

## A MIXED-METRIC ROUTING WITH OPTIMAL FLOW DISTRIBUTION FOR MPLS NETWORKS

---

**G. A. F. M. KHALAF**

*Associate Professor at Communications Department, Faculty of Engineering, Helwan, Helwan University, Egypt.*

*(Received September 20, 2005, Accepted October 30, 2005)*

*This paper proposes a new routing system for IP network traffic engineering. In this system, the link cost metrics are chosen to reflect the congestion level in terms of the number of flows traversing a given link as well as its residual bandwidth. This cost is used in the design of a dynamic routing for MPLS networks. At the same time, an optimal distribution of network flows is found, and used to impose a balanced traffic loading overall network links. This in turns, have led to a bounded network delay performance. Results have shown robustness of the proposed routing technique over networks with arbitrary topologies and heavy traffic demands.*

**KEYWORDS:** *MPLS Networks, Mixed-Metrics, Link Cost, Load Balancing.*

### I. INTRODUCTION

Traditional Internet Protocol (IP) networks typically support only best effort service. As a consequence, advanced network scenarios require the introduction of technologies able to provide users with quality of service (QoS), and network operator with more dynamic, and flexible utilization of the network resources, i.e., the capability to perform traffic engineering (TE) [1]. Traffic engineering, means the process of mapping traffic demands into the network topology, and realizing such mapping via routing protocols so that a predefined performance objective is achieved. From the network point of view, the objective is a balanced loaded network with minimum congestion , i.e., there should not be unnecessarily over utilized links while others are underutilized.

Current IGP (Interior Gateway Protocol) routing protocols provide little capability for solving the above mentioned TE-problems. Open Shortest Path First (OSPF) is one such IGP routing protocols. In the most frequent implementation of OSPF, a unit cost is allocated to network links, leading to pick the least number of hops as the shortest path. Problems arise when:

1. multiple traffic streams converge on specific links or nodes.
2. a traffic stream is routed through a link or node that lack enough bandwidth to service it [2].

Extensions such as those proposed in [3], and [4], to support QoS routing based on OSPF have been proposed to take into consideration both aspects (i.e., 1, and 2 above) of the problem. OSPF [4], is proposed to support QoS by flooding the network with the available and used link resources. The problem is that OSPF generates a substantial amount of communication overheads especially in the context of bursty Internet traffic wherein available resources changes rapidly. QoS-OSPF [3], tries to minimize this overhead by using a trigger mechanism that fires every period,  $T$ , or when a link resource changes by a given percentage. A potential problem occurs when QoS-OSPF measures underutilized but allocated resources. These resources could be reallocated, causing packet drops once the client starts using his/here bandwidth allocation.

These TE-problems can be effectively addressed by the evolving Multi Protocol Label Switching (MPLS) technology [5-8]. In an MPLS cloud, connection request is accepted by setting up a route (Label Switched Path (LSP)) that will service that connection with necessary resources [6], [10]. With MPLS, routes between source/destination routers are calculated at source routers (ingress router), which takes into account not only the network topology but also traffic oriented constraint (such as bandwidth, hop count) according to some optimization criteria [9], [14]. The main limitation of the MPLS approach is the choice of the best optimization criteria which is not completely independent of the network load [11], and at the same time guarantees an even (balanced) loaded network.

This paper proposes a hybrid of TE, and MPLS routing that actively uses MPLS tunnels (LSPs) to avoid potential network congestion. Section 2 introduces a routing system based on a mixed-metric link cost function. An optimal distribution for the LSPs over the network links is presented in section 3. Results, and the performance evaluations of the proposed systems are given in section 4. Section 5, summarizes the findings in this paper, and provides a future research plan.

## II. A MIXED-METRIC ROUTING SYSTEM

This section addresses the above mentioned problems by proposing a routing technique that ensures even utilization of links in an MPLS cloud. Since path costs reflect the network congestion level, we use this cost to design a dynamic routing technique that avoids the over (under) utilized links problem inherited in the OSPF algorithm.

Consider a directed (for simplicity) graph  $G(N, L)$ , where  $N$  is the set of nodes (e.g., MPLS routers), and  $L$  is the set of links (arcs). Assuming,

$f_{ij}$ : traffic demands between ingress-egress router pairs  $(i, j)$ , where  $i=1, 2, \dots$

$N-1$ , and  $j=2, 3, \dots, N$ , in a directed graph,

$P_{ij}$ : set of feasible paths (LSPs) connecting router pairs  $(i, j)$ ,

$f_l$ : traffic flow traversing link  $l$ ,  $l=1, 2, \dots, L$ ,

$C_l$ : capacity of link  $l$ ,  $l=1, 2, \dots, L$ ,

$L_l$ : length (cost) of link  $l$ ,  $l=1, 2, \dots, L$ ,

$L_p$  : length of path  $p$ ,  $p \in P_{ij}$  for all  $i, j$ ,

$L_{ps}$  : length of shortest path,

$n_l$  : number of LSPs traversing link  $l$ ,  $l=1, 2, \dots, L$ .

