On the Performance of Averaged Optimal Routing

Nithin Michael, Ao Tang and G. Edward Suh
School of Electrical and Computer Engineering, Cornell University, Ithaca, NY
Email:{nm373,at422,gs272}@cornell.edu

***Abstract*—Traffic uncertainty makes designing optimal routing protocols for many networks a difficult problem. A good way to capture uncertainty is via an *uncertainty set* - a collection of traffic matrices with associated probabilities of occurrence. Averaged optimal routing, where the performance metric is averaged over the potential traffic matrices, allows us to incorporate an uncertainty set into an optimization framework. We derive bounds on the performance of averaged optimal routing where the objective is to minimize maximum channel load, a commonly used metric. Furthermore, we provide examples of networks and uncertainty sets where the bounds are tight. These bounds can be used to quickly check whether averaged optimal routing is an effective routing protocol for a given network and uncertainty set. Numerical evaluations of the performance of the bounds for uncertainty sets drawn from random graphs are presented and some interesting trends are noted. Also, simulations performed on a cycle level on-chip network simulator demonstrate that averaged optimal routing outperforms general-purpose routing algorithms.**

I. Introduction

Optimal routing for a given network and traffic pattern has been an important subject of research since it was first studied in the seventies [2,3]. But in practice, while the network structure might be known, we cannot typically predict what traffic pattern might occur on the network at any point. Real world applications where this is the case include internet traffic engineering and the rapidly evolving field of on-chip networks. In the first case, the difficulty arises from the infeasibility of characterizing dynamic intra-domain traffic with a single traffic matrix. On the other hand, on-chip networks have dynamically changing application phases that generate different traffic patterns and the routes that are configured *a priori* need to handle all such application phases.

As a result, most routing algorithms that are implemented are general-purpose algorithms, since they are designed to perform well over a wide-range of traffic patterns. Some network-on-chip algorithms like dimension ordered routing, ROMM [6] and O1TURN [8] are completely oblivious to the network's traffic pattern while others like minimal adaptive routing [4] and GOAL [10] adapt to the network traffic through indirect local information about the network's global performance. Similarly, OSPF which is used for intra-domain routing adapts to internet traffic conditions by choosing the least congested paths available.

A drawback of these heuristic algorithms is that they lack provable guarantees on their performance when compared to the optimal routes that the traffic could take at any instant. But before attempting an optimization-based approach, it is important to have a good quantitative characterization of the network traffic patterns. We begin by noting that it is difficult to specify one traffic pattern that captures the behavior of a general network's traffic. But by studying the traffic carefully, we might be able to capture the behavior of the network traffic using a set of traffic matrices for which we can specify associated probabilities of occurrence. We will call this set an *uncertainty set*. However, as we will show, even with this information we cannot solve an optimal routing problem that will necessarily yield the best routes for every pattern in the uncertainty set.

One natural idea is to try and solve a combined optimal routing problem where the objective function is the averaged value of the objective functions of the individual optimal routing problems. The advantage that this approach offers is that unlike the heuristic routing schemes discussed earlier, it helps us prove a guarantee on the performance of the routes obtained. Also, as we will show in the simulations section, taking an optimal routing approach gives better performance than the other heuristic routing algorithms.

The main contribution of this paper is that we provide a performance guarantee on averaged optimal routing which is important as a natural approach to dealing with traffic uncertainty. Aside from its value as a good tool for evaluating when averaged optimal routing would be a good routing algorithm, we believe that the approach we present to derive the performance guarantees is of independent theoretical interest in the area of robust optimization.

II. Related Work

The difficult reality of dealing with traffic uncertainty when formulating the optimal routing problem has only begun to receive attention over the last few years. Algorithms like COPE [11] approach the problem by trying to minimize the worst case performance of the routing scheme within an uncertainty set. On the other hand, the problem of finding optimal routes by minimizing the expected cost over a set of traffic patterns has been studied previously in the context of intra-domain routing on the Internet [12]. However, the focus was on setting up the problem and extending the prior results [3] to develop a distributed solution method. Importantly, there was no guarantee provided for how well the averaged optimal routes performed with respect to the specialized optimal routes for the traffic matrices in the uncertainty set.

