TN99-003

Braess-like Paradoxes in Distributed Computer Systems

Hisao Kameda, Eitan Altman, Takayuki Kozawa, and Yoshihisa Hosokawa

H. Kameda is with the Institute of Information Sciences and Electronics, University of Tsukuba, Tsukuba Science City, Ibaraki 305-8573, Japan. Tel: +81-298-53-5539 Fax: +81-298-53-5206 E-mail: kameda@is.tsukuba.ac.jp .

E. Altman is with INRIA B.P. 93, 06902 Sophia Antipolis Cedex, France. Tel: +33-4-92-38-77-86 Fax: +33-4-92-38-77-65 E-mail: Eitan.Altman@sophia.inria.fr

T. Kozawa is with Doctoral Program in Engineering, University of Tsukuba, Tsukuba Science City, Ibaraki 305-8573, Japan. Tel: +81-298-53-5156 Fax: +81-298-53-5206 E-mail: kozawa@osdp.is.tsukuba.ac.jp .

Y. Hosokawa is with Graduate School of Systems and Information Engineering, University of Tsukuba, Tsukuba Science City, Ibaraki 305-8573, Japan. Tel: +81-298-53-5156 Fax: +81-298-53-5206 E-mail: hosokawa@osdp.is.tsukuba.ac.jp .

Abstract

We consider optimal distributed decisions in distributed computer systems. We identify a Braess-like paradox in which adding capacity to the system may degrade the performance of all users. Unlike the original Braess paradox, we show that this behavior occurs only in the case of finitely many users and not in the case of infinite number of users.

Keywords

Braess paradox, Nash equilibrium, Wardrop equilibrium, performance optimization, distributed computer system, load balancing.

I. Introduction

In many systems including communication networks, distributed computer systems, transportation flow networks, etc., we have several distinct objectives for performance optimization. Among them, we have three typical objectives or optima:

(1) the system-optimum, overall optimum, or social optimum, where a certain overall and single measure like the total cost or the overall average response time over all the users is to be optimized. We call it the *overall optimum* here.

(2) the individual optimum, Wardrop equilibrium, or user optimum by some people; where each of infinitely many individuals, users, packets, jobs, or vehicles cannot receive any benefit by changing its own decision. Infinitely many users individually seek their own optimization. It is further assumed that decisions of a single individual have a negligible impact on the performance of other individuals. We call it the *individual optimum* or *Wardrop equilibrium* here.

(3) the class optimum, Nash noncooperative equilibrium, or user optimum by some other people, where each of a finite number of users, classes, or players cannot receive any benefit by changing its decision. Infinitely many individuals, packets, jobs, or vehicles are divided into a finite number ($N(>1)$) of classes or groups each of which is managed by one user or player. A finite number of classes, players, or users seek their own optimization noncooperatively. The decisions of a single class may have a nonnegligible impact on the performance of other classes. We call it the *class optimum* or *Nash equilibrium* here.

Actually, (3) is reduced to (1) when the number of classes reduces to 1 ($N = 1$) and approaches (2) and the number of classes becomes infinitely many ($N \to \infty$) [4].

We can think that the total processing capacity of a system will increase when the capacity of a part of the system increases, and so we expect improvements in performance objectives accordingly in that case. The famous Braess paradox tells us that this is not always the case; i.e., increased capacity of a part of the system may sometimes lead to the degradation in the benefits of all users in a Wardrop equilibrium [1], [2], [3], [4]. We can expect that, in the Nash equilibrium, a similar type of paradox occurs (with large N), i.e., increased capacity of a part of the system may lead to the degradation in the benefits of *all classes* in a *Nash equilibrium*, whenever it occurs for the Wardrop equilibrium ($N \to \infty$). We call it the *Braess-like paradox.* Indeed, Korilis et al. found examples wherein the Braess-like paradox appears in a Nash equilibrium where all user classes are identical in the same topology for which the original Braess paradox (for the Wardrop equilibrium) was in fact obtained [11], [12].