The objective of the routing algorithm is to find the shortest path  $L_{ps}$ ,  $p_s \in P_{ij}$ , that satisfies the following constraint,

$$f_{ij} \leq \min_{l \in p_s} (C_l - f_l) \quad (1)$$

$$L_p = \sum_{l \in p} L_l \quad (2)$$

$$L_{ps} = \min_{p \in P_{ij}} L_p \quad (3)$$

$$n_{ls} = \min_{l \in p} n_l \quad (4)$$

Equation (1), insures that all LSPs (i.e.,  $f_{ij}$ ), are routed over links having maximum residual bandwidth ( $C_l - f_l$ ), equation (4), adds on a load balancing constraint by choosing shortest paths that are comprised of links carrying the minimum number of flows,  $n_l$ . Combining equations (1) to (4), gives the following routing system:

Find the shortest path  $L_{ps} = \min_{p \in P_{ij}} L_p$  under the following link metric,

$$L_l = \frac{n_l}{C_l - f_l}, \quad l=1, 2, \dots, L \quad (5)$$

Of important here to note that, our routing algorithm does not rely on minimum number of hop-paths as did in the OSPF. However, it does guarantee a fairly, even link utilization among all network links which is a significantly valuable feature. This should be true as it would, consequently, leads to a significant improvement in the overall network (delay) performance.

### III. OPTIMAL DISTRIBUTION OF LSP-FLOWS OVER NETWORK LINKS

In previous section, placement (routing) of LSPs on network links is reduced to setting up a number of MPLS-tunnels between some source-destination routers such that the following mixed-metric (equation (5)), is satisfied,

$$L_l = \frac{n_l}{C_l} \left( \frac{1}{1 - u_l} \right)$$

where  $u_l$  represents the link utilization,  $u_l = \frac{f_l}{C_l}$ . As can be seen, without bounding

the link flows,  $f_l$ , the link cost,  $L_l$ , goes to infinity i.e., when the link utilization approaches unity. This means that, an incrementally small link-traffic change may results in a significantly large change in its cost metric. This is known to lead to instability of routing system when used as a congestion indication especially in a heavily loaded networks. An approach to avoid this problem is to use a fraction of the link capacity  $C'_l = \alpha C_l$ ,  $\alpha < 1$ , and the remaining capacity  $(1 - \alpha)C_l$ , is left to absorb any unexpected traffic variations. Certainly, choosing  $\alpha$  would depend largely on the degree of robustness required to be achieved against large variations in the utilization of a given link.

In the following, we present an approach to solving these problems. First, an optimal distribution for the total LSP-flows is obtained. Second, the routing system presented in previous section is constrained to establish LSPs using the set of links which does not violate the optimum link-flows. The objective is to set a lower bound on the link flows, hence, guaranteeing a (optimal) lower bound on the network delay performance.

More specifically, the mean packet delay is chosen as our objective function, and we assume  $M^X \setminus M \setminus 1 \setminus \infty$  queues with batch Poisson arrivals. Although it is simple, the  $M^X \setminus M \setminus 1 \setminus \infty$  model can achieve an acceptable fit for the busiest periods of network traffics [11]. Assuming exponential distribution for packet lengths with mean  $\tau$ , and a geometric distribution for the number of packets in a batch with mean  $X$ . The mean service rate at a given router,  $i$ , is  $\mu_i = \frac{C_i}{\tau}$ , and the batch flow through a given

link,  $i$ , is  $\lambda_i$ . The expected delay,  $E(T_i)$ , for an  $M^X \setminus M \setminus 1 \setminus \infty$  queue is given by [12],

$$E(T_i) = \frac{\rho_i}{\lambda_i(1 - \rho_i)} \quad (6)$$

where  $\rho_i = \frac{X\lambda_i}{\mu_i}$  designates the traffic intensity. The mean network delay,  $T_n$  is the

combination of delays due to the queues,  $T_i$  ( among other delay sources which we assume negligible along this paper), and is given by,

$$T_n = \frac{1}{\gamma} \sum_{i=1}^L X \lambda_i E(T_i)$$

$$T_n = \frac{X}{\gamma} \sum_{i=1}^L \frac{X \lambda_i}{(\mu_i - X \lambda_i)} \quad (7)$$

where  $\gamma = \sum_{i=1}^{N-1} \sum_{j=2}^N f_{ij}$  is the total flow at the network input. The optimum distribution

of network flows,  $\lambda_i, i=1, 2, \dots, L$ , can be found by finding the set of flows that

minimizes  $T_n$ , subject to the constrain  $\lambda_T = \sum_{i=1}^L \lambda_i$ , where  $\lambda_T$  is the total link

flows. Certainly,  $\lambda_T$  is not yet known. However, if we set a given "average number,

k," of hops as a target in our optimization. Then, a good estimate of  $\lambda_T$  would be

$\lambda_T = k \gamma$ . Now, finding the optimum distribution of the network flows can be found using the Lagrangean technique, and the following Kuhn-Tucker conditions [13], for  $i=1, \dots, L$

$$\frac{\partial}{\partial \lambda_i} \left( \sum_{i=1}^L \frac{X \lambda_i}{(\mu_i - X \lambda_i)} - \alpha \left( \sum_{i=1}^L \lambda_i - \lambda_T \right) \right) = 0 \tag{8}$$

where  $\alpha$ , is the lagrange multiplier. Solving the following set of Kuhn-Tucker conditions for  $\lambda_i$ , gives,

$$\frac{X \lambda_i}{(\mu_i - X \lambda_i)^2} - \alpha = 0 \tag{9}$$

In order to check that  $\lambda_i$  obtained from equation (9) will result in a minimum of  $T_n$ , we differentiate (8) once more. This gives,

