Proceedings of the Fourth IASTED International Conference
Communications, Internet, and Information Technology
October 31 - November 2, 2005, Cambridge, USA

BRAESS PARADOX IN DYNAMIC ROUTING FOR THE COHEN-KELLY NETWORK

Atsushi Inoie
Doctoral Program in Computer Science,
University of Tsukuba,
Tsukuba Science City,
Ibaraki,305-8573, Japan.
email: inoie@osdp.cs.tsukuba.ac.jp

Hisao Kameda and Corinne E. Touati
Department of Computer Science,
University of Tsukuba,
Tsukuba Science City,
Ibaraki,305-8573, Japan.
email: {kameda, corinne}@osdp.cs.tsukuba.ac.jp

ABSTRACT
A paradox of dynamic routing in the Cohen-Kelly network is studied. Intuitively, we expect that adding capacity to a network improves the performance of the users. Braess showed an example wherein the opposite occurs. We refer to such a situation as a paradox. In this paper, we deal with a class of queueing networks in which Cohen and Kelly discovered a paradox in static routing. We consider the dynamic routing problems in the above mentioned class of networks, and show the existence of a paradox in dynamic routing of the above mentioned class of networks through simulation experiments analogous to what Cohen and Kelly showed.

KEY WORDS
Braess paradox, Queueing network, Routing, Wardrop equilibrium.

1 Introduction

An important problem in current high-speed and large-scale computer and communication networks (e.g., GRID, Internet) is to provide all users with satisfactory network performance. Intuitively, we expect that the total performance of a network will increase in its capacity. In non-cooperative optimization (which results in a Nash [14] or Wardrop [16] equilibrium), however, adding capacity or a link to a network may sometimes degrade the performance for all users as exemplified by the Braess paradox [2]. That is, Braess [2] considered a network routing problem in which all the costs of the links are linear. In a Wardrop equilibrium, he discovered a paradoxical phenomenon as described above. The Braess paradox attracted attention of many researchers, and was found in various types of network optimization problems, for example, [1], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [15].

Cohen and Kelly [5] studied static routing in two queueing networks, called the *initial* and *augmented* networks. The initial network has two paths, and the augmented network has an additional path. An individual user arriving at the network only knows the expected transit time of each path in the network. In that situation, they showed the existence of a paradox. Also, in [5], the authors raised the question of whether the paradox also occurs in networks where users have knowledge not only on the expected transit time of each path in the network, but also on the instantaneous length of each queue.

Calvert et al. [3] considered dynamic routing in the same networks as introduced by Cohen and Kelly [5]. In their study, individual users have full knowledge of all the instantaneous queue lengths of the network servers and they can make use of that knowledge in their dynamic routing decisions. They analytically derived the recursive equations that provide a dynamic routing decision. Through some simulation experiments, they contended that they showed that a Braess paradox occurred in dynamic routing analogous to what Cohen and Kelly showed.

To show the existence of the paradox, however, they just compared between the overall mean transit times of jobs in the initial and augmented networks. That is, they did not show that the mean transit time of jobs routed through *any* path was higher in the augmented network than in the initial network.

Furthermore, we note that the system used in the simulation study performed by Calvert et al. [3] does not seem to be identical to the one they used in their analytical study. We particularly note that, in the analytical study, they assumed that the total number of jobs entering the system is finite while in their simulation study, the total number of jobs entering the system is infinite.

In this paper, therefore, we deal with dynamic routing for the Cohen-Kelly network. Based on the assumptions and formulations of the network in [3], we derive dynamic routing decisions in the initial and augmented networks. We make experiment on the simulation system that reflects exactly the analytical model given in [3]. Through the simulation experiments, we show a case where the paradox occurs for all users. Also, we compare our results with the results by Calvert et al. [3], and point out that their results are not sufficient to show that a paradox occurs in the network, in our sense.

This paper is organized as follows. In section 2, we describe the Cohen-Kelly network and derive the dynamic routing decision. In section 3, we describe our simulation method, and show some simulation results. In section 4, we conclude this paper.

2 The Model and Assumption

We consider the two networks as shown in Figures 1 and 2. We call the former the *initial network* and the latter the *augmented network*. Both networks consist of an entrance (node 0), four queues (nodes 1, 2, 3 and 4), and an exit (node 5). Node 1 (respectively 4) is a first-come-first-served (FCFS) queueing system that has an exponential server with mean $1/\mu_1$ (respectively $1/\mu_4$). Nodes 2 and 3 are queueing systems with infinite servers (IS) whose service times are exponential distributed with means $1/\mu_2$ and $1/\mu_3$, respectively. Let $x_1(t)$, $x_2(t)$, $x_3(t)$ and $x_4(t)$ denote the numbers of jobs in nodes 1, 2, 3 and 4 at time t, respectively.

