An optimisation model for traffic distribution forecasting in packet-switching networks

Faramak Zandi

Faculty of Technology and Engineering,
Alzahra University,
Vanak, Tehran 19938-91176, Iran
E-mail: zandi@iust.ac.ir

Madjid Tavana*

Lindback Distinguished Chair of Information Systems,
La Salle University,
Philadelphia, PA 19141, USA
Fax: +1 267 295 2854
E-mail: tavana@lasalle.edu
*Corresponding author

Abstract: Traffic distribution forecasting is an essential step in network planning for packet-switching networks. It is frequently necessary to forecast and develop an optimal network configuration to meet the requirements of new traffic demands or changes in the existing demands. Several methods have been developed to forecast network configurations. While these studies have revealed some interesting traffic characteristics, little progress has been made in developing good models for the purpose of traffic engineering and performance prediction. We propose a novel multi-objective optimisation model for traffic distribution forecasting in packet-switching networks by mapping these networks into multi-commodity networks. Initially, the radial basis function (RBF) network is used to monitor and learn the current real-traffic distribution. Next, a quadratic model is used to calibrate these functions for a precise traffic distribution. The proposed multi-objective optimisation method can effectively and efficiently forecast the traffic distribution of packet-switching networks in both crisp and fuzzy environments. A numerical example is presented to demonstrate the application and effectiveness of this model.

Keywords: fuzzy; multi-objective optimisation; RBF network; multi-commodity network; traffic distribution matrix.

Reference to this paper should be made as follows: Zandi, F. and Tavana, M. (2010) 'An optimisation model for traffic distribution forecasting in packet-switching networks', *Int. J. Mathematics in Operational Research*, Vol. 2, No. 5, pp.515–539.

Biographical notes: Faramak Zandi is an Assistant Professor of Information Systems and Chairman of the Industrial Engineering Department at Alzahra University in Iran. He holds a PhD in Industrial Engineering. His research interests include information technology, enterprise architectures, decision making, quality management systems and transportation planning. He has published in the *Int. J. Business Information Systems*, *Int. J. Information*

Copyright © 2010 Inderscience Enterprises Ltd.

Technology and Management, *Int. J. Mathematics in Operational Research, IEEE Computer Society*, *Journal of Tehran University*, *Amirkabir Journal of Science and Technology* and *Review Quarterly Journal of Educational Innovations*. He has presented in the *International Conferences on Business and Information (BAI2008)*, *Urban Transport and the Environment in the 21st Century in Transactions of the Wessex Institute*, *Information Communication Technology Management*, *Information Technology Project Management* and *International Conference of Iranian Operations Research Society*.

Madjid Tavana is a Professor of Management Information Systems and Decision Sciences and the Lindback Distinguished Chair of Information Systems at La Salle University where he served as Chairman of the Management Department and Director of the Center for Technology and Management. He has been a distinguished faculty fellow at NASA's Kennedy Space Center, NASA's Johnson Space Center, Naval Research Laboratory – Stennis Space Center and Air Force Research Laboratory. He was awarded the prestigious Space Act Award by NASA in 2005. He holds an MBA, a PMIS and a PhD in Management Information Systems. He has received his post-doctoral diploma in strategic information systems from the Wharton School of the University of Pennsylvania. He is the Editor-in-Chief for the *Int. J. Applied Decision Sciences* and the *Int. J. Strategic Decision Sciences*. He has published in journals such as *Decision Sciences*, *Interfaces*, *Information Systems*, *Information and Management*, *Computers and Operations Research*, *Journal of the Operational Research Society* and *Advances in Engineering Software*, among others.

1 Introduction

Packet switching is a network communications method that groups all transmitted data into suitably sized blocks called packets. The network over which packets are transmitted is a shared network where any host connecting to the network can, in theory, send packets to any other host. Each packet contains address information that identifies the sending computer and intended recipient. Using these addresses, network switches and routers determine how best to transfer the packet to its destination. The principal goals of packet switching are to optimise utilisation of available link capacity and to increase the robustness of communication. Since packet networks often demand a capacity review, a network vision and organised plan should be determined to accommodate future growth in capacity. The planning process requires information about the network capacity and utilisation. Traffic matrices could be used to determine network access resource requirements and details such as location, size, operation, protocol support, performance characteristics and device type. The traffic distribution matrix represents the traffic volume between the origin/destination (O/D) pairs in a network. Traffic distribution forecasting is an essential step in network planning for packet-switching networks.

2 Literature review and background

Network planning and forecasting is a common problem in traffic engineering and capacity planning of packet-switching networks. It is frequently necessary to forecast and develop an optimal network configuration to meet the requirements of new traffic demands or changes in the existing demands. Such configurations may involve changing the existing topology, resizing network elements and adding or removing switches and links. Network planning and forecasting requires specifications of the existing and projected O/D traffic demands (Conway and Li, 2003). The demand data derive network design and are the primary input parameters in most commercially available network capacity planning tools.

In practice, the total traffic generated at a particular node may be known or easily measurable, however, the O/D traffic demands in an existing packet-switching network may not be easily measured. In addition, as the size of a network increases, the problem of measuring traffic demands is compounded since the number of proportions increases exponentially with the number of nodes. Feldmann et al. (2001) have proposed a model of traffic demands that captures

1 the volume of data

2 the entry point into the network

3 the destination reachability information.

They faced challenges in their demand model, such as working with different datasets, ambiguity of ingress points and dynamics of the egress points.

In recent years, several methods have been developed to forecast traffic matrices. Zhang et al. (2003b) presented a new approach to traffic estimation and demonstrated with real data that their method is fast, accurate, flexible and robust. They also acknowledged that there is still considerable work to do in this area. Teixeira et al. (2005) used the integer programming and multi-commodity algorithm proposed in Teixeira et al. (2004) to show that

1 although the likelihood of large traffic fluctuations is small, big changes do sometimes occur

2 most routing changes do not cause much variation in the traffic matrix

3 routing changes are responsible for many of the large traffic shifts.

Zhang et al. (2005) later used fast algorithms based on modern convex optimisation theory for traffic matrix estimation using a regularisation based on 'entropy penalisation'. Their solution chose the traffic matrix consistent with the measured data that was information theoretically closest to a model, in which S/D pairs were stochastically independent.

The performance of estimated traffic matrices in traffic engineering have been studied by Zhang et al. (2003a) and Roughan et al. (2003). Zhang et al. (2003a) draw from ideas in 'gravity modelling' (Lam et al., 1997; Medina et al., 2002; Roughan et al., 2002) and

'tomographic methods' (Adams et al., 2000; Cao et al., 2000; Coates et al., 2002) to compute traffic matrix estimates from widely available data: link load and network routing and configuration data. Zhao et al. (2006) showed that dirty data can contaminate a traffic matrix and proposed a comprehensive solution for robust traffic matrix estimation by using multiple readily available but imperfect data. Their algorithm estimated traffic matrices more accurately upon topology and routing changes. They also showed that their algorithm reduces the errors in traffic matrix estimation by more than an order of magnitude. Liu et al. (1999) proposed a hierarchical model with multiple levels each of which contained several classes of applications. At the packet level, they used deterministic and exponential packet interarrival times within the flows. They used simulation and demonstrated that their approach provides quality performance prediction and forecasting. All the above mentioned methods have focused on computer networks as single commodity networks.

While these studies have revealed some interesting traffic characteristics, little progress has been made in developing good models for the purpose of traffic engineering and performance prediction. We propose a new traffic distribution forecasting model that considers same size packets as a commodity and maps packet-switching networks into multi-commodity networks. A radial basis function (RBF) network is used to monitor and learn the current real-traffic distribution and quadratic modelling is used to calibrate these functions for precise traffic distribution forecasting. This paper is organised into five sections. This novel approach enables us to utilise the conventional methods in modelling multi-commodity networks and provides traffic distribution forecasting classifications based on the packet size. Section 3 presents a detailed description of the proposed method followed by a numerical example in Section 4 and conclusions and futures research directions in Section 5.