$$2X(\mu_i - X \lambda_i) \tag{10}$$

As we know, for a stable network  $\mu_i > X \lambda_i$  for all links, otherwise the delay will be infinite. This makes equation (10) positive, thus insuring a minimum network delay performance. Equation (9), gives us the optimum link flows but in terms of  $\alpha$ . So we should find  $\alpha$  first. Rewriting equation (9) gives,

$$\begin{aligned} \frac{X}{\gamma} \sum_{i=1}^L \frac{X \lambda_i}{(\mu_i - X \lambda_i)} &= X^2 \sum_{i=1}^L \frac{\lambda_i}{\gamma} \sqrt{\frac{\alpha}{X \mu_i}} \\ &= T_n \end{aligned}$$

from which  $\alpha$  is obtained. Note that,  $T_n$  here is the mean network delay performance to be achieved.

$$\sqrt{\alpha} = \frac{T_n \gamma + L X}{\sum_{i=1}^L \sqrt{X \mu_i}} \tag{11}$$

finally, the set of optimum link flows,  $\lambda_i^o$ ,  $i=1,2,\dots,L$ , are obtained from equation (9),

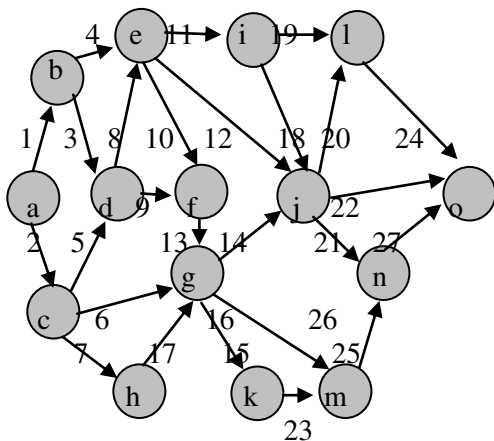
$$\lambda_i^o = \frac{\mu_i}{X} + \sqrt{\frac{\mu_i}{\alpha X}}. \quad (12)$$

As mentioned earlier, constraining the routing algorithm such that no link flow should exceed its pre-computed level,  $\lambda_i^o$ , would enforce a lower bound on overall network delay performance.

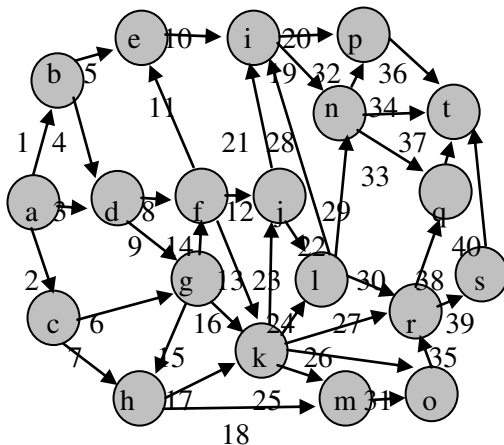
#### IV. PERFORMANCE EVALUATION

The mixed-metric routing technique, and the optimum flow distribution system are both implemented as a traffic engineering tool for setting up LSPs between ingress-egress routers in an MPLS cloud. In this section, we study the effectiveness of applying this, analytical, tool. Two network graphs with arbitrary topologies are chosen. The first one is comprised of 15-MPLS-router (node), interconnected using 27-link. The second consists of 20-MPLS-router, with 40-links interconnecting them. In all scenarios, we have assumed equal traffic demands between any ingress-egress router pairs as well as equal link capacities. This would help us gaining a better understanding of the behavior of our algorithms. Short (LSPs) paths between MPLS-routers are obtained using the well known Dijkstra algorithm. The network topologies chosen are shown in **figures 1**, and **2** below. As shown, each router is identified by an alphabetic letter, while links are identified using numeric "titles".

##### Phase I: Validation of the mixed-metric routing



**Figure 1:** 15-router MPLS cloud.



**Figure 2:** 20-router MPLS cloud.

**Figures 3**, and **4**, show the number of LSPs sharing the capacity of each link in the two networks. The light-shaded histogram is obtained after running the OSPF algorithm, and the heavily-shaded one results after running the mixed-metric routing technique (equation (5)). As can be seen, the OSPF algorithm tends to over load some