A flow of jobs arrives at the system according to a Poisson process with rate λ. As assumed in [3], the total number of jobs that enter the network is finite, and we denote it by N. Also, we denote the number of jobs before arriving at node 0 by x_0.

Let D denote the set of states in the whole system using dynamic routing, that is,

$$D = \{\boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{N}^4 \text{ and } x_0 + x_1 + x_3 + x_4 \leq N\}, \quad (1)$$

where $\boldsymbol{x} = (x_0, x_1, x_3, x_4)$ is the state vector for dynamic routing decision, and $\mathbb{N} = \{0, 1, 2, \ldots\}$. Note that we can ignore x_2 in any routing decisions since the expected transit time of jobs routed through any path is not affected by x_2.

2.1 Initial Network

The initial network (Figure 1) has two paths, $0-1-2-5$ and $0-3-4-5$, and a job dynamically chooses one of these paths in order to minimize its own expected transit time from the entrance (node 0) to the exit (node 5). Let $T_1^I(\boldsymbol{x})$ and $T_3^I(\boldsymbol{x})$ denote the expected transit times of jobs routed through nodes 1 and 3 when the job sees at node 0 that the system state is $\boldsymbol{x}$ in the initial network, respectively.

We denote the two subsets of states in the whole system in the initial network by D_1^I and D_3^I. D_1^I and D_3^I are given as follows:

$$D_1^I = \{\boldsymbol{x} \mid \boldsymbol{x} \in D \text{ and } T_1^I(\boldsymbol{x}) \leq T_3^I(\boldsymbol{x})\}, \quad (2)$$

and

$$D_3^I = \{\boldsymbol{x} \mid \boldsymbol{x} \in D \text{ and } T_1^I(\boldsymbol{x}) > T_3^I(\boldsymbol{x})\}. \quad (3)$$

Note that $D_1^I \cap D_3^I = \phi$ and $D_1^I \cup D_3^I = D$. If a job sees at node 0 that the system state is in D_1^I, then the job is routed

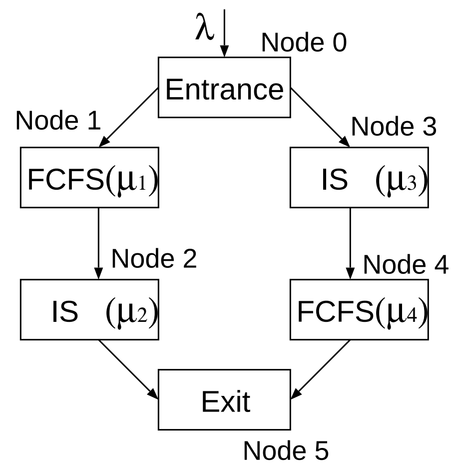

Figure 1. Initial Network: The network has two paths (0-1-2-5 and 0-3-4-5).

through node 1, and otherwise (i.e., the system state is in D_3^I) it is routed through node 3.

From [3], $T_1^I(\boldsymbol{x})$ and $T_3^I(\boldsymbol{x})$ are given by the following equations:

$$T_1^I(\boldsymbol{x}) = \frac{x_1 + 1}{\mu_1} + \frac{1}{\mu_2}, \ \boldsymbol{x} \in D, \quad (4)$$

and

$$\begin{aligned} T_3^I(\boldsymbol{x}) = & \frac{1}{I_{X_0}\lambda + I_{X_1}\mu_1 + (x_3+1)\mu_3 + I_{X_4}\mu_4}\Big[1 \\ & + \ I_{X_0}\lambda T_3^I(x_0 - 1, x_1 + I_{C_1}, x_3 + 1 - I_{C_1}, x_4) \\ & + \ I_{X_1}\mu_1 T_3^I(x_0, x_1 - 1, x_3, x_4) + \frac{\mu_3(x_4+1)}{\mu_4} \\ & + \ x_3\mu_3 T_3^I(x_0, x_1, x_3 - 1, x_4 + 1) \\ & + \ I_{X_4}\mu_4 T_3^I(x_0, x_1, x_3, x_4 - 1)\Big], \ \boldsymbol{x} \in D, \end{aligned} \quad (5)$$

where $X_i = \{x_i > 0\}$ and $C_1 = \{(x_0-1, x_1, x_3+1, x_4) \in D_1^I\}$, and

$$I_X = \begin{cases} 1, & \text{if } X \text{ is true,} \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

