HRVOJE GOLD, D.Sc.
ZVONKO KAVRAN, B.Eng.,
GORDANA ŠTEFANČIĆ, D.Sc.
Fakultet prometnih znanosti
Zagreb, Vukelićeva 4.

Traffic Planning
Preliminary Communication
U.D.C. 656.132.03:621.391:612.8
Accepted: May 4, 1998
Approved: May 6, 1998

BUS TICKETS SALES FORECASTING USING NEURO-GENETIC METHODS

SUMMARY

This paper describes the applications of an enhanced neural network and genetic algorithm for bus tickets sales forecasting. The proposed approach has several significant advantages over conventional prediction methods. The major advantage of the approach is that no assumptions need to be made about the underlying function or model, since the neural network is able to extract hidden information from the historical data. Although neural networks represent a promising alternative for forecasting, the problem of network design remains and could impair widespread applications in practice. The genetic algorithm is used to evolve neural network architectures automatically, thus eliminating the pitfalls associated with human engineering approach.

1. INTRODUCTION

Bus stations make complex systems in terms of technology and management, whose efficient operations set up the requirements of keeping pace with transport technology advancement [1]. One of the main tasks is to determine the optimum aspect of technological modernisation, which will in conjunction with certain quality of transport service, result in such capacities of basic transport service and technology requiring minimal investments and costs of maintenance and utilisation.

Passenger flows at the bus stations are of stochastic character. With application of forecasting techniques, stochastic passenger flows can be determined to a major extent in advance. On this basis, it is possible to predict the necessary capacities for effective bus tickets sales service. Without looking at the results of forecasting, the demand may exceed the capacities, which will lead to saturation of the service.

A forecast is an event that is going to happen in the future based on its present surrounding events. Forecasting assumes that future occurrences are based, at least partly, on the currently observable or past events. Moreover, it assumes that some aspects of the past patterns will continue into the future. Past relationships can then be discovered through study and observations.

Forecasting techniques can be categorised into three groups [2]. The first is called qualitative, where all the information and judgement relating to an item are used to forecast the items demands. This technique is often used when little or no demand history is available. The second group is called causal, where cause-effect type of relation is sought. The forecaster seeks a relation between the item demands and other factors. The relationship is used to forecast the future demands of the item. The third group is called the time series prediction, where a statistical analysis on past pattern in the historical data series is used to extrapolate that pattern into the future. Classical time series methods are auto-regression (the Box-Jenkins method, or ARIMA method) and exponential smoothing.

In general, a time series can be broken down into four components: secular trend, cyclical variation, seasonal fluctuation and irregular fluctuation. Secular trends are long-term, slow-moving, low-frequency components. Cyclical variations usually take a few years to complete a cycle, while seasonal variations are completed within a year. Irregular fluctuations are random in nature and usually difficult to predict. The irregular, random components can be further divided into two sub-components: the deterministic chaotic behaviour and the stochastic noise.

The passenger flow, expressed in the number of sold tickets, can be presented as a time series, which will in general include the previously mentioned components.

2. FORECASTING BASED ON THE NEURAL NETWORKS AND GENETIC ALGORITHMS

Neural networks non-linear learning and interpolative smoothening capabilities have proven superior to conventional methods in time series forecasting [3]. Neural networks are not only capable of decoding deterministic chaos, they also gave much superior predictions of certain types of non-linear, dynamic systems compared to several conventional methods based on adaptive filtering and polynomial curve-fitting, such as auto-regressive model which is a linear model and the exponential smoothing method which only carries out piecewise linear approximation. Specifically, for time series with long memory, the neural network produces comparable results. However, neural networks perform much better for the short-memory series.

Neural networks are computational models that are capable of identifying complex non-linear relationships between input and output data sets [4]. They consist of a number of interconnected neurones or processing elements. The structure of a network is determined by the arrangement and the strengths of the inter-neurone connections. The strengths of the connections are adjusted or trained to achieve a desired overall behaviour of the network by corresponding learning algorithm. Neural networks can be classified according to their structures and learning algorithms.

