 ms for monitoring and control purpose. In ideal situation, these data should be collected in the control room every 200 ms which is the cycle time. In order to know the status of the plant, these plant data needs to be processed in such a small cycle time. For this purpose, the system or model has to be interpretable. But this degree of interpretability does not guarantee the specified accuracy. To increase the accuracy, complexity of the system needs to be increased. Hence, in return this again reduces the interpretability of the system. Therefore, a tradeoff between interpretability and accuracy is necessary for systems where both are of prime importance (Fazzolari et al., 2013a; Ishibuchi and Nojima, 2007).

In this paper, it is explained that a simple FRBCS with optimized selection of input variables can be used as an online event classification method for a system which has a very short cycle time. There are complex systems like NPP undergoing continuous monitoring of system features with a cycle time of 200 ms. Within this short cycle time, the FRBCS output may not be conducive for any type of conclusion on online event classification based on data from this system. In such case, monitoring the FRBCS output for certain number of cycles produces an evident on the occurrence of an event which is more authenticated. In this paper, Section 2 explains about the two basic FRBSs being used more often i.e. Mamdani-type FRBS and Takagi–Sugeno–Kang (TSK) type FRBS. Section 3 is about the experiment conducted on FRBCS using datasets collected from the PFBR Operator Training Simulator (OTS) for classifying the occurrence of an event. This section analyses the behavior and the response time of three uniquely designed FRBCS towards different set of input datasets. Section 4 compares the performance of these three FRBCS. Section 5 concludes the paper inferring the feasibility of a FRBCS in PFBR along with the best FRBCS which gives a better result.

2. Fuzzy rule based system

A fuzzy rule based system (FRBS) is a data driven systems in which a relationship is being established between the input and the output, based on a set of IF-THEN rules or otherwise called as fuzzy rules. The set of desired input–output numerical data pairs are represented as:

$$\left(x_1^{(1)}, x_2^{(1)}; y^{(1)}\right), \left(x_1^{(2)}, x_2^{(2)}; y^{(2)}\right), \cdots\cdots, \left(x_1^{(n)}, x_2^{(n)}; y^{(n)}\right) \quad (1)$$

where x_1 and x_2 are inputs and y is the output of a two inputs and one output system. The fuzzy rules are generated from these input output pairs which determines a mapping f: $(x_1, x_2) \rightarrow y$ (Wang and Mendel, 1992). An example of a fuzzy rule in this kind of system is of the form

$$R^1 : \text{IF } x_1^{(1)} \text{ is } F_1^l \text{ and } x_2^{(1)} \text{ is } F_2^l \text{ THEN } y^{(1)} \text{ is } G^l \quad (2)$$

Here F_k^l represent the linguistic variable and $F_k^l \subset f(X_k)$, $k = 1, 2, 3, \ldots n$, are the antecedent membership functions, and $G^l \subset f(Y)$ is the consequent membership function (Mendel and Mouzouris, 1997). The input linguistic variables are denoted by x_k, $k = 1, 2, 3, \ldots n$ and the output linguistic variable is denoted by y. For classification using FRBS, the consequent part of the rule can be categorized into three varieties (Cordon et al., 1999). In this paper, the fuzzy rules with the class in consequent have been used for classification purpose. This approach was chosen because the FRBS would take data from a safety critical system which needs faster and interpretable outcomes. The other two categories can certainly be tested in future.

2.1. Membership function (MF)

Universe of Discourse is the set that contains all the sets of interest for the given context problem. The membership function ($\mu_A(x)$) of the fuzzy set (A) maps the universe of discourse (X) on to the numerical values in the range [0, 1].

$$\mu_A(x) : X \rightarrow [0, 1] \quad (3)$$

$$A = \{(x, \mu_A(x)); x \in X, \mu_A(x) \in [0, 1]\} \quad (4)$$

2.2. Mamdani-type FRBS

There are mostly two types of FRBS. The most popularly used among them is the Mamdani-type FRBS as shown in Fig. 1. The following steps are involved in Mamdani-type FRBS approach. These are

Step 1. Segregation of the universe of discourse into fuzzy regions.
Step 2. Generation of the fuzzy rules and rule base.
Step 3. Processing of input output numerical data pairs.
Step 4. Defuzzification procedure.

2.2.1. Step 1 – Segregation of the universe of discourse into fuzzy regions

This is the initial step where the domain interval of each fuzzy region is decided based on the input numerical data. The membership function for each input domain and output domain is constructed based on this segregation. The shape of the membership functions depends on the system designer based on the problem