2.2 Augmented Network

The augmented network (Figure 2) has a path $0-1-4-5$ in addition to the paths that the initial network has. Therefore, there exist two decision-making points (nodes 0 and 1) in the augmented network. At node 0, a job chooses a path by using dynamic routing decision. If the job is routed through node 1, then it needs to choose between the two paths (nodes 2 and 4). Note that if a job at node 1 sees that the current state is (x_0, x_1, x_3, x_4), then the expected

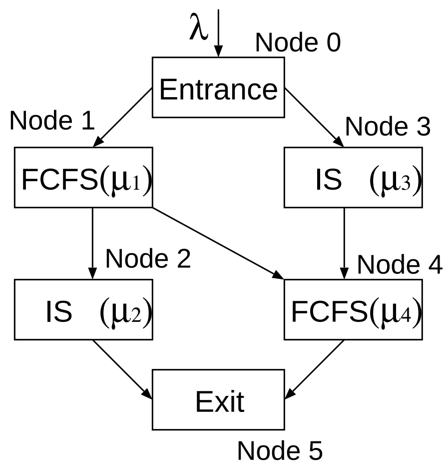

Figure 2. Augmented Network: The network has three paths (0-1-2-5, 0-3-4-5 and 0-1-4-5).

delays at nodes 2 and 4 are $1/\mu_2$ and $(x_4+1)/\mu_4$, respectively. Therefore, if $1/\mu_2 \le (x_4+1)/\mu_4$, then the job is routed through node 2. Otherwise, it is routed through node 4.

Let $T_1^A(\boldsymbol{x})$ and $T_3^A(\boldsymbol{x})$ denote the expected transit times of jobs routed through nodes 1 and 3 when the job sees at node 0 that the system state is $\boldsymbol{x}$ in the augmented network, respectively. Then, similarly to the initial network, we define two subsets of D, D_1^A and D_3^A as follows:

$$D_1^A = \{\boldsymbol{x} |\ \boldsymbol{x} \in D \text{ and } T_1^A(\boldsymbol{x}) \le T_3^A(\boldsymbol{x})\}, \quad (7)$$

and

$$D_3^A = \{\boldsymbol{x} |\ \boldsymbol{x} \in D \text{ and } T_1^A(\boldsymbol{x}) > T_3^A(\boldsymbol{x})\}. \quad (8)$$

Note also that $D_1^A \cap D_3^A = \phi$ and $D_1^A \cup D_3^A = D$.

As to $T_1^A(\boldsymbol{x})$, we note that the expected transit time of a job at node 1 is affected by the number of jobs waiting behind. We denote the number of jobs behind the job at node 1 by x_1', the state vector including x_1' by $\boldsymbol{x}' = (x_0, x_1, x_1', x_3, x_4)$, and the set of $\boldsymbol{x}'$ by

$$D' = \{\boldsymbol{x}' \mid \boldsymbol{x}' \in \mathbb{N}^5 \text{ and } x_0+x_1+x_1'+x_3+x_4 \le N\}. \quad (9)$$

Then, from [3] the expected transit time of the job, $T_1'^A(\boldsymbol{x}')$ is given by the following equation:

$$\begin{aligned} T_1'^A(\boldsymbol{x}') = & \frac{1}{I_{X_0}\lambda + \mu_1 + x_3\mu_3 + I_{X_4>0}\mu_4}\Big[1 \\ & + \ I_{X_0}\lambda T_1^A(x_0-1, x_1, x_1'+I_{C_3}, x_3+1-I_{C_3}, x_4) \\ & + \ \mu_1\big(I_{X_1}T_1^A(x_0, x_1-1, x_1', x_3, x_4+I_{C_2}) \\ & + \ I_{x_1=0}\min(\frac{1}{\mu_2}, \frac{x_4+1}{\mu_4})\big) \qquad (10) \\ & + \ x_3\mu_3 T_1^A(x_0, x_1, x_1', x_3-1, x_4+1) \\ & + \ I_{X_4}\mu_4 T_1^A(x_0, x_1, x_1', x_3, x_4-1)\Big], \ \boldsymbol{x}' \in D' \end{aligned}$$

where $C_2 = \{1/\mu_2 > (x_4+1)/\mu_4\}$ and $C_3 = \{(x_0-1, x_1+1, x_3, x_4) \in D_1\}$. Note that $T_1^A(x_0, x_1, x_3, x_4) = T_1'^A(x_0, x_1, 0, x_3, x_4)$.