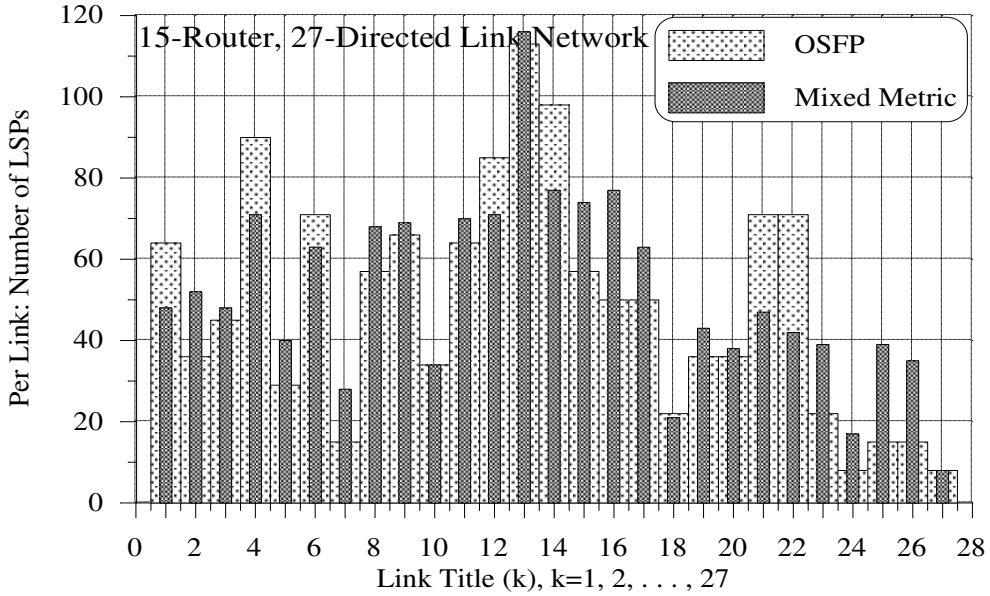


Fig.(3), Per Link LSP Count

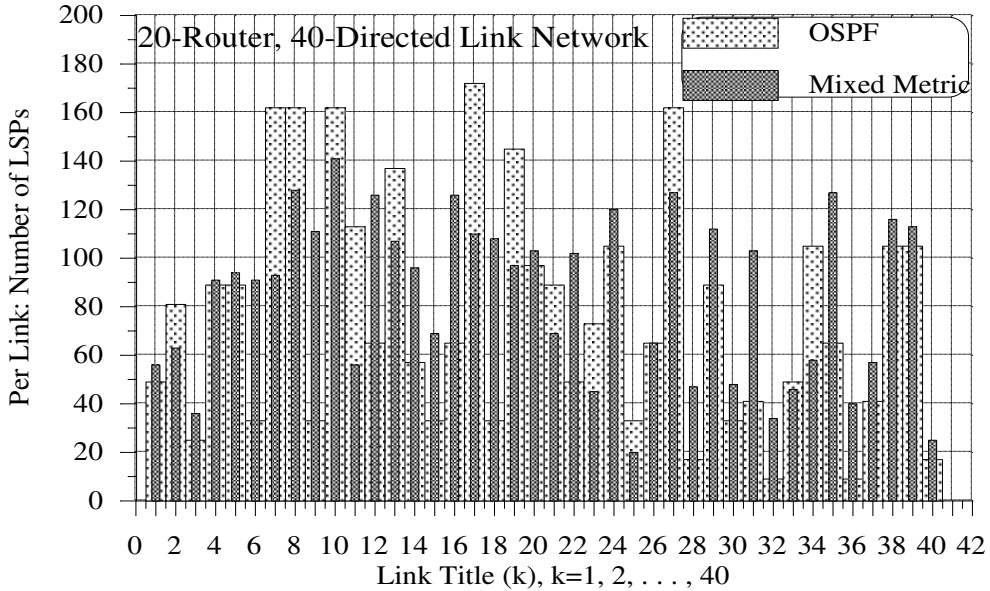


Fig.(4), Per Link LSP Count

links, while under load others. On the other hand, the mixed-metric algorithm is seen to be load balancing most of the network links. It does so by reducing the number of LSPs sharing a heavily loaded (congested) link (equation (4)). This observation can, also, be seen from **Figures 5**, and **6** but in terms of the per link traffic utilization. Now looking at **Figures (3-6)** collectively shows that even with our mixed-metric system, there still, however, some over utilized links,  $u_l > 0.7$ , (e.g., link #14, #16 in the first network, and #8, #10, #24 in the second) which would, in turn, results in a relatively

high network delay performance (see equation (6)). This should not leads us to a wrong conclusion. As we know the routing process is a fairly network-topology depended process, that is, it depends largely on route availability in the physical network. Moreover, this very observation have encouraged us towards finding the optimal flow bounds.

