Available online at http://ijim.srbiau.ac.ir

Int. J. Industrial Mathematics Vol. 2, No. 1 (2010) 1-7

The Application of Data Envelopment Analysis and Queuing Models to Large Scale Computer Networks

N. Malekmohammadi*, H. Zainuddin

Institute of Mathematical Research, Putra Malaysia University, Serdang, Malaysia.

Received 27 February 2010; accepted 22 June 2010.

Abstract

This paper considers a technique for evaluating the operational efficiency of large-scale computer networks via data envelopment analysis and queuing models. The technique consists of two stages. In the second stage, a target DEA model is used which yields the advantages of the proposed technique over the previous one. Numerical illustration is provided to show the improvement of our aspect.

Keywords : Data envelopment analysis; Queuing theory; Large-scale problems

1 Introduction

Queuing theory modeling and the implementation of operational research techniques are means of evaluating the overall network efficiency and target setting through the study of the relative efficiency levels of its main constituent parts (i.e., the processing nodes). A network node may be thought of as a semiautonomous processing unit that accepts inputs from its environment and, by utilizing its own resources, produces outputs according to rules implemented in the unit's structural system [1].

Giokas and Pentzaropoulos [1] established two-staged method for evaluating the relative operational efficiency of large-scale computer networks, in the first stage analytical results for the main performance indicators are obtained by a queuing model (M/M/I/K) of a typical network and the results are used, in the second stage, as a starting point for the application of data envelopment analysis (DEA) procedure for obtaining the relative efficiency and improving the efficiency level of inefficient nodes, via input-oriented CCR dual model [10]. This technique just considers input-oriented CCR dual model which is a

*Corresponding author. Email address: malekmm@inspem.upm.edu.my;n.malekmohammadi@gmail.com Tel:+60173020975

radial model concentrating on the decrease of inputs.

In this paper in the second stage we use a target DEA model that considers both input reduction and output expansion together. Our main aim is to decrease the total input consumption and increase the total output production which results in solving one rather than n mathematical programming problems. Numerical illustration is provided to show some advantages of using the DEA target model over output-oriented CCR dual model.

This paper proceeds as follows, in section 2 queuing network modeling is introduced. In section 3, data envelopment analysis and our proposed method are presented. Numerical illustration and conclusion are provided in section 4 and 5 respectively.

2 Queuing network modeling

Suppose that in the general network topology we have N nodes, where $N > 1$ and the links between the nodes are integer. To connect a specified number of nodes for such a network, we define a virtual path to be a series of links for example, $\pi_{n\{1,\ldots,i,j,k,\ldots N\}}$ represents an end-to-end virtual path which extends from node 1 (source) to node N(destination) also including some intermediate nodes like i, j, k etc.

2.1 Network operational characteristics

Suppose that λ_i is the total flow into node i from all adjacent nodes; this is the sum of all flow rates of the links going through node i and denoted by $[\gamma_{jk}]$ of the network traffic matrix. Therefore we have

$$\lambda_i = \sum_{1<j,k}^{N} [\gamma_{jk}] \quad \text{where} \quad j,k : L_{jk\epsilon\pi_{n(i)}} \tag{2.1}$$

that is, the link (channel) L_{jk} carries traffic converging to node i. Each node in the network will be represented by the $M/M/I/K$ queuing model, i.e., a system with exponential inter arrival and service times and finite waiting room. Previous researches were done by [1, 6, 7]. Now suppose that K is taken as finite, considering the situation of nodes with memory constraints and containing a limited number of buffers to accommodate the incoming messages. We assume that the service policy at node i is to be in order of arrival, i.e., FCFS, and also $1/\mu_i$ represents the mean service time (in seconds) at the same node [1]. A general measure of traffic intensity at node i is given by $r_i = \lambda_i/\mu_i$. In this proceeding analysis, we can also get to these formulas which is an adjusted traffic intensity expression:

$$1/\mu_i\,(h) = (1/\mu_i) + (hR_i) \tag{2.2}$$

$$r_i\,(h) = \lambda_i\,(w)\,/\mu_i\,(h) = \lambda_i l\,(p)\,\{(1/\mu_i) + (hR_i)\} \tag{2.3}$$

