Paradoxes in Distributed Decisions on Optimal Load Balancing for Networks of Homogeneous Computers

Hisao Kameda* and Odile Pourtallier†‡

Abstract

In completely symmetric systems that have homogeneous nodes (hosts, computers, or processors) with identical arrival processes, an optimal static load balancing scheme does not involve the forwarding of jobs among nodes. Using an appropriate analytic model of a distributed computer system, we examine the following three decision schemes for load balancing: completely distributed, intermediately distributed, and completely centralized. We show that there is no forwarding of jobs in the completely centralized and completely distributed schemes, but that in an intermediately distributed decision scheme, mutual forwarding of jobs among nodes is possible, leading to degradation in system performance for every decision maker. This result appears paradoxical, because by adding communication capacity to the system for the sharing of jobs between nodes, the overall system performance is degraded. We characterize conditions under which such paradoxical behavior occurs, and we give examples in which the degradation of performance may increase without bound. We show that the degradation reduces and finally disappears in the limit as the intermediately distributed decision scheme tends to a completely distributed one.

keywords Braess paradox, distributed decision, homogeneous distributed system, load balancing, Nash equilibrium, non-cooperative game, performance optimization, Wardrop equilibrium.

1 Introduction

We consider systems that consist of a finite number of facilities and of threads or flows of customers that arrive from infinite sources and are served by the facilities. For example, distributed

*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
†O. Pourtallier is with INRIA B.P. 93, 06902 Sophia Antipolis Cedex, France. Tel: +33/0-492-387-826 Fax: +33/0-492-387-858 E-mail: Odile.Pourtallier@sophia.inria.fr
‡The present work was mostly done during her visit to the University of Tsukuba. This paper is to appear in Journal of the ACM. Copyright © ACM, Inc. Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Publications Dept, ACM Inc., fax +1 (212) 869-0481, or permissions@acm.org.
computer systems have continuing arrivals of jobs or transactions to be processed by computers, communication networks have flows of packets or calls to pass through communication links, transportation flow networks have incoming threads of vehicles to drive through roads, etc. In the following, we use the term system representing the set of a finite number of facilities such as distributed computer systems, communication networks, or transportation flow networks, and the term job representing each of continuously arriving customers such as jobs, packets, vehicles, etc. Thus, for a period of infinite length, infinitely many jobs pass through the system, and each job alone can have only an infinitesimal impact on the overall system performance.

We may have various issues on optimization, depending on the degree of distribution of the decisions. Among them, we have three typical optima corresponding to three typical decision schemes:

(A) [Completely centralized decision scheme]: All jobs are regarded to belong to one group that has only one decision maker. The decision maker seeks to optimize a single performance measure such as the total cost over all jobs (e.g., the mean response time of the entire system). In the literature, the corresponding solution concept is referred to as a system optimum, overall optimum, cooperative optimum or social optimum. In this paper, we shall refer to it as the overall optimum. This may reflect the situation where the entire system is used by a single unified organization.

(B) [Completely distributed decision scheme]: Each of infinitely many jobs (or the user of each) optimizes its own cost (e.g., its own expected response time), independently of the others. In this optimizing situation, each job cannot expect any further benefit by changing its own decision. It is also assumed that the decision of a single job has a negligible impact on the performance of the entire system. In the literature, the corresponding solution concept is referred to as an individual optimum, Wardrop equilibrium, or user optimum. In this paper, we shall refer to it as the individual optimum.

(C) [Intermediately distributed decision scheme]: Infinitely many jobs are classified into a finite number ($N > 1$) of classes or groups, each of which has its own decision maker and is regarded as one player or user. Each decision maker optimizes non-cooperatively its own cost (e.g., the expected response time) over only the jobs of its own class. The decision of a single decision maker of a class has a non-negligible impact on the performance of other classes. In this optimizing situation, each of a finite number of classes or players cannot receive any further benefit by changing its decision. In the literature, the corresponding solution concept is referred to as a class optimum, Nash equilibrium, or user optimum. In this paper, we shall refer to it as the class optimum. This may reflect the situation where the system is shared by a finite number of mutually independent organizations each of which is totally unified. We may have different levels in intermediately distributed optimization.

Note that (C) is reduced to (A) when the number of players reduces to 1 ($N = 1$) and approaches (B) when the number of players becomes infinitely many ($N \to \infty$) (Haurie and Marcotte, 1985).

Intuitively, 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., adding capacity to the system may sometimes lead to the degradation in the benefits of all users in an individual optimum (Braess, 1968; Murchland, 1970; Frank, 1981; Cohen and Kelly, 1990; Calvert et al., 1997; Cohen and Jeffries, 1997). The Braess paradox has attracted the attention of many scholars, such as the economist Paul Samuelson (1992).
It has also attracted the attention of researchers in the field of software multi-agent systems (see Arora and Sen (1997) for example) and in the theory of computing (see Roughgarden and Tardos (2000) for example).

We can expect that, in a class optimum (i.e., Nash equilibrium) a similar type of paradox may occur (with large N), whenever it occurs for the individual optimum (N → ∞). Indeed, Korilis et al. (1995) found examples wherein the Braess-like paradox appears in a class optimum where all user classes are identical in the same topology for which the original Braess paradox (for the individual optimum) was in fact obtained. Furthermore, Korilis et al. (1999) obtained a sufficient condition under which the Braess paradox should not occur in the more general model that has one source-destination pair and identical user classes.

As it is known that the class optimum converges to the individual optimum as the number of classes becomes large (Haurie and Marcotte, 1985), it is natural to expect the same type of paradox in the class optimum context (for a large number of classes), whenever it occurs for the individual optimum, although it never occurs in the overall optimum where the total cost is minimized.

Kameda et al. (2000) have obtained, however, numerical examples where a paradox similar to Braess’s appears in the class optimum but does not occur in the individual optimum in the same environment. These cases look quite strange if we note that such a paradox should never occur in the overall optimum and if we regard the class optimum as an intermediate between the overall optimum and the individual optimum. In models of distributed computer systems with class and individual optimization, Kameda et al. (1997b; 1997a) have already observed such (weaker) paradoxical cases that increased capacity of a part of a system may lead to increase in an overall measure such as the system-wide overall mean response time. (As to using overall measures in evaluating class optima, see also Koutsoupias and Papadimitriou (1999) for example.) The methods and algorithms for obtaining the optima and the equilibria are described by Kameda et al. (1997a; 2001b), Kim and Kameda (1990), Li and Kameda (1998), and Tantawi and Towsley (1985). Some related results on class optima are given by e.g., Orda et al. (1993).

Static load balancing in distributed computer systems means the balancing of loads, i.e., jobs, over nodes, on the basis of statistically averaged performance values of nodes and interconnection networks. For formulation and analytic treatment of static load balancing in distributed computer systems, see, e.g., Tantawi and Towsley (1985). Its counterpart exists for packet routing in communication networks like the Internet. On the other hand, dynamic load balancing means the balancing of loads based on the information on performance values of nodes etc. at each instant of time. Apparently, dynamic load balancing would seem to outperform static load balancing, but would not easily be realizable, because it would require large overheads. For example, in the Internet and on-line transaction processing systems, respectively, many packets and small transactions move around within the network and systems. Therefore, dynamic load balancing would require frequent transfer of information from nodes etc. in order to react to ceaseless changes in their performance values that result from the moves of many packets and transactions. The frequent transfer of information would lead to large overheads, and then would result in overall performance degradation. Thus, load balancing based on the performance values of nodes etc., averaged over each time period of a certain length, may be reasonable. This is quasi-static load balancing. If the system is in a statistical equilibrium for a very long time period, quasi-static load balancing may be close to static load balancing. For static vs. dynamic load balancing, the reader is referred to Kameda et al. (1997a) for example. Note that the expressions of ‘static’ and ‘dynamic’ in computer networking, however, have
somewhat different meanings from those in load balancing in distributed computer systems.

In this paper, we present an analytic study of a model of static load balancing among homogeneous nodes each of which has an identical arrival process. It would be naturally anticipated that optimal static load balancing would lead to no forwarding of jobs among nodes and thus would bring neither benefit nor harm. We can confirm that this is actually true both in the completely centralized decision scheme and in the completely distributed decision scheme of optimal load balancing. In an intermediated distributed decision scheme for load balancing, however, we can find paradoxical cases where adding communication capacity among nodes to the system leads to mutual forwarding of jobs among nodes and thus brings the degradation of system performance for every decision maker. We refer to a system where the nodes are constituted by the same kind of equipment (e.g., servers with the same characteristics), each of which has an identical arrival process as a system of symmetrical nodes. We investigate a system of symmetrical nodes here because Kameda et al. (2001b; 2001a) have observed that in numerical and analytical investigations on systems of asymmetrical nodes, such counter-intuitive phenomena appear most strongly in systems of symmetrical nodes.

In the model studied in this paper, each node (or processor) has at its disposition a communication means, which it may use to forward to other nodes an arbitrary portion of its job arrival stream. We show that in the overall and individual optima, there is no forwarding of jobs between nodes, but that in the class optimum, forwarding of jobs may or may not occur, depending on the values of certain system parameters. We thus characterize conditions under which the paradoxical behavior appears. As a measure of performance degradation, we consider the ratio of the mean response time for a decision maker when the communication means is available, to that when the means is not available at all. The situation where the ratio is not less than 1 for all decision makers and is greater than 1 for some decision makers, is considered paradoxical. In some class optimization where the values of parameters of all classes are identical, we see the existence of the paradoxical cases where the ratio of the performance degradation can increase without bound, and also see that the ratio of performance degradation decreases and finally disappears in the limit as the number of classes tends to infinity. We also obtain the solution of the class optimum when the paradox occurs. The analysis is performed under general assumptions on nodes and the interconnection network except for the symmetricity of nodes.