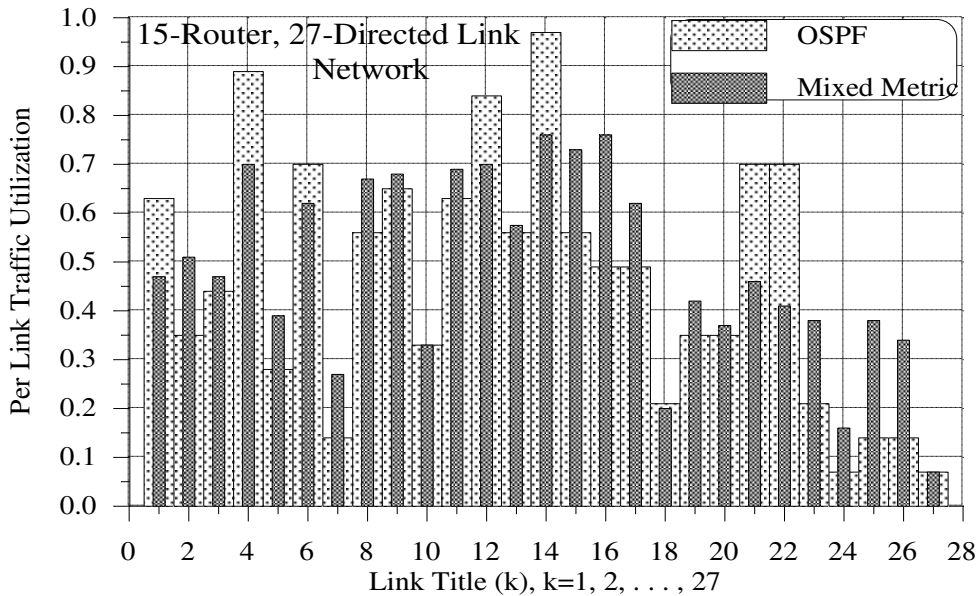


Fig.(5), Per Link Traffic Utilization

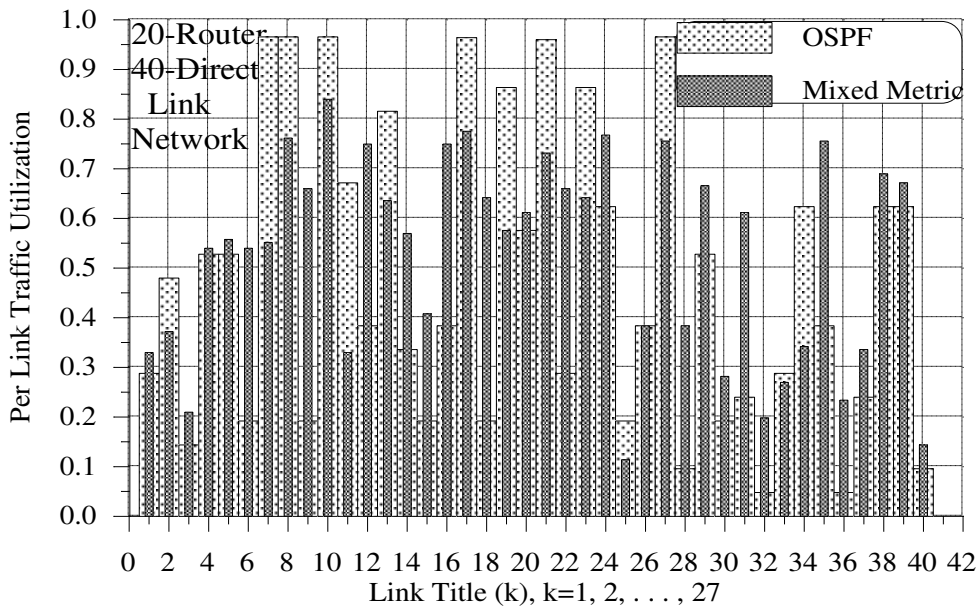


Fig.(6), Per Link Traffic Utilization



Figures 7, and 8, completes our discussion by showing the LSP-lengths in terms of the number of hops (links) traversed. As can be seen, under the mixed-metric routing, some LSPs had to traverse longer paths which are found to be less congested. This might indicate that LSPs under the mixed-metric system would result in a longer network delay. This, in fact, is not true. Simply because traversing a relatively longer paths comprised of less congested links would not necessarily results in a longer path delays. In order to validate this argument, Table I, gives a sample delay performance results for the per link, and the end-to-end (ETE) delays (measured m. Sec.). Of importance here to note that the numbers given in Table I, presents a set of figures we have used for comparison sake apart from the implications of their actual values.