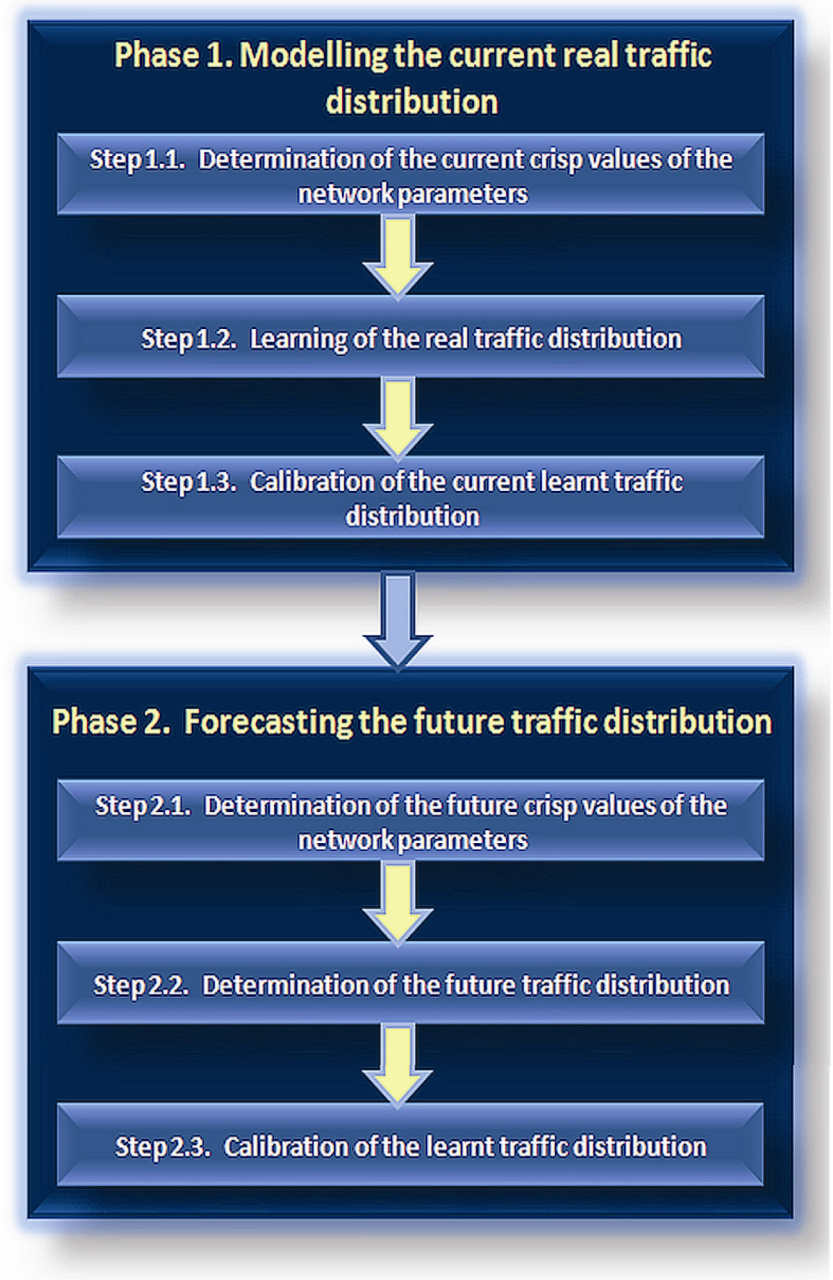
3 The proposed method

The proposed method depicted in Figure 1 is composed of two phases. Each phase is comprised three steps. In phase 1, we model the current real-traffic distribution; and in phase 2, we forecast the future traffic distribution.

3.1 Phase 1: modelling the current real-traffic distribution

To model the current traffic distribution, we characterise the packet-switching networks as multi-commodity networks, where any packet with size k is viewed as a commodity. Phase 1 is divided into three steps. In Step 1.1, we determine the current crisp values of the network parameters. In Step 1.2, we learn the current real-traffic distribution and in Step 1.3, we calibrate the learnt current traffic distribution.

Figure 1 The proposed framework for determination of the future traffic distribution (see online version for colours)



Step 1.1 Determination of the current crisp values of the network parameters

Let us define the network parameters for a packet-switching network with Q nodes and R arcs as follows:

$\tilde{T}_{ij}^{k}$ is the current number of packets (with size k) transmitted from node i to node j (the current real-traffic distribution)

$\hat{T}_{ij}^{k}$ is the current learnt traffic distribution

$\tilde{P}_i$ is the total current data sent from node i

$\tilde{A}_j$ is the total current data received at node j

$\tilde{D}_{ij}$ is the current physical distance between nodes i and j

$\tilde{L}_{ij}^{k}$ is the maximum rate at which the sender is planning to send packets (for packets with size k)

a_{ij} is the current packet size transferred from node i to node j

C_{ij} is the current upper bound of the capacity of the arc from node i to node j

Let us further assume that the elements of input pattern vectors $\underline{\tilde{X}}_{ij}^{k} = [\tilde{P}_i, \tilde{A}_j, \tilde{D}_{ij}, \tilde{L}_{ij}^{k}]$, $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$, are non-negative fuzzy numbers of trapezoidal form as follows:

$$\begin{aligned}
\tilde{P}_i &= \left(P_{i1}, P_{i2}, P_{i3}, P_{i4}\right) \\
\tilde{A}_j &= \left(A_{j1}, A_{j2}, A_{j3}, A_{j4}\right) \\
\tilde{L}_{ij}^{k} &= \left(L_{ij1}^{k}, L_{ij2}^{k}, L_{ij3}^{k}, L_{ij4}^{k}\right) \\
\tilde{D}_{ij} &= \left(D_{ij1}, D_{ij2}, D_{ij3}, D_{ij4}\right)
\end{aligned} \tag{1}$$

We use the Fuller method (Zimmermann, 2001) to defuzzify the fuzzy elements of these vectors to the following crisp elements:

$$\begin{aligned}
P_i &= \frac{1}{3}\left[P_{i1} + P_{i3} + \frac{1}{2}\left(P_{i2} + P_{i4}\right)\right] \\
A_j &= \frac{1}{3}\left[A_{j1} + A_{j3} + \frac{1}{2}\left(A_{j2} + A_{j4}\right)\right] \\
L_{ij}^{k} &= \frac{1}{3}\left[L_{ij1}^{k} + L_{ij3}^{k} + \frac{1}{2}\left(L_{ij2}^{k} + L_{ij4}^{k}\right)\right] \\
D_{ij} &= \frac{1}{3}\left[D_{ij1} + D_{ij3} + \frac{1}{2}\left(D_{ij2} + D_{ij4}\right)\right] \\
T_{ij}^{k} &= \frac{1}{3}\left[T_{ij1}^{k} + T_{ij3}^{k} + \frac{1}{2}\left(T_{ij2}^{k} + T_{ij4}^{k}\right)\right]
\end{aligned} \tag{2}$$

Therefore, we have $\underline{X}_{ij}^{k} = [P_i, A_j, D_{ij}, L_{ij}^{k}]$

Next, we use linear scale transformation to convert the above crisp values into comparable scale-less parameters by dividing their value by their maximum possible value:

$$P(s)_i = \frac{P_i}{\mathrm{Max}_i P_i},\ A(s)_i = \frac{A_i}{\mathrm{Max}_i A_i} \text{ and } L(s)_{ij}^{k} = \frac{L_{ij}^{k}}{\mathrm{Max}_i L_{ij}^{k}} \tag{3}$$

For the negative criterion $D(s)_{ij}$, the minimum possible value is divided by its value:

$$D(s)_{ij} = \frac{\mathrm{Min}_i D_{ij}}{D_{ij}} \tag{4}$$

Therefore, the input scale-less pattern vectors can be written in the following form:

$$\underline{X}_{ij}^{k}(s) = \left[P(s)_i, A(s)_j, D(s)_{ij}, L(s)_{ij}^{k} \right] \tag{5}$$

This linear scale transformation is needed to ensure that the relative order of the magnitudes for all parameters remains equal.

Step 1.2 Learning of the real-traffic distribution

According to the RBF (Haykin, 1998; Poggio and Girosi, 1990; Tikhonov, 1973), the input pattern vectors

$$\underline{X}_{ij}^{k}(s) = \left[P(s)_i, A(s)_j, D(s)_{ij}, L(s)_{ij}^{k} \right]$$

are used to compute corresponding output patterns $\hat{T}_{ij}^{k}$. Then, these computed output patterns are compared to the desired output patterns T_{ij}^{k} and the errors $e_{ij} = T_{ij}^{k} - \hat{T}_{ij}^{k}$ are determined. Next, in order to learn the current real-traffic distribution, we begin by minimising the sum of the square errors between the current real-traffic distribution and the current learnt traffic distribution:

$$\mathrm{SSE} = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m} \left(T_{ij}^{k} - \hat{T}_{\mathrm{ij}}^{k} \right)^2 \tag{6}$$

The following approximation function will minimise Equation (5) in large-scale packet-switching networks based on N input training pattern vectors:

$$\hat{T}_{ij}^{k} = \sum_{f=1}^{N} w_f^k G\left(\underline{X}_{ij}^{k}(s), \underline{X}_{f}^{k}(\mathrm{SA}) \right) \tag{7}$$

with $\underline{X}_{f}^{k}(\mathrm{SA})$ as a set of input training pattern vectors (training data) for the RBF neural network, we define the Gaussian exponential functions $G(\underline{X}_{ij}^{k}(s), \underline{X}_{f}^{k}(\mathrm{SA}))$ with centroid $\underline{X}_{f}^{k}(\mathrm{SA})$ as:

$$G\left(\underline{X}_{ij}^{k}(s), \underline{X}_{f}^{k}(\mathrm{SA})\right) = \mathrm{e}^{-\left(\left\|\underline{x}_{ij}^{k}(s)-\underline{x}_{f}^{k}(\mathrm{SA})\right\|\right)^{2}} \tag{8}$$

where w_f^k is the fth element of the weight vector $\underline{W}^k(\mathrm{SA}) = \left[w_1^k, w_2^k, \ldots, w_N^k\right]^{\mathrm{T}}$