The network originally studied by Braess (1968) and the system we analyzed are both regarded as concrete examples of paradoxical cost degradation for all decision makers and, in more general terms, of Pareto inefficiency in non-cooperative games. We present here some differences, between these two paradoxes, other than what has been stated above, i.e., the point that our paradox is observed only for class optimization but not for individual optimization. It has been shown that in the Braess network and in extended Braess networks (Cohen and Kelly, 1990; Cohen and Jeffries, 1997), the ratio of the performance degradation is bounded and less than 2 (see Kameda (2002)). (To the best of our knowledge, there has not been reported any case where the ratio of the performance degradation can increase without bound.) This is contrastive to the existence of paradoxical cases where the ratio of the performance degradation can increase without bound in the system we studied. Furthermore, in the system of symmetrical nodes studied in this paper, adding means of job forwarding looks apparently ineffective because the load on each node is already equalized without job forwarding. Thus, it is, in particular, counter-intuitive, that, in some class optimum, adding the means causes mutual job forwarding among nodes and thus brings about the strongest paradox, in the case of symmetrical nodes. This is contrastive to the Braess network in which adding capacity itself does not
look ineffective and may benefit all users even in some symmetrical cases.

The model considered in this paper may also be regarded as a model of communication networks. For example, assume the following situation. There are two continents, \( A \) and \( E \). In continent \( A \), there are port cities \( A_1, A_2, \ldots, A_m \), each of which is connected by a separate optical cable to the corresponding one of port cities \( E_1, E_2, \ldots, E_m \) in continent \( E \). Cable \( i \) connects cities \( A_i \) and \( E_i \), and is managed by an independent organization \( C_i \). The cables are of almost equal length. Consider the demand of sending packets from continent \( A \) to \( E \). Almost the same rate of demand of sending packets comes to each organization \( C_i \). Within each continent, there is a communication means which every organization, say \( C_i \), may use to forward to another organization, say, \( C_j \) \((j \neq i)\) an arbitrary portion of its demand and send it back. For example, the section of \( C_i \) in city \( A_i \) sends a portion of its demand from \( A_i \) to \( A_j \) through the communication means in continent \( A \), which is sent through cable \( i \) to city \( E_j \) and finally back to the section of \( C_i \) in city \( E_i \) through the communication means in continent \( E \). This is regarded as a kind of bypath, which may be effective in the case where the demand to each organization \( C_i \) highly fluctuates and/or each cable may breakdown at times. Our results imply that if each organization would strive to optimize the mean packet transmission time only for its demand, paradoxical cost degradation for all organizations might occur.

Section 2 presents the description of the model studied in this paper and the assumptions used in the analysis. Section 3 gives the solutions of overall, individual, and two class optima, and their proofs. Section 4 gives examples of solutions. Section 5 gives numerical examples. Section 6 gives concluding remarks. The Appendix shows a glossary of symbols used in this paper.

## 2 The Model and Assumptions

We consider a system with \( m \) nodes (host computers or processors) connected with a communication means. Jobs that arrive at each node \( i, i = 1, 2, \ldots, m \), are classified into \( n \) types \( k \), \( k = 1, 2, \ldots, n \). Consequently, we have \( mn \) different job classes \( R_{ik} \). Each class \( R_{ik} \) is distinguished by the node \( i \) at which its jobs arrive and by the type \( k \) of the jobs. We call such a class local class, or simply class.

We assume that each node has an identical arrival process and identical processing capacity. Jobs of type \( k \) arrive at each node with node-independent rate \( \phi_k \). We denote the total arrival rate to the node by \( \phi (= \sum_k \phi_k) \), and without loss of generality, we assume a time scale such that \( \phi = 1 \).

We also consider what we call global class \( J_k \) that consists of the collection of local class \( R_{ik} \), i.e., \( J_k = \cup_i R_{ik} \). \( J_k \) thus consists of all jobs of type \( k \). Whereas, for local class \( R_{ik} \), all the jobs arrive at the same node \( i \), the arrivals of the jobs of global class \( J_k \) are equally distributed over all nodes \( i \).

The average processing (service) time (without queueing delays) of a type-\( k \) job at any node is \( 1/\mu_k \) and is, in particular, node-independent. We denote \( \phi_k/\mu_k \) by \( \rho_k \) and \( \rho = \sum_k \rho_k \).

Out of type-\( k \) jobs arriving at node \( i \), the ratio \( x_{ijk} \) of jobs is forwarded upon arrival through the communication means to another node \( j \) \((\neq i)\) to be processed there. The remaining ratio \( x_{iik} = 1 - \sum_{j(\neq i)} x_{ijk} \) is processed at node \( i \). Thus \( \sum_j x_{ijk} = 1 \). That is, the rate \( \phi_k x_{ijk} \) of type-\( k \) jobs that arrive at node \( i \) is forwarded through the communication means to node \( j \), while the rate \( \phi_k x_{iik} \) of local-class \( R_{ik} \) jobs is processed at the arrival node \( i \). We have \( 0 \leq x_{ijk} \leq 1 \), for
all $i$, $j$, $k$. Within these constraints, a set of values for $x_{ik}$ ($i = 1, 2, \ldots, m, k = 1, 2, \ldots, n$) are chosen to achieve optimization, where

$$x_{ik} = (x_{i1k}, \ldots, x_{imk})$$

is an $m$-dimensional vector and called ‘local-class $R_{ik}$ strategy’. We define a global-class $J_k$ strategy as the $mn$-dimensional vector

$$x_k = (x_{1k}, x_{2k}, \ldots, x_{mk}).$$

We will also denote, by an $mmn$-dimensional vector $x$, the vector of strategies concerning all local classes,

$$x = (x_{11}, x_{12}, \ldots, x_{1n}, x_{21}, \ldots, x_{2n}, \ldots, x_{m1}, \ldots, x_{mn}),$$

or $x = (x_1, x_2, \ldots, x_n)$.

We call $x$ the strategy profile.

For a strategy profile $x$, the load $\beta_i$ on node $i$ is

$$\beta_i = \beta_i(x) = \sum_{j,k} \rho_k x_{jik}. \quad (1)$$

The contribution $\beta_i^{(k)}$ on the load of node $i$ by type-$k$ jobs is

$$\beta_i^{(k)} = \beta_i^{(k)}(x) = \sum_j \rho_k x_{jik}, \quad (2)$$

and clearly $\beta_i = \beta_i^{(1)} + \beta_i^{(2)} + \cdots + \beta_i^{(n)}$.

Clearly from (1)

$$\sum_i \beta_i = \sum_{i,p,q} \rho_k x_{jik} = \sum_k \rho_k \sum_i \sum_j x_{jik} = \sum_k \rho_k m = mp. \quad (3)$$

We denote the set of $x$‘s that satisfy the constraints (i.e., $\sum_i x_{jik} = 1, x_{jik} \geq 0$, for all $i, j, k$) by $C$. Note that $C$ is a compact set.

We have the following assumptions:

**Assumption Π1** We assume that the expected processing (including queueing) time of a type-$k$ job that is processed at node $i$ (or the cost function at node $i$), is a strictly increasing, convex and continuously differentiable function of $\beta_i$, denoted by $\mu_k^{-1} D(\beta_i)$ for all $i, k$.

**Assumption Π2** We assume that the mean communication delay (including queueing delay) or the cost for forwarding type-$k$ jobs arriving at node $i$ to node $j$ ($i \neq j$), denoted by $G_{ijk}(x)$, is a positive, nondecreasing, convex and continuously differentiable function of $x$. We assume that $G_{ijk}(x) = 0$. We assume further that each job is forwarded at most once.

**Example 1** We may consider the following simple functions for mean processing time and mean communication delay. For the mean processing time:

$$1/\mu_k D(\beta_i) = \frac{1/\mu_k}{1 - \beta_i} \quad \text{for } \beta_i < 1, \text{ otherwise it is infinite.} \quad (4)$$
For the mean communication delay:

\[ G_{ijk}(\mathbf{x}) = t. \] (5)

Equation (4) holds, e.g., if we have a simple assumption of the external time-invariant Poisson arrival for each local class, and the mean service time (without queueing delays) for each type-\(k\) jobs is \(\mu_k^{-1}\) at each node \(i\), and if the service discipline is ‘processor sharing’ or ‘preemptive-resume last-come first-served’ (see, e.g., Kleinrock (1976)). When \(\mu_k = \mu\) for all \(k\) and when no forwarding of jobs occurs, the mean processing time is, simply, \(1/(\mu - 1)\).

Equation (5) means that the expected communication time of a job arriving at node \(i\) and being processed at node \(j \neq i\) is expressed simply as \(t\), i.e., independent of the traffic and of the job class, with no queueing delay. It holds, e.g., in the case where one communication line that has traffic-independent delay is provided commonly for sending jobs from one node to another, or in the case where one communication line that has the same traffic-independent delay is provided separately for sending jobs from one node to another.

We refer to the length of time between the instant when a job arrives at a node and the instant when it leaves one of the nodes, after all processing and communication, if any, are over as the response time for the job. The expected response time of a local-class \(R_{ik}\) job that arrives at node \(i\), \(T_{ik}(\mathbf{x})\), is expressed as,

\[ T_{ik}(\mathbf{x}) = \sum_j x_{ijk} T_{ijk}(\mathbf{x}), \] (6)

where

\[ T_{ik}(\mathbf{x}) = \mu_k^{-1} D(\beta_i(\mathbf{x})), \quad \text{and} \]

\[ T_{ijk}(\mathbf{x}) = \mu_k^{-1} D(\beta_j(\mathbf{x})) + G_{ijk}(\mathbf{x}) = T_{ik}(\mathbf{x}) + G_{ijk}(\mathbf{x}), \text{ for } j \neq i. \] (8)

Using the fact that all nodes have the same arrival process, the expected response time of a global-class \(k\) job is

\[ T_k(\mathbf{x}) = \frac{1}{m} \sum_i T_{ik}(\mathbf{x}). \] (9)

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

\[ T(\mathbf{x}) = \sum_k \phi_k T_k(\mathbf{x}) = \frac{1}{m} \sum_{i,k} \phi_k T_{ik}(\mathbf{x}). \] (10)

**Remark 2.1** Note that as a consequence of Assumptions \(\Pi 1\) and \(\Pi 2\), the functions \(T(\cdot)\), \(T_{ik}(\cdot)\) and \(T_k(\cdot)\) are convex and differentiable with respect to the strategy profile \(\mathbf{x}\).