From the perspective of potential applications of averaged optimal routing, the problem has been considered in the context of network-on-chip where varying application phases and limited reconfigurability presents a natural setting that involves

978-1-4673-3140-1/12/$31.00 ©2012 IEEE

traffic uncertainty [7]. We vary from that work in that we consider maximum channel load as an objective function since it makes it easier to develop insights into how well averaged optimal routing works. As noted earlier, another area where traffic uncertainty plays an important role is in internet traffic engineering. In fact, prior work [1] has looked at studying the fundamental tradeoffs in dealing with uncertain traffic demands. But while it is hard to estimate a representative traffic matrix for autonomous systems on the Internet, it is still possible to approximate it with a suitable uncertainty set. Our work provides a framework for evaluating and deploying averaged optimal routing as an effective routing protocol for both these problems as well as others where traffic uncertainty is common.

III. Problem Formulation

In order to introduce the notation used in the paper, we first describe the linear programming formulation of the optimal routing problem when the objective is to minimize the maximum channel load for a single traffic pattern. Following that, the more general combined optimal routing problem for multiple traffic patterns and its formulation as a linear program is presented. As stated earlier, we wish to provide performance guarantees for the routes generated by the combined problem with respect to the routes generated by the specialized problem for each traffic matrix in an uncertainty set.

A. Specialized Optimal Routing Problem

We consider a network graph $\mathbf{G}$ with N nodes and L links. The following terms will help with the mathematical formulation of the optimal routing problem on the graph $\mathbf{G}$.

Traffic Matrix/Pattern (D) **–** The traffic matrix $D \in \mathbb{R}^{N\times N}$ specifies the traffic requirements of the application. Each entry $D(s,d)$ represents the desired rate of data transfer from node s to node d and each such source-destination pair is said to constitute a network flow. We suppose that there are F non-zero flows in each traffic matrix, and we label the flow from s to d as the tuple $\langle s,d\rangle$.

Incidence Matrix (A) **–** The flow constraints imposed by the topology of the network are captured by its incidence matrix $A \in \mathbb{R}^{N\times L}$ which is defined as follows,

$$A(i,j) = \begin{cases} +1, & \text{if link } j \text{ is directed to node } i \\ -1, & \text{if link } j \text{ is directed away from node } i \\ 0, & \text{otherwise.} \end{cases}$$

Link Rates (Y) **–** $Y \in \mathbb{R}^{L\times F}$ represents the rate on each link due to each flow in the traffic matrix. It is easy to see that solving for the link rates for each flow specifies the route the flow takes through the network. We also define $\gamma = \sum_{j=1}^{F} Y^j$ as the vector of the total rate on each link required by the traffic matrix where Y^j represents the link rates corresponding to the flow j.

Maximum Channel Load (W) **–** $W = \max \gamma$ is a useful metric since minimizing it makes sure that we are using the network's resources as well as we can and also because when the channels are not heavily loaded we get better performance in terms of latency and throughput which are important performance metrics that are used in many networks.

With the above notation, the specialized optimal routing problem for a given traffic matrix D can be formulated as follows,

$$\begin{aligned} \underset{Y,W}{\text{minimize}} \quad & W \\ \text{subject to} \quad & AY = \overline{D}, \\ & \sum_{j=1}^{F} Y^j \leq We, \\ & Y \geq 0. \end{aligned} \tag{1}$$

where the matrix $\overline{D} \in \mathbb{R}^{N\times F}$ is obtained from the traffic matrix D as follows,

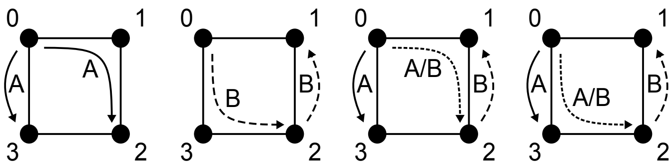
$$\overline{D}(l,sd) = \begin{cases} +D(s,d), & \text{if } l = d \text{ for the flow } \langle s,d\rangle \\ -D(s,d), & \text{if } l = s \text{ for the flow } \langle s,d\rangle \\ 0, & \text{otherwise.} \end{cases}$$

and $e \in \mathbb{R}^L$ is a vector of 1s.