also, $\lambda_i\,(w) = \lambda_i l\,(p)$ where $l\,(p)$ is the mean packet length(in Kbits), which denotes the mean work load (in Kbits/sec) embeding i; h is the average network overhead; R_i is a random number in the interval $[0, 1]$ which is supposed to distribute h among the network nodes in a non-uniform way. $1/\mu_i$ and h may be calculated from known network characteristics. The probability that all buffers of node are full at a given time is estimated as follow:

$$p_k\,(i) = \{[1 - r_i\,(h)]\, r_i\,(h)^K\}/[1 - r_i\,(h)^{K+1}] \tag{2.4}$$

where $\lambda_i \neq \mu_i$. As we have space limitation (number of buffers available) the actual traffic entering node $i, \lambda_i\,(e)$ and the actual server (node i) utilization, ρ_i will be:

$$\lambda_i(e) = \lambda_i(w)\{1 - p_k(i)\} = \lambda_i l(p)\{1 - p_k(i)\} \tag{2.5}$$

$$\rho_i = \lambda_i(e)/\mu_i(h) = \lambda_i(w)\{1 - p_k(i)\}/\mu_i(h) = \{1 - p_k(i)\}r_i(h) \tag{2.6}$$

respectively. ρ_i represents that the fraction of time node i is busy. One important measure in a network with limited storage capacity is the number of packets that are turned away from node i , because all buffers are full at the time of arrival of the next packet. This may be expressed in percentage form by taking into account the probability $p_k(i)$ as follows:

$$F_i = 100p_k(i) \tag{2.7}$$

and high values of F_i indicate a relative storage inefficiency of node i. Because of the possibly large values of arrival rates for nodes, we may get to large queue sizes, which are estimated as:

$$Q_i = \{r_i(h)/[1 - r_i(h)]\} - \{[(K+1)r_i(h)^{K+1}]/(1 - r_i(h)^{K+1})\} \tag{2.8}$$

also, the mean waiting time W_i and a useful measure to characterize processing efficiency in performance comparisons of various computer systems in a network [8], the stretch factor S_i of node i, can be obtained by:

$$W_i = Q_i/\lambda_i(e) \tag{2.9}$$

$$S_i = W_i/\{(1/\mu_i + (hR_i)\}. \tag{2.10}$$

We have to mention that S_i shows how many times (over some required minimum time for service) a stream of packets may be delayed as a result of a large queue size. A relative processing efficiency of node i decreases as S_i increases. More information about this section are considered in [1].
Giokas and Pentzaropoulos [1] characterized a network node i as follows:
(I_1) λ_i, total flow from all adjacent nodes (in packets/sec);
(I_2) $1/\mu_i$, mean service time (in seconds);
(I_3) Ki, memory capacity (in number of buffers available);
(O_1) ρ_i, mean node utilization (fraction between 0 and 1);
(O_2) F_i, packets turned away (%);
(O_3) S_i, mean stretch factor (real number above 1).

3 Data envelopment analysis and target setting

Data envelopment analysis (DEA) is a mathematical programming approach for evaluating the relative efficiency of decision-making units (DMUs). In 1978, the first DEA model was proposed by Charnes et al. which is a non-linear fractional mathematical programming model, known as the CCR model. The objective function in this model is considered to reach the best set of weights for the single ratio of the weighted outputs to the weighted inputs for a particular DMU denoted by DMUo. In this model, along with evaluating the efficiency, all the DMUs will be projected to the efficient frontier separately.

In primary DEA models, the major goal was evaluating the efficiency of the DMUs. Since it is important to know whether the DMU projected onto the efficient frontier is acceptable and ideal for the decision makers (DMs) or not, many researches have been started under the names of target setting and resource allocation. Since then, DEA has been used for future programming of organizations and the responses of different policies. Golany [11], Thanassoulis and Dyson [3] and Athanassopoulos [4,5] have introduced models for assessing targets and allocating resources based on data envelopment analysis.