We analyze several decision schemes.

(A) In the completely centralized decision scheme, a single decision maker makes forwarding decision over all jobs. His strategy is therefore the choice of the optimal \(m m n\)-dimensional vector \(\bar{x}\), with components \(\bar{x}_{ijk}\). This optimized situation is the overall optimum.
(B) At the opposite, i.e., in the completely distributed decision scheme, we consider that each single job chooses the node to be processed. Thus the system has infinitely many decision makers. We denote by \( \hat{x}_{ijk} \) the resulting optimal ratio of jobs of local class \( R_{ik} \) that choose the node \( j \) to be processed. This optimized situation is the individual optimum. We denote the individually optimal strategy profile, i.e., the vector of components \( \hat{x}_{ijk} \), by \( \hat{x} \).

(C) For intermediately distributed decision schemes, we consider two slightly different schemes, C-I and C-II.

(C-I) In the first intermediately distributed decision scheme, each local class \( R_{ik} \) has its own decision maker (\( ik \)). The amount of forwarding for local-class \( R_{ik} \) jobs is chosen by the corresponding decision maker (\( ik \)). The optimal strategy for decision maker (\( ik \)), or equivalently local-class job \( R_{ik} \), is denoted by the \( m \)-dimensional vector,

\[
\hat{x}_{ik} = (\hat{x}_{i1k}, \hat{x}_{i2k}, \cdots, \hat{x}_{imk}),
\]

and an optimal strategy profile, that we will denote by \( \hat{x} \), is the collection of strategies \( \hat{x}_{ik} \). We call this scheme the intermediately distributed decision scheme I, and this optimized situation the local-class optimum.

(C-II) In the second intermediately distributed decision scheme, jobs of local classes \( R_{ik} \) for all \( i \) are united into one global class \( J_k \) that has a single decision maker (\( k \)). Each decision maker (\( k \)) of global class \( J_k \) chooses the amount of job forwarding for the \( m \) local classes, \( R_{1k}, R_{2k}, \ldots, R_{mk} \). The optimal strategy for decision maker \( k \) is consequently an \( mm \)-dimensional vector

\[
\check{x}_k = (\check{x}_{1k}, \check{x}_{2k}, \cdots, \check{x}_{mk}).
\]

An optimal strategy profile is an \( mmn \)-dimensional vector \( \check{x} = (\check{x}_1, \check{x}_2, \ldots, \check{x}_n) \). We call this scheme the intermediately distributed decision scheme II, and this optimized situation the global-class optimum.

3 Results

(A) Completely centralized decision scheme: Overall optimization

The overall optimum is given by \( \check{x} \) that satisfies the following,

\[
T(\check{x}) = \min T(x) \quad \text{with respect to} \quad x \in \mathbb{C}. \tag{11}
\]

We define \( x_{-(ijk)} \) to be an \( m(m-1)n \)-dimensional vector such that all elements \( x_{ijk} \), for all \( i,k \), are excluded from the \( mmn \)-dimensional vector \( x \) whereas all its elements are the same as the remaining \( m(m-1)n \) elements of \( x \).

Solution: The solution \( \check{x} \) is unique and given as follows:

\[
\check{x}_{-(ijk)} = 0, \quad \text{i.e.,} \quad \check{x}_{ijk} = 0, \text{ and } \check{x}_{iik} = 1, \text{ for all } i,j(\neq i),k. \tag{12}
\]

The mean response time is

\[
T_k(\check{x}) = T_{ik}(\check{x}) = \mu^{-1}_k D(\rho), \text{ for all } i,k, \quad T(\check{x}) = \rho D(\rho). \tag{13}
\]
We refer to this solution as the **solution of no job forwarding**.

**Proof:** This solution and its uniqueness are clear from the assumptions on $D$ and $G_{ijk}$. □

(B) **Completely distributed decision scheme: Individual optimization**

An individual optimum is given by $\hat{x}$ that satisfies the following for all $i, k$,

$$T_{ik}(\hat{x}) = \min_j \{T_{ijk}(\hat{x})\} \text{ and } \hat{x} \in C. \quad (14)$$

**Solution:** The same as the solution of no job forwarding, i.e., the solution is unique and given by (12), and the mean response time by (13), both with $\hat{x}$ being replaced by $\hat{x}$.

**Proof:** This can be easily seen in the following way. The solution $\hat{x}$ for (14) is characterized as follows: For all $i, j, k$ we have

$$T_{ijk}(\hat{x}) = \hat{\alpha}_{ik}, \quad \hat{x}_{ijk} \geq 0, \quad (15)$$

$$T_{ijk}(\hat{x}) > \hat{\alpha}_{ik}, \quad \hat{x}_{ijk} = 0, \quad \sum_{j'} \hat{x}_{ijk} = 1, \quad (16)$$

where $\hat{\alpha}_{ik} = \min_{j'} \{\mu_{k}^{-1}D(\beta_{j'}(\hat{x}))\}$. By noting (8) and (7), We can see that these relations are satisfied if $\hat{x}_{ijk} = 0$ for all $i, j(\neq i), k$. Then, we have $\beta_{i} = \rho$ for all $i$.

We can see the uniqueness of solution $\hat{x}$ as follows. Assume that there exists another solution $\hat{x}'$. Assume that for $\hat{x}'$, there exists $i'$ such that $\beta'_{i'} \neq \rho$. Then, by noting (3), there exists $i$ such that $\beta'_{i} \leq \beta'_{j'}$ for all $j (\neq i)$, and thus $\beta'_{i} < \rho$. By noting (7) and (8),

$$T_{ik}(\hat{x}') \leq T_{jik}(\hat{x}') < T_{jik}(\hat{x}') + G_{ijk}(\hat{x}') = T_{jik}(\hat{x}')', \quad \text{for all } j (\neq i), k.$$

Then, from (15) and (16), $\hat{x}'_{ijk} = 0$, for all $j (\neq i), k$, which implies $\beta'_{i} \geq \rho$, which contradicts $\beta'_{i} < \rho$ that is shown above. Therefore, $\beta'_{i} = \rho$ for all $i$ and $T_{ik}(\hat{x}') = T_{jik}(\hat{x}')$ for all $i, j (\neq i), k$. Then, $T_{ik}(\hat{x}') < T_{jik}(\hat{x}') + G_{ijk}(\hat{x}') = T_{jik}(\hat{x}''),$ and from (15) and (16), we must have $\hat{x}'_{ijk} = 0$, which is the same as the solution $\hat{x}$. Thus, we have seen the uniqueness of the solution. □

We suppose in the following (C-I) scheme and in the subsequent (C-II) scheme that the following assumption holds true:

**Assumption I3** We assume that $G_{ijk}(x)$ is one of the following functions, where $\omega_k$ and $\sigma_k$ are
constant, and \(G(x)\) is a nondecreasing, convex, and differentiable function of \(x\) and \(G(0) = 1\).

**Type G-I**

\[
G_{ijk}(x) = \omega_k^{-1}G(\sigma_k x_{ijk})
\]

(one dedicated line for each combination of a pair of origin and destination nodes, and a local class: i.e., \(m(m - 1)n\) lines in total),

**Type G-II(a)**

\[
G_{ijk}(x) = \omega_k^{-1}G\left(\sum_{p \neq p, q \neq q}^{} \sigma_k x_{pqk}\right)
\]

(one bus line for each global class: i.e., \(n\) bus lines in total),

**Type G-II(b)**

\[
G_{ijk}(x) = \omega_k^{-1}G\left(\sum_{p \neq p, k}^{} \sigma_k x_{pqk}\right)
\]

(one common bus line for the entire system: i.e., 1 bus line),

**Remark 3.1** \(\omega_k^{-1}\) can be regarded as the mean communication time (without queueing delays) for forwarding a Type-\(k\) job from the arrival node to another processing node. \(\sigma_k x_{ijk}\) \((j \neq i)\) is the traffic intensity of the communication line through which local-class \(R_{ik}\) jobs are forwarded to node \(j\).

\[
\Box
\]

**Example 2** We use the same definition (4) for the mean processing time as in Example 1. We define \(G_{ijk}(x)\), the mean communication delay, as follows. We assume \(\omega_k = \theta\) and \(\sigma_k = \phi_k/\theta\), for all \(k\), and set

\[
G_{ijk}(x) = \frac{1}{\theta - \sum_{p \neq p, k}^{} \phi_k x_{pqk}} \text{ for } \sum_{p \neq p, k}^{} \sigma_k x_{pqk} < 1, \text{ and otherwise infinite.}
\]

This is identical to:

\[
G_{ijk}(x) = \frac{1}{\theta - \sum_{p \neq p, k}^{} \phi_k x_{pqk}} \text{ for } \sum_{p \neq p, k}^{} \phi_k x_{pqk} < \theta, \text{ and otherwise infinite.} \tag{17}
\]

This holds, for example, if we assume that one bus-type communication line is provided commonly for all nodes to be used for forwarding of jobs to other nodes in the same way as in Example 1, except that the mean communication delay is traffic dependent, the transmission time without queueing delay is exponentially distributed with mean \(\theta^{-1}\), and the scheduling discipline is ‘First-Come-First-Served’. Thus, the expected communication time of a job arriving at node \(i\) and being processed at node \(j\) \((\neq i)\) is expressed as \(1/(\theta - \sum_{p \neq p, k}^{} \phi_k x_{pqk})\), i.e., independent of the job type and of the origin and destination nodes.

\[
\Box
\]

**(C-I) Intermediately distributed decision scheme I: Local-class optimization**

The local-class optimum is given by \(\bar{x}\) that satisfies the following for all \(i, k\),

\[
T_{ik}(\bar{x}) = \min_{x_{ik}} T_{ik}(\bar{x}_{-(ik)}, x_{ik}) \text{ with respect to } x_{ik} \text{ such that } (\bar{x}_{-(ik)}), x_{ik}) \in C,
\]
where \((\bar{x}_{-ik}, \bar{x}_{ik})\) denotes an \(mmn\)-dimensional vector in which the elements corresponding to \(\bar{x}_{ik}\) have been replaced by \(x_{ik}\).