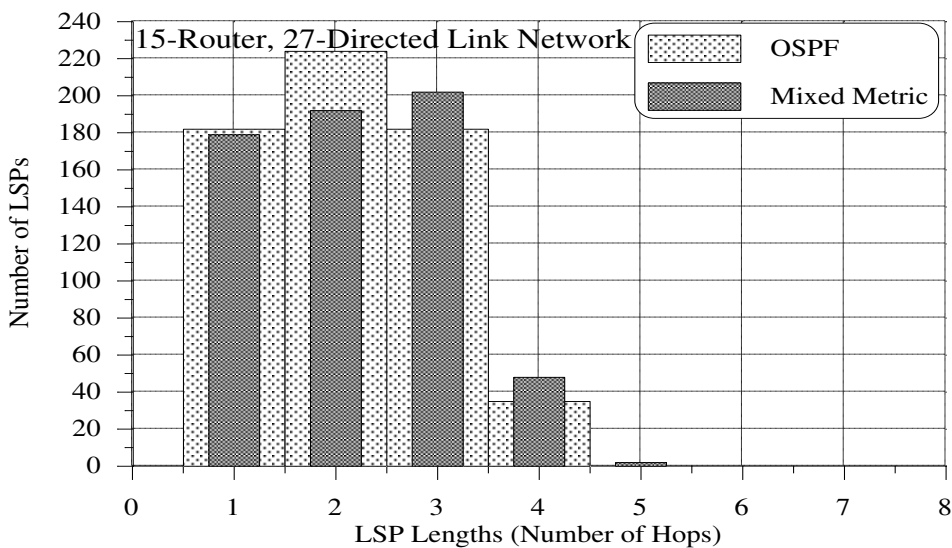


Fig.(7),LSP Length in Hops

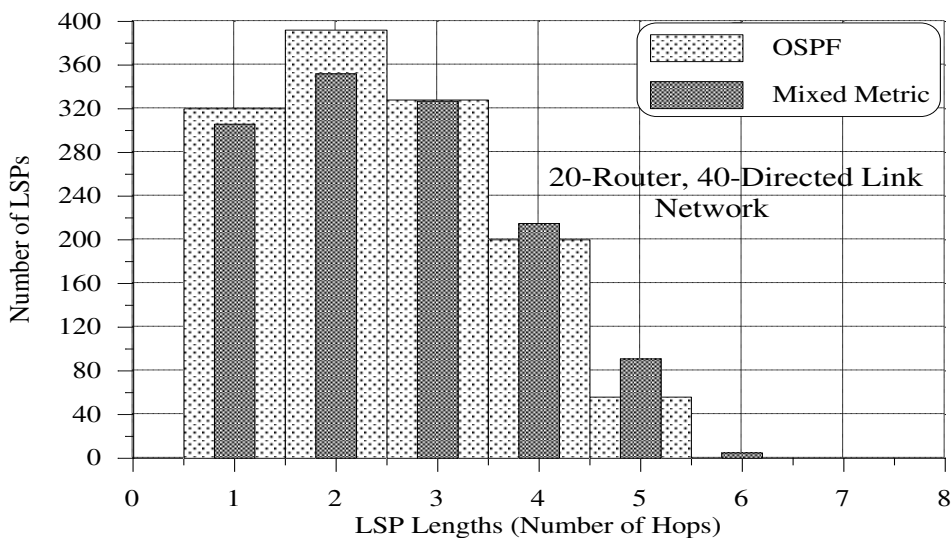


Fig.(8),LSP Length in Hops

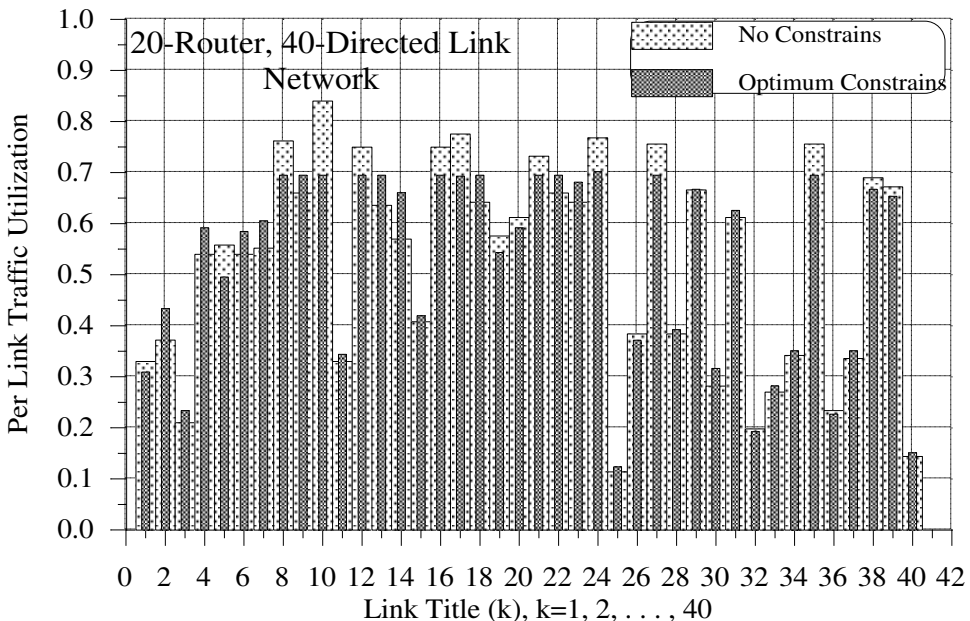
**Table I:** Mean Delay Performance.

| Performance,   | 15-Router, 27-Link |              | 20-Router, 40-Link |              |
|----------------|--------------------|--------------|--------------------|--------------|
|                | OSPF               | Mixed-Metric | OSPF               | Mixed-Metric |
| Per Link Delay | 25.23              | 12.64        | 50.71              | 15.68        |
| ETE Delay      | 63.08              | 37.92        | 152.1              | 54.88        |
| Mean Hops      | 4.691              | 4.456        | 4.231              | 3.832        |

### **Phase II: Validation of the optimum flow distribution**

In this phase, the network traffic demands,  $f_{ij}$ , are adjusted to get a heavily loaded network. The objective is to examine the behavior of the optimum load balancing system. In this respect, we shall restrict the discussion to the, larger sized (i.e., 20-router) network for it has a, relatively, larger number of, alternate paths.

At the beginning, the optimal link flows (equation (12)) are calculated based on the expected network's traffic matrix. Then, the mixed-metric algorithm searches for links that violate the optimum link flows. Such links are eliminated by assigning them infinity-costs. Then the shortest paths are found using Dijkstra algorithm. **Figure 9**, shows the per-link utilization with, and without the optimal link flow restrictions. As can be seen, utilization of all congested links (light-shaded) are restricted to the pre calculated, optimal link utilization as imposed by the network's traffic matrix we have chosen. Once again, this has been achieved at the price of some LSPs had to traverse, relatively, more number of hops as shown in **Figure 10**.

