Research On Rough Set Approach To Traffic Flow Prediction

Ms M. A. Deshpande

Research Scholar

G.H.R.C.E, Nagpur

Guide:

Dr. P.R.Bajaj

Prof. G.H.R.C.E Nagpur

Abstract

Rough set theory is a mathematical method which analyses and treats vagueness and uncertainty, and offers an effective method to traffic flow prediction system. This paper compares performance of rough set theory combined with support vector machine to neural network performance. All of them have great advantages on dealing with various imprecise and incomplete data. However, there exists essential difference among them. Firstly, this method uses the rough set theory for data reduction pre-treatment, and then constructs the traffic flow prediction model based on support vector machine according to the information structure. The results of the model are better than the BP Neural network and single support vector machine model. Besides, the combined prediction model not only has fault tolerant and anti-jamming capability, but also can shorten the operation time and improve the speed of the system and also forecast accuracy. Hence, it can be used to forecast real-time traffic flow.

Introduction

As one of the most important traffic information, traffic flow plays a very significant role in ITS. So, the forecast traffic is the key of transportation controlling and traffic guidance to achieve. Traffic is formed by tens and thousands of travel group behaviour, which have high degree of variability, nonlinear and uncertainty. If traffic flow is random, then the traffic flow can't be predicted according to the meaning of predictability, we can just describe this with probability theory. Traffic flow at different times has a certain similarity at the same checkpoint; waveform is similar, which shows that traffic flow has a periodic repetition. This shows that despite the dramatic changes in traffic flow, seemingly unsystematic, traffic flow time series has its cyclical and self-similarity in actually. Thus, short-term traffic flow is predictable.

As compared to SVM, neural network has several weaknesses such as slow learning rate, difficult convergence, complex network structure and unambiguous meaning of network. With the development of research, there appear many new theories in information processing and knowledge discovery since 1980's. In these theories, rough sets theory and support vector machine (SVM) are the most attractive. Advantages have been shown in rough sets theory dealing with various imprecise, incomplete information and in neural networks. However, there

exists essential difference between them. Rough sets simulate abstract logic mind of our human being while neural networks simulate intuition mind. Rough sets theory express logic rules based on indiscernibility relation and knowledge reduction while neural networks state relation between input and output by using nonlinear mapping. In general, neural networks can not reduce dimensions of inputs. More complex structures and training cost required in neural networks of a higher input dimensions. This paper mainly introduce application of the rough sets on traffic flow prediction. In the result, we can find that the performance of system achieve much improvement by using the rough sets. SVM is a new machine learning theory proposed by Vapnik et al. in mid 1990s. It is a universal method to solve multidimensional function. It has been applied some areas such as function simulation, pattern recognition and data classification and obtained a perfect result. There exist some defects in neural network such as determination of network structure, local minima problems, under learning and over learning. All of them restrict the application of neural network. SVM has advantages in solving the problems of non-linear, pattern selected, high dimension, small specimen, which is good complementary with neural network. This paper is organized as follows. In section II, basic theories of three methods are briefly reviewed. In section III, we apply three methods into the traffic flow prediction respectively and give the comparative results. Section IV is the analysis of the three methods in which advantage and disadvantage of three methods are analyzed and some suggestion for future research are also presented. The last section is conclusion.

In this paper, we utilize the rough set combined with support vector machine to overcome the lack of a single prediction. We also analyse the various factors affecting future traffic flow, and considered the historical data, real-time data and weather conditions. First, choosing those data through the integration form this target data. Then using the rough set to do table filled, discretion, attribute reduction for knowledge of the composition of the expression of target data. Finally, using the decision-reduced table as an SVM input for normalization, kernel function selection and finding the optimal parameters, establish short term traffic flow prediction model, obtaining the predicted results. Finally, after analyzing of results, and compared with BP neural network and simple support vector machines, verify rough sets-support vector machine model has a high precision and generalization ability.

2. Review of Neural network, Rough Sets and Support Vector Machine

2.1 Neural network

Since neural network researches revived in 1980s, substantial progress has been achieved in application as well as in theory. Neural networks have been widely applied in pattern recognition, control optimization, predicting management and so on. In the field of artificial intelligence, neural networks have been combined with genetic algorithm, fuzzy sets [9]. Classification is a very important task in area of information processing and knowledge discovery. Classification of neural network is a supervised training algorithm. It has a high tolerance capability and self-organization performance. Lots of work have been done and large numbers of literatures have been introduced in the field of neural networks. Presently, most methods of neural network in traffic flow prediction use BP learning algorithm for supervised learning classification. BP network is a feedforward network which is in fact a nonlinear criterion function.