Let us define \(\bar{g}_{ijk}(\cdot)\) as

\[
\bar{g}_{ijk}(x) = \frac{\partial}{\partial x_{ijk}} [\phi_k \sum_{p \neq i} x_{ipk} G_{ipk}(x)].
\]

Define \(v_k = \phi_k/\omega_k\). By Assumption \(\Pi 3\), we have

\[
\bar{g}_{ijk}(x) = \nu_k [G(x) + \sigma_k (1 - x_{ijk}) G'(x)], \text{ for type } G-I, \quad \bar{g}_{ijk}(x) = \nu_k [G(x) + \sigma_k x_{ijk} G'(x)], \text{ for type } G-II,
\]

where \(x = \sum_{p,q,p} \sigma_k x_{pqk}\) for type \(G-II(a)\), and

\[
\bar{x} = \sum_{p,q,p,k} \sigma_k x_{pqk} \text{ for type } G-II(b).
\]

We see that under Assumption \(\Pi 3\), functions \(G_{ijk}(x)\) satisfy for all \(i, j \neq i, j' \neq i, k,\)

\[
\bar{g}_{ijk}(x) \geq \bar{g}_{ij'k}(x) \text{ if } x_{ijk} > x_{ij'k}.
\]

If Assumption \(\Pi 3\) holds, for \(x\) such that \(x_{ijk} = x_k\) for all \(i, j \neq i, k\), we denote

\[
G_k(x) = G_{ijk}(x) \text{ and } \bar{g}_k(x) = \bar{g}_{ijk}(x).
\]

In particular, for \(x\) such that \(x_{ijk} = x\) for all \(i, j \neq i, k\), we denote

\[
G(x) = G_k(x) = G_{ijk}(x) \text{ and } \bar{g}(x) = \bar{g}_k(x) = \bar{g}_{ijk}(x).
\]

**Solution:** We denote \(\Gamma_k = \rho_k^2 \nu_k^{-1}\). The solution \(\bar{x}\) is unique and is given as follows:

For Types G-I and G-II(a)

(i) For local class \(R_{ik}\) such that \(\rho_k^2 D'(<) = \bar{g}_k(0) = \nu_k\), i.e., \(\Gamma_k D'(>) \leq 1\), the solution has the same value as the solution of no job forwarding, i.e., the solution \(\bar{x}\) is unique and given by (12), and the mean response time by (13), both with \(\bar{x}\) being replaced by \(\bar{x}\).

(ii) For local class \(R_{ik}\) such that \(\rho_k^2 D'(>) = \bar{g}_k(0) = \nu_k\), i.e., \(\Gamma_k D'(>) > 1\), the solution is given as follows:

\[
\bar{x}_{ijk} = \bar{x}_k, \text{ for all } i, j \neq i,
\]

where \(\bar{x}_k\) is a unique solution of

\[
\rho_k^2 (1 - m \bar{x}_k) D'(>) = \bar{g}_k(\bar{x}_k)
\]

\[
= \nu_k [G(m(m-1)\sigma_k \bar{x}_k) + \sigma_k (m-1) \bar{x}_k G'(m(m-1)\sigma_k \bar{x}_k)].
\]

The mean response time is

\[
T_k(\bar{x}) = T_k(\bar{x}_k) = \mu_k^{-1} D(>) + (m-1)\bar{x}_k G_k(\bar{x}_k), \text{ for all } i.
\]
For Type G-II(b)

The solution is given as follows. We first change the numbering of \( k \) such that \( \Gamma_1 \geq \Gamma_2 \geq \cdots \geq \Gamma_k \geq \cdots \geq \Gamma_n \). The following three situations can occur:

\[
\text{We can find } K \text{ such that } \Gamma_k D'(\rho) > 1 \text{ and } \Gamma_{k+1} D'(\rho) \leq 1, \quad (25)
\]

or \( \Gamma_n D'(\rho) > 1 \) (i.e., \( K = n \)), \( (26) \)

or \( \Gamma_1 D'(\rho) \leq 1 \). \( (27) \)

When (27) holds, we have a unique solution of \( \tilde{x}_k = 0 \) for all \( k \). When (25) or (26) holds, we can find a unique solution as follows. Let us define function \( F_k(X) \) as

\[
F_k(X) = \left\{ \sum_{i=1}^{k} \frac{\sigma_i [\Gamma_i D'(\rho) - G(X)]}{m \Gamma_i D'(\rho) + (m-1) \sigma_i G'(X)} \right\} - \frac{X}{m(m-1)}. \quad (28)
\]

We obtain the largest \( k = \tilde{k} \leq K \) and \( X = \tilde{X}_k(>0) \) that satisfies \( F_k(\tilde{X}_k) = 0 \) and \( \Gamma_k D'(\rho) - G(\tilde{X}_k) > 0 \). Then by using

\[
\sigma_i [\Gamma_i D'(\rho) - G(\tilde{X}_k)] = \sigma_i \tilde{x}_i [m \Gamma_i D'(\rho) + (m-1) \sigma_i G'(\tilde{X}_k)], \quad (29)
\]

for \( k = 1, 2, \cdots, \tilde{k} \), we obtain a unique set of values such that \( \tilde{x}_k > 0, k = 1, 2, \cdots, \tilde{k} \), and that \( \tilde{x}_{k+1} = \tilde{x}_{k+2} = \cdots = \tilde{x}_n = 0 \) that satisfies the above relation, which is a unique solution. Then, the mean response time is given by (24).

**Proof:** We define

\[
\tilde{t}_{ik}(x) = \phi_k \frac{\partial}{\partial x_{ij,k}} T_{ik}(x). \quad (30)
\]

Because \( T_{ik} \) are convex functions and \( C \) is a convex set, the solution \( \tilde{x} \) of the problem exists (see, Rosen (1965)). From the Kuhn-Tucker conditions, it is characterized by the relations (see, e.g., Shapiro (1979)):

\[
\tilde{t}_{ijk}(\tilde{x}) = \tilde{\alpha}_{jk}, \quad \tilde{x}_{ijk} > 0,
\]

\[
\geq \tilde{\alpha}_{jk}, \quad \tilde{x}_{ijk} = 0,
\]

\[
\sum_j \tilde{x}_{ijk} = 1, \quad \text{for all } i, k, \quad (31)
\]

where \( \tilde{\alpha}_{jk} \) are the Lagrange multipliers. From Definitions (1), (6) to (8), (18), and (30), we have

\[
\tilde{t}_{ik}(X) = \phi_k \frac{\partial T_{ik}(x)}{\partial x_{ij,k}} = \rho_k [D(\beta_i(x)) + \rho_k x_{ijk} D'(\beta_i(x))], \quad (32)
\]

\[
\tilde{t}_{ijk}(X) = \phi_k \frac{\partial T_{ik}(x)}{\partial x_{ij,k}} = \rho_k [D(\beta_j(x)) + \rho_k x_{ijk} D'(\beta_j(x)) + \tilde{g}_{ijk}(x)] \text{ for } j \neq i. \quad (33)
\]

We assume that there exists a local-class optimum strategy profile, \( \tilde{x} \). Define \( \tilde{\beta}_i = \beta_i(\tilde{x}) \).

(1) First, we show by contradiction that \( \tilde{\beta}_j = \tilde{\beta}_{j'} \) for every pair \((j, j')\), and consequently, \( \tilde{\beta}_i = \rho \) for all \( i \).
We define
\[ \Xi_{ijk'}(x) = \tilde{t}_{ijk}(x) - \tilde{t}_{jk'}(x). \] (34)

Assume that \( \tilde{\gamma}_j > \tilde{\gamma}_{j'} \) for some \( j \) and \( j' \).

(1-1) Assume \( \tilde{x}_{ijk} > \tilde{x}_{ij'}k \) for some \( i \neq j, j', k \). Then, we have \( \tilde{g}_{ijk}(\tilde{x}) \geq \tilde{g}_{ij'k}(\tilde{x}) \) by (21). From Equation (33) and Definition (34) we have
\[ \Xi_{ijk';i'j'k}(x) = \rho_k[D(\beta_j(x)) - D(\beta_{j'}(x))] + \rho_k^2[x_{ijk}D'(\beta_j(x)) - x_{ij'k}D'(\beta_{j'}(x))] + \tilde{g}_{ijk}(x) - \tilde{g}_{ij'k}(x). \] (35)

Together with the fact that \( D \) and \( D' \) are, respectively, increasing and nondecreasing (II1), it follows that \( \Xi_{ijk';i'j'k}(x) > 0 \). However, from (31) we must have
\[ \Xi_{ijk';i'j'k} \leq 0, \] (36)
which contradicts the above. Consequently, we have \( \tilde{x}_{ijk} \leq \tilde{x}_{ij'k} \) for all \( i \neq j, j', k \).