The above problem can be solved efficiently for a given traffic matrix D to obtain optimal routes for it using the standard techniques developed for linear programs. In the rest of the paper, feasible and optimal solutions to the above problem will be called specialized routes and specialized optimal routes respectively. Next we describe the combined optimal routing problem when we are given an uncertainty set $\mathbb{D} = \{D_1, \ldots, D_M\}$ of M traffic patterns that occur with probabilities $p_1, \ldots, p_M$ instead of a single traffic matrix D. In order to keep notation consistent, we will use the sub-index i to indicate quantities associated with traffic matrix D_i. Also, feasible solutions to the combined problem will be called combined routes and the optimal solution will be called combined optimal routes.

B. Combined Optimal Routing Problem

The key challenge in setting up the combined optimal routing problem is creating a unified optimization framework that can determine *both* the specialized routes for each traffic pattern *and* the way these specialized routes interact to determine combined optimal routes. The approach that we adopt is to design an optimization problem that minimizes the expected maximum channel load across all traffic matrices in the uncertainty set. In addition, the combined routes have to satisfy the requirements of every traffic matrix in the uncertainty set simultaneously. This means that if there is a flow that is shared across multiple traffic matrices, i.e. for a given source-destination pair multiple traffic matrices have a corresponding non-zero entry, then the route computed for it should be the same for each of those traffic matrices. From [7], we know that this problem can be formulated as,



(a) (b) (c) (d)

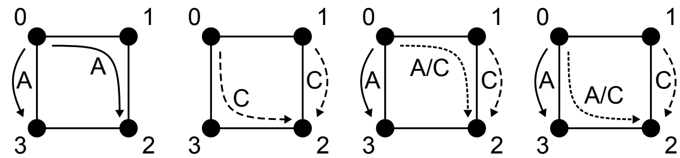
Fig. 1. **Specialized and Combined Routes for Traffic Matrices D_A and D_B** – $D_A(i,j) = D_B(i,j) = 0$ except for $D_A(0,3) = D_A(0,2) = D_B(0,2) = D_B(2,1) = 1$.

$$
\begin{aligned}
\underset{\substack{Y_1,Y_2,\ldots,Y_M\\ W_1,W_2,\ldots,W_M}}{\text{minimize}} \quad & \sum_{i=1}^{M} p_i W_i \\
\text{subject to} \quad & AY_i = \overline{D}_i, i = 1,2,\ldots,M \\
& \sum_{j=1}^{F_i} Y_i^j \leq W_i e, i = 1,2,\ldots,M \qquad (2)\\
& Y_i \geq 0, i = 1,2,\ldots,M \\
& Y_i^{\langle s,d\rangle}/D_i(s,d) = Y_j^{\langle s,d\rangle}/D_j(s,d) \\
& \text{if flow } \langle s,d\rangle \text{ is in both } \overline{D}_i \text{ and } \overline{D}_j.
\end{aligned}
$$

It is possible that we can find combined routes that are the same as the specialized optimal routes for each traffic matrix in the uncertainty set. Solving the above optimization problem is guaranteed to find this solution if it exists. Suppose that there exists such a solution for an uncertainty set $\mathbb{D} = \{D_1, \ldots, D_M\}$ with probabilities $p_1, \ldots, p_M$ and that the solution to Problem 2 does not correspond to these routes. We can then show that a contradiction results, since selecting the aforementioned solution will further decrease the cost function in Problem 2 as W_i corresponding to each D_i is minimized by the specialized optimal routes by definition.

As a simple example of when combined routes are the same as specialized optimal routes consider Fig. 1. Here each network link represents two unidirectional channels. So there are two specialized optimal routes that work for the flow$\langle 0,2\rangle$ in traffic matrix D_B. The combined routes shown in Fig. 1c are also specialized optimal for both D_A and D_B while the routes shown in Fig. 1d though feasible are neither combined optimal nor specialized optimal for D_A.

But there are still cases as in Fig.2 when it is simply not possible to find combined routes that are also specialized optimal for every traffic matrix in the uncertainty set. Here choosing combined routes that are specialized optimal for one traffic matrix results in a sub-optimal solution for the other traffic matrix. In such cases we would like to provide bounds on how much performance is lost by using the combined optimal routes instead of the specialized optimal routes for the traffic matrix that needs to routed. We would also like to know how good the bounds in question are in practice. In the next section we present answers to these questions.