2.2 Rough Set Theory

The Rough Sets Theory (RS) was proposed by Zdzislaw Pawlak in 1982 as a mathematical model to represent knowledge and to treatment of uncertainty. To define rough sets mathematically, we begin by defining an information system $S = (U,A)$, where U and A are finite and non-empty sets that represent the data objects and attributes respectively. Every attribute $a \in A$ has a set of possible values Va. Va is called the domain of a. A subset of A say B will determine a binary relation $I(B)$ on U, which is called the indiscernibility relation. The relation is defined as follows: $(x, y) \in I(B)$ if and only if $a(x) = a(y)$ for every a in B, where $a(x)$ denotes the value of attribute a for data object x [10]. $I(B)$ is an equivalence relation. All equivalence classes of $I(B)$ as $U/I(B)$. An equivalence class of I(B) containing x is denoted as $B(x)$. If (x,y) belong to $I(B)$ they are said to be indiscernible with respect to B. All equivalence classes of the indiscernibility relation, $I(B)$, are referred to as B-granules or B-elementary sets [10].

In the information system defined above, we define as in [10]:

$$X \subseteq U \qquad (1)$$

And,

$$B \subseteq A \qquad (2)$$

We now define the two operators assigned to every (1) two sets called the upper and lower approximation of X. The two sets are defined as follows [10]:

$$B_l(X) = \cup_{x \in U} \{B(X): B(X) \subseteq X\} \qquad (3)$$

And,

$$B_u(X) = \cup_{x \in U} \{B(X): B(X) \cap X \neq \varphi\} \qquad (4)$$

Thus, the lower approximation is the union of all B elementary sets that are included in the target set, whilst the upper approximation is the union of all B-elementary sets that have a non-empty intersection with the target set. The difference between the two sets is called the boundary of region of X.

$$BR = B_u(X) - B_l(X) \qquad (5)$$

If the boundary region is an empty set then X is crisp with respect to B, if however the boundary region is non-empty then X is rough with respect to B. Accordingly, the set is said to be rough if it cannot be defined exactly from the available data. The set of attributes that is sufficient to represent the entire equivalence class structure is called the *reduct*. The reduct of the information system is not unique. There are potentially many subsets of attributes which preserve the equivalence class structure. The set of attributes common to all reducts is called the *core*. The core can be regarded as the indispensable attribute of the information system. However, in practical applications where the information system contains thousands or possible tens of thousands of objects it is seldom that a core exists as shown in [1].

Rough Selection provides a means by which discrete or real valued noisy data (or a mixture of both) can be effectively reduced without the need for user-supplied information. Additionally, this technique can be applied to data with continuous or nominal decision attributes, and as such can be applied to regression as well as classification datasets.

2.3 Support Vector Machine

The support vector machine (SVM) is a new algorithm developed from the machine learning community. In machine learning, support vector machines (SVMs, also support vector networks[2]) are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. The basic SVM takes a set of input data and predicts, for each given input, which of two possible classes forms the output, making it a non-probabilistic binary linear classifier. The strategy in this technique is to map the

input vectors into a high dimension feature space corresponding to a kernel, and construct a linear decision function in this space to separate the dataset with maximum margin Via the freedom to utilize different types of kernel, the linear decision functions in the feature space are equivalent to a variety of non-linear decision functions in the input space. Although SVM is a good tool for non-linear, high dimension data mining, it still does not work well when the input data is massive, noisy and missing. Rough set was developed for processing coarse information by Pawlak in 1982[1]. Until now, it has been conceived to conceptualize, organize and analyse various types of data, in particular, to deal with inexact, uncertain or vague knowledge in applications related to Artificial Intelligence. By combining rough set and SVM, it becomes possible to eliminate the redundant traffic data and reduce the scale of the network. Most importantly, it also can improve the accuracy of travel time prediction on urban network.

Our target is to build a criterion function to separate the two classes. If there exists a hyperplane w.x+b=0 which makes

$$\{(w.x_i) + b \geq 1, y_i = 1\}$$

Equivalent to

$$(w.x_i) + b \leq 1, y_i = -1\}$$

$$y_i[(w.x_i) + b] - 1 \geq 0$$

We can obtain the optimizing problem as follow

$$min \frac{1}{2} \parallel w \parallel^2$$

Then Lagrange function can be defined as below,

$$L(w, b, \propto) = \frac{1}{2}(w.w)$$
$$- \sum_{i=1}^{n} \propto_i \{y_i[(w.x_i) + b]$$
$$- 1\}$$

where $\propto_i$ are the Lagrange multipliers. (10) can be transformed to its dual problem in order to minimize the equation. According to Kühn-Tucker condition, we can obtain the optimal classification function,

$$f(x) = sgn\{(w^*.x) + b^*\} = sgn\{\sum_{i=1}^{n} \propto^* y_i(x_i.x) + b^*\}$$