(1-2) From the assumption \( \tilde{\gamma}_j > \tilde{\gamma}_{j'} \), we have at least for some \( k \),
\[ \tilde{x}_{ijk} + \tilde{x}_{ij'k} > \tilde{x}_{ij'k} + \tilde{x}_{ij'k}. \] (37)

(1-2-1) If \( \tilde{x}_{ij'k} = 0 \), we have
\[ \tilde{x}_{ijk} > \tilde{x}_{ij'k} \quad \text{and} \quad \tilde{x}_{ij'k} \geq \tilde{x}_{ij'k} \quad \text{(Condition I)}. \]

(1-2-2) If \( \tilde{x}_{ij'k} > 0 \), similarly as in (1-1), we see that if \( \tilde{x}_{ij'k} > \tilde{x}_{ij'k} \), then \( \Xi_{ijk';i'j'k}(\tilde{x}) > 0 \), which contradicts (36). Thus we have \( \tilde{x}_{ij'k} \leq \tilde{x}_{ij'k} \) and \( \tilde{x}_{ijk} > 0 \). From (31), (32), and (33),
\[ \tilde{t}_{ijk}(x) = \rho_k[D(\tilde{\gamma}_j) + \rho_k\tilde{x}_{ijk}D'(\tilde{\gamma}_j)] = \tilde{t}_{jk}, \]
\[ \tilde{t}_{jk'}(x) = \rho_k[D(\tilde{\gamma}_{j'}) + \rho_k\tilde{x}_{ij'k}D'(\tilde{\gamma}_{j'})] + \tilde{g}_{jk'}(\tilde{x}) = \tilde{t}_{jk}. \]

Then, by adding the last two equations, we have,
\[ \rho_k[2D(\tilde{\gamma}_j) + \rho_k(\tilde{x}_{ijk} + \tilde{x}_{ij'k})D'(\tilde{\gamma}_j)] + \tilde{g}_{jk}(\tilde{x}) = \tilde{t}_{jk} + \tilde{t}_{jk}. \] (38)

Note that we have from (31), (32), and (33),
\[ \tilde{t}_{ijk}(\tilde{x}) + \tilde{t}_{jk'}(\tilde{x}) = \rho_k[2D(\tilde{\gamma}_j) + \rho_k(\tilde{x}_{ijk} + \tilde{x}_{ij'k})D'(\tilde{\gamma}_{j'})] + \tilde{g}_{jk'}(\tilde{x}) \geq \tilde{t}_{jk} + \tilde{t}_{jk}. \] (39)

Since \( D \) is increasing, \( D' \) is nondecreasing, \( \tilde{\gamma}_j > \tilde{\gamma}_{j'} \) and \( \tilde{x}_{ijk} + \tilde{x}_{ij'k} > \tilde{x}_{ij'k} + \tilde{x}_{ij'k} \) by assumption, the only possibility for (38) and (39) not to contradict each other is
\[ \tilde{g}_{jk}(\tilde{x}) > \tilde{g}_{jk'}(\tilde{x}). \] (40)

Therefore, in the special case where \( \tilde{g}_{jk}(\tilde{x}) = \tilde{g}_{jk'}(\tilde{x}) \), these two relations contradict each other. For the other cases, we investigate in the following (1-2-2-1) and (1-2-2-2). Recall equations (19) and (20), respectively, for Type G-I and Type G-II cases. For Type G-I case, from equation
(19), we see that \( \tilde{g}_{ijk} \) depends only on, and increases in, \( x_{ijk} \), which will be discussed in (1-2-2-1). On the other hand, for Type G-II case, \( \tilde{g}_{ijk} \) does not depend directly on \( x_{ijk} \), and from equation (20), we see that
\[
\tilde{g}_{jfk}(\bar{x}) - \tilde{g}_{jfk}(\bar{x}) = v_k \sigma_k (x_{jfk} - x_{jjk}) G'(x),
\]
which will be discussed in (1-2-2-2).

(1-2-2-1) Consider Type G-I case. From the above relations (40) and (41) on \( \tilde{g}_{ijk} \), we have \( \tilde{x}_{jjk} > \tilde{x}_{jfk} \), from which \( \tilde{x}_{jjk} > \tilde{x}_{jfk} \) follows by noting relation \( \tilde{x}_{jjk} + \tilde{x}_{jfk} > \tilde{x}_{jfk} + \tilde{x}_{jfk} \) (37) at the beginning of (1-2). We thus have
\[
\tilde{x}_{jfk} > \tilde{x}_{jk} \quad \text{and} \quad \tilde{x}_{jk} > \tilde{x}_{jfk},
\]
which satisfies Condition I given in (1-2-1).

(1-2-2-2) Consider Type G-II case. From the above relations (40) and (41) on \( \tilde{g}_{ijk} \), we have \( \tilde{x}_{jfk} > \tilde{x}_{jk} \), from which \( \tilde{x}_{jfk} > \tilde{x}_{jfk} \) follows by noting \( \tilde{x}_{jjk} + \tilde{x}_{jfk} > \tilde{x}_{jfk} \) (37) at the beginning of (1-2). We thus have
\[
\tilde{x}_{jfk} > \tilde{x}_{jk} \quad \text{and} \quad \tilde{x}_{jk} > \tilde{x}_{jfk} \quad \text{(Condition II)}.
\]

Therefore, we see that, from the assumption \( \tilde{\beta}_j > \tilde{\beta}_j' \), either Condition I or Condition II holds.

(1-3) Now we examine Conditions I and II, respectively, in the following sections (1-3-1) and (1-3-2), and will show that both lead to contradictions.

(1-3-1) Consider the case where Condition I holds. Since
\[
\tilde{t}_{jk}(\bar{x}) = \rho_k [D(\tilde{\beta}_j) + \rho_k \tilde{x}_{jjk} D'(\tilde{\beta}_j)] = \tilde{\alpha}_{jk},
\]
\[
\tilde{t}_{jk}(\bar{x}) = \tilde{\rho}_k [D(\tilde{\beta}_j) + \rho_k \tilde{x}_{jfk} D'(\tilde{\beta}_j)] \geq \tilde{\alpha}_{jfk},
\]
we have \( \tilde{\alpha}_{jk} > \tilde{\alpha}_{jfk} \), because \( D \) is increasing, \( D' \) is nondecreasing and \( \tilde{\beta}_j > \tilde{\beta}_j' \) by assumption.

We next show by contradiction that \( \tilde{x}_{jk} \geq \tilde{x}_{jlk} \), for all \( l \neq j, j' \). Assume \( \tilde{x}_{jlk} < \tilde{x}_{jlk} \). Then \( \tilde{x}_{jlk} > 0 \), and we have from (31) and (33),
\[
\tilde{t}_{jlk}(\bar{x}) = \rho_k [D(\tilde{\beta}_l) + \rho_k \tilde{x}_{jlk} D'(\tilde{\beta}_l)] + \tilde{g}_{jlk}(\bar{x}) = \tilde{\alpha}_{jlk},
\]
\[
\tilde{t}_{jlk}(\bar{x}) = \rho_k [D(\tilde{\beta}_l) + \rho_k \tilde{x}_{jlk} D'(\tilde{\beta}_l)] + \tilde{g}_{jlk}(\bar{x}) \geq \tilde{\alpha}_{jk} > \tilde{\alpha}_{jfk},
\]
which contradicts the assumption, as we see by noting that here for G-I, \( \tilde{g}_{jlk}(\bar{x}) \leq \tilde{g}_{jlk}(\bar{x}) \). Therefore we must have
\[
\tilde{x}_{jlk} \geq \tilde{x}_{jlk}.
\]

From this and Condition I, it follows
\[
\tilde{x}_{jlk} > \tilde{x}_{jfk}, \quad \tilde{x}_{jlk} \geq \tilde{x}_{jlk} \quad \text{for all} \quad l \neq j, j'.
\]
This implies
\[ 1 = \sum_l \tilde{x}_{ijk} > \sum_l \tilde{x}_{jlk} = 1, \]
which is impossible. That is, the assumption leads to a contradiction.

(1-3-2) Consider the case where Condition II holds. This implies \( \tilde{x}_{jk} > 0 \) and we have
\[
\begin{align*}
\tilde{t}_{jk}(\tilde{x}) &= \rho_k[D(\tilde{\beta}_j) + \rho_k \tilde{x}_{jk} D'(\tilde{\beta}_j)] + \tilde{g}_{jk}(\tilde{x}) = \tilde{\alpha}_{jk}, \\
\tilde{t}_{jk}(\tilde{x}) &= \rho_k[D(\tilde{\beta}_j) + \rho_k \tilde{x}_{jk} D'(\tilde{\beta}_j)] + \tilde{g}_{jk}(\tilde{x}) \geq \tilde{\alpha}_{jk}.
\end{align*}
\]
Since \( D \) is increasing, \( D' \) is nondecreasing, \( \beta_j > \beta_f \) and \( \tilde{x}_{jk} > \tilde{x}_{jfk} \), we have
\[
\tilde{\alpha}_{jk} - \tilde{g}_{jk}(\tilde{x}) > \tilde{\alpha}_{jk} - \tilde{g}_{jfk}(\tilde{x}).
\]
By noting that for type G-II, we have \( \tilde{g}_{jk}(\tilde{x}) = \tilde{g}_{jfk}(\tilde{x}) \) and \( \tilde{g}_{jk}(\tilde{x}) = \tilde{g}_{jfk}(\tilde{x}) \) for any \( l (\neq j, j') \), and from the Kuhn-Tucker conditions, we have
\[
\begin{align*}
\rho_k[D(\tilde{\beta}_j) + \rho_k \tilde{x}_{jk} D'(\tilde{\beta}_j)] &= \tilde{\alpha}_{jk} - \tilde{g}_{jk}(\tilde{x}) = \tilde{\alpha}_{jk} - \tilde{g}_{jfk}(\tilde{x}), \quad \text{for} \quad \tilde{x}_{jk} > 0, \\
\rho_k[D(\tilde{\beta}_j) + \rho_k \tilde{x}_{jk} D'(\tilde{\beta}_j)] &= \tilde{\alpha}_{jk} - \tilde{g}_{jk}(\tilde{x}) = \tilde{\alpha}_{jk} - \tilde{g}_{jfk}(\tilde{x}), \quad \text{for} \quad \tilde{x}_{jk} = 0, \\
\rho_k[D(\tilde{\beta}_j) + \rho_k \tilde{x}_{jk} D'(\tilde{\beta}_j)] &= \tilde{\alpha}_{jk} - \tilde{g}_{jk}(\tilde{x}) = \tilde{\alpha}_{jk} - \tilde{g}_{jfk}(\tilde{x}), \quad \text{for} \quad \tilde{x}_{jk} > 0, \\
\rho_k[D(\tilde{\beta}_j) + \rho_k \tilde{x}_{jk} D'(\tilde{\beta}_j)] &= \tilde{\alpha}_{jk} - \tilde{g}_{jk}(\tilde{x}) = \tilde{\alpha}_{jk} - \tilde{g}_{jfk}(\tilde{x}), \quad \text{for} \quad \tilde{x}_{jk} = 0,
\end{align*}
\]
which holds only when \( \tilde{x}_{jk} \leq \tilde{x}_{jfk} \) for all \( l (\neq j, j') \).