The convenient network structure is the feed-forward network. It has been found that it has high performances in input-output function approximation as well as in time series forecasting. In a feed-forward network, the neurones are grouped into layers. Signals flow from the input layer through the internal-hidden layers to the output layer via unidirectional connections. The neurones are connected from one layer to the next, but not within the same layer. Back-propagation network is the convenient example of the feed-forward neural network. The adjustment of the strengths or weights of the inter-neurone connections corresponding to a given input are produced according to the difference between the desired and actual network outputs. The adjustment requires a supervisor or a teacher to provide the desired or targeted output signals. A frequently used supervised learning algorithm is the back-propagation algorithm.

Genetic algorithms have been increasingly applied in conjunction with neural networks. Two primary areas of activity are topology optimisation and genetic training algorithms. In topology optimisation, the genetic algorithm is used to select a pattern of connectivity for the neural network that is in turn trained by using some fixed training method, most commonly back-propagation. In genetic training algorithms, the learning of a neural network is formulated as a weights optimisation problem.

Analysis of the natural adaptive systems shows that natural evolution embodies an elegant generate and test strategy that can rapidly identify and exploit regularities in the environment [5]. Even in the large and complicated search spaces, given certain conditions on the problem domain, genetic algorithms would tend to converge on solutions that were globally optimal or nearly so.

There are four stages in the genetic search process: initialisation, selection, crossover and mutation. In the initialisation stage, a population of genetic structures that are randomly distributed in the solution space is selected as the starting point of the search. After the initialisation stage, each structure is evaluated using a user-defined fitness function and assigned a utility value. Based on their relative utility values, structures in the current population are selected for reproduction. A stochastic procedure ensures that the expected number of offspring associated with a given structure is related to the ratio of observed performance of the structure to the average performance of all structures in the current population. Thus, structures with high performance may be chosen for replication several times while poor performing structures may not be chosen at all. In the absence of other mechanisms, such a selective pressure would cause the best performing structures in the initial population to occupy an increasingly larger proportion of the population over time.

The selected structures from the previous stage are recombined using crossover, which operates by swapping the corresponding segments of a string representation of the parents. Crossover serves two complementary search functions. First, it provides new points for further testing of schemata, i.e. a subset of strings in a population with similarities at certain string positions, already present in the population and second, it introduces instances of new schema into the population.

However, crossover draws only on the information present in the solutions of the current population in generating new solutions for evaluation. If specific information is missing, then crossover is unable to produce new structures that contain this piece of information. A mutation operator, which arbitrarily alters one or more components of a selected structure, provides the means for introducing new information into the population. A mutation operator functions as a background operator with a very low probability of application. The presence of mutation ensures that the probability of reaching any point in the search space is never zero.

3. BUS TICKETS SALES FORECASTING

The set of real data is employed as an example in the investigation. Excerpt from data set is shown in Table 3.1. The data in Table 3.1 are the numbers of bus tickets sold daily within the period from January 1, 1997 to March 20, 1997. There are 79 data in Table 3.1. The data set was taken from the database of the bus tickets sold at the Zagreb bus station.

In the simulation, the data set of the bus tickets sold during the past 79 days have been used as input to

Table 3.1 The number of bus tickets sold daily from Jan. 1, 1997 to March 20, 1997