(a) (b) (c) (d)

Fig. 2. **Specialized and Combined Routes for Traffic Matrices D_A and D_C** – $D_A(i,j) = D_C(i,j) = 0$ except for $D_A(0,3) = D_A(0,2) = D_C(0,2) = D_C(1,2) = 1$.

IV. Performance Evaluation of Combined Optimal Routes

We will first use a deterministic construction to get a bound on the maximum channel load resulting from using the combined optimal routes for a traffic matrix D_i. Then, after giving examples where this bound is tight, we will use it to derive a bound that also accounts for the probabilities associated with the traffic matrices in a given uncertainty set. Throughout this section we will use the superscript * to represent specialized optimal solutions.

Theorem 1. *Suppose that $\overline{Y_i}$ represents the routes that are generated for a traffic matrix D_i by solving the combined optimal routing problem. Then $\overline{W_i}$, the maximum channel load corresponding to these routes, can be bounded as follows,*

$$\overline{W_i} \leq W_i^* + \sum_{\substack{j=1\\ j\neq i}}^{M} \sum_{\langle s,d\rangle^j} D_i(s,d)$$

where $\langle s,d\rangle^j$ represents the flows that are shared between D_i and D_j but which are not already shared with D_k where $k < j$ and which do not have the same routes for both Y_i^* and $Y_j^*, j \neq i$.

Proof: Let us construct a solution $\hat{Y_i}$ with corresponding maximum channel load $\hat{W_i}$. We can do this by selecting the routes $Y_i^{\langle s,d\rangle *}$ for all $\langle s,d\rangle$ that are not shared with any of the other traffic matrices and adding to them the routes $D_i(s,d)Y_j^{\langle s,d\rangle *}/D_j(s,d)$ for the flows that are shared with traffic matrix D_j. In the case where flows are shared with multiple traffic matrices, we select the routes specified by the first Y_j^* encountered.

Note that any constructed solution $\hat{W_i}$ can be upper bounded by the sum of the demands generated by the shared flows between the traffic matrices D_i and D_j where $j = 1, \ldots, M, j \neq i$ and W_i^*, the maximum channel load seen by the network when routing using Y_i^*. This is because the most heavily utilized link's load can be increased by at most the demands of the shared flows when routing using the solution for another matrix. Of course, here it is important to not double count flows. Also, it is not necessary to include all shared flows in the summation since the flows which are shared yet share the same specialized optimal routes will not contribute to this worst case load. The same observation is true about the maximum channel load resulting from implementing the combined optimal routes for a traffic matrix D_i. This is because in $\overline{Y_i}$ we use the optimal routes for the flows that are not shared with the other traffic matrices, and the worst case increase in maximum channel load can be captured as,

$$\overline{W_i} \leq W_i^* + \sum_{\substack{j=1 \\ j \neq i}}^{M} \sum_{\langle s,d \rangle^j} D_i(s,d)$$

■

At this point a natural observation is that this bound can be very bad if there are multiple shared flows of large value. Regardless, it is possible to show that we need to account for them in bounding the channel load. The following simple example illustrates that the bound derived above can be tight, i.e., achieve equality. More cases where the bound achieves equality can be constructed by extending the ideas in the following construction. Consider the network shown in Fig. 3. Suppose that the uncertainty set consists of two traffic matrices D_A and D_B with $\langle s, d \rangle$ pairs and their specialized optimal routes as shown. Also, suppose that the two traffic patterns have associated probabilities of occurrence $p_A = p_B = 1/2$.