From this and Condition II,
\[
\begin{align*}
\tilde{x}_{jk} &< \tilde{x}_{jfk}, \\
\tilde{x}_{jfk} &< \tilde{x}_{jfk}, \\
\tilde{x}_{jk} &\leq \tilde{x}_{jfk} \quad \text{for all} \quad l (\neq j, j').
\end{align*}
\]
This implies
\[ 1 = \sum_l \tilde{x}_{ijk} < \sum_l \tilde{x}_{jfk} = 1, \]
which is impossible. That is, the assumption leads to a contradiction.

Thus we see that the assumption \( \tilde{\beta}_j > \tilde{\beta}_f \) leads to either Condition I [(1-2-1) and (1-3-1)] or Condition II [(1-3-2)], both of which lead to contradictions.

Therefore, we must have \( \tilde{\beta}_j = \tilde{\beta}_f \), and consequently \( \tilde{\beta}_l = \rho \) for all \( i \).

(2) Hence for all \( i, j (\neq i), j' (\neq i), k \),
\[
\Xi_{ijk;jfk}(\tilde{x}) = \rho_k^2(\tilde{x}_{ijk} - \tilde{x}_{jfk}) D'(\rho) + \tilde{g}_{ijk}(\tilde{x}) - \tilde{g}_{jfk}(\tilde{x}). \tag{42}
\]
Thus, if \( \tilde{x}_{ijk} > \tilde{x}_{jfk} \) for some \( i, j (\neq i), j' (\neq i), k \), we have \( \Xi_{ijk;jfk} > 0 \) since \( D'(\rho) > 0 \), which contradicts (36). Therefore, we must have
\[ \tilde{x}_{ijk} = \tilde{x}_k \text{ for all } i, j (\neq i), k. \tag{43} \]
(3) We note that since $\sum_j x_{ijk} = 1$, from the assumption on the arrival ratio of each local-class job, $\bar{x}_k$ has to belong to the interval $[0, 1/(m-1)]$. We discuss the case for Types G-I and G-II(a) and the case for Type G-II(b), separately.

**Types G-I and G-II(a):**

We have

$$\Xi_{ijk,ijk}(\bar{x}) = -\rho_k^2(1 - m\bar{x}_k)D'(\rho) + \tilde{g}_k(\bar{x}_k).$$

where

$$\tilde{g}_k(\bar{x}_k) = v_k[G(\sigma_k\bar{x}_k) + \sigma_k\bar{x}_k\hat{G}'(\sigma_k\bar{x}_k)] \quad \text{(Type G-I), or}$$

$$\tilde{g}_k(\bar{x}_k) = v_k[G(m(m-1)\sigma_k\bar{x}_k) + (m-1)\sigma_k\bar{x}_k\hat{G}'(m(m-1)\sigma_k\bar{x}_k)] \quad \text{(Type G-II(a)).}$$

Let us define function $H_k$ as

$$H_k(x) = -\rho_k^2(1 - mx)D'(\rho) + \tilde{g}_k(x).$$

Clearly, $H_k$ is continuous and monotonically increasing in $x$.

(i) For local class $R_{ik}$ such that $\rho_k^2D'(\rho) \leq \tilde{g}_k(0) = v_k$, i.e., $\Gamma_kD'(\rho) \leq 1$, we have $H_k(x) > H_k(0) \geq 0$ for any $x > 0$, which proves that $x = \bar{x}_k = 0$ is a unique optimal solution.

Therefore, for local class $R_{ik}$ such that $\rho_k^2D'(\rho) \leq \tilde{g}_k(0) = v_k$, the solution has the same value as the solution of no job forwarding, i.e., the solution $\bar{x}$ is unique and given by (12), and the mean response time by (13), both with $\bar{x}$ being replaced by $\bar{x}$.

(ii) For local class $R_{ik}$ such that $\rho_k^2D'(\rho) > \tilde{g}_k(0) = v_k$, i.e., $\Gamma_kD'(\rho) > 1$, the optimal solution is uniquely given as follows:

$$\tilde{x}_{ijk} = \bar{x}_k, \text{ for all } i, j \neq i \quad \text{(eq. (22))},$$

where $\bar{x}_k$ is a unique solution of (23). The left and right hand sides of (23) are continuous and respectively, decreasing and non-decreasing, in $\bar{x}_k \geq 0$, since $\tilde{g}_k(\cdot)$ is non-decreasing (see (44) and (45)). If $\bar{x}_k = 0$, the left hand side of (23) is greater than the right hand side, whereas if $\bar{x}_k = 1/m$, the former is less than the latter. Thus we see that there is a unique $\bar{x}_k$ that satisfies (23), and $0 < \bar{x}_k < 1/m$.

We note that eq. (23) assures that $0 < \bar{x}_k < 1/m < 1/(m-1)$. Therefore, the mean response time is given by (24).

**Type G-II(b):**

We have

$$\Xi_{ijk,ijk}(\bar{x}) = -\rho_k^2(1 - m\bar{x}_k)D'(\rho) + \tilde{g}_k(\bar{x}).$$

In the following, we will show that there exists a unique solution, $\bar{x}_k$, $k = 1, 2, \cdots, n$, for the following system of relations:

$$\rho_k^2(1 - m\bar{x}_k)D'(\rho) = \tilde{g}_k(\bar{x}) \quad \text{and} \quad \bar{x}_k \geq 0,$$

$$\rho_k^2D'(\rho) \leq \tilde{g}_k(\bar{x}) \quad \text{and} \quad \bar{x}_k = 0, \quad \rho_k^2D'(\rho) = \tilde{g}_k(\bar{x}), \quad 0 \leq \bar{x}_k \leq 1/(m-1),$$

(47)
where \( \tilde{g}_k(\bar{x}) = v_k[G(m(m-1) \sum_k \sigma_k \bar{x}_k) + \sigma_k(m(m-1) \bar{x}_k \bar{G}'(m(m-1) \sum_k \sigma_k \bar{x}_k))]. \)

The relations (47) are equivalent to the following:

\[
\begin{align*}
\sigma_k[\Gamma D'({\rho}) - G(\bar{X})] &= \sigma_k \bar{x}_k[m\Gamma D'({\rho}) + (m-1)\sigma_k \bar{G}'(\bar{X})] \quad \text{and} \quad \bar{x}_k \geq 0, \\
\sigma_k[\Gamma D'({\rho}) - G(\bar{X})] &\leq 0 \quad \text{and} \quad \bar{x}_k = 0, \\
0 &\leq \bar{x}_k \leq 1/(m-1),
\end{align*}
\]

where we recall that \( \Gamma_k = \rho_k^2 v_k^{-1} \) and we define \( \bar{X} = m(m-1) \sum_i \sigma_i \bar{x}_i/\sigma_i \). Thus if we denote by \( \tilde{K}(\bar{X}) \) the set of \( k \) such that \( \bar{x}_k > 0 \) (i.e., \( \Gamma_k D'({\rho}) - G(\bar{X}) > 0 \)), we have

\[
\sum_{k \in \tilde{K}(\bar{X})} \frac{\sigma_i[\Gamma D'({\rho}) - G(\bar{X})]}{m\Gamma D'({\rho}) + (m-1)\sigma_i \bar{G}'(\bar{X})} = \frac{\bar{X}}{m(m-1)}. \tag{49}
\]

We easily see that we can change the numbering of \( k \) such that \( \Gamma_1 \geq \Gamma_2 \geq \cdots \geq \Gamma_k \geq \cdots \geq \Gamma_n \). Thus, the three situations given by relations (25), (26), and (27), can occur.

Consider the case (27) holds. Then, we have \( \Gamma_k D'({\rho}) \leq 1 \), for all \( k \). Note, from Assumption \( \Pi 3 \), that \( G(0) = 1 \). Thus, when (27) holds, we can find a unique solution of \( \bar{x}_k = 0 \) for all \( k \).

When (25) or (26) holds, we can find a unique solution as follows. Recall that \( \tilde{K}(X) = \{k|\Gamma_k D'({\rho}) - G(X) > 0\} \). Consider the following function.

\[
F(X) = \left\{ \sum_{k \in \tilde{K}(X)} \frac{\sigma_i[\Gamma D'({\rho}) - G(X)]}{m\Gamma D'({\rho}) + (m-1)\sigma_i \bar{G}'(X)} \right\} - \frac{X}{m(m-1)}. \tag{50}
\]

Denote by \( X_i \) such a value of \( X \) that satisfies \( \Gamma_k D'({\rho}) = G(X) \). Since \( \Gamma_1 \geq \Gamma_2 \geq \cdots \geq \Gamma_k \geq \cdots \geq \Gamma_n \), we have \( X_1 \geq X_2 \geq \cdots \geq X_{k} > 0 \). Even if \( K < n \), we cannot obtain \( X_{k+1} \), since \( \Gamma_{k+1} D'({\rho}) < 1 = G(0) \leq G(X) \) for \( X \geq 0 \). Then, define \( X_{k+1} = 0 \).