01.01. 1.037	17.01. 3.833	02.02. 2.032	18.02. 2.614	06.03. 2.913
02.01. 3.642	18.01. 2.687	03.02. 2.714	19.02. 2.679	07.03. 3.923
03.01. 3.683	19.01. 1.996	04.02. 2.503	20.02. 2.901	08.03. 2.625
04.01. 3.229	20.01. 2.560	05.02. 2.560	21.02. 3.663	09.03. 2.194
05.01. 2.337	21.01. 2.367	06.02. 2.938	22.02. 2.675	10.03. 2.724
06.01. 2.474	22.01. 2.490	07.02. 4.094	23.02. 2.102	11.03. 2.591
07.01. 2.994	23.01. 2.799	08.02. 2.891	24.02. 2.724	12.03. 2.748
08.01. 3.278	24.01. 3.777	09.02. 1.992	25.02. 2.670	13.03. 3.174
09.01. 3.372	25.01. 2.500	10.02. 2.727	26.02. 2.861	14.03. 3.829
10.01. 4.045	26.01. 2.108	11.02. 2.552	27.02. 2.900	15.03. 3.003
11.01. 3.081	27.01. 2.457	12.02. 2.651	28.02. 3.977	16.03. 2.246
12.01. 2.262	28.01. 2.297	13.02. 2.869	01.03. 2.949	17.03. 3.028
13.01. 2.593	29.01. 2.373	14.02. 3.651	02.03. 2.263	18.03. 2.901
14.01. 2.427	30.01. 2.779	15.02. 2.740	03.03. 2.641	19.03. 3.062
15.01. 2.520	31.01. 3.779	16.02. 2.072	04.03. 2.467	20.03. 3.527
16.01. 2.852	01.02. 2.614	17.02. 2.654	05.03. 2.651	

the network. Since we want to make prediction for a period in the future, we extend the predicted data to a period in the future. The network is trained to predict the number of tickets that will be sold the next day. From the 79 available data, 40 data are used for training and 39 for testing (the training set: data 1, 3, 5, ..., 79; the testing set: data 2, 4, 6, ..., 78). Setting up the training process in this way will instruct the program to interleave the training and testing data.

On the basis of input-output data pairs the Neuro-Genetic Optimizer program [6] generated, as the best network, the more general form of back-propagation network named Time Delay Neural Network. It employs the fast back-propagation technique for setting weights between neurones. The first layer has one processing element, one internal layer has five processing elements and the output layer has one processing element.

In the learning process the minimum network training passes for each network were 20. Average Absolute Error on the training set was 235.7281. Average Absolute Error has been calculated between the expected and the actual neural outputs and is averaged across all output neurones. Minimal Average Absolute Error on the test set was 208.1303.

The generated prediction results are shown in Figure 3.1.

Figure 3.1 shows that the generated neural network gives reasonable prediction results for the data from the data set.

Figure 3.1 also shows that, if there are abrupt changes in a training data set, while the major part of the data set is normal, the trained network performs satisfactorily for the normal part of the data set. However, it gives poor results for the abnormal data. One possibility to enable the trained network to accommodate data with unforeseen changes, is to superimpose small disturbance in the form of step changes with small magnitude on the data set and after that modification, to use the data set for training. The network trained in this way, can work not only with normal data but with abnormal data as well.

4. CONCLUSION

Neural and genetic methods of predictions are applicable to data forecasting. Though conventional methods are still in use and show their potential, their weakness is the relatively complex mathematical treatment. The major advantage of the neural approach is that no assumptions need to be made about the underlying function or model. The problem of network design can be solved by using the genetic algorithm.

The results of application of the neuro-genetic methods for the generation of forecasting model for the number of bus tickets sold at the bus station for the stated period on the daily basis, suggest that the generated and tested neural network is an acceptable prediction tool.

SAŽETAK

PREDVIĐANJE PRODAJE AUTOBUSNIH KARATA PRIMJENOM NEURO-GENETSKIH METODA

U radu je opisana primjena neuronske mreže i genetskog algoritma u predviđanju prodaje autobusnih karata. U odnosu na postojeće metode predviđanja predloženi pristup ima određene prednosti. Izgrađena neuronska mreža samostalno uspostavlja funkcije odnosno modelira odnose između varijabli promatrane pojave. Šira primjena neuronskih mreža u