Next, we rewrite this equation in a matrix form

$$\underline{T}^k(\mathrm{SA}) = G^k \cdot \underline{W}^k(\mathrm{SA}) \tag{9}$$

where $\underline{T}^k(\mathrm{SA}) = [T_1^k, T_2^k, \ldots, T_N^k]^{\mathrm{T}}$ is the current vector of the desired output pattern for learning traffic distribution.

$$\begin{aligned} G^k &= \begin{bmatrix} \mathrm{e}^{-\left(\left\|\underline{X}_{11}^{k}(s)-\underline{X}_{1}^{k}(\mathrm{SA})\right\|\right)^{2}} & \cdots & \mathrm{e}^{-\left(\left\|\underline{X}_{11}^{k}(s)-\underline{X}_{N}^{k}(\mathrm{SA})\right\|\right)^{2}} \\ \vdots & & \vdots \\ \mathrm{e}^{-\left(\left\|\underline{X}_{R}^{k}(s)-\underline{X}_{1}^{k}(\mathrm{SA})\right\|\right)^{2}} & \cdots & \mathrm{e}^{-\left(\left\|\underline{X}_{R}^{k}(s)-\underline{X}_{N}^{k}(\mathrm{SA})\right\|\right)^{2}} \end{bmatrix} \\ &= \begin{bmatrix} G\left(\underline{X}_{11}^{k}(s), \underline{X}_{1}^{k}(\mathrm{SA})\right) & \cdots & G\left(\underline{X}_{11}^{k}(s), \underline{X}_{N}^{k}(\mathrm{SA})\right) \\ \vdots & & \vdots \\ G\left(\underline{X}_{R}^{k}(s), \underline{X}_{1}^{k}(\mathrm{SA})\right) & \cdots & G\left(\underline{X}_{R}^{k}(s), \underline{X}_{N}^{k}(\mathrm{SA})\right) \end{bmatrix} \end{aligned} \tag{10}$$

Since the matrix G^k is not a square matrix, the approximation functions (9) can be calculated by the following weight vector:

$$\underline{W}^K(\mathrm{SA}) = (G^{k^{\mathrm{T}}} \cdot G^k)^{-1} \cdot G^{k^{\mathrm{T}}} \cdot \underline{T}^k(\mathrm{SA}) \tag{11}$$

Since $(G^{k^{\mathrm{T}}} \cdot G^k)^{-1} \cdot G^{k^{\mathrm{T}}}$ is the pseudo-inverse of the matrix G^k, the current real-traffic distribution is learnt as follows:

$$\underline{\hat{T}}^k = G^k \cdot \underline{W}^k(\mathrm{SA}) \tag{12}$$

where $\underline{\hat{T}}^k$ is the vector of the computed output pattern for the current learning traffic distribution.

Step 1.3 Calibration of the current learnt traffic distribution

Next, we propose the following fuzzy multi-objective model for calibrating the current learnt traffic distribution in Step 1.2:

Objective function 1: using the conservation of flow (Bazaraa et al., 1990), the sum of data sent from node i should be equal to $P_i(s)$:

$$\sum_{j=1}^{m} \hat{T}_{ij}^k \cdot a_{ij} \cong P_i \quad i = 1, 2, \ldots, n \tag{13}$$

Here $\cong$ denotes the fuzzified version of = and has a linguistic interpretation of 'essentially equal to'. On the other hand, with regards to the quality of service parameter

of cell loss ratio, the sum of the data sent from node i do not necessarily equal to P_i. Therefore, $\hat{T}_{ij}^k$ must be calibrated by the calibration coefficients x_i^k, x_j^k as follows:

$$\sum_{j=1}^{m} \hat{T}_{ij}^k \cdot x_i^k \cdot x_j^k \cdot a_{ij} \cong P_i \quad i = 1, 2, \ldots, n \tag{14}$$

or

$$f_i^k = P_i - \sum_{j=1}^{m} \hat{T}_{ij}^k \cdot x_i^k \cdot x_j^k \cdot a_{ij} \cong 0 \quad i = 1, 2, \ldots, n \tag{15}$$

Objective function 2: similarly, the sum of the data received for destination node j in the network should be equal to A_j. On the other hand, with regards to the quality of service parameter of cell misinsertion rate, the sum of the data received for destination nodes j do not necessarily equal to A_j. Therefore, $\hat{T}_{ij}^k$ must be calibrated by the calibration coefficients x_i^k, x_j^k as follows:

$$\sum_{i=1}^{n} \hat{T}_{ij}^k \cdot x_i^k \cdot x_j^k \cdot a_{ij} \cong A_j \quad j = 1, 2, \ldots, m \tag{16}$$

or

$$f_j^k = A_j - \sum_{i=1}^{n} \hat{T}_{ij}^k \cdot x_i^k \cdot x_j^k \cdot a_{ij} \cong 0 \tag{17}$$

Objective function 3: the sum of the data received at intermediate node should be equal to the sum of the data sent from this node*:*

$$\sum_{H=1}^{h} \hat{T}_{iH}^k \cdot a_{ih} \cong \sum_{H=1}^{h} \hat{T}_{hj}^k \cdot a_{Hj} \tag{18}$$

On the other hand, with regards to the quality of service parameter of severely errored cell block ratio (the fraction of *N*-cell blocks of *M* or more cells containing an error), the sum of the data received by each intermediate node does not necessarily equal to the sum of the data sent from it, $\hat{T}_{ih}^k$ and $\hat{T}_{hj}^k$ must be calibrated as follows:

$$\sum_{H=1}^{h} \hat{T}_{iH}^k \cdot x_i^k \cdot x_H^k \cdot a_{iH} \cong \sum_{H=1}^{h} \hat{T}_{Hj}^k \cdot x_H^k \cdot x_j^k a_{Hj} \tag{19}$$

or

$$f_H^k = \sum_{H=1}^{h} \hat{T}_{iH}^k \cdot x_i^k \cdot x_H^k \cdot a_{iH} - \sum_{H=1}^{h} \hat{T}_{Hj}^k \cdot x_H^k \cdot x_j^k a_{Hj} \cong 0 \tag{20}$$

where h denotes the number of intermediate nodes between source hosts and destination hosts.

Objective function 4: in addition, the distribution of the total number of calculated transmitted bits vs. the physical distance for data should be equal to the distribution of the total number of real transmitted bits vs. the physical distance for data (Medina et al., 2002):

$$\sum_{D} \hat{T}_{ij}^{k} \cdot a_{ij} \cdot x_i^k \cdot x_j^k - \sum_{D} T_{ij}^{k} \cdot a_{ij} \cong 0 \tag{21}$$

or

$$f_D^k = \sum_{D} \hat{T}_{ij}^{k} \cdot x_i^k \cdot x_j^k - \sum_{D} T_{ij}^{k} \cong 0 \tag{22}$$

Constraints: the total transmitted data through each arc cannot exceed the arc capacity. This implies that:

$$\sum_{k} \hat{T}_{ij}^{k} \cdot x_i^k \cdot x_j^k \cdot a_{ij} \le C_{ij} \tag{23}$$

The above objective functions and constraints collectively form a fuzzy multi-objective decision-making model. Next, we convert the fuzzy objective function into a corresponding constraint using the method proposed by Hwang (1989), Hwang et al. (1989) and Zimmermann (2001):

$$\text{Find } x_i^k, x_j^k \tag{24}$$

s.t.

$$P_i - \sum_{j=1}^{m} \hat{T}_{ij}^{k} \cdot x_i^k \cdot x_j^k \cdot a_{ij} \cong 0$$

$$A_j - \sum_{i=1}^{n} \hat{T}_{ij}^{k} \cdot x_i^k \cdot x_j^k \cdot a_{ij} \cong 0$$

$$\sum_{H=1}^{h} \hat{T}_{iH}^{k} \cdot x_i^k \cdot x_H^k \cdot a_{iH} - \sum_{H=1}^{h} \hat{T}_{Hj}^{k} \cdot x_H^k \cdot x_j^k a_{Hj} \cong 0$$

$$\sum_{D} \hat{T}_{ij}^{k} \cdot x_i^k \cdot x_j^k - \sum_{D} T_{ij}^{k} \cong 0$$

$$\sum_{k} \hat{T}_{ij}^{k} \cdot x_i^k \cdot x_j^k \cdot a_{ij} \le C_{ij}$$

$$x_i^k, x_j^k, x_h^k \ge 0$$

Using the above formulation, the optimal solution can be found by solving a single objective fuzzy model as the model simultaneously calibrates all of the approximation

functions. Next, we use the following approach suggested by Zimmermann (2001) to find the optimal solution of the above fuzzy model:

$$\text{Minimise}\left[\text{Min}\left[\mu_i\left(\hat{T}_{ij}^k\right),\mu_j\left(\hat{T}_{ij}^k\right),\mu_D\left(\hat{T}_{ij}^k\right),\mu_H\left(\hat{T}_{ij}^k\right)\right]\right] \tag{25}$$