In parallel to these studies, it has been observed that increased capacity of a part of a system may lead to the degradation of the overall performance measure, in a model of distributed computer system [7], [6], [18] in Wardrop and Nash equilibria. We call it the *weaker Braess-like paradox.* Furthermore, Kameda et al. [6] found a seemingly anomalous case where in a Nash equilibrium each of two processing nodes (servers) forwards the same type of jobs mutually to be processed by the other node, thus incurring additional communication delays, although such mutual forwarding should never occur in the overall optimum and even in the Wardrop equilibrium in the same model.

In this paper, however, we present cases where a Braess-like paradox appears in the Nash equilibrium but does not occur in the Wardrop equilibrium in the same environment, in a load balancing problem for distributed computer systems. These cases look quite strange if we note that such a paradox should never occur in the overall optimum and if we regard the Nash equilibrium as an intermediate between the overall optimum and the Wardrop equilibrium. Our model has asymmetric classes; i.e., classes are not identical. In the models considered, the equilibria and the optimum exist and are unique [5].

II. The Model and Assumptions

We consider a model consisting of two nodes (hosts) and a communication means that connects both nodes. Nodes are numbered 1 and 2. Each node consists of a single exponential server with service rate μ_i ($i = 1, 2$). Node i has the external Poisson arrival with rate ϕ_i, out of which the rate x_{ii} of jobs are processed at node i. The rate x_{ij} ($i \neq j$) of jobs are forwarded through the communication means to the other node j to be processed there, and the results of those jobs are returned back through the communication means to node i. Then we have $x_{ii} + x_{ij} = \phi_i$ $(i \neq j)$, $x_{ij} \geq 0, \quad i, j = 1, 2$. We denote the vector $(x_{11},\, x_{12},\, x_{21},\, x_{22})$ by $\boldsymbol{x}$. We denote the set of $\boldsymbol{x}$'s that satisfy the constraints by $\boldsymbol{C}$ and let $\Phi = \phi_1 + \phi_2$. Within these constraints, a set of values of x_{ij} $(i, j = 1, 2)$ are chosen to achieve optimization. Thus the load on node i is $x_{ii} + x_{ji}$ $(i \neq j)$ and is denoted by β_i. Then, the expected processing (including queueing) delay of a job that is processed at node i, is

$$D_i(\beta_i) = \frac{1}{\mu_i - \beta_i} \text{ if } \beta_i < \mu_i, \text{ otherwise it is infinite.}$$

As to the communication means, we consider two alternatives. See Figure 1.

(A) The one is a single-channel communication line that is used commonly in forwarding and sending back jobs that arrive at both nodes. We assume that the expected time length of forwarding and sending back a job is

$$G(\lambda) = \frac{1}{\theta - (x_{12} + x_{21})}$$

if $x_{12} + x_{21} < \theta$, and is otherwise infinite, where $\lambda = x_{12} + x_{21}$ is the network traffic. That is, we assume the communication channel is modeled by a processor sharing server with service rate θ; *i.e.*, the mean communication (without queueing) time is θ^{-1}, and thus the capacity of the communication line is θ.

(B) The other consists of two-way communication lines 1 and 2. One two-way line i is used for forwarding of a job that arrives at node i (and for sending back the processed result of the job). The assumption on the line is the same as (A) except that there are two lines each of which is used only for jobs arriving at one node and is not used in common by two nodes. Thus the expected communication (with queueing) delay of a job arriving at node i and being processed at node j $(\neq i)$ is expressed as

$$G_i(x_{ij}) = \frac{1}{\theta - x_{ij}}$$

if $x_{ij} < \theta$ (and is otherwise infinite). Clearly, the communication capacity for case (B) is greater than or equal to that for case (A).

We refer to the length of time between the instant when a job arrives at a node and the instant when a job leaves the node, where it has been processed, after all processing and communication, if any, are over, as *the response time* of the job.