3.1 The proposed model

Now we propose the DEA target model, which is in the frame work of previously introduced model in [2], as follows:
Let $j, r = 1, ..., n$, be the indices for decision-making units (DMUs) while each unit uses input quantities $X \in R^m_+$ to deliver output quantities $Y \in R^s_+$. We can also consider the indices sets of inputs, $I = \{1, ..., m\}$ and outputs, $O = \{1, ...s\}$. The vector $(\lambda_{1r}, \lambda_{2r}, ..., \lambda_{nr})$, such that $\sum_{j=1}^{n} \lambda_{jr} = 1,\ r = 1, ..., n$ is imposed for convex combination between inputs or outputs for n DMUs. θ_i ,Z_k are the contraction rate of input i and expansion rate of output k, respectively (characterization of *non-radial* models).

$$\max_{\lambda_j, Z_k, \theta_i} \frac{\sum_{k\in O} 1/s \quad Z_k}{\sum_{i\in I} 1/m \quad \theta_i}$$

$$\begin{aligned} s.t. \quad \sum_{r=1}^{n}\sum_{j=1}^{n} \lambda_{jr} x_{ij} &= \theta_i \sum_{j=1}^{n} x_{ij}, \qquad i \in I \\ \sum_{r=1}^{n}\sum_{j=1}^{n} \lambda_{jr} y_{kj} &= Z_k \sum_{r=1}^{n} y_{kr}, \qquad k \in O \\ \sum_{j=1}^{n} \lambda_{jr} &= 1, \qquad r = 1, ..., n \end{aligned}$$

$$\lambda_{jr} \geq 0, j, r = 1, ..., n, \theta_i \quad \text{free}, i \in I, Z_k \quad \text{free}, k \in O, \tag{3.11}$$

where x_{ij} is the quantity of input i of unit j ; $y_{kr}$the quantity of output k of unit$r$;$1/m$ and $1/s$ have been used instead of the user specified constants reflecting the decision makers' preferences over the improvement of input/output components in [2]. After computing model (3.11), we can obtain target input and output from the below formulas:

$$x^*_{ij} = \theta^*_i x_{ij}, \qquad i = 1, ..., m; j = 1, ..., n \tag{3.12}$$

$$y^*_{kr} = Z^*_k y_{kr}, \qquad k = 1, ..., s; r = 1, ..., n \tag{3.13}$$

Giokas and Pentzaropoulos [1] collected $(I_1), (I_2)$ and (I_3) as inputs and $(O_1), (O_2)$ and (O_3) as outputs (Section 2) to be considered for the input-oriented CCR-DEA model [9]. In our approach we used model (3.11) that considers both inputs reduction and outputs expansion simultaneously. Our main aim is to decrease total input consumption and increase total output production that results in solving one mathematical programming problem instead of n ones. In the numerical illustration, some of the advantages of using model (3.11) over output-oriented CCR dual model has been shown.

4 Numerical illustration

In this section we will apply model (3.11) to evaluate and set a target for computer networks. The data set is collected from Giokas and Pentzaropoulos [1]. The limited-scale sub network is considered which is shown in Table 1. For each node, the individual value of $\mu_i(h)$ from equation (2.2) is obtained because of the presence of h which includes network overheads. The inverse expressions of these values, i.e., $\mu_i(h)$ give the corresponding mean service rate values [1]. By considering these assumptions we can compute typical values of input flows (λ_i) and buffer sizes (K_i) for the 11 nodes and then the estimation for ρ_i, F_i and S_i from equations (2.6), (2.7), and (2.10), respectively. In Table 2 we computed the target input and target output from model (3.11) and formulations (3.12) and (3.13) (is solved by Lingo 11.0). In the last three rows total target of input consumption and total target of output production are considered.

Table 1
Input/output performance parameters used for DEA.

	Input	Input	Input	Output	Output	Output
Node	λ_i	K_i	$\mu_i(h)$	ρ_i	F_i	S_i
N1	7	10	24.390	0.590	0.213	2.393
N2	6	6	22.727	0.534	1.163	2.023
N3	10	7	23.256	0.815	7.790	3.511
N4	9	9	22.727	0.783	3.237	3.671
N5	9	7	22.727	0.772	5.559	3.219
N6	8	7	31.250	0.524	0.539	2.035
N7	11	11	30.303	0.734	0.987	3.440
N8	10	8	31.250	0.655	1.296	2.659
N9	8	8	28.571	0.566	0.475	2.230
N10	14	15	30.303	0.903	3.716	6.830
N11	10	10	29.412	0.686	0.783	2.984

Source:Giokas and Pentzaropoulos [1].