Introducing λ as a new variable, we have:

$$\lambda = \text{Min}\left[\mu_i\left(\hat{T}_{ij}^k\right),\mu_j\left(\hat{T}_{ij}^k\right),\mu_D\left(\hat{T}_{ij}^k\right),\mu_H\left(\hat{T}_{ij}^k\right)\right] \tag{26}$$

Let us define the membership functions as:

$$\mu_j(\hat{T}_{ij}^k) = \begin{cases} 1+\dfrac{f_j^k}{r_j} & -r_j \le f_j^k \le 0 \\ 1-\dfrac{f_j^k}{r_j} & 0 \le f_j^k \le r_j \\ 0 & \text{otherwise} \end{cases} \tag{27}$$

or equivalently, substituting Equation (27) in Equation (26), we get:

$$\begin{cases} (\lambda-1)q_i \le f_i^k \le (1-\lambda)q_i \\ (\lambda-1)r_j \le f_j^k \le (1-\lambda)r_j \\ (\lambda-1)s_H \le f_H^k \le (1-\lambda)s_H \\ (\lambda-1)v_D \le f_D^k \le (1-\lambda)v_D \end{cases} \tag{28}$$

where the decision-maker can subjectively associate the tolerances r_j in Equation (28) and q_i, s_H and v_D in similar membership functions as constants for admissible violations of the objective functions. Consequently, the new proposed model (*P*) is obtained from Equations (24), (25) and (28):

Min λ Model (*P*)

s.t.

$$(\lambda-1)q_i \le P_i - \sum_{j=1}^{m}\hat{T}_{ij}^k(I)\cdot a_{ij} \le (1-\lambda)q_i$$

$$(\lambda-1)r_j \le A_j - \sum_{i=1}^{n}\hat{T}_{ij}^k(I)\cdot a_{ij} \le (1-\lambda)r_j$$

$$(\lambda-1)s_H \le \sum_{H=1}^{h}\hat{T}_{iH}^k(I)\cdot a_{iH} - \sum_{H=1}^{h}\hat{T}_{Hj}^k(I)\cdot a_{Hj} \le (1-\lambda)s_H$$

$$(\lambda-1)v_D \le \sum_D \hat{T}_{ij}^k(I) - \sum_D T_{ij}^k \le (1-\lambda)v_D$$

$$\sum_k \hat{T}_{ij}^k(I)\cdot a_{ij} \le C_{ij}$$

$$\hat{T}_{ij}^k(I) = \hat{T}_{ij}^k \cdot x_i^k \cdot x_j^k$$

$$\hat{T}_{iH}^k(I) = \hat{T}_{iH}^k \cdot x_i^k \cdot x_H^k$$

$$\hat{T}_{Hj}^k(I) = \hat{T}_{Hj}^k \cdot x_H^k \cdot x_j^k$$

$$x_i^k, x_j^k, x_h^k, \lambda, \hat{T}_{ij}^k(I), \hat{T}_{iH}^k(I), \hat{T}_{Hj}^k(I) \ge 0$$

The optimal solution to model (*P*) is $\lambda^*, x_i^{k*}, x_j^{k*}, x_h^{k*}, \hat{T}_{ij}^{k*}(I), \hat{T}_{iH}^{k\,*}(I), \hat{T}_{Hj}^{k\,*}(I)$.

Therefore, this model (*P*) calibrates the functions (14) as follows:

$$\begin{aligned} \hat{T}_{ij}^{k*}(I) &= \hat{T}_{ij}^k \cdot x_i^{k*} \cdot x_j^{k*} \\ \hat{T}_{iH}^{k\,*}(I) &= \hat{T}_{iH}^k \cdot x_i^{k*} \cdot x_H^{k\,*} \\ \hat{T}_{Hj}^{k\,*}(I) &= \hat{T}_{Hj}^k \cdot x_H^{k\,*} \cdot x_j^{k*} \end{aligned} \tag{29}$$

Fortunately, since most of the constraints in model (*P*) are quadratic (convex), Lingo software (Lindo Systems, 2004) can be used for solving this model (Winston, 1987).

3.2 Phase 2: forecasting the future traffic distribution

In this phase, we propose the following three steps. In Step 2.1, we determine future crisp values of the network parameters. In Step 2.2, we specify the future traffic distribution and in Step 2.3, we calibrate the learnt future traffic distribution as follows:

Step 2.1 Determination of the future crisp values of the network parameters

Let us model the packet-switching network in phase 1 as a packet-switching network with Q' nodes and R' arcs in the future with the following network parameters:

$\tilde{P}_i'$ is the number of packets (with size k) sent from node *i* in the future

$\tilde{A}_j'$ is the number of packets (with size k) received at node *j* in the future

$\tilde{D}_{ij}'$ is the physical distance between nodes *i* and *j* in the future

$\tilde{L}_{ij}^{k\prime}$ is the average transit time from node i to node j in the future (for packets with size k)

a_{ij}' is the packet size transferred from node i to node j in the future

$\hat{T}_{ij}^{k\prime}$ is the traffic distribution in the future

C_{ij}' is the upper bound of the capacity of the arc from node i to node j in the future.

Let us assume that the elements of input vectors $\underline{\tilde{X}}_{ij}^{k\prime} = [\tilde{P}_i^{k\prime}, \tilde{A}_j^{k\prime}, \tilde{D}_{ij}{}', \tilde{L}_{ij}^{k\prime}]$, in the future, $i = 1, 2, \ldots, n'$ and $j = 1, 2, \ldots, m'$, are non-negative fuzzy numbers of trapezoidal form as follows:

$$\begin{aligned}
\tilde{P}_i' &= \left(P_{i1}^{k\prime}, P_{i2}^{k\prime}, P_{i3}^{k\prime}, P_{i4}^{k\prime}\right) \\
\tilde{A}_j{}' &= \left(A_{j1}^{k}{}', A_{j2}^{k}{}', A_{j3}^{k}{}', A_{j4}^{k}{}'\right) \\
\tilde{L}_{ij}^{k\prime} &= \left(L_{ij1}^{k}{}', L_{ij2}^{k}{}', L_{ij3}^{k}{}', L_{ij4}^{k}{}'\right) \\
\tilde{D}_{ij}{}' &= \left(D_{ij1}{}', D_{ij2}{}', D_{ij3}{}', D_{ij4}{}'\right)
\end{aligned} \tag{30}$$

We use the Fuller method (Zimmermann, 2001) to defuzzify the fuzzy elements of these vectors in the future to the following crisp elements:

$$\begin{aligned}
P_i' &= \frac{1}{3}\left[P_{i1}{}' + P_{i3}{}' + \frac{1}{2}\left(P_{i2}{}' + P_{i4}{}'\right)\right] \\
A_j{}' &= \frac{1}{3}\left[A_{j1}{}' + A_{j3}{}' + \frac{1}{2}\left(A_{j2}{}' + A_{j4}{}'\right)\right] \\
L_{ij}^{k\prime} &= \frac{1}{3}\left[L_{ij1}^{k}{}' + L_{ij3}^{k}{}' + \frac{1}{2}\left(L_{ij2}^{k}{}' + L_{ij4}^{k}{}'\right)\right] \\
D_{ij}{}' &= \frac{1}{3}\left[D_{ij1}{}' + D_{ij3}{}' + \frac{1}{2}\left(D_{ij2}{}' + D_{ij4}{}'\right)\right] \\
T_{ij}^{k\prime} &= \frac{1}{3}\left[T_{ij1}^{k\prime} + T_{ij3}^{k}{}' + \frac{1}{2}\left(T_{ij2}^{k}{}' + T_{ij4}^{k}{}'\right)\right]
\end{aligned} \tag{31}$$

Therefore, we have: $\underline{X}_{ij}^{k\prime} = \left[P_i', A_j{}', D_{ij}{}', L_{ij}^{k\prime}\right]$

Next, we use linear scale transformation to convert the above crisp values in the future into comparable scale-less parameters by dividing their value by their maximum possible value of:

$$P(s)_i{}' = \frac{P_i'}{\mathrm{Max}_i P_i'},\ A(s)_i{}' = \frac{A_i'}{\mathrm{Max}_i A_i'} \text{ and } L(s)_{ij}^{k\prime} = \frac{L_{ij}^{k\prime}}{\mathrm{Max}_i L_{ij}^{k\prime}} \tag{32}$$

For the negative criterion $D(s)_{ij}{}'$, the minimum possible value is divided by its value:

$$D(s)_{ij}{}' = \frac{\mathrm{Min}_i D_{ij}{}'}{D_{ij}{}'} \tag{33}$$

Therefore, the future input scale-less vectors can be written in the following form:

$$\underline{X}_{ij}^{k}(s)' = \left[P(s)_i{}', A(s)_j{}', D(s)_{ij}{}', L(s)_{ij}^{k\prime}\right] \tag{34}$$

Step 2.2 Determination of the future traffic distribution

Using Equations (12) and (29), the future traffic distribution ($\hat{T}_{ij}^{k\prime}$) can be calculated as:

$$\begin{aligned}
\hat{T}_{ij}^{k\prime}(I) &= \hat{T}_{ij}^{k\prime} \cdot x_i^{k*} \cdot x_j^{k*} \\
\hat{T}_{iH}^{k\prime}(I) &= \hat{T}_{iH}^{k\prime} \cdot x_i^{k*} \cdot x_H^{k*} \\
\hat{T}_{Hj}^{k\prime}(I) &= \hat{T}_{Hj}^{k\prime} \cdot x_H^{k*} \cdot x_j^{k*}
\end{aligned} \tag{35}$$

where

$$\underline{\hat{T}}^{k\prime} = G^{k\prime} \cdot \underline{W}^k(\mathrm{SA})$$

and

$\underline{\hat{T}}^{k\prime}$ is the computed output vector for forecasting future traffic distribution.

$$G^{k\prime} = \begin{bmatrix} G\left(\underline{X}_{11}^k(s)', \underline{X}_1^k(\mathrm{SA})\right) & \cdots & G\left(\underline{X}_{11}^k(s)', \underline{X}_N^k(\mathrm{SA})\right) \\ \vdots & & \vdots \\ G\left(\underline{X}_{R'}^k(s)', \underline{X}_1^k(\mathrm{SA})\right) & \cdots & G\left(\underline{X}_{R'}^k(s)', \underline{X}_N^k(\mathrm{SA})\right) \end{bmatrix} \tag{36}$$

In Equation (35), we have used the calibration coefficients obtained in Step 1.3 to find the traffic distribution in future.