Also, similarly to the initial network, $T_3^A(\boldsymbol{x})$ is given by the following equation:

$$\begin{aligned} T_3^A(\boldsymbol{x}) + & \frac{1}{I_{X_0}\lambda + \mu_1 I_{X_1} + (x_3+1)\mu_3 + I_{X_4}\mu_4}\Big[1 \\ & + \ I_{X_0}\lambda T_3^A(x_0-1, x_1+I_{C_1}, x_3+1-I_{C_1}, x_4) \\ & + \ I_{X_1}\mu_1 T_3^A(x_0, x_1-I_{x_1>0}, x_3, x_4+I_{C_2}) \qquad (11) \\ & + \ \frac{\mu_3(x_4+1)}{\mu_4} + x_3\mu_3 T_3^A(x_0, x_1, x_3-1, x_4+1) \\ & + \ I_{X_4}\mu_4 T_3^A(x_0, x_1, x_3, x_4-1)\Big], \ \boldsymbol{x} \in D. \end{aligned}$$

3 Simulation Experiments

3.1 The Method

In this section, we describe our simulation method. Note that it is different from the simulation method of Calvert et al. [3].

We programmed the simulator using Microsoft Visual Studio 2003 (C++ Language), and used Mersenne Twister [13] as a pseudo-random number generator. We ran simulator on several personal computers, each of which has an AMD Athlon64 FX-51 (2.2GHz) CPU, and 2GB memory.

We can obtain dynamic routing decisions in the initial (D_1^I, D_3^I) and augmented (D_1^A, D_3^A) networks by solving (4) and (5) for $\boldsymbol{x} \in D$ (in the augmented network, (10) for $\boldsymbol{x}' \in D'$ and (11) for $\boldsymbol{x} \in D$) recursively, and comparing between $T_1^I(\boldsymbol{x})$ and $T_3^I(\boldsymbol{x})$ (in the augmented network, $T_1^A(\boldsymbol{x})$ and $T_3^A(\boldsymbol{x})$) for $\boldsymbol{x} \in D$, respectively.

We perform the following simulation in both the initial and augmented networks. First, we set the initial state of the network.

$$(x_0(0), x_1(0), x_2(0), x_3(0), x_4(0)) = (N, 0, 0, 0, 0),$$

at time 0. We start each simulation at the initial state, and the simulation continues until the state becomes $(0,0,0,0,0)$. We regard it as one simulation cycle. In all the results shown in next section, we set the parameter value as follows: $N = 100$. We then discard the first 20 jobs since initial states are not always steady-states. We obtain the mean value of transit times of the remaining 80 jobs. We then compute the expectation of the mean transit times obtained by 50000 simulation cycles.

We repeat the above procedure 20 times using random numbers of different seeds, and calculate the 95% confidence interval for each value of arrival rate λ. Note, however, that the lengths of the confidence intervals are so short

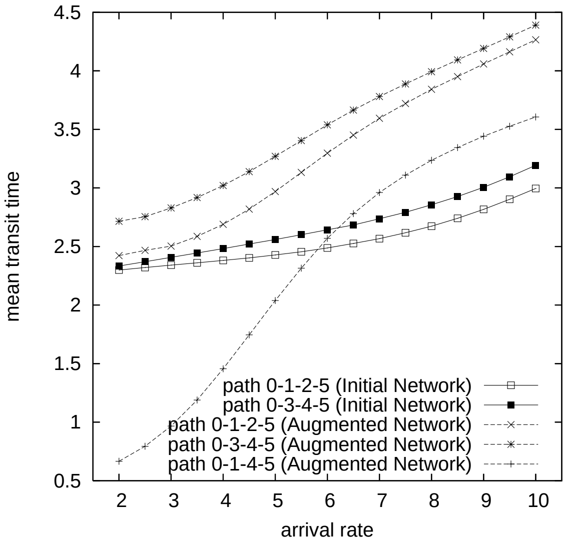

Figure 3. Mean transit time of jobs routed through each path in the initial and augmented networks with dynamic routing for each value of arrival rate λ ($\mu_1 = \mu_4 = 5.0$, and $\mu_2 = \mu_3 = 0.5$).

