

Chapter 14

Passenger Train Delay Classification

Masoud Yaghini

Iran University of Science and Technology, Iran

Maryam Setayesh Sanai

Iran University of Science and Technology, Iran

Hossein Amin Sadrabady

Research and Training Center of Iranian Railways, Iran

ABSTRACT

One of the most popular data mining areas, which estimate future trends of data, is classification. This research is dedicated to predict Iranian passenger train delay with high accuracy over Iranian railway network. A hybrid method based on neuro-fuzzy inference system and Two-step clustering is used for this purpose. The results indicate that the hybrid method is superior over the other common classification methods. The result can be used by train dispatcher to accurate schedule trains to diminish train delay average.

1. INTRODUCTION

Delay propagation can impact consequence of train delay in the railway network. Accurate predicting of train delay can help dispatcher in effective timetable. The main idea of this paper is to present a classification method by hybridizing Adaptive Neuro Fuzzy Inference System (ANFIS) and Two-step clustering to predict Iranian passenger train delay.

Artificial Neural Net (ANNs) and Fuzzy Inference System (FISs) have some fault solely which combination of them (neuro-fuzzy systems) has been popularized as a powerful alternative approach

to develop adaptive soft computing systems. Nauck and Kruse (1999) discussed advantages and disadvantages of both methods in-depth. Merging the neural network with the fuzzy system has resulted in neuro fuzzy systems which use artificial neural networks (ANNs) theory in order to determine their fuzzy sets and fuzzy rules by processing data samples. A hybrid system termed Adaptive-Network-based fuzzy inference system or Adaptive Neuro-Fuzzy Inference System (ANFIS) is a kind of neuro fuzzy which has been proposed by (Jang, 1992). In ANFIS, the membership function parameters are extracted from a data set that describes the system behavior.

DOI: 10.4018/978-1-4666-8473-7.ch014

The ANFIS learns from the data set and adjusts the system parameters according to a given error criterion (Jang, 1992; Jang, 1993). Among neuro fuzzy methods, ANFIS has high speed of training and the most effective learning algorithm (Jang, 1995). Neuro-fuzzy systems can learn system's behavior from a sufficiently large data set and automatically generate fuzzy rules and fuzzy sets to a pre-specified accuracy level. Other advantage of this combination is capability of generalization and overcoming the key disadvantages of the FISs, i.e. self-learning inability and lack of generalization capability. Neural network part creates fuzzy logic rules and membership during the period of learning automatically. Even after the learning, neural network continues to modify the membership function and fuzzy rules (Kurtulus, 2010).

Nowadays, ANFIS is one of the most efficient methods in modeling and prediction (Nayak et al., 2004; Kermani et al., 2008; Ata & Kocyigit, 2010). ANFIS has been successfully used for many applications such as database management, system designing and water resource forecasting (Silva et al., 1999; Chen et al., 2006; Chang & Chang, 2006; Akil et al., 2007; Özgür, 2006). Another successful implementations area of ANFIS in biomedical is engineering which is reported for classification (Ubeyli & Guler, 2005). The synergism of fuzzy logic systems and neural network has produced a functional system capable of learning high-level thinking and reasoning (Zadeh, 1977). It is an improved tool for determining the behavior of imprecisely defined complex systems. The purpose of a neuro-fuzzy system is to apply neural learning techniques to identify the parameters and/or structure of neuro-fuzzy systems. The fuzzy systems have received the growing attention and interest in bioinformatics applications, decision making studies, pattern recognition, and data analysis (Kissi et al., 2004; Marsala et al., 2000; Ros et al., 2002).

In this paper, a classification model of Iranian passenger train delay by using hybrid clustering and adaptive neuro-fuzzy inference system is presented. This method has two main steps. In the

first step, by using Two-step method, data clustering will be done and the next step is dedicated to design appropriate ANFIS model for each cluster. The rest of the paper is organized as follows. Section 1 outlines a brief description of data set of passenger train delay. Section 2 is assigned to explore the concept of the classification method. The result of the model implementation is proposed in section 3. The conclusion of this study is the topic of the final section.

2. PASSENGER TRAIN DELAY

To illustrate the practical application of the classification model, the Iranian passenger train delay is used as a case study. In Iranian railway, the data of passenger train delay is registered and recorded every day. At the weekend, these records are merged to create weekly delay data. Finally, at the end of each month, these weekly records are merged to create monthly delay data. In this research, monthly records from 2004 to end of 2008 are used. The passenger train delay is recorded as a real numbers in minutes. In this study, 179,982 data instances are used for train delay classification. The common approach is to divide the data set into training and testing data set. The training subset is used to build the networks and adjust the connected weights of the constructed network while the test data set is assigned to be utilized for the validation of the model. The data of 2004- 2007 is considered as training data set and data of 2008 as test one.

According to the information of the database, the yearly delay average is 1,058,705 minutes per year and 29.4 minutes for each passenger train. Number of records in each year represents number of dispatched trains in that year. Table 1 summarizes this database. The first column of the table stands for year and the second one indicates the number of trains dispatching in that year. Summation and average of train delay in term of minutes is shown in the last two columns.

8 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/passenger-train-delay-classification/128670

Related Content

An Integrated Digital Authentication Mechanism for Intrusion Detection System

Ch Rupa (2019). *Big Data Analytics for Smart and Connected Cities* (pp. 158-169).

www.irma-international.org/chapter/an-integrated-digital-authentication-mechanism-for-intrusion-detection-system/211746

Emerging Challenges and Opportunities for U.S. High-Speed Rail Development

Francis P. Banko and Jackson H. Xue (2016). *Emerging Challenges and Opportunities of High Speed Rail Development on Business and Society* (pp. 81-123).

www.irma-international.org/chapter/emerging-challenges-and-opportunities-for-us-high-speed-rail-development/152052

Adding Electric Vehicle Modeling Capability to an Agent-Based Transport Simulation

Rashid A. Waraich, Gil Georges, Matthias D. Galus and Kay W. Axhausen (2015). *Transportation Systems and Engineering: Concepts, Methodologies, Tools, and Applications* (pp. 1563-1600).

www.irma-international.org/chapter/adding-electric-vehicle-modeling-capability-to-an-agent-based-transport-simulation/128736

The Role of Software in Construction Management

(2021). *Managing Business in the Civil Construction Sector Through Information Communication Technologies* (pp. 69-97).

www.irma-international.org/chapter/the-role-of-software-in-construction-management/264280

Numerical Modeling of Historic Masonry Structures

Panagiotis G. Asteris, Vasilis Sarhosis, Amin Mohebkhah, Vagelis Plevris, L. Papaloizou, Petros Komodromos and José V. Lemos (2016). *Civil and Environmental Engineering: Concepts, Methodologies, Tools, and Applications* (pp. 27-68).

www.irma-international.org/chapter/numerical-modeling-of-historic-masonry-structures/144492