Step 2.3 Calibration of the future learnt traffic distribution

Next, we propose the following fuzzy multi-objective model for calibrating the future learnt traffic distribution in Step 2.2 by feeding the sub-optimal solution for the traffic distribution into the following model to improve the calibration coefficients and obtain the calibrated optimal solution.

$$\text{Min } \lambda' \qquad \text{Model } (F)$$

s.t.

$$(\lambda' - 1)q_i' \le P_i' - \sum_{j=1}^{m'} \hat{T}_{ij}^{k\prime}(I) \cdot a_{ij}' \le (1-\lambda')q_i'$$

$$(\lambda' - 1)r_j' \le A_j' - \sum_{i=1}^{n'} \hat{T}_{ij}^{k\prime}(I) \cdot a_{ij}' \le (1-\lambda')r_j'$$

$$(\lambda' - 1)s_H' \le \sum_{H=1}^{h'} \hat{T}_{iH}^{k\prime}(I) \cdot a_{iH}' - \sum_{H=1}^{h'} \hat{T}_{Hj}^{k\prime}(I) a_{Hj}' \le (1-\lambda')s_H'$$

$$\sum_k \hat{T}_{ij}^{k\prime}(I) \cdot a_{ij}' \le C_{ij}'$$

$$\hat{T}_{ij}^{k\prime}(I) = \hat{T}_{ij}^{k\prime} \cdot x_i^{k\prime} \cdot x_j^{k\prime}$$

$$\hat{T}_{iH}^{k\prime}(I) = \hat{T}_{iH}^{k\prime} \cdot x_i^{k\prime} \cdot x_H^{k\prime}$$

$$\hat{T}_{Hj}^{k\prime}(I) = \hat{T}_{Hj}^{k\prime} \cdot x_H^{k\prime} \cdot x_j^{k\prime}$$

$$x_i^{k\prime}, x_j^{k\prime}, x_h^{k\prime}, \lambda', \hat{T}_{ij}^{k\prime}(I), \hat{T}_{iH}^{k\prime}(I), \hat{T}_{Hj}^{k\prime}(I) \geq 0$$

Finally, the values of $\hat{T}_{ij}^{k\prime *}(I), \hat{T}_{iH}^{k\prime *}(I)$ and $\hat{T}_{Hj}^{k\prime *}(I)$ will forecast the future traffic distribution where the vector ($\lambda'^*, x_i^{k\prime *}, x_j^{k\prime *}, x_h^{k\prime *}, \hat{T}_{ij}^{k\prime *}(I), \hat{T}_{iH}^{k\prime *}(I), \hat{T}_{Hj}^{k\prime *}(I)$) is the optimal solution to the quadratic model (*F*).

Considering the network with Q nodes, R arcs and k commodities; the maximum size of model (*P*) will be $(Q+R)\cdot k$ constraints and $Q \cdot R$ variables (which are manageable for large problems); and, the size of the model (*F*) will be $Q' \cdot R'$ constraints and variables. The optimal solution of model (*F*) can be obtained by Lingo software (Lindo Systems, 2004). We illustrate the proposed model with a numerical example presented in Section 4.

4 Numerical example

Consider the following packet-switching network that includes two well-known frame relays and switched multi-megabit data service (SMDS) protocols with different packet sizes. The real-traffic distribution of this network is shown in Figure 2. Considering a scenario, where the goal is to develop the packet-switched network presented in Figure 3, we forecast the future traffic distribution of this newly developed network.

Figure 2 The current packet-switching network (see online version for colours)

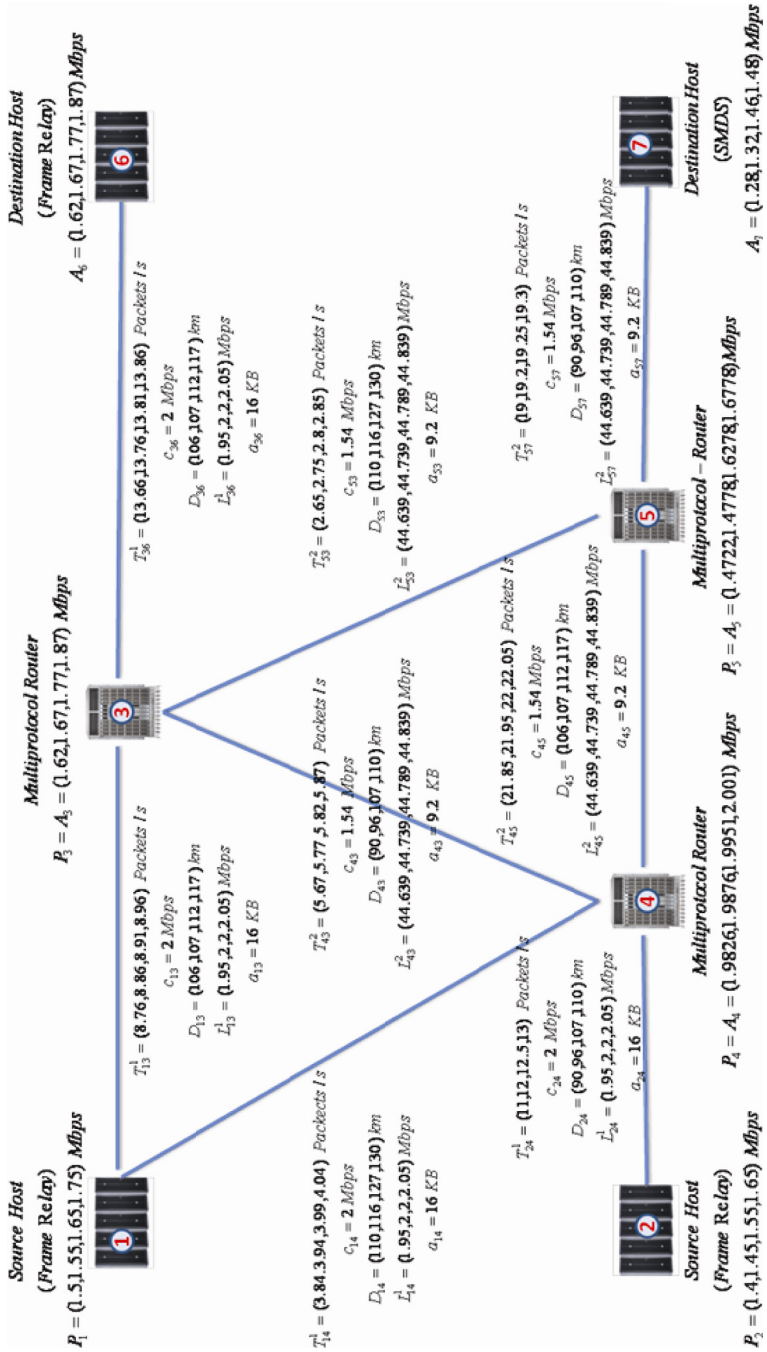


Figure 3 The developed packet-switching network (see online version for colours)