Recall the function \( F_k(X) \) defined by (28). We can see that \( \tilde{F}(X) = F_k(X) \) for \( X_{k+1} \leq X < X_{k-i+1} \) when \( \Gamma_{k+i} > \Gamma_{k-i+1} = \Gamma_k > \Gamma_{k+1} \). For \( X \geq 0 \), \( \tilde{F}(X) \) is continuous and monotonically decreasing in \( X \) as shown below. For \( X \geq 0 \), the member of the set \( \tilde{K}(X) \) is \( \{1, 2, \cdots , K\} \). For interval \( [X_{k+1}, X_{k-i+1}] \) with \( \Gamma_{k+i} > \Gamma_{k-i+1} = \Gamma_k > \Gamma_{k+1} \), \( \tilde{K}(X) \) remains the same as \( \{1, 2, \cdots , k\} \). Then, as \( X \) increases and passes over \( X_{k-i+1} \), members \( k - i + 1, \cdots , k \) go out of the set, but no jump occurs in the value of \( \tilde{F}(X) \). Thus, \( \tilde{F}(X) \) is continuous and monotonically decreasing in \( X \) (although it may not be differentiable).

It is also easily seen that \( \tilde{F}(0) > 0 \) and \( \tilde{F}(X) < 0 \) for sufficiently large \( X \).

Thus there exists a unique value, \( \tilde{X} (> 0) \), of \( X \) that satisfies \( \tilde{F}(X) = 0 \). Therefore, for the case where (25) or (26) holds, we can find the largest \( k = \bar{k} \) such that \( \Gamma_k D'({\rho}) - G(\bar{X}) > 0 \) (and \( \Gamma_{\bar{k}+1} D'({\rho}) - G(\bar{X}) \leq 0 \)). Then given \( \tilde{X}_k (= \tilde{X}) \) from (29), we can obtain a unique set of values for \( \tilde{x}_k \), \( 1 \leq k \leq \bar{k} \).

From (29), by noting that \( G \geq 1 \) and \( G' \geq 0 \), we easily see that \( 0 < \tilde{x}_k \leq 1/m < 1/(m-1) \), \( 1 \leq k \leq \bar{k} \). The set of values, \( \tilde{x}_k, k = 1, 2, \cdots , \bar{k} \), obtained above and \( \tilde{x}_k = 0, k = \bar{k} + 1, \bar{k} + 2, \cdots , n \), is a unique solution.

Therefore, we see that we have a unique solution as stated at the end of the above solution. That is, we can obtain a unique set of values such that \( \tilde{x}_k > 0, k = 1, 2, \cdots , \bar{k} \), and \( \tilde{x}_{\bar{k}+1} = \cdots = \tilde{x}_n = 0 \), which satisfies the above relation.

The mean response time is obtained as (24) by noting the definitions (1), (6), (7), (8), and (9). \( \square \)
(C-II) Intermediately distributed decision scheme: Global-class optimization

The global-class optimum (class optimum for another set of decision makers) is given by \( \tilde{x} \) that satisfies the following for all \( i, k \),

\[
T_k(\tilde{x}) = \min T_k(\tilde{x}_{-(k)}; x_k), \text{ with respect to } x_k \text{ such that } (\tilde{x}_{-(k)}; x_k) \in C,
\]

where \( (\tilde{x}_{-(k)}; x_k) \) denotes the \( mmn \)-dimensional vector in which the elements corresponding to the coordinates of \( \tilde{x}_k \) have been replaced by the vector \( x_k \). We note that

\[
\phi_k m T_k(x) = \sum_i \beta_i^k(x) D(\beta_i(x)) + \sum_{i,j \neq i} \phi_k x_{ijk} G_{ijk}(x).
\]

Note that we have the assumption \( \Pi_3 \) on the function \( G_{ijk} \).

We define \( \tilde{g}_{ijk}(x) \) as

\[
\tilde{g}_{ijk}(x) = \frac{\partial}{\partial x_{ijk}} \left( \sum_{p,q \neq p} \phi_k x_{pqk} G_{pqk}(x) \right).
\]

By Assumption \( \Pi_3 \), we have

\[
\tilde{g}_{ijk}(x) = \tilde{\alpha}_{ik} \text{ if } x_{ijk} > 0,
\]

\[
\sum_{i,j \neq i} \tilde{\alpha}_{ik} = 1, \text{ for all } i, k,
\]

Solution: The solution has the same value as the solution of no job forwarding, i.e., the solution \( \tilde{x} \) is unique and given by (12), and the mean response time by (13), both with \( \bar{x} \) being replaced by \( \tilde{x} \).

Proof: We define

\[
\tilde{t}_{ijk}(x) = m \phi_k \frac{\partial}{\partial x_{ijk}} T_k(x).
\]

Again, because \( T_k \) is a convex function and \( C \) is compact, the solution \( \tilde{x} \) of the problem exists (see Rosen (1965)). From the Kuhn-Tucker conditions, it is characterized by the relations (see, e.g., Shapiro (1979)):

\[
\tilde{t}_{ijk}(\tilde{x}) = \tilde{\alpha}_{ik} \text{ for } \tilde{x}_{ijk} \text{ such that } \tilde{x}_{ijk} > 0,
\]

\[
\sum_j \tilde{x}_{ijk} = 1, \text{ for all } i, k,
\]
where \( \bar{\alpha}_{ik} \) are the Lagrange multipliers. From the definitions (1) to (9), (53), and (55), we have

\[
\tilde{t}_{ik}(x) = m\phi_k \frac{\partial T_k}{\partial x_{ik}} = \rho_k [D(\beta_i) + \beta_i^{(k)} D'(\beta_i)], \tag{57}
\]

\[
\tilde{t}_{ijk}(x) = m\phi_k \frac{\partial T_k}{\partial x_{ijk}} = \rho_k [D(\beta_j) + \beta_j^{(k)} D'(\beta_j)] + \check{\rho}_{ijk}, \quad \text{for } j \neq i. \tag{58}
\]

We define

\[
\check{\Xi}_{ijk,ij'k} = \tilde{t}_{ijk}(x) - \tilde{t}_{ij'k}(x). \tag{59}
\]

From (58) we have

\[
\check{\Xi}_{ijk,ij'k} = \rho_k [D(\beta_j) - D(\beta_{j'})] + \rho_k [\beta_j^{(k)} D'(\beta_j) - \beta_{j'}^{(k)} D'(\beta_{j'})] + \check{\rho}_{ijk} - \check{\rho}_{ij'k}. \tag{60}
\]

Let \( \check{x} \) be a global-class optimum. Denote \( \check{\beta}_i = \beta_i(\check{x}) \).

(1) We first show by contradiction that \( \check{\beta}_j = \check{\beta}_{j'} \) for every pair of \( (j, j') \), which implies that \( \check{\beta}_i = \rho \) for all \( i \).

(1-1) Suppose that \( \check{\beta}_j > \check{\beta}_{j'} \) for some \( j \) and \( j' \). Then there must exist \( k \) such that \( \check{\beta}_j^{(k)} > \check{\beta}_{j'}^{(k)} \). From (56), (57), and (58),

\[
\tilde{t}_{f'k}(x) = \rho_k [D(\check{\beta}_{j'}) + \check{\beta}_{j'}^{(k)} D'(\check{\beta}_{j'})] = \check{\alpha}_{f'k}, \quad \text{for } \check{x}_{f'k} > 0,
\]

\[
\tilde{t}_{f'k}(x) = \rho_k [D(\check{\beta}_{j'}) + \check{\beta}_{j'}^{(k)} D'(\check{\beta}_{j'})] \geq \check{\alpha}_{f'k}, \quad \text{for } \check{x}_{f'k} = 0,
\]

\[
\tilde{t}_{f'k}(x) = \rho_k [D(\check{\beta}_{j'}) + \check{\beta}_{j'}^{(k)} D'(\check{\beta}_{j'})] + \check{\rho}_{f'k}(x) = \check{\alpha}_{f'k}, \quad \text{for } \check{x}_{f'k} > 0,
\]

\[
\tilde{t}_{f'k}(x) = \rho_k [D(\check{\beta}_{j'}) + \check{\beta}_{j'}^{(k)} D'(\check{\beta}_{j'})] + \check{\rho}_{f'k}(x) \geq \check{\alpha}_{f'k}, \quad \text{for } \check{x}_{f'k} = 0.
\]

Therefore, from the fact that \( D \) and \( D' \) are, respectively, increasing and nondecreasing functions (Assumption P1) and from Property (54), we have \( \check{x}_{f'k} = 0 \) and consequently \( \check{x}_{f'k} \leq \check{x}_{ij'k} \).