Thus the expected response time of a job that arrives at node i is

$$\begin{aligned} T_i(\boldsymbol{x}) &= \frac{1}{\phi_i}\sum_k x_{ik}T_{ik}(\boldsymbol{x}), \\ \text{where} \qquad T_{ii}(\boldsymbol{x}) &= \frac{1}{\mu_i-\beta_i} \\ \text{and for} \quad j\neq i, \qquad T_{ij}(\boldsymbol{x}) &= \frac{1}{\mu_j-\beta_j}+\frac{1}{\theta-(x_{ij}+x_{ji})}, \qquad \text{for case (A)}, \\ &= \frac{1}{\mu_j-\beta_j}+\frac{1}{\theta-x_{ij}}, \qquad \text{for case (B)}. \end{aligned}$$

(The above expressions hold, again, only for positive values of denominators, and are otherwise infinite.)

Then, the overall expected response time of a job that arrives at the system is

$$T(\boldsymbol{x}) = \frac{1}{\Phi}\sum_i \phi_i T_i.$$

We have three optima, the overall, the individual, and the class.

(1) The overall optimum is given by such $\bar{\boldsymbol{x}}$ as satisfies the following,

$$T(\bar{\boldsymbol{x}}) = \min T(\boldsymbol{x}) \quad \text{with respect to} \quad \boldsymbol{x}\in\boldsymbol{C}.$$

(2) The individual optimum is given by such $\hat{\boldsymbol{x}}$ as satisfies the following for all i,

$$T_i(\hat{\boldsymbol{x}}) = \min\{T_{ii}(\hat{\boldsymbol{x}}), T_{ij}(\hat{\boldsymbol{x}})\}\ (i\neq j) \quad \text{such that} \quad \hat{\boldsymbol{x}}\in\boldsymbol{C}$$

(3) The class optimum is given by such $\tilde{\boldsymbol{x}}$ as satisfies the following for all i,

$$\begin{aligned} T_i(\tilde{\boldsymbol{x}}) &= \min T_i(x_{ii}, x_{ij}, \tilde{x}_{ji}, \tilde{x}_{jj}) \quad \text{with respect to } x_{ii},\ x_{ij}\ (j\neq i), \\ &\quad \text{such that } (x_{ii}, x_{ij}, \tilde{x}_{ji}, \tilde{x}_{jj})\in\boldsymbol{C}. \end{aligned}$$

We are sure of the existence and the uniqueness of the overall, individual, and class optima for the model given here (provided that the natural stability conditions hold, i.e., there are enough resources in the system to process all loads; if they do not hold then the costs are infinite). For the existence and uniqueness of those optima see [5].

Remark II.1: Note that there should be no mutual forwarding in overall and individual optima. That is, in overall and individual optima, either one of x_{ij} $(i\neq j)$ must be zero in case (A) due to [17] or to Section 2.2.2 of [7] and in case (B) due to [13]. And thus, when one of x_{ij} $(i\neq j)$, say x_{ij}, is non zero, $T_i(\boldsymbol{x})$ decreases and $T_j(\boldsymbol{x})$ increases with the increase of θ as shown in Theorems 2.5 and 2.7 of [7].

III. The numerical experiments

We examined the cases of the following parameter values of the model given in Section II.

$$\phi_1 = 110(\text{jobs/sec}), \qquad \phi_2 = 20(\text{jobs/sec})$$
$$\mu_1 = 120(\text{jobs/sec}), \qquad \mu_2 = 25(\text{jobs/sec})$$

The value of the communication (without queueing) time parameter θ^{-1} is varied from 0 (sec) till 1 (sec) in steps of 0.005 (sec), i.e., $\theta^{-1} = 0,\ 0.005,\ 0.010,\ 0.015,\ \cdots,\ 1$.

The algorithms used to obtain the overall and individual optima are based on the algorithms given in [9], [10]. The algorithm for the class optimum is obtained similarly as above.

IV. Results and Discussion

Figure 2 shows the values of class optimal T_1 (Figure 2 left) and T_2 (Figure 2 right) with various values of θ^{-1} for the cases (A) and (B). We can observe the Braess-like paradoxes for class optima in three ways.