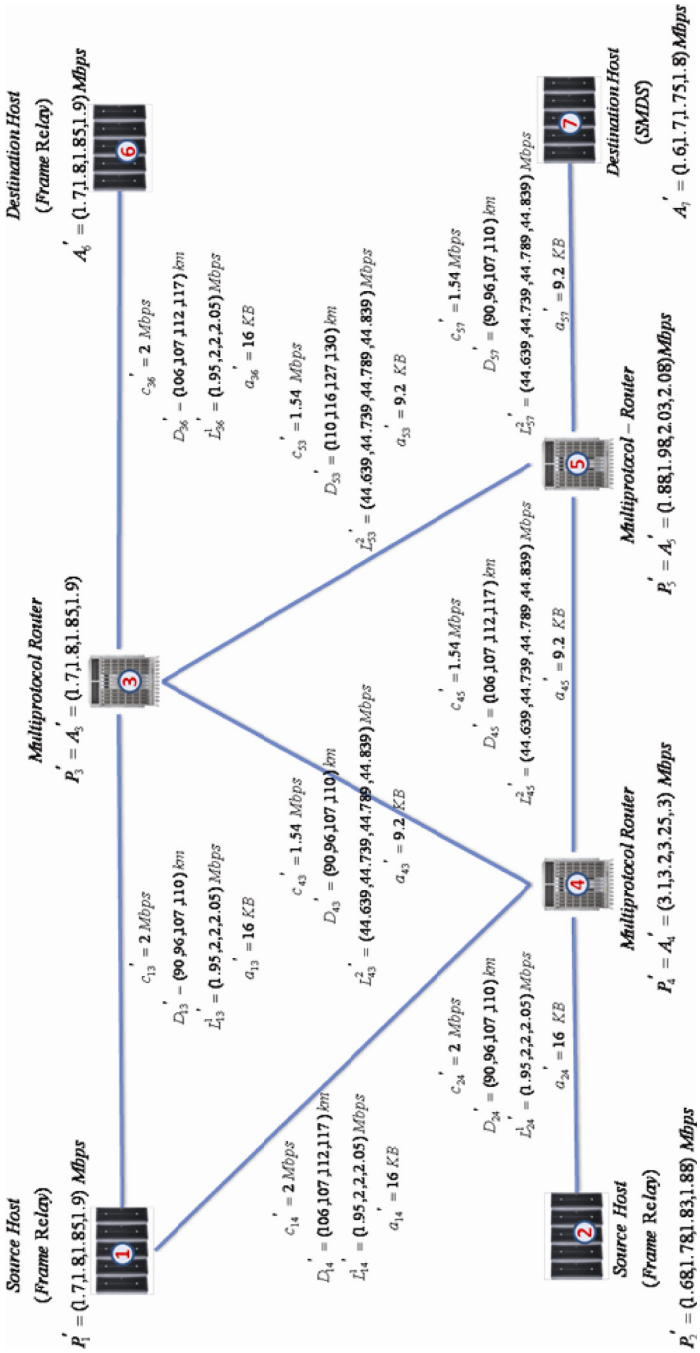


Table 1 The input vectors with the fuzzy elements in the current packet-switched network

Packet size	*Input vectors*
16 kB (k = 1)	$\underline{\tilde{X}}_{13}^{1} = [(1.5,1.55,1.65,1.75), (1.62,1.67,1.77,1.87), (106,107,112,117), (1.95,2,2,2.05)]$
	$\underline{\tilde{X}}_{14}^{1} = [(1.5,1.55,1.65,1.75), (1.9826,1.9876,1.9951,2.0026), (110,116,127,130), (1.95,2,2,2.05)]$
	$\underline{\tilde{X}}_{24}^{1} = [(1.4,1.45,1.55,1.65), (1.9826,1.9876,1.9951,2.0026), (90,96,107,110), (1.95,2,2,2.05)]$
	$\underline{\tilde{X}}_{36}^{1} = [(1.62,1.67,1.77,1.87), (1.62,1.67,1.77,1.87), (106,107,112,117), (1.95,2,2,2.05)]$
9.2 kB (k = 2)	$\underline{\tilde{X}}_{43}^{2} = [(1.9826,1.9876,1.9951,2.0026), (1.62,1.67,1.77,1.87), (90,96,107,110), (44.630,44.739,44.789,44.839)]$
	$\underline{\tilde{X}}_{45}^{2} = [(1.9826,1.9876,1.9951,2.0026), (1.4722,1.4778,1.6278,1.6778), (106,107,112,117), (44.639,44.739,44.789,44.839)]$
	$\underline{\tilde{X}}_{53}^{2} = [(1.4722,1.4778,1.6278,1.6778), (1.62,1.67,1.77,1.87), (110,116,127,130), (44.639,44.739,44.789,44.839)]$
	$\underline{\tilde{X}}_{57}^{2} = [(1.4722,1.4778,1.6278,1.6778), (1.28,1.32,1.46,1.48), (90,96,107,110), (44.639,44.739,44.789,44.839)]$

Step 1.1 The input vectors with the fuzzy elements in the current packet-switched network are given in Table 1. These fuzzy elements are then defuzzifed to the crisp elements, as shown in Table 2. Next, we find the input vectors with the scale-less crisp elements given in Table 3.

Table 2 The input vectors with the defuzzified elements in the current packet-switched network

16 kB (k = 1) packet size	*9.2 kB (k = 2) packet size*
$\underline{X}_{13}^{1} = [1.6,1.72,110,2]$	$\underline{X}_{43}^{2} = [1.992587,1.72,100,44.739]$
$\underline{X}_{14}^{1} = [1.6,1.992587,120,2]$	$\underline{X}_{45}^{2} = [1.992587,1.5778,110,44.739]$
$\underline{X}_{24}^{1} = [1.5,1.992587,100,2]$	$\underline{X}_{53}^{2} = [1.5778,1.72,120,44.739]$
$\underline{X}_{36}^{1} = [1.6,1.72,110,2]$	$\underline{X}_{57}^{2} = [1.5778,1.38,100,44.739]$

Table 3 The input vectors with the scale-less crisp elements in the current packet-switched network

16 kB (k = 1) packet size	*9.2 kB (k = 2) packet size*
$\underline{X}_{13}^{1}(s) = [0.8030, 0.8632, 0.9091, 0.0447]$	$\underline{X}_{43}^{2}(s) = [1, 0.8030, 1, 1]$
$\underline{X}_{14}^{1}(s) = [0.8030, 1, 0.8333, 0.0447]$	$\underline{X}_{45}^{2}(s) = [1, 0.7918, 0.9091, 1]$
$\underline{X}_{24}^{1}(s) = [0.7528, 1, 1, 0.0447]$	$\underline{X}_{53}^{2}(s) = [0.7918, 0.8030, 0.8333, 1]$
$\underline{X}_{36}^{1}(s) = [0.8030, 0.8632, 0.9091, 0.0447]$	$\underline{X}_{57}^{2}(s) = [0.7918, 0.6926, 1, 1]$

Step 1.2 According to the following matrices G^1, G^2 and the following vectors $\underline{T}^1(\text{SA})$, $\underline{T}^2(\text{SA})$, the pseudo-inverse of the matrices G^1, G^2 and the weight vectors are calculated using MATLAB software as follows:

$$G^1 = \begin{bmatrix} 1 & 0.9772 & 0.9723 & 0.9997 \\ 0.9772 & 1 & 0.9701 & 0.9734 \\ 0.9723 & 0.9701 & 1 & 0.9709 \\ 0.9997 & 0.9734 & 0.9709 & 1 \end{bmatrix}$$

$$G^2 = \begin{bmatrix} 1 & 0.9916 & 0.9314 & 0.9460 \\ 0.9916 & 1 & 0.9521 & 0.9404 \\ 0.9314 & 0.9521 & 1 & 0.9608 \\ 0.9460 & 0.9404 & 0.9608 & 1 \end{bmatrix}$$

$$\underline{T}^1(\text{SA}) = \begin{bmatrix} 8.86 & 3.94 & 12 & 13.76 \end{bmatrix}^{\text{T}}$$

$$\underline{T}^2(\text{SA}) = \begin{bmatrix} 5.77 & 21.95 & 2.75 & 19.2 \end{bmatrix}^{\text{T}}$$

$$(G^1)^{-1} = \begin{bmatrix} 3,578.7 & -288.7 & 205 & -3,316.4 \\ -288.7 & 47.3 & -12.1 & 254.3 \\ 20.5 & -12.1 & 22.1 & -30.1 \\ -3,316.4 & 254.3 & -30.1 & 3,098.1 \end{bmatrix}$$

$$(G^2)^{-1} = \begin{bmatrix} 127.7756 & -132.0043 & 46.0273 & -40.9619 \\ -132.0043 & 148.1426 & -54.9868 & 38.3941 \\ 46.0273 & -54.9868 & 34.2884 & -24.7765 \\ -40.9619 & 38.3941 & -24.7765 & 27.4494 \end{bmatrix}$$

$$\underline{W}^1(\text{SA}) = \begin{bmatrix} -126.04 & 9.82 & -0.15 & 138.87 \end{bmatrix}^{\text{T}}$$

$$\underline{W}^2(\text{SA}) = \begin{bmatrix} -282.01 & 307.6 & -132.28 & 106.53 \end{bmatrix}^{\text{T}}$$

Next, the computed output vectors as the current learning traffic distributions are calculated as follows:

$$\underline{\hat{T}}^1 = \begin{bmatrix} 8.86 & 3.94 & 12 & 13.76 \end{bmatrix}^{\mathrm{T}}$$

$$\underline{\hat{T}}^2 = \begin{bmatrix} 5.77 & 21.95 & 2.75 & 19.2 \end{bmatrix}^{\mathrm{T}}$$

With regard to Table 4, the vectors $\underline{\hat{T}}^1$ and $\underline{\hat{T}}^2$ are equal to the vectors $\underline{T}^1$ and $\underline{T}^2$ and these vectors are not calibrated.