Now it is easy to show that the optimal averaged maximum channel load for this example has value 2. Let us first focus attention on D_A. One combined optimal solution is to use the specialized optimal routes for the unshared flows, $D_A(4,3)$ and $D_B(4,3)$ and to route $D_A(4,6)$ the same way that $D_B(4,6)$ is routed. Thus, the maximum channel load in case traffic pattern B occurs will still be 2 but if traffic pattern A occurs it increases to 2 from 1. The bound is tight since from Theorem 1 for D_A it can be computed as,

$$W_A^* + D_A(4,6) = 1 + 1 = 2$$

Next we can focus on traffic matrix D_B. Another combined optimal solution to problem (2) would be to once again use the specialized optimal routes for $D_A(4,3)$, $D_B(4,3)$ and the unshared flows and to modify the route taken by $D_B(4,6)$ to that taken by $D_A(4,6)$. Now the maximum channel load when traffic pattern B occurs increases to 3 but the maximum channel load when traffic pattern A occurs is still 1. We can see that the bound is tight since for D_B the bound can be computed as,

$$W_B^* + D_B(4,6) = 2 + 1 = 3$$

The above example shows that for two traffic patterns the bound is tight. It is easy to construct examples for any number of traffic patterns by using similar networks and traffic patterns. For instance, another example is presented in Fig. 4. Now there are three traffic matrices D_A, D_B and D_C with $\langle s, d \rangle$ pairs and corresponding specialized optimal routes as shown. The probabilities of ocurrence are $p_A = p_B = p_C = 1/3$. The optimal averaged maximum channel load for this uncertainty set is $7/3$. One set of routes that achieves this average maximum channel load can be obtained by using the specialized optimal routes for D_A and D_B and modifying the routes used by $D_C(7,5)$ and $D_C(1,3)$ accordingly. Then we can show that the bound achieves equality for traffic pattern C since,

$$W_C^* + D_C(7,5) + D_C(1,3) = 1 + 1 + 1 = 3 = \overline{W_C}$$

Another set of routes that achieves the optimal value can be specified by using the specialized optimal routes for D_C and modifying the routes taken by $D_A(1,3)$ and $D_B(7,5)$ accordingly. Again, the bound can be shown to be tight for both traffic patterns A and B since,

$$W_A^* + D_A(1,3) = 2 + 1 = 3 = \overline{W_A}$$
$$W_B^* + D_B(7,5) = 2 + 1 = 3 = \overline{W_B}$$

So far we have described a deterministic bound on the averaged optimal channel load obtained by using the combined

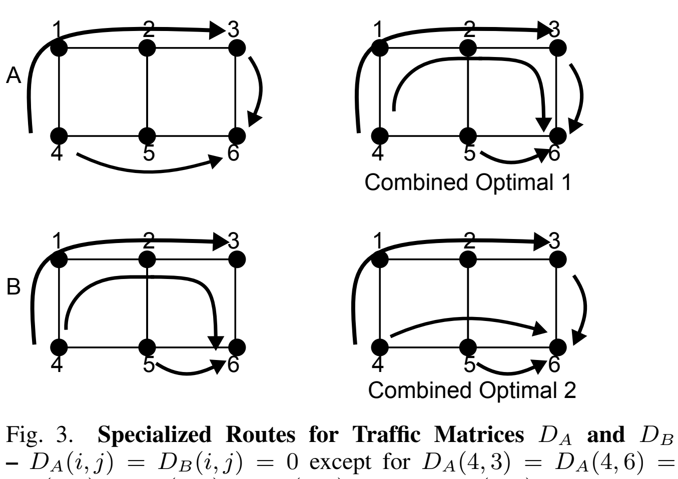
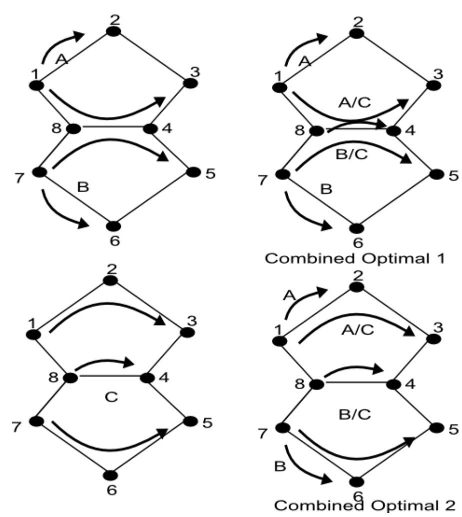


Fig. 3. **Specialized Routes for Traffic Matrices D_A and D_B** – $D_A(i,j) = D_B(i,j) = 0$ except for $D_A(4,3) = D_A(4,6) = D_A(3,6) = D_B(4,3) = D_B(4,6) = 1$ and $D_B(5,6) = 2$.