(I) For case (A) and (II) for case (B) both with the increase of the line capacity θ (i.e., with the decrease in the mean communication time θ^{-1}).

(III) With the increase of the line capacity from case (A) to case (B).

We can see the situations where both T_1 and T_2 increase with the increase of the communication line capacity as in (I), (II) and (III).

Paradox (I) for case (A)

In Figure 2, for $0.1 \leq \theta^{-1} \leq 0.8$, we see that there are sets of two values of θ^{-1} for which T_1 and T_2 with the smaller value of θ^{-1} are greater than T_1 and T_2 with the larger value of θ^{-1}, which looks paradoxical. For example, consider the case with $\theta = 1.25$ ($\theta^{-1} = 0.8$). Then we have

$$T_{11} = 0.0993..., \quad T_{12} = 1.0989..., \quad T_{21} = 0.9956..., \quad T_{22} = 0.2027...,$$
$$x_{11} = 109.899..., \quad x_{12} = 0.1004..., \quad x_{21} = 0.0337..., \quad x_{22} = 19.966...$$
$$T_1 = 0.10025..., \quad T_2 = 0.20404...,$$

by noting that $T_1 = (x_{11}T_{11} + x_{12}T_{12})/\phi_1, \quad T_2 = (x_{21}T_{21} + x_{22}T_{22})/\phi_2,$

Furthermore, consider the case with $\theta = 10$ ($\theta^{-1} = 0.1$). Then

$$T_{11} = 0.09879..., \quad T_{12} = 0.5513..., \quad T_{21} = 0.4451..., \quad T_{22} = 0.20498...,$$
$$x_{11} = 106.382..., \quad x_{12} = 3.617..., \quad x_{21} = 3.495..., \quad x_{22} = 16.504....$$
$$T_1 = 0.11367..., \quad T_2 = 0.24696...$$

Therefore, we have

$$T_1 = 0.10025..., \quad T_2 = 0.20404..., \quad \text{for} \quad \theta = 1.25, \quad \text{and}$$
$$T_1 = 0.11367..., \quad T_2 = 0.24696..., \quad \text{for} \quad \theta = 10,$$

and we see that in the case where the line capacity θ is increased from 1.25 to 10, both T_1 and T_2 increase, although at least one of them is expected to decrease, which is a Braess-like paradox.

Paradox (II) for case (B)

Similarly as above, in Figure 2, for $0.1 \leq \theta^{-1} \leq 0.8$, we see that there are sets of two values of θ^{-1} for which T_1 and T_2 with the smaller value of θ^{-1} are greater than T_1 and T_2 with the larger value of θ^{-1}, which looks again paradoxical. For example, consider the case with $\theta = 1.25$ ($\theta^{-1} = 0.8$). Then

$$T_{11} = 0.0995..., \quad T_{12} = 1.088..., \quad T_{21} = 0.9537..., \quad T_{22} = 0.2017...,$$
$$x_{11} = 109.87, \quad x_{12} = 0.1219..., \quad x_{21} = 0.0792..., \quad x_{22} = 19.920....$$
$$T_1 = 0.1006..., \quad T_2 = 0.2047...$$

Furthermore, consider the case with $\theta = 10$ ($\theta^{-1} = 0.1$). Then

$$T_{11} = 0.0990..., \quad T_{12} = 0.4608..., \quad T_{21} = 0.3496..., \quad T_{22} = 0.2039...,$$
$$x_{11} = 103.892..., \quad x_{12} = 6.107..., \quad x_{21} = 6.009..., \quad x_{22} = 13.990....$$
$$T_1 = 0.1191..., \quad T_2 = 0.2477...$$

Therefore, we have

$$T_1 = 0.1006..., \quad T_2 = 0.2047..., \quad \text{for} \quad \theta = 1.25, \quad \text{and}$$
$$T_1 = 0.1191..., \quad T_2 = 0.2477..., \quad \text{for} \quad \theta = 10,$$

and we see that in the case where the line capacity θ is increased from 1.25 to 10, both T_1 and T_2 increase, although at least one of them is expected to decrease, which is again a Braess-like paradox.