**Fig.(9), Optimum Link Utilization**

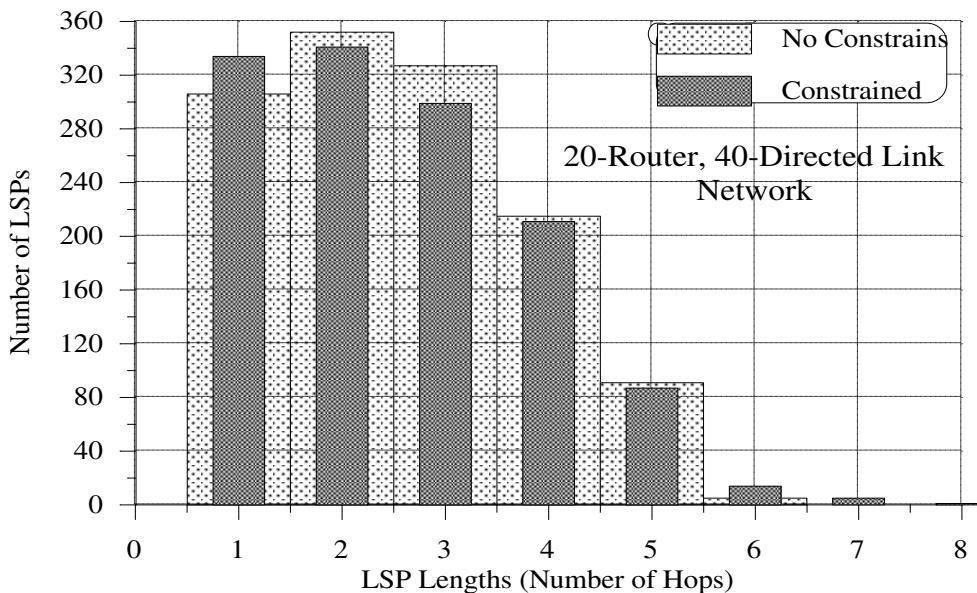


Fig.(10),LSP Length in Hops

Another results are shown in **Figures 11**, and **12**, but for a heavily loaded, randomized network traffic matrix. As can be seen, the optimal flow constrains results in a semi equalized link loadings. This would in turn, leads to a better, and bounded mean network delay performance. **Table II** quantifies the network delay performance.

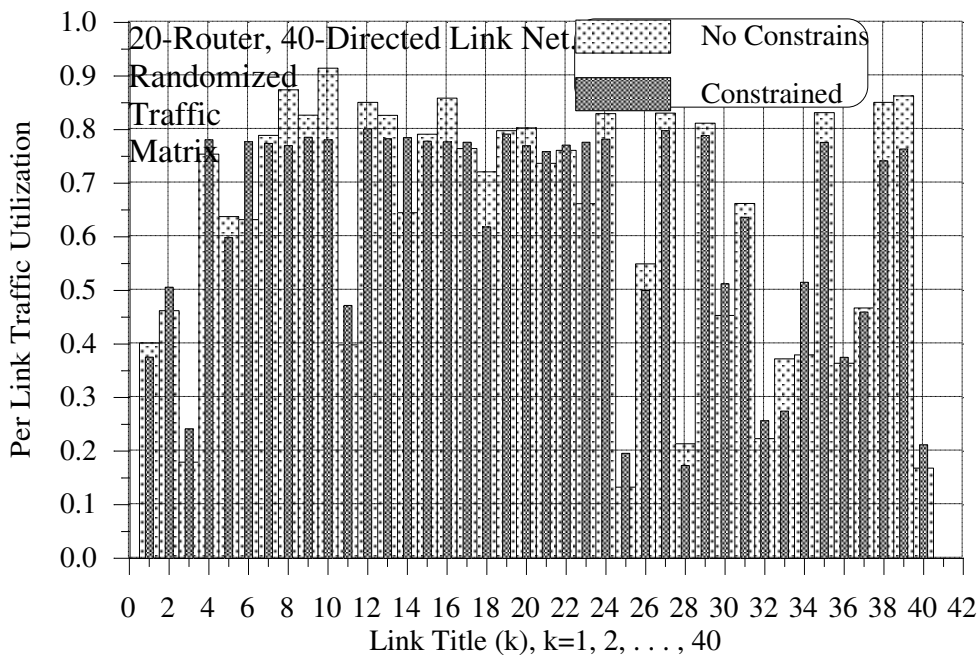


Fig.(11), Optimum Link Utilization, Randomized Traffic Matrix.

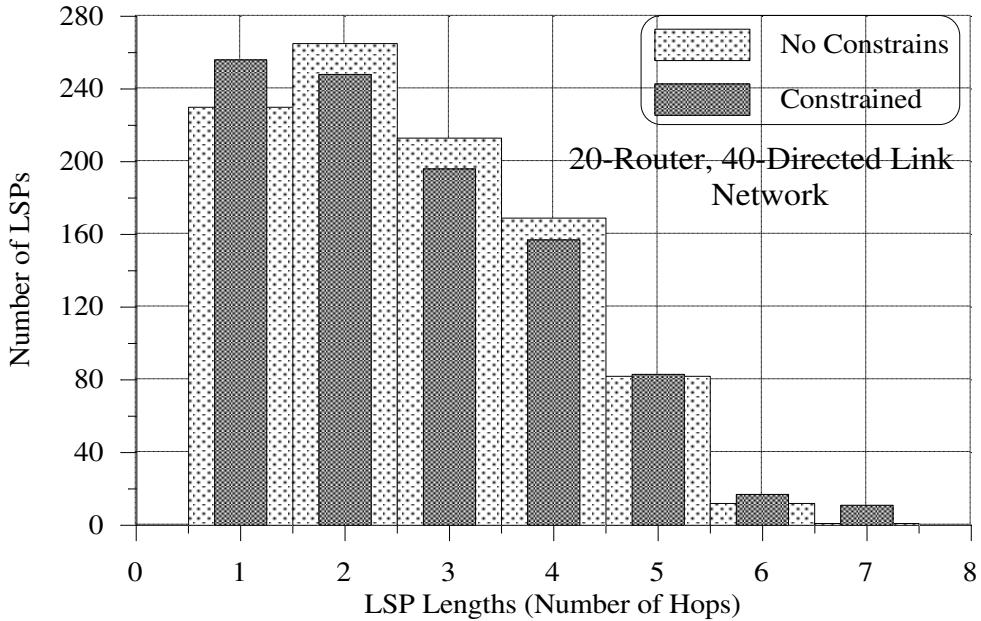


Fig.(12),LSP Length in Hops

Table II: Mean Delay Performance.