.

Fig. 4. **Specialized Routes for for Traffic Matrices D_A, D_B and D_C** – $D_A(i,j) = D_B(i,j) = D_C(i,j) = 0$ except for $D_A(1,2) = D_B(7,6) = 2$ and $D_A(1,3) = D_B(7,5) = D_C(1,3) = D_C(7,5) = D_C(8,4) = 1$.

optimal routes. But depending on the parameters, accounting for the probabilities associated with the traffic matrices could improve the bound. Consequently, we also present a bound on the averaged optimal channel load that incorporates the probabilities of occurrence of the different traffic matrices.

Theorem 2. *Suppose that $\overline{Y_i}$ represents the routes that are generated for a traffic matrix D_i by solving the combined optimal routing problem. Then $\overline{W_i}$, the maximum channel load corresponding to these routes, can be bounded as follows,*

$$\overline{W_i} \leq W_i^* + \frac{1}{p_i} \sum_{\substack{j=1 \\ j \neq i}}^{M} p_j \sum_{\substack{k=1 \\ k \neq j}}^{M} \sum_{\langle s,d \rangle^k} D_j(s,d)$$

where $\langle s,d \rangle^k$ represents the flows that are shared between D_j and D_k but which are not already shared with D_l where $l < k$ and which do not have the same routes for both Y_i^* and $Y_j^*, j \neq i$.

Proof: We will rely on constructing a feasible solution to the combined optimal routing problem in order to establish the result. Note that a feasible solution can be constructed by taking the routes specified by Y_i^* for the flows in D_i, taking the routes specified by Y_j^* for the flows in D_j (where $j \neq i$) that are not shared with each other and adding to them the routes specified by Y_i^* for the flows that are shared with D_i and choosing routes arbitrarily for the flows that are shared among matrices not including D_i. Then for each traffic matrix D_i and corresponding maximum channel load $\overline{W_i}$, the result holds because of the following chain of inequalities.

$$\sum_{\substack{j=1 \\ j \neq i}} p_j \sum_{\substack{k=1 \\ k \neq j}}^{M} \sum_{\langle s,d \rangle^k} D_j(s,d) + \sum_{j=1}^{M} p_j W_j^* \geq \sum_{\substack{j=1 \\ j \neq i}}^{M} p_j \hat{W}_j + p_i W_i^*$$

$$\geq \sum_{j=1}^{M} p_j \overline{W_j} \geq \sum_{\substack{j=1 \\ j \neq i}} p_j W_j^* + p_i \overline{W_i}$$

Here the first inequality follows from the bound from Theorem 1 applied to each $\hat{W}_j$, the second inequality is true since a feasible solution has a higher average maximum channel load than the optimal solution to problem (2) and the last inequality holds since the optimal maximum channel loads are less than or equal to the channel loads obtained by using combined optimal routes. ∎

V. Numerical and Simulation Results

Till now we have focused on providing performance bounds for the combined optimal routes. Of course, this is useful only if averaged optimal routing is a good approach for problems where there is traffic uncertainty. As noted in the introduction, on-chip networks are a natural setting for combined optimal routing. Also, the key metrics that are of interest in a network are communication latency and network throughput. In previous work [7] it was shown that when the objective is to minimize the total traffic on the network, combined optimal routes perform very well compared to specialized routes with respect to these metrics. A natural question is whether this is still the case when the objective is changed to minimizing the maximum channel load.

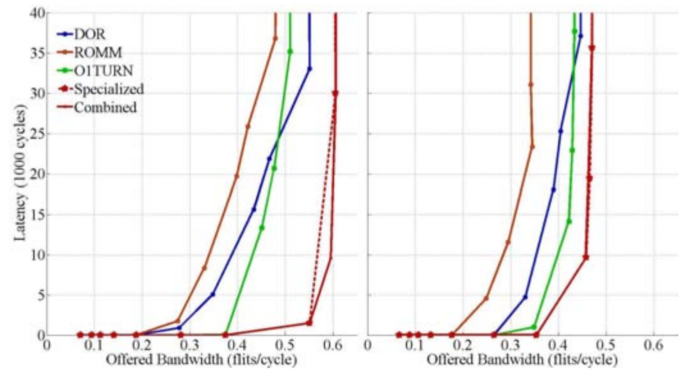


Fig. 5. **Latency vs. Offered Bandwidth for Optimal Combined Routes –** Theoretical and numerical analysis predicts the combined routes should be able to achieve the same throughput as the specialized routes.