Paradox (III) between cases (A) and (B)

In Figure 2, we can see for $0.1 \leq \theta^{-1} \leq 0.8$, both T_1 and T_2 of case (A) are smaller than T_1 and T_2 of case (B), respectively.

Clearly, this looks paradoxical. That is, although the system in case (B) has a larger communication capacity than in case (A), the expected response times in case (B) for both nodes are larger than the corresponding values in case (A) at the same time. We show the details of some case as an example. For example, consider the case where $\theta = 10$ ($\theta^{-1} = 0.1$). As given in (I) and (II) above, we have

$$\begin{aligned} \text{For (A)}: \quad & T_1 = 0.11367..., \quad T_2 = 0.24696..., \\ \text{For (B)}: \quad & T_1 = 0.1191..., \quad T_2 = 0.2477.... \end{aligned}$$

Note again that the communication line capacity for each node in case (B) is greater than or equal to that for the node in case (A). Thus, we see that while the communication of case (B) is greater than that of case (A), both T_1 and T_2 of case (B) are greater than those of case (A), although at least one of them is expected to decrease, which is again a Braess-like paradox.

We have many other sets of parameter values for which Braess-like paradoxes appear, but they are not shown here due to limitation of space.

Remark IV.1: As we saw in the last few statements given in Section II, for individual optima, such a Braess-like paradox as given here would not occur. Figure 3 shows the values of individually optimal T_1 and T_2 with various values of θ^{-1}. These values are identical for both cases (A) and (B); and we can see that T_1 and T_2 for a value of θ^{-1} cannot be greater than for another value of θ^{-1} at the same time, which is not paradoxical.

For reference, we present relevant data in Figure 4. Figure 4 (left) shows the overall mean response time with various values of the mean communication time parameter θ^{-1} for overall, class, and individual optima for cases (A). Note that the overall and individual optima must be identical for cases (A) and (B) as we recall what we noted at the end of the previous section. Only class optima for cases (A) and (B) may be different from each other, which is shown in Figure 4 (right). We see that as the line capacity θ increases from 10 (θ^{-1} decreases from 0.1) the overall mean response time of the individual optimum increases, which is a weaker paradox and is not a real Braess-like paradox.

V. Concluding Remarks

In this paper, we have shown the existence of some paradoxical and awkward behaviors of the Nash equilibrium in a model of load balancing for distributed computer systems. We have seen that the Nash noncooperative decision may sometimes lead to the degradation of performance for every member. Adding capacity to direct links between a source and a destination cannot cause Braess-type paradox, as has been proved in Proposition 4.3 in ([12]). Our setting shows that it suffices that one of the ends of a link, to which capacity is added, is not the source or destination, that the Braess paradox can occur.

Braess type paradoxes have been known in the context of road-traffic and of communication networks. We have shown that it might also occur in the context of load balancing in distributed computing. Our example seems to be simpler than the previously known paradoxes in that it has less nodes and links. In the original Braess paradox, the delay of all users increased when capacity was added to a link which was *interior*, in the network. In some sense, the paradoxical behavior seems more astonishing in our context, since the capacity is added to links *directly connected to the sources.*

Such a paradoxical behavior does not occur for the overall and Wardrop optimum in the same setting of the model. That may imply that the Nash equilibrium may have more complicated characteristics than the overall optimum and the Wardrop equilibrium.

ACKNOWLEDGMENTS

We would like to thank Professor Colman Altman for suggestions to improve the presentation of this paper. We also appreciate constructive comments made by anonymous referees.

This research has been supported in part by Japan Society for Promotion of Science and by Telecommunications Advancement Organization of Japan.

REFERENCES

[1] J. E. Cohen and F. P. Kelly, "A Paradox of Congestion in a Queuing Network," *J. Appl. Prob. 27*, pp.730-734, 1990.