Step 2.1 The input vectors with the fuzzy elements in the developed packet-switched network are given in Table 5. These fuzzy elements are then defuzzifed to the crisp elements, as shown in Table 6. Next, we find the input vectors with the scale-less crisp elements given in Table 7. These input vectors are used to forecast the future traffic distribution in Step 2.2.

Table 4 The real and computed traffic distributions in the current packet-switched network

The current real-traffic distribution		*The current computed traffic distribution*	
9.2 kB (k = 2) packet size	16 kB (k = 1) packet size	9.2 kB (k = 2) packet size	16 kB (k = 1) packet size
$T_{43}^2 = 5.77$	$T_{13}^1 = 8.85$	$\hat{T}_{43}^{2*}(I) = 5.77$	$\hat{T}_{13}^{1*}(I) = 8.85$
$T_{45}^2 = 21.95$	$T_{14}^1 = 3.94$	$\hat{T}_{45}^{2*}(I) = 21.95$	$\hat{T}_{14}^{1*}(I) = 3.94$
$T_{53}^2 = 2.75$	$T_{24}^1 = 12.00$	$\hat{T}_{53}^{2*}(I) = 2.75$	$\hat{T}_{24}^{1*}(I) = 12.00$
$T_{57}^2 = 19.20$	$T_{36}^1 = 13.76$	$\hat{T}_{57}^{2*}(I) = 19.20$	$\hat{T}_{36}^{1*}(I) = 13.76$

Table 5 The input vectors with the fuzzy elements in the developed packet-switched network

Packet size	*Input vectors*
16 kB (k = 1)	$\underline{\tilde{X}}_{13}^{1\,\prime} = [(1.7,1.8,1.85,1.9), (1.7,1.8,1.85,1.9), (90,96,107,110), (1.95,2,2,2.05)]$
	$\underline{\tilde{X}}_{14}^{1\,\prime} = [(1.7,1.8,1.85,1.9), (3.1,3.2,3.25,3.3), (106,107,112,117), (1.95,2,2,2.05)]$
	$\underline{\tilde{X}}_{24}^{1\,\prime} = [(1.68,1.78,1.83,1.88), (3.1,3.2,3.25,3.3), (90,96,107,110), (1.95,2,2,2.05)]$
	$\underline{\tilde{X}}_{36}^{1\,\prime} = [(1.7,1.8,1.85,1.9), (1.7,1.8,1.85,1.9), (106,107,112,117), (1.95,2,2,2.05)]$
9.2 kB (k = 2)	$\underline{\tilde{X}}_{43}^{2\,\prime} = [(3.1,3.2,3.25,3.3), (1.7,1.81,1.85,1.9), (90,96,107,1100), (44.639,44.739,44.789,44.839)]$
	$\underline{\tilde{X}}_{45}^{2\,\prime} = [(3.1,3.2,3.25,3.3), (1.88,1.98,2.03,2.08), (106,107,112,117), (44.639,44.739,44.789,44.839)]$

Table 5 The input vectors with the fuzzy elements in the developed packet-switched network (continued)

Packet size	*Input vectors*
	$\underline{\tilde{X}}_{53}^{2\prime} = [(1.88, 1.98, 2.03, 2.08), (1.7, 1.8, 1.85, 1.9), (116, 127, 130)0,11, (44.639, 44.739, 44.789, 44.839)]$
	$\underline{\tilde{X}}_{57}^{2\prime} = [(1.88, 1.98, 2.03, 2.08), (1.6, 1.7, 1.75, 1.8), (90, 96, 107, 110), (44.639, 44.739, 44.789, 44.839)]$

Table 6 The input vectors with the defuzzified elements in the developed packet-switched network

16 kB (k = 1) packet size	*9.2 kB (k = 2) packet size*
$\underline{X}_{13}^{1\prime} = [1.8, 1.8, 100, 2]$	$\underline{X}_{43}^{2\prime} = [3.2, 1.8, 100, 44.739]$
$\underline{X}_{14}^{1\prime} = [1.8, 3.2, 110, 2]$	$\underline{X}_{45}^{2\prime} = [3.2, 1.98, 110, 44.739]$
$\underline{X}_{24}^{1\prime} = [1.78, 3.2, 100, 2]$	$\underline{X}_{53}^{2\prime} = [1.98, 1.8, 120, 44.739]$
$\underline{X}_{36}^{1\prime} = [1.8, 1.8, 110, 2]$	$\underline{X}_{57}^{2\prime} = [1.98, 1.7, 100, 44.739]$

Table 7 The input vectors with the sale less crisp elements in the developed packet-switched network

16 kB (k = 1) packet size	*9.2 kB (k = 2) packet size*
$\underline{X}_{13}^{1}(s)' = [0.5625, 0.5625, 1, 0.0447]$	$\underline{X}_{43}^{2}(s)' = [1, 0.5625, 1, 1]$
$\underline{X}_{14}^{1}(s)' = [0.5625, 1, 0.9091, 0.0447]$	$\underline{X}_{45}^{2}(s)' = [1, 0.6187, 0.9091, 1]$
$\underline{X}_{24}^{1}(s)' = [0.5562, 1, 1, 0.0447]$	$\underline{X}_{53}^{2}(s)' = [0.6187, 0.5625, 0.8333, 1]$
$\underline{X}_{36}^{1}(s)' = [0.5625, 0.5625, 0.9091, 0.0447]$	$\underline{X}_{57}^{2}(s)' = [0.6187, 0.5312, 1, 1]$

Step 2.2 According to the calculated weight vectors in Step 1.2, the following matrices $G^{1\prime}$, $G^{2\prime}$, are used to forecast the future traffic distributions:

$$G^{1\prime} = \begin{bmatrix} 0.8551 & 0.7581 & 0.7964 & 0.8597 \\ 0.9264 & 0.9385 & 0.9565 & 0.8551 \\ 0.9159 & 0.9151 & 0.9620 & 0.9159 \\ 0.8623 & 0.7750 & 0.7899 & 0.8623 \end{bmatrix}$$

$$G^{2\prime} = \begin{bmatrix} 0.9438 & 0.9409 & 0.8791 & 0.9416 \\ 0.9586 & 0.97 & 0.4360 & 0.9445 \\ 0.7937 & 0.8157 & 0.9159 & 0.9280 \\ 0.8031 & 0.8012 & 0.8766 & 0.9455 \end{bmatrix}$$

$$\underline{\hat{T}}^{1\prime} = \begin{bmatrix} 18.942 & 11.065 & 20.601 & 18.562 \end{bmatrix}^{\mathrm{T}}$$

$$\underline{\hat{T}}^{2\prime} = \begin{bmatrix} 7.2811 & 70.9807 & 4.7826 & 4.7344 \end{bmatrix}^{\mathrm{T}}$$

Step 2.3 Assuming the tolerances $r_j' = q_i' = 0.2$, $s_H' = 0.05$, the following model (F) is used for calibration of the future traffic distributions:

$\text{Min}\ \lambda'$

$$0.2(\lambda'-1) \le 1.8 - 0.1250\left(\hat{T}_{13}^{1}{}'(I) + \hat{T}_{14}^{1}{}'(I)\right) \le 0.2(1-\lambda')$$

$$0.2(\lambda'-1) \le 1.8 - 0.1250\hat{T}_{36}^{1}{}'(I) \le 0.2(1-\lambda')$$

$$0.2(\lambda'-1) \le 1.78 - 0.1250\hat{T}_{24}^{1}{}'(I) \le 0.2(1-\lambda')$$

$$0.2(\lambda'-1) \le 1.7 - 0.0719\hat{T}_{57}^{2}{}'(I) \le 0.2(1-\lambda')$$

$$0.05(\lambda'-1) \le 16\left(\hat{T}_{13}^{1}{}'(I) + \hat{T}_{43}^{2}{}'(I)\right) + 9.2\hat{T}_{53}^{2}{}'(I) - 16\hat{T}_{36}^{1}{}'(I) \le 0.05(1-\lambda')$$

$$0.05(\lambda'-1) \le 16\left(\hat{T}_{14}^{1}{}'(I) + \hat{T}_{24}^{1}{}'(I)\right) - 9.2\left(\hat{T}_{43}^{2}{}'(I) + \hat{T}_{45}^{2}{}'(I)\right) \le 0.05(1-\lambda')$$

$$0.05(\lambda'-1) \le 9.2\left[\hat{T}_{45}^{2}{}'(I) - \left(\hat{T}_{53}^{2}{}'(I) + \hat{T}_{57}^{2}{}'(I)\right)\right] \le 0.05(1-\lambda')$$