Table 2
Target input/output from model (3.11)

	Target Input	Target Input	Target Input	Target Output	Target Output	Target Output
Node	$\theta_1^*\lambda_i$	$\theta_2^*K_i$	$\theta_3^*\mu_i(h)$	$Z_1^*\rho_i$	$Z_2^*F_i$	$Z_3^*S_i$
N1	7	7.9	20.975	0.708	0.701	2.632
N2	6	4.74	19.545	0.641	3.838	2.225
N3	10	5.53	20.000	0.978	25.707	3.862
N4	9	7.11	19.545	0.940	10.682	4.038
N5	9	5.53	19.545	0.926	18.345	3.541
N6	8	5.53	26.875	0.629	1.779	2.239
N7	11	8.69	26.061	0.881	3.257	3.784
N8	10	6.32	26.875	0.786	4.277	2.925
N9	8	6.32	24.571	0.679	1.568	2.453
N10	14	11.85	26.061	1.084	12.263	7.513
N11	10	7.9	25.294	0.823	2.581	3.282
Total Targets						
$\sum_{j=1}^{n}\theta_i^* x_{ij}$	102	77.42	255.347			
$\sum_{r=1}^{n} Z_k^* y_{kr}$				9.075	84.998	38.494
Existing	102	98	296.916	7.562	25.758	34.995

As it is clear we considered both total input increase and total output decrease together that results in solving 1 mathematical programming and also model (3.11) is a *non-radial* model which reduces or increases all inputs and outputs respectively in the different measure. Giokas and Pentzaropoulos [1] considered input-oriented CCR dual model which is a *radial* model that concentrates on inputs decrease and also it is needed to solve 11 linear programming problems.

5 Conclusion

This paper deals with the development of the existing technique for evaluating the operational efficiency of large-scale computers communications networks. The proposed technique consists of two-stages, in the second stage, a *non-radial* DEA target model is used that considers both decrease of total input consumption and increase of total output production that leads to computing only one mathematical programming problem instead of n ones. This approach yields to the improvement of the proposed technique over the previous one.

Acknowledgments

This work is supported by Institute of Mathematical Research, Putra Malaysia Universitiy under Science Fund Grant Scheme No: 06-01-04-SF0256.

References

[1] D.I. Giokas and G.C. Pentzaropoulos, Evaluating the relative operational efficiency of large-scale computer networks: an approach via data envelopment analysis, Applied Mathematical Modeling 19 (1995).

[2] N. Malekmohammadi, F. Hosseinzadeh Lotfi, Azmi B Jaafar, Data envelopment scenario analysis with imprecise Data, Central European Journal of Operations Research, Published online: 13 December 2009.

[3] E. Thanassoulis, R. Dyson, Estimating preferred targets input-output levels using data envelopment analysis, European Journal of Operational Research 56 (1992) 80-97.

[4] A.D. Athanassopoulos, Goal programming and Data Envelopment Analysis (GoDEA) models for multi level multi unitorganizations: An application to Greek local authorities, European Journal of Operational Research 87 (1995) 535-555.

[5] A.D. Athanassopoulos, Assessing the comparative spatial disadvantage of European regions using non-radial data envelopment analysis models, European Journal of Operational Research 94 (1996) 439-452.

[6] J.F. Hayes, Modeling and Analysis of Computer Communications Networks, Plenum Press, New York, 1984.

[7] J.N. Daigle, Queuing Theory for Telecommunications. Addison-Wesley, Reading, MA, 1992.

[8] G.C. Barney, G.C. Pentzaropoulos and W. Swindells, Performance comparisons within a computer network of Prime systems. Proc. IEEE Melecon '83 (No A506), Athens, Greece 2426 May 1983.

[9] A. Charnes, W.W. Cooper , E. Rhodes, Measuring the efficiency of decision making units, European Journal of Operational Research 2 (6) (1978) 429-444.

[10] W.W. Cooper, L.M. Seiford, K. Tone, Data envelopment analysis. A comprehensive text with models, applications, references and DEA-solver software. Boston, NA: Kluwer Academic Publishers, 2000.

[11] B. Golany, An interactive MOLP procedure for the extension of DEA to effectiveness analysis, Journal of the Operational Research Society 39 (1988) 725-73.