Where sgn is the symbolic function. Given a kernel function $K(x_i, x_j) = \varphi(x_i^T)\varphi(x_j)$ a decision function is given as $y(x) = \sum_{i=1}^{n} a_i K(x, x_{i)+b}$ (4)

As one of the most important traffic information, traffic flow plays a very significant role in ITS. So, the forecast of traffic is the key of transportation controlling and traffic guidance to achieve. Traffic is formed by tens and thousands of travel group behavior, which have high degree of variability, nonlinear and uncertainty. But for a particular observation point on the road, when observing long scale, the statistical characteristic behaviors of traffic volume show strong certainty, and shows increasing or decreasing tend gradually within a certain time, so using general prediction model can get accurate predict results. For short-term traffic flow prediction, traffic is influenced by random factors with the reduced scale of observation not only because of nature (e.g., season, climate), but also from man-made causes (e.g. emergency, psychological state of the driver). The statistical behavior is not a fixed length, periodic or quasi-cycle even is a pure random behavior.

METHODOLOGY

Conceptual Framework of Rough Set:

Structure of prediction model travel time prediction model on urban network released in this article are composed by two parts, rough set data pre-processor and SVM regression. Firstly, the rude traffic data which contains travel time, traffic volume and occupancy ratios is pre-processed by rough set. Then, the results are used as input samples to train the SVM for travel time prediction. The model encompasses four successive steps, i.e., establish decision table and attributes reduction, select input samples for SVM, train the SVM, predict the travel time on urban way.

Step 1. Establish decision table using traffic data. The number of condition attributes are 3n+2 in the decision table; ie, time t; road grade π; $Q_i(i = 1,2, ..., n)$; $\bar{D}_i(i = 1,2, ..., n)$; $\bar{v}_i(i = 1,2, ..., n)$. Q_i refers to traffic volume on the I road , $\bar{D}_i$ refers to average delay in the ith intersection, $\bar{v}_i$ refers to average link speed on the I road. The decision attribute is link travel time T.

Step 2. Select input samples for SVM.
The redundant data and the inconsistent data will be removed by attributes reduction. The reserved attributes arranged follow the importance degree are the input samples for SVM.

Step 3. Train the SVM.
The prediction model based on SVM regression will be developed by choosing appropriated kernel function and be trained with train samples. The training will not be stopped until the total of errors E meet the requirement of accuracy.

Step 4. Predict the travel time.
After the training has been done, the model can be used for travel time prediction. The travel time T can be calculated by the equation (4)

Conclusion

The rough set can be conceived as a tool to conceptualize, organize and analyse various types of traffic data. By the attribute reduction and decision reduction, the redundant and inconsistent traffic data will be eliminated effectively, which can provide basic traffic data of good quality for transportation information system. Rough set and SVM are complementary to process the traffic data. The integration of two models can predict travel time effectively. The accuracy of the error bars may therefore further be improved by changing the prior distributions and/or the likelihood function from RBF to other shapes that can be derived through the collected data, for example by investigating the probability density function of the noise that is removed by the de-noising procedure.

References:

1)Pawlak Z., Rough sets theory and its applications to data analysis, Cybernetics and Systems, Vol.29, No.3, pp.661-668, 1998.
2) Heermann P D., Khazenie Classification of multispectral remote sensing data using a back propagation neural network, IEEE Trans on Geoscience and Remote Sensing, Vol.30, No1 1992, pp. 81-88.
[3] Logi F, Ritchiei S.G., Development and evaluation of a knowledge-based system for traffic congestion management and control, TransportationResearch Part C, Vol. 9, No. 6, pp.433-459, 2001.
[4] Pi Xiaoliang, Wang Zheng, Han Hao, Application research of traffic state classification method based on collected information from loop detector,Journal of Highway and Transportation Research and Development,Vol.23, No.4, pp.115-119, 2006.

[5] SisiopikuI V, Rouphail N, Santiago A., Analysis of Correlation between Arterial Travel Time and Detector Data from Simulation and Field Studies, Transportation Research Record, Vol. 1457, pp.166-173, 1994.

[6] Van Lint, Reliable Travel Time Prediction for Free-ways, Delft, the Netherlands, Delft University Press, 2004.

[7] van Hinsbergen, C.P.I. and van Lint, J.W.C., Bayesian combination of travel time prediction models, the 87th Annual Meeting of the Transportation Research Board, Washington DC, USA, 254-261, 2008.

[8] van Hinsbergen, C.P.I., van Lint, J.W.C. and Sanders, F.M, Short term traffic prediction models, In Proceedings of the 14th ITS World Congress,Beijing, China, 164-170, 2007.

[9] Chua, C.G. and Goh, A.T.C., A hybrid Bayesian back-propagation neural