$$0.1250\hat{T}_{13}^{1}{}'(I) \le 1.2$$

$$0.1250\hat{T}_{14}^{1}{}'(I) \le 2$$

$$0.1250\hat{T}_{24}^{1}{}'(I) \le 2$$

$$0.1250\hat{T}_{36}^{1}{}'(I) \le 2$$

$$0.0719\hat{T}_{43}^{2}{}'(I) \le 1.54$$

$$0.0719\hat{T}_{45}^{2}{}'(I) \le 1.54$$

$$0.0719\hat{T}_{53}^{2}{}'(I) \le 1.54$$

$$0.0719\hat{T}_{57}^{2}{}'(I) \le 1.54$$

$$\hat{T}_{13}^{1\prime}(I) = 18.942x_1^{1\prime} \cdot x_3^{1\prime}$$

$$\hat{T}_{14}^{1\prime}(I) = 11.065x_1^{1\prime} \cdot x_4^{1\prime}$$

$$\hat{T}_{24}^{1\prime}(I) = 20.601x_2^{1\prime} \cdot x_4^{1\prime}$$

$$\hat{T}_{36}^{1\prime}(I) = 18.562x_3^{1\prime} \cdot x_6^{1\prime}$$

$$\hat{T}_{43}^{2\prime}(I) = 7.2811x_4^{2\prime} \cdot x_3^{2\prime}$$

$$\hat{T}_{45}^{2\prime}(I) = 70.9807x_4^{2\prime} \cdot x_5^{2\prime}$$

$$\hat{T}_{53}^{2\prime}(I) = 4.7826x_5^{2\prime} \cdot x_3^{2\prime}$$

$$\hat{T}_{57}^{2\prime}(I) = 4.7344x_5^{2\prime} \cdot x_7^{2\prime}$$

$$\lambda', \hat{T}_{13}^{1\prime}(I), \hat{T}_{14}^{1\prime}(I), \hat{T}_{24}^{1\prime}(I), \hat{T}_{36}^{1\prime}(I), \hat{T}_{43}^{2\prime}(I), \hat{T}_{45}^{2\prime}(I), \hat{T}_{53}^{2\prime}(I), \hat{T}_{57}^{2\prime}(I),$$
$$x_1^{1\prime}, x_2^{1\prime}, x_3^{1\prime}, x_4^{1\prime}, x_6^{1\prime}, x_3^{2\prime}, x_4^{2\prime}, x_5^{2\prime}, x_7^{2\prime} \geq 0$$

The optimal solution of model (*F*) solved by optimisation software Lingo 9.0 (Lindo Systems, 2004) is given in Table 8.

Table 8 The future traffic distribution for the developed packet-switching network

16 kB (k = 1) packet size	*9.2 kB (k = 2) packet size*
$\hat{T}_{13}^{1\prime *}(I) = 9.600$	$\hat{T}_{43}^{2\prime *}(I) = 6.397$
$\hat{T}_{14}^{1\prime *}(I) = 3.200$	$\hat{T}_{45}^{2\prime *}(I) = 21.145$
$\hat{T}_{24}^{1\prime *}(I) = 12.640$	$\hat{T}_{53}^{2\prime *}(I) = 0.000$
$\hat{T}_{36}^{1\prime *}(I) = 16.000$	$\hat{T}_{57}^{2\prime *}(I) = 21.151$

5 Conclusions

In this study, we proposed a novel traffic distribution forecasting model for packet-switching networks. The continual utilisation of the proposed model allows for network design and capacity planning. In the proposed method, the traffic distribution forecasting problem is formulated using a fuzzy multi-objective optimisation and the current traffic distribution is learnt according to the real-traffic distribution. Next, this multi-objective model is converted into a single objective model using the approach proposed by Zimmermann (2001). The proposed method has the following advantages:

1 the proposed method can be used to forecast the traffic distribution of packet-switching networks in both crisp and fuzzy environments

2 the proposed fuzzy multi-objective optimisation model incorporates all calibrations simultaneously to save the execution time

3 the special structure of the proposed model is similar to quadratic programming where its convexity is confirmed

4 the proposed method allows traffic distribution forecasting from single commodity to multi-commodity

5 in large-scale packet-switching networks, the modelling of real-traffic distribution can be learnt by the RBF neural network

6 dimension consideration verifies the efficiency of this method for large-scale packet-switching networks.

The current model could be extended for data distribution forecasting in packet-switching networks based on Markovian queuing systems. One strategy is to study transmission of packets independent of their original sequence and the other one is to maintain the packet order. Both alternatives could be analysed through exact or approximate Markov queuing systems.

Acknowledgements

The authors would like to thank the anonymous reviewers and the editor for their insightful comments and suggestions.

References

Adams, A., Bu, T., Caceres, R., Duffield, N., Friedman, T., Horowitz, J., Presti, F.L., Moon, S., Paxson, V. and Towsley, D. (2000) 'The use of end-to-end multicast measurements for characterizing internal network behavior', *IEEE Communications Magazine*, Vol. 38, No. 5, pp.152–158.

Bazaraa, M.S., Jarvis, J.J. and Sherali, H.D. (1990) *Linear Programming and Network Flows* (2nd ed.). New York, NY: John Wiley and Sons, Inc.

Cao, J., Davis, D., Wiel, S.V. and Yu, B. (2000) 'Time-varying network tomography', *Journal of the American Statistical Association*, Vol. 95, pp.1063–1075.

Coates, M., Hero, A., Nowak, R. and Yu, B. (2002) 'Internet tomography', *IEEE Signal Processing Magazine*, Vol. 19, No. 3, pp.47–65.

Conway, A.E. and Li, M. (2003) 'Estimating the traffic distribution matrix of a packet-switched network using aggregate traffic measurements', *Telecommunication Systems*, Vol. 24, No. 1, pp.9–28.

Feldmann, A., Greenberg, A., Lund, C., Reingold, N., Rexford, J. and True, F. (2001) 'Deriving traffic demands for operational IP networks: methodology and experience', *IEEE/ACM Transactions on Networking*, Vol. 9, No. 3, pp.265–279.

Haykin, S. (1998) *Neural Networks – A Comprehensive Foundation* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Hwang, C.L. (1989) *Fuzzy Mathematical Programming, Methods and Application*. Berlin: Spinger-Verlag.

Hwang, C.L., Masud, A.S.M., Paidy, S.R. and Yoon, K. (1989) Multiple Objective Decision Making, Methods and Applications: A State of the Art Survey. Berlin: Springer-Verlag.

Lam, D., Cox, D. and Widom, J. (1997) 'Teletraffic modeling for personal communications services', *Wireless and Broadband Networks*, Vol. 35, No. 2, pp.79–87.

Lindo Systems (2004) Lindo Systems. LINGO (Release Version: 9.0). Lindo Systems, Inc.

Liu, D., Huebner, F. and Levy, Y. (1999) 'A hierarchical multi-class traffic model for data networks', *Proceedings of ITC-16*, pp.1221–1229.

Medina, A., Taft, N., Salamatian, K., Bhattacharyya, S. and Diot, C. (2002) 'Traffic matrix estimation: existing techniques and new directions', *Proceedings of ACM SIGCOMM*, pp.161–174.

Poggio, T. and Girosi, F. (1990) 'Networks for approximation and learning', *Proceedings of the IEEE*, Vol. 78, No. 9, pp.1481–1497.

Roughan, M., Greenberg, A., Kalmanek, C., Rumsewicz, M., Yates, J. and Zhang, Y. (2002) 'Experience in measuring backbone traffic variability: models, metrics, measurements and meaning', *Proceedings of ACM SIGCOMM Internet Measurement Workshop*, pp.91–92.

Roughan, M., Thorup, M. and Zhang, Y. (2003) 'Traffic engineering with estimated traffic matrices', *Proceedings of SIGMETRICS*, pp.206–217.

Teixeira, R., Duffield, N., Rexford, J. and Roughan, M. (2005) 'Traffic matrix reloaded: impact of routing changes', *Proceedings of Passive and Active Measurement (PAM) Conference*, pp.251–264.

Teixeira, R., Shaikh, A., Griffin, T. and Rexford, J. (2004) 'Dynamics of hot-potato routing in IP networks', *Proceedings of ACM SIGMETRICS*, pp.307–319.

Tikhonov, A.N. (1973) 'On regularization of ill-posed problems', *Doklady Akademii. Nauk USSR*, Vol. 153, pp.49–52.

Winston, W.L. (1987) *Operations Research Applications and Algorithms*. Boston: Duxbury Press.

Zhang, Y., Roughan, M., Duffield, N. and Greenberg, A. (2003a) 'Fast accurate computation of large-scale IP traffic matrices from link loads', *Proceedings of ACM SIGMETRICS*, pp.206–217.

Zhang, Y., Roughan, M. and Lund, C. (2003b) 'An information-theoretic approach to traffic matrix estimation', *Proceeding of ACM SIGCOMM*, pp.301–312.

Zhang, Y., Roughan, M., Lund, C. and Donoho, D.L. (2005) 'Estimating point-to-point and point-to-multipoint traffic matrices: an information-theoretic approach', *IEEE/ACM Transactions on Networking*, Vol. 13, No. 5, pp.301–312.

Zhao, Q., Zihui, G., Wang, J. and Xu, J. (2006) 'Robust traffic matrix estimation with imperfect information: making use of multiple data sources', *Proceedings of ACM SIGMETRICS*, pp.133–144.

Zimmermann, H.J. (2001) *Fuzzy Set Theory and Its Applications* (4th ed.). Kluwer Academic Publishers.