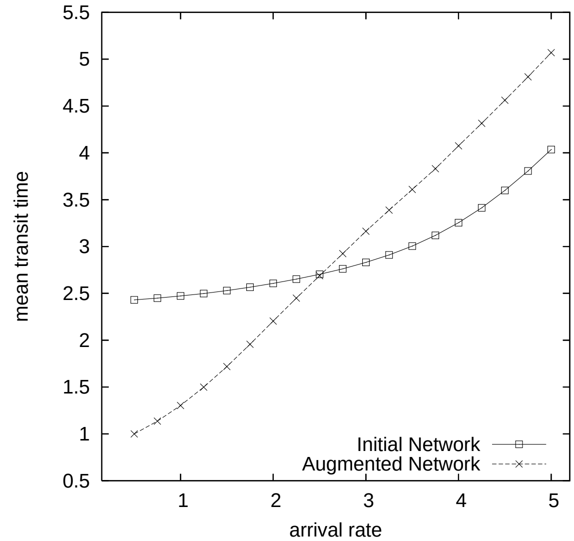

Figure 4. Overall mean transit time of jobs routed through all paths in the initial and augmented networks with dynamic routing for each value of arrival rate λ ($\mu_1 = \mu_4 = 2.5$, and $\mu_2 = \mu_3 = 0.5$).

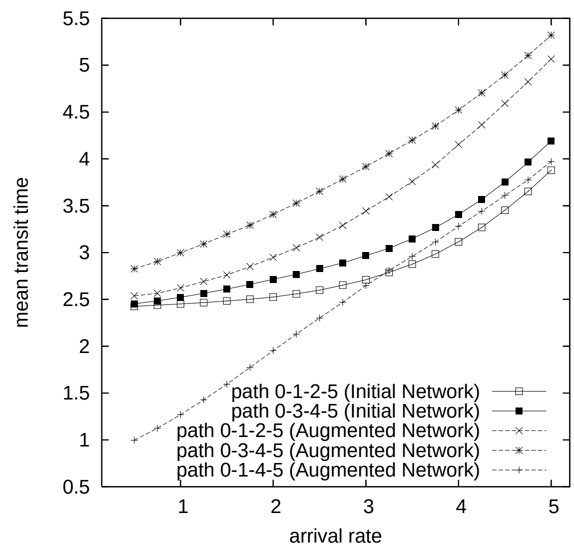

Figure 5. Mean transit time of jobs routed through each path in the initial and augmented networks with dynamic routing for each value of arrival rate λ ($\mu_1 = \mu_4 = 2.5$, and $\mu_2 = \mu_3 = 0.5$).

that the errorbars in the graphs are invisible. Also, note that the capacity of the initial and augmented networks is $(\mu_1 + \mu_4)$.

Because of the recursive nature of the equations giving the expected transit time, the size of memory needed for the computation increases rapidly in the total number of jobs. More precisely, from (10), for a given N, the program stores in memory the different values of T_1^A corresponding to the different cases of $\boldsymbol{x}'$, which is a subset of $\mathbb{N}^5$. We therefore limit the parameter value N to be 100 in our simulations.

3.2 The Results

An example in which the paradox occurs: Figure 3 illustrates the mean transit times of jobs routed through the two paths, 0-1-2-5 and 0-3-4-5 in the initial network, and three paths, 0-1-2-5, 0-3-4-5 and 0-1-4-5 in the augmented network with dynamic routing for each value of arrival rate λ where the parameter values are given as follows: $\mu_1 = \mu_4 = 5.0$, and $\mu_2 = \mu_3 = 0.5$. Note that these networks are symmetric. For static routing, the worst-case paradox has been shown to occur in a case where both the initial and augmented networks are symmetric [8].

We say that "the Braess paradox occurs" if the mean transit time of the jobs routed through *any* path in the augmented network is larger than the mean transit time of the jobs routed through *any* path in the initial network. In Figure 3 we observe that the Braess paradox occurs in the case where the range of the value of the arrival rate is $6.5 \leq \lambda \leq 10$.

Note that if we consider static routing in the initial and augmented networks described in [5], the mean transit time of jobs routed through any path may have the same value. On the other hand, in Figure 3, we observe that the mean transit time of jobs routed through 0-1-4-5 is different from the mean transit times of jobs routed through 0-1-2-5 and 0-3-4-5 in the augmented network. Therefore, the behavior of individual users in dynamic routing may not be similar to that of individual users in static routing.