| Performance,   | Deterministic-Traffic Matrix |             | Randomized-Traffic Matrix |             |
|----------------|------------------------------|-------------|---------------------------|-------------|
|                | No Constrains                | Constrained | No Constrains             | Constrained |
| Per Link Delay | 15.68                        | 14.02       | 29.45                     | 23.64       |
| ETE Delay      | 54.88                        | 63.17       | 117.8                     | 94.57       |
| Mean Hops      | 3.838                        | 3.849       | 3.737                     | 3.71        |

As can be seen, even under heavily loaded network scenarios, the traffic engineering tool proposed in this paper, still capable of optimally distributing the traffic load among all "possible" network links as long as there is a "feasible" physical path in the directed topology which we have chosen. On the other hand, it is not difficult to see that some links (e.g., #3, and #25 in all results obtained for the 20-router network) have inherently a low utilization profile. This is due to the physical structure of the network topology which constrains the availability of a given route.

## V. SUMMARY AND CONCLUSIONS

This paper have presented a mixed-metric routing system with optimal LSP-flow distribution for traffic engineering in MPLS networks. The mixed-metric system routes LSPs over links having maximum residual bandwidth, and minimum number of flows. This have led to a significant improvement in the overall network resource (link) utilization as seen from the results presented in previous section. Moreover, an

optimal LSP-flow distribution among network links is found, and its effectiveness is tested in the validation phase. Results have indicated that, when optimal flow distribution is imposed on the routing process. It results in a lower bounded network delay performance. Such a feature is considered mandatory for a variety of applications especially those with a real time relationships among contiguous parts of their traffic streams. Typically, voice, and video traffic streams are a good examples of such an applications.

For future work, we will investigate a range of representative network topologies, and traffic models as well as other algorithms for optimizing the MPLS tunnel placement to allow for the placement of backup tunnels in case of network's link failures.

## VI. REFERENCES

- [1] D. Awduche, A. Chin, A. Elwalid, I Widjaja, and Xiao, "Overview, and principles of Internet traffic engineering," IETF RFC 3272, 2002.
- [2] D. O. Awduche, J. Malcolm, J. Agogbua, M. O'Dell, and J. McManus, "Requirements for traffic engineering over MPLS," Internet Draft, June 1999.
- [3] A. Apostolopoulos, R. Guerin, S. Kamat, Orda, T.Przygienda, and D. Williams, "QoS routing mechanisms and OSPF extensions," Internet Draft, April 1998.
- [4] Z. Zhang, C. Sanchez, B. Salkewicz, and E. Crawley, "Quality of Service Extensions to OSPF or Quality of service path first routing," Internet Draft, September 1997.
- [5] D. Awduche et al, "Requirements for traffic engineering over MPLS," IETF RFC 2720.
- [6] E. Rosen et al, "Multiprotocol Label Switching architecture," IETF RFC3031, <http://ietf.org/rfc/rfg3031.txt>, 2001.
- [7] X. Xiao, A. Hannan, B. Baily, and L. M. Ni, "Traffic engineering with MPLS in the Internet," IEEE Network magazine, March/April 2000.
- [8] G. Swallow, "MPLS advantages for traffic engineering," IEEE Communication magazine, December 1999.
- [9] J. Boyle et al, "Application statement for traffic engineering with MPLS,' IETF RFC 3346, 2002.
- [10] C. Casetti, G. Favalessa, M. Mellia, and M. Munafo, "An adaptive routing algorithm for best-effort in integrated-services networks," presented at the 16<sup>th</sup> Int. Teletraffic Congress (ITC-16), Edinburgh, UK, June 1999.
- [11] W. E. Leland, and D. V. Wilson, "High time-resolution measurements, and analysis of LAN traffic: Implementations for LAN interconnection," Proc. IEEE INFOCOM, 1991, pp. 1360-1366.
- [12] D. Gross, and C. M. Harris, "Fundamentals of Queuing Theory", 2<sup>nd</sup> edition. New York: Wiley, 1985.
- [13] F. S. Hillier, and G. J. Lieberman, "Introduction to Operations Research", 4<sup>th</sup> Ed. New York: Holden-Day, 1986.
- [14] A. Mameli, E. Manconi, and F. Ubaldi, "Edge distributed admission control in MPLS networks," IEEE Communi. Letters, Vol. 7, No. 2, Febr. 2003.

## تحديد المسارات باستخدام القيم المتعددة مع التوزيع الأعظم للحركة في شبكات MPLS

يقدم هذا البحث نظام ديناميكي جديد لتحديد المسارات في شبكات MPLS يعتمد على استخدام القيم المتعددة مع التوزيع الأعظم للحركة. يهدف هذا البحث إلى وضع حد أدنى لمستوى الأداء في هذا النوع من الشبكات. وقد أظهرت النتائج قدرة النظام المقترح على توزيع أحمال الشبكة بطريقة مثالية حتى عند الأحمال العالية.