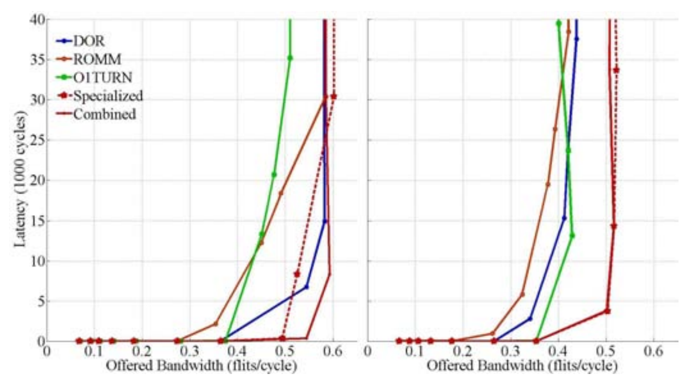


Fig. 6. **Latency vs. Offered Bandwidth for Sub-Optimal Combined Routes –** Theoretical analysis shows that combined routes that are specialized optimal for both traffic matrices are infeasible, but numerical analysis predicts that the combined routes should still perform close to the specialized routes.

The simulator that we used was DARSIM [5], a cycle-level on-chip network simulator. All the simulations were performed on a 6×6 two-dimensional mesh network. The simulator was given a warm-up period of 20,000 cycles after which performance statistics were collected over 100,000 cycles in order to ensure the accuracy of the results. The primary performance criteria that we measured were throughput and latency. The data rates are expressed in flits/cycle and each packet is divided

p	0.1	0.3	0.5	0.7	0.9
0.1	1.78	3.67	6.13	8.06	10.71
0.3	2.12	3.88	6.23	10.27	13.57
0.5	2.37	4.55	6.51	9.00	12.94
0.7	2.41	5.18	6.90	8.30	10.90
0.9	2.53	5.69	7.96	8.90	9.10

TABLE I
Expected performance ratio as calculated by the deterministic bound for 4x4 meshes. The $(i, j)th$ entry is the expected performance ratio for the matrix drawn from the family with probability specified by row i.

into 8 flits. Also the simulator was configured so that each physical channel was divided into 6 virtual channels with 8 flits of buffering each. The capacity of the physical channel was set to be 1 flit/cycle. In the simulator, virtual channels are pre-allocated to the different flows once the routes are computed so that deadlock is avoided according to the static virtual channel allocation scheme described in [9].

The simulations were performed with a view to incorporating affects like buffering and flow control so that we can evaluate the combined optimal routing model better. We present two examples, one where the combined optimal routes match the performance of the specialized optimal routes and one where they do not. Also plotted are throughput-latency curves associated with commonly used on-chip network routing algorithms like ROMM, O1TURN, and DOR.

The results in Fig. 5 and Fig. 6 show that combined optimal routes perform well even when the objective is changed to minimizing the maximum channel load. The graphs show two cases of randomly generated uncertainty sets with two matrices. In Fig. 5, the uncertainty set had combined optimal routes that were also specialized optimal for both matrices in the set. But in Fig. 6 we see that the combined optimal routes are no longer specialized optimal though they still perform very well. Lastly, note that combined optimal routing with respect to maximum channel load outperforms the application-oblivious routing algorithms.

Another experiment that we performed was to try and quantify how well the bound performed when the uncertainty set consisted of two matrices drawn from random graphs with different probabilities of links (of weight 1) occuring for a 4x4 mesh. The results were evaluated numerically and for each pair of random graph families the expected bound was averaged over 100 runs. The results are presented in Table I.