[2] J. E. Cohen and C. Jeffries, "Congestion Resulting from Increased Capacity in Single-Server Queueing Networks," *IEEE/ACM Trans. on Networking 5*, 2, April pp.1220-1225, 1997.

[3] B. Calvert, W. Solomon and I. Ziedins, "Braess's Paradox in a Queueing Network with State-Dependent Routing," *J. Appl. Prob. 34,* pp.134-154, 1997.

[4] A. Haurie and P. Marcotte, "On the Relationship between Nash-Cournot and Wardrop Equilibria," *Networks 15*, 1985, pp.295-308.

[5] H. Kameda, E. Altman and T. Kozawa, "Braess-like Paradoxes of Nash Equilibria for Load Balancing in Distributed Computer Systems," ISE Technical Report ISE-TR-98-157, Institute of Information Sciences and Electronics, University of Tsukuba, 1998.

[6] H. Kameda, T. Kozawa and J. Li, "Anomalous Relations among Various Performance Objectives in Distributed Computer Systems," *Proc. 1st World Congress on Systems Simulation*, IEEE, 1997, pp.459-465.

[7] H. Kameda, J. Li, C. Kim and Y. Zhang, *Optimal Load Balancing in Distributed Computer Systems*, Springer, 1997.

[8] H. Kameda and Y. Zhang, "Uniqueness of the Solution for Optimal Static Routing in Open BCMP Queueing Networks," *Mathematical and Computer Modelling 22*, 10-12, pp.119-130, 1995.

[9] C. Kim and H. Kameda, "An Algorithm for Optimal Static Load Balancing in Distributed Computer Systems," *IEEE Trans. Comput. 41*, 3, 1990, pp.381-384.

[10] C. Kim and H. Kameda, "Optimal Static Load Balancing of Multi-class Jobs in a Distributed Computer System," *Proc. 10th Intl. Conf. Distributed Comput. Syst.*, IEEE, 1990, pp.562-569.

[11] Y. A. Korilis, A. A. Lazar and A. Orda, "Architecting Noncooperative networks", *IEEE Journal on Selected Areas in Communications 13*, 7, September 1995, pp.1241-1251. *IEEE Trans. Automatic Control 42*, 3, March 1997, pp.309-871.

[12] Y. A. Korilis, A. A. Lazar and A. Orda, "Avoiding the Braess Paradox in Noncooperative Networks", *Proc. IEEE Conference on Decision & Control (San Diego)*, 1997, pp.864-878.

[13] J. Li and H. Kameda, "Load Balancing Problems for Multiclass Jobs in Distributed/Parallel Computer Systems," *IEEE Trans. Comput. 47*, 3, 1998, pp.322-1998.

[14] A. Orda, R. Rom and N. Shimkin, "Competitive Routing in Multiuser Communication Networks," *IEEE/ACM Trans. on Networking 1*, 1993, pp.614-627.

[15] J. B. Rosen, "Existence and Uniqueness of Equilibrium Points for Concave N-person Games", *Econometrica 33*, pp. 153-163, 1965.

[16] J. F. Shapiro, *Mathematical Programming, Structures and Algorithms*, J. Wiley and Sons, 1979.

[17] A. N. Tantawi and D. Towsley, "Optimal Static Load Balancing in Distributed Computer Systems," *J. ACM 32*, 2, pp.445-465, Apr, 1985.

[18] Y. Zhang, H. Kameda and K. Shimizu, "Parametric Analysis of Optimal Load Balancing in Distributed Computer Systems," *Journal of Information Processing 14*, 4, 1992, pp.433-441.