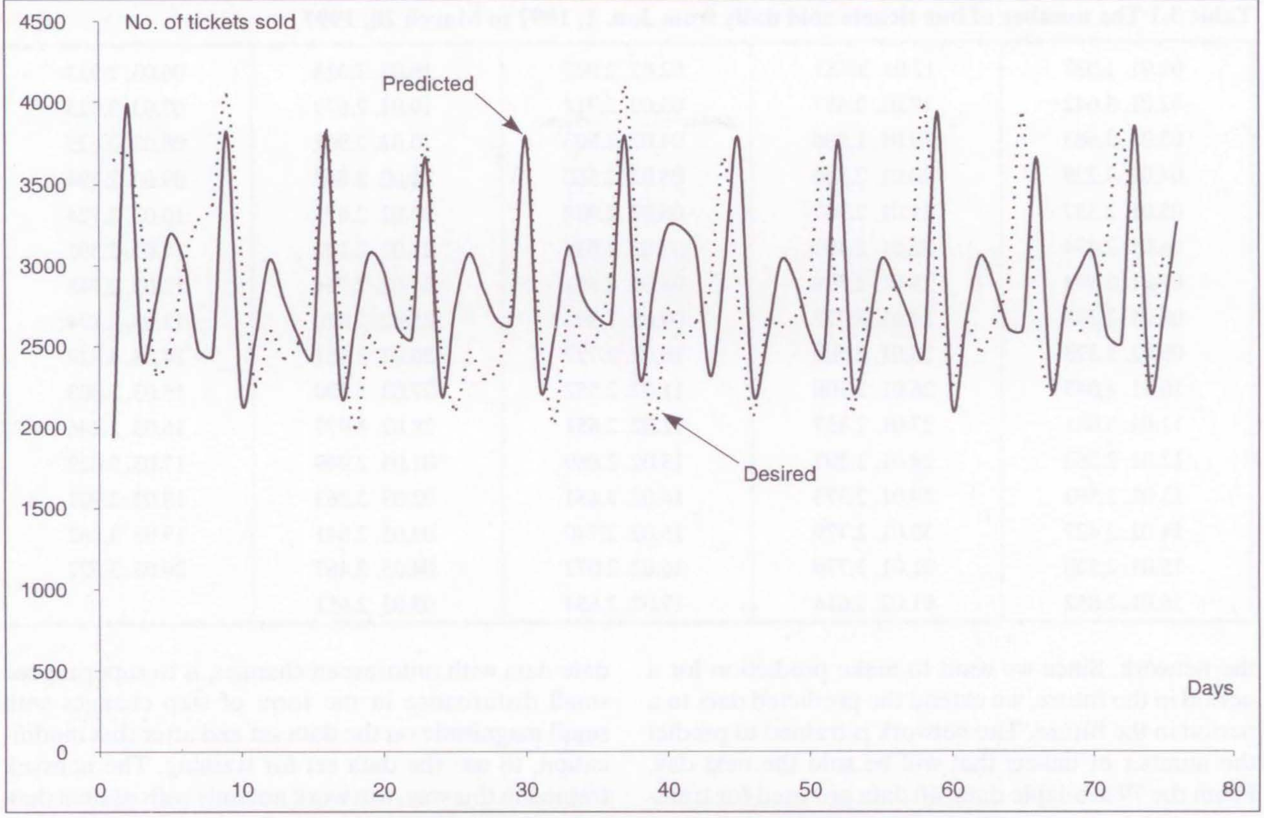

Figure 3.1 - Generated network prediction results for the data set

postupku predviđanja uvjetovana je određenom složenošću oblikovanja njezine arhitekture. Poteškoće pri oblikovanju neuronske mreže od strane čovjeka izbjegavaju se korištenjem genetskog algoritma.

REFERENCES:

[1] **G. Štefančić**: *The Model of Advancement of the Traffic Technology on the Bus Station*. PhD thesis, Faculty of Traffic Sciences, Zagreb, 1993. (in Croatian)

[2] **N.T. Thomopoulos**: *Applied Forecasting Methods*. Prentice Hall, Englewood Cliffs, N.J., 1980.

[3] **D.T. Pham, X. Liu**: *Neural Networks for Identification, Prediction and Control*. Springer Verlag, London, 1995.

[4] **D.E. Rumelhart, J.L. McClelland**: *Parallel Distributed Processing*. Vol.1, MIT Press, Cambridge, MA, 1986.

[5] **J.H. Holland**: *Genetic Algorithms*. Scientific American, Vol. 267, pp. 66-72, 1992.

[6] *NeuroGeneticOptimizer*, BioComp Systems, Inc., 1997.