It is interesting to note the trend that the more dense demand matrix has a better predicted performance bound. This is because it determines how most of the traffic will be routed in the averaged case as it will have a higher maximum channel load. On the other hand, the predicted performance for the sparser matrix is worse since it has less influence in determining the routes in the averaged optimal solution. Another observation is that when both demand matrices are relatively sparse the resulting performance bounds are good since there is more freedom for routes to be reconfigured without affecting the optimal channel load. A surprising observation was that, on average, at least for the case of two randomly generated demand matrices, the performance bound was not as large as its definition indicated it should be.

Of course, it should be noted that in many of the cases that we tested, the averaged optimal routes performed as well as their specialized counterparts despite the prediction of the bound. But this was expected since the bound was designed to capture worst case scenarios which do not necessarily occur when randomly generating uncertainty sets.

VI. Summary and Future Work

In this paper we derived a deterministic and a probabilistic bound on the performance of average optimal routing given an uncertainty set and also provided examples where the deterministic performance bound is tight. We showed that averaged optimal routing where the objective is to minimize maximum channel load performs better than other routing schemes that are typically implemented to deal with traffic uncertainty. From an engineering perspective, the performance guarantees are useful as tools to quickly evaluate whether averaged optimal routing would be a good routing algorithm for a given uncertainty set. From a theoretical perspective, the bounds help improve our understanding of averaged optimal routing as a robust optimization approach. As part of future work, we would like to construct examples where the probabilistic bounds are tight as well as further explore the structure of the averaged optimal routing approach to see if these bounds are the best performance guarantee that we can provide.

References

[1] D. Applegate and E. Cohen, "Making intra-domain routing robust to changing and uncertain traffic demnds: Understanding fundamental tradeoffs," *ACM SIGCOMM*, 2003.

[2] D. Cantor and M. Gerla, "Optimal Routing in a Packet-Switched Computer Network," *IEEE Trans. on Computers*, vol. 23, no. 10, pp. 1062–1069, 1974.

[3] R. Gallager, "A Minimum Delay Routing Algorithm Using Distributed Computation," *IEEE Transactions on Communications*, vol. 25, no. 1, pp. 73–85, Jan 1977.

[4] L. Gravano, G. D. Pifarre, P. E. Berman, and J. L. C. Sanz, "Adaptive Deadlock- and Livelock-Free Routing with All Minimal Paths in Torus Networks," *IEEE Trans. on Parallel and Distributed Systems*, vol. 5, no. 12, pp. 1233–1252, Dec 1993.

[5] M. Lis, K. S. Shim, M. H. Cho, P. Ren, O. Khan, and S. Devadas, "DARSIM: A Parallel Cycle-Level NoC Simulator," *Workshop on Modeling, Benchmarking, and Simulation*, Jun 2010.

[6] T. Nesson and S. Johnsson, "ROMM Routing on Mesh and Torus Networks," *Symp. on Parallelism in Algorithms and Architecture*, Jul 1995.

[7] N.Michael, M. Nikolov, A. Tang, G. E. Suh, and C. Batten, "Analysis of Application-Aware On-Chip Routing under Traffic Uncertainty," *Int'l Symp. on Networks-on-Chips*, May 2011.

[8] D. Seo, A. Ali, W. Lim, and N. Rafique, "Near-Optimal Worst-Case Throughput Routing for Two-Dimensional Mesh Networks," *Int'l Symp. on Computer Architecture*, Jun 2005.

[9] K. S. Shim, M. H. Cho, M. Kinsy, T. Wen, M. Lis, G. E. Suh, and S. Devadas, "Static Virtual Channel Allocation in Oblivious Routing," *Int'l Symp. on Networks-on-Chips*, May 2009.

[10] A. Singh, W. Dally, A. Gupta, and B. Towles, "GOAL: A Load-Balanced Adaptive Routing Algorithm for Torus Networks," *Int'l Symp. on Computer Architecture*, Jun 2003.

[11] H. Wang, H. Xie, L. Qiu, Y. Yang, Y. Zhang, and A. Greenberg, "COPE: Traffic Engineering in Dynamic Networks," *SIGCOMM*, Sep 2006.

[12] C. Zhang, Y. Liu, W. Gong, J. Kurose, R. Moll, and D. Towsley, "On Optimal Routing with Multiple Traffic Matrices," *INFOCOM*, Mar 2005.