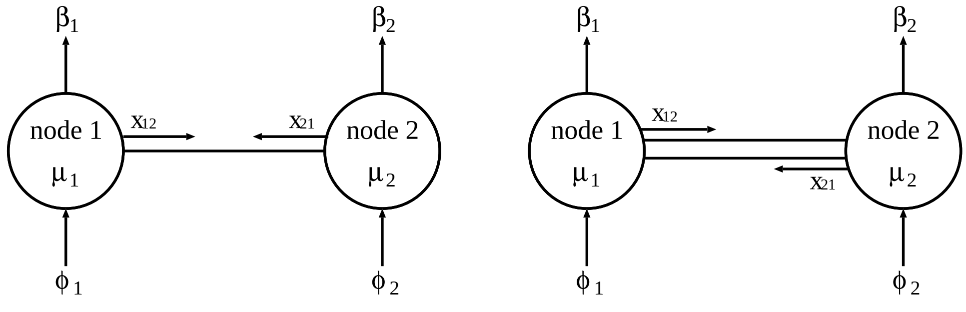
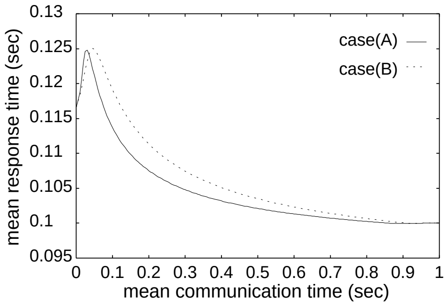


Fig. 1. The system model, case (A) (**left**) and case (B) (**right**).



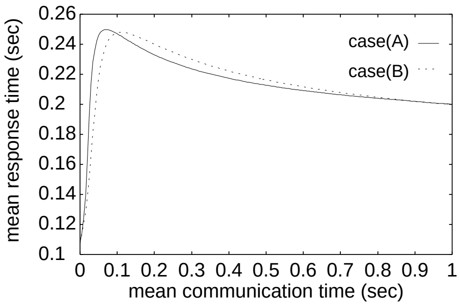


Fig. 2. **Left**: The mean response time T_1 of a job that arrives at node 1 in class optima (cases (A) and (B)) for the various values of the mean communication time θ^{-1}. **Right**: The mean response time T_2 of a job that arrives at node 2 in class optima (cases (A) and (B)) for the various values of the mean communication time θ^{-1}.

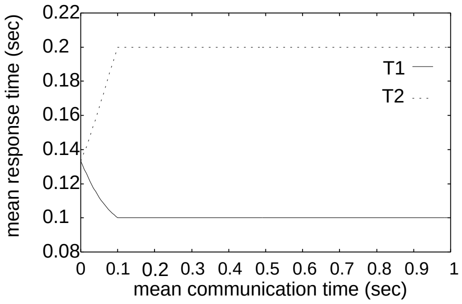
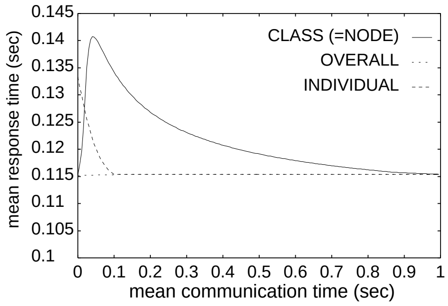


Fig. 3. The mean response times, T_1 and T_2, of a job that arrives at nodes 1 and 2 with various values of θ^{-1} in individual optima, respectively. They have the same values for cases (A) and (B).



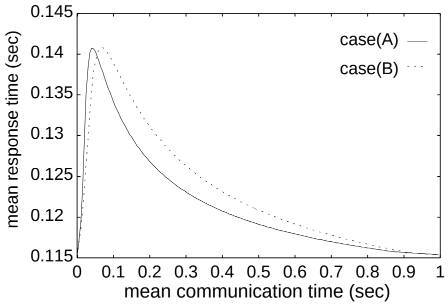


Fig. 4. **Left**: The overall mean response time in overall, individual, and class optima (case (A)) for the various values of the mean communication time θ^{-1}. The values are the same for overall and individual optima (case (B)). **Right**: The overall mean response time in class optima (case (A) and (B)) for the various values of the mean communication time θ^{-1}.

List of Figures