Discussion of a result by Calvert et al. [3]: Figure 4 illustrates the overall mean transit times of jobs in the initial and augmented networks with dynamic routing for each value of arrival rate λ where the parameter values are given as follows: $\mu_1 = \mu_4 = 2.5$, and $\mu_2 = \mu_3 = 0.5$.

Note that the values of λ and μ_i are the same as the values in [3], and the simulation method is different from in [3]. When λ is larger than 2.5, the overall mean transit time of jobs in the augmented network is larger than the overall mean transit time of jobs in the initial network. This behavior is very similar to the result shown in [3]. Calvert et al. [3] said, "Braess paradox appears for λ greater than the crossover value, i.e. about $\lambda = 2.65$." Figure 5 shows that the mean transit time of jobs routed through each path in the initial and augmented networks with dynamic routing. The parameter values are same as those of Figure 4.

In Figure 5, we observe that the mean transit time of jobs routed through the path 0-1-4-5 in the augmented network is smaller than the mean transit time of jobs routed through the path 0-3-4-5 in the initial network. We think that the Braess paradox occurs when by "adding capacity (a link) to a network degrades the costs for *all* users." Therefore, Figure 4 shows no paradox in our sense.

4 Conclusion

In this paper, we have studied the existence of the Braess paradox in dynamic routing for the Cohen-Kelly network. Based on the analytical study by Calvert et al., we have obtained the dynamic routing decisions in the initial and augmented networks. We have compared the performance of the initial and augmented networks by simulation, and have found a paradoxical phenomenon similar to the one that Cohen-Kelly showed under static routing. We have touched on the results of Calvert et al.

References

[1] N. G. Bean, F. P. Kelly, & P. G. Taylor. Braess' paradox in a loss network. *J. Appl. Prob., 34*, 1997, 155–159.

[2] D. Braess. Uber ein Paradoxien aus der Verkehrsplanung. *Unternehmensforschung, 12*, 1968, 258–268.

[3] B. Calvert, I. Solomon, & W. Ziedins. Braess's paradox in a queueing network with state dependent routing. *J. Appl. Prob., 34*, 1997, 134–154.

[4] J. E. Cohen & C. Jeffries. Congestion resulting from increased capacity in single-server queueing networks. *IEEE/ACM Trans. Netw., 5*(2), 1997, 305–310.

[5] J. E. Cohen & F. P. Kelly. A paradox of congestion in a queueing network. *J. Appl. Prob., 27*, 1990, 730–734.

[6] S. F. El-Zoghdy, H. Kameda, & J. Li. Numerical studies on a paradox for non-cooperative static load balancing in distributed computer systems. *Computer and Operations Research, 33*(2), 2006, 345–355.

[7] A. Inoie, H. Kameda, & C. Touati. A paradox in optimal flow control of M/M/n queues. *Computers and Operations Research, 33*(2), 2006, 356–368. (A shorter version appeared in *Proc. of the IEEE CDC'04*, 2004, 2768–2773).

[8] H. Kameda. How harmful the paradox can be in the Braess/Cohen-Kelly-Jeffries networks. In *IEEE INFOCOM*, 2002, 437–445.

[9] H. Kameda, E. Altman, T. Kozawa, & Y. Hosokawa. Braess-like paradoxes in distributed computer systems. *IEEE Trans. Automatic Contr., 45*, 2000, 1687–1691.

[10] H. Kameda & O. Pourtallier. Paradoxes in distributed decisions on optimal load balancing. *J. ACM, 32*, 2002, 407–433.

[11] Y. A. Korilis, A. A. Lazar, & A Orda. Architecting noncooperative networks. *IEEE J. Sel. Areas Commun., 15*(7), 1995, 1241–1251.

[12] Y. A. Korilis, A. A. Lazar, & A Orda. Avoiding the Braess paradox in non-cooperative networks. *J. Appl. Prob., 36*, 1999, 211–222.

[13] M. Matsumoto & T. Nishimura. Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Trans. Modeling and Comp. Simula., 8*(1), 1998, 3–30.

[14] J. F. Nash. Non-cooperative games. *Ann. Math., 54*, 1951, 286–295.

[15] T. Roughgarden & E. Tardos. How bad is selfish routing. *J. ACM, 49*(2), 2002, 236–259.

[16] J. G. Wardrop. Some theoretical aspects of road traffic research. *Proc. Inst. Civ. Eng., 1*, 1952, 325–378.