(1-2) Suppose we have \( \check{x}_{ijk} > \check{x}_{ij'k} \) for some \( i \neq j, j' \), then necessarily \( \check{\rho}_{ijk} \geq \check{\rho}_{ij'k} \) by Property (54). Since, by (P1), \( D \) and \( D' \) are, respectively, increasing and nondecreasing, \( \check{\Xi}_{ijk,ij'k}(\check{x}) > 0 \). However, from (56), since \( \check{x} \) is a global-class optimum we have

\[
\check{\Xi}_{ijk,ij'k} \leq 0, \tag{61}
\]

which contradicts the above. Consequently,

\[
\check{x}_{ijk} \leq \check{x}_{ij'k} \quad \text{for all } i.
\]

Therefore, from \( \check{\beta}_j^{(k)} > \check{\beta}_{j'}^{(k)} \), we must have

\[
\check{x}_{jj'k} + \check{x}_{f'k} > \check{x}_{f'k} + \check{x}_{ij'k}.
\]

Thus from (1-1)

\[
\check{x}_{f'k} \geq \check{x}_{f'k} \text{ and } \check{x}_{jj'k} > \check{x}_{ij'k}.
\]

(1-3) Since

\[
\rho_k [D(\check{\beta}_{j'}) + \check{\beta}_{j'}^{(k)} D'(\check{\beta}_{j'})] = \check{\alpha}_{jk}.
\]
\[ \rho_k [D(\check{\beta}_f) + \check{\beta}^{(k)}_f D'(\check{\beta}_f)] \geq \check{\alpha}_{f,k}, \]

we have \( \check{\alpha}_{jk} > \check{\alpha}_{jk}. \)

We next show that \( \check{x}_{jk} \geq \check{x}_{f_{jk}} \), for all \( l \neq j, j' \), by contradiction. Assume \( \check{x}_{jk} < \check{x}_{f_{jk}}. \) Then \( \check{x}_{f_{jk}} > 0 \), and we have from (56), (57), and (58),

\[ \rho_k [D(\check{\beta}_l) + \check{\beta}^{(k)}_l D'(\check{\beta}_l)] + \check{g}_{f_{jk}}(\check{\xi}) = \check{\alpha}_{f,k}, \]

\[ \rho_k [D(\check{\beta}_l) + \check{\beta}^{(k)}_l D'(\check{\beta}_l)] + \check{g}_{j_{jk}}(\check{\xi}) \geq \check{\alpha}_{jk} > \check{\alpha}_{f,k}, \]

which contradicts the assumption, as we see by noting that \( \check{g}_{j_{jk}}(\check{\xi}) \leq \check{g}_{f_{jk}}(\check{\xi}) \) for both of G-I and G-II. Therefore we must have

\[ \check{x}_{jk} \geq \check{x}_{f_{jk}}. \]

From this and (1-2),

\[ \check{x}_{jk} > \check{x}_{f_{jk}}, \]
\[ \check{x}_{j_{jk}} \geq \check{x}_{f_{jk}}, \]
\[ \check{x}_{jk} \geq \check{x}_{f_{jk}} \text{ for all } l(\neq j, j'). \]

This implies

\[ 1 = \sum_l \check{x}_{jk} > \sum_l \check{x}_{f_{jk}} = 1, \]

which is impossible. Thus we see that the above assumption leads to a contradiction. Therefore necessarily \( \check{\beta}_j = \check{\beta}_f \), and consequently \( \check{\beta}_l = \rho \) for all \( i \).

(2) We next show, by contradiction, that \( \check{\beta}^{(k)}_j = \check{\beta}^{(k)}_f \) for every pair \( (j, j') \), which implies that \( \check{\beta}^{(k)}_j = \rho_k \) for all \( i, k \).

From (60) we have for all \( i, j(\neq i), j'(\neq i), k \),

\[ \check{\xi}_{ij_{jk}:j_{fjk}}(\check{\xi}) = \rho_k (\check{\beta}^{(k)}_j - \check{\beta}^{(k)}_f) D'(\check{\beta}_f) + \check{g}_{ij_{jk}}(\check{\xi}) - \check{g}_{j_{fjk}}(\check{\xi}). \] (62)

Assume \( \check{\beta}^{(k)}_j > \check{\beta}^{(k)}_f \) for some \( (j, j') \). We can follow the same line of logic as (1-1), (1-2), and (1-3) above,

even though \( \check{\beta}_l = \rho \) for all \( i \). We see that the above assumption leads to a contradiction. Therefore necessarily \( \check{\beta}^{(k)}_j = \check{\beta}^{(k)}_f \), and consequently \( \check{\beta}^{(k)}_f = \rho_k \) for all \( i, k \).

(3) Now from (60) we have for all \( i, j(\neq i), j'(\neq i), k \),

\[ \check{\xi}_{ij_{jk}:j_{fjk}}(\check{\xi}) = \check{g}_{ij_{jk}}(\check{\xi}) - \check{g}_{j_{fjk}}(\check{\xi}). \] (63)

Thus, if \( \check{x}_{ijk} > \check{x}_{i_{fjk}} \) for some \( i, j(\neq i), j'(\neq i), k \), we have \( \check{\xi}_{ij_{jk}:j_{fjk}} \geq 0 \), which contradicts relation (61). Therefore,

\[ \check{x}_{ijk} = \check{x}_k \text{ for all } i, j(\neq i), k, \]

and from (58) and (60) we have for all \( i, j(\neq i), k \),

\[ \check{\xi}_{ij_{jk}:i_{fjk}}(\check{\xi}) = \check{g}(\check{\xi}) > 0. \]
Consequently, using (56) we have \( \bar{x}_k = 0 \) for all \( k \). That is,

\[
\bar{x}_{-iik} = 0, \text{ i.e., } \bar{x}_{ijk} = 0, \text{ and } \bar{x}_{iik} = 1, \text{ for all } i, j(\neq i), k.
\]

The mean response time is obtained by noting the definitions (1), (6), (7), (8), (9), and (10). □

Now we have the following property.

**Theorem 3.1** Consider a symmetric system that satisfies all the assumptions given in this paper.

No Braess-like paradox occurs for decision schemes (A), (B) and (C-II).

In decision scheme (C-I), a necessary and sufficient condition for the appearance of a Braess-like paradox is that there exists a job type-\( k \) for which \( \Gamma_k D'(\rho) > 1 \).

**Remark 3.2** We see that under local-class optimization (C-I), whether or not the Braess-like paradox occurs depends on the system load \( \rho = \sum \rho_k \), and the value of \( \Gamma_k = \left( \rho_k^2 / \nu_k = \phi_k \omega_k / \mu_k^2 \right) \) for each local-class \( R_{ik} \), where \( \phi_k \) is the arrival rate, \( \mu_k^{-1} \) is the processing time requirement, and \( \omega_k^{-1} \) is the communication time requirement for local-class \( R_{ik} \). Whether or not the paradox occurs is independent of the number, \( m \), of nodes and of the shape of mean communication time function, \( G_{ijk}(\cdot) \), except for \( \omega_k \) and its monotonicity, convexity and continuous differentiability. When the paradox occurs, however, the shape of mean communication time function, \( G_{ijk}(\cdot) \) is needed to obtain the solution of the decision problem with the present analysis. □

## 4 Examples

We consider here Examples 1 and 2 introduced in Sections 2 and 3, respectively. We restrict ourselves to the local-class optimization (C-I) where \( \mu_k = \mu \) and \( \phi_k = 1/n \), for all \( k \). We have \( \rho = 1/\mu \).

### 4.1 Example 1

In this case we have \( \omega_k^{-1} = t, \tilde{g}_k(x) = t/n \), and \( D(\rho) = 1/(1 - \rho) \). We therefore note that

\[
\Gamma_k D'(\rho) - 1 = \frac{\phi_k \omega_k}{\mu_k^2} D'(\rho) - 1 = \frac{1}{n(\mu - 1)^2} - 1.
\]

(i) If \( \Gamma_k D'(\rho) \leq 1 \), i.e., \( t \geq 1/[n(\mu - 1)^2] \), then the solution \( \bar{x} \) is unique and given by

\[
\bar{x}_{-iik} = 0, \text{ i.e., } \bar{x}_{ijk} = 0, \quad \bar{x}_{iik} = 1, \quad \text{for all } i, j(\neq i), k.
\]

The mean response time is

\[
T(\bar{x}) = T_{ik}(\bar{x}) = \mu_k^{-1} D(\rho) = \frac{1}{\mu - 1}, \quad i = 1, 2, \cdots, m, \quad k = 1, 2, \cdots, n.
\]

In particular it is the same for the overall, individual, and global-class optima.
(ii) If \( \Gamma_k D' (\rho) > 1 \), i.e., \( t < 1/(n(\mu - 1)^2) \), from (22) and (23), the solution \( \tilde{x} \) is given by

\[
\tilde{x}_{ijk} = \frac{1}{m} [1 - nt(\mu - 1)^2], \quad \tilde{x}_{iik} = \frac{1}{m} [1 + (m - 1)nt(\mu - 1)^2], \quad \text{for all } i, j(\neq i), k.
\]

The mean response time is

\[
T(\tilde{x}) = T_0(\tilde{x}) = \frac{1}{\mu - 1} + \frac{m - 1}{m} t[1 - nt(\mu - 1)^2], \quad \text{for all } i, k.
\]

For some parameters \((\mu, m, n)\), \( \tilde{T} = T(\tilde{x}) \) attains its maximum in \( t \) (i.e., the worst performance), that we denote \( \tilde{T}_{\max}(\mu, m, n) \) for \( t_{\max} = 1/2n(\mu - 1)^2 \).

We have

\[
\tilde{T}_{\max}(\mu, m, n) = \frac{1}{\mu - 1} [1 + \frac{m - 1}{4mn(\mu - 1)^2}].
\]

Thus, if we add the communication lines with delay \( t_{\max} = 1/[2n(\mu - 1)^2] \) to the system that has had no communication means, the mean response time, \( T_0(\tilde{x}) \), for each local class, increases in the amount of \( \frac{m - 1}{4mn(\mu - 1)^2} \) (i.e., the performance degrades). This is a Braess-like paradox. We define the worst ratio of the performance degradation \( \Delta(\mu, m, n) \) in the paradox for given \( \mu, n \) as

\[
\Delta(\mu, m, n) = \frac{\tilde{T}_{\max}(\mu, m, n) - T_0(\mu)}{T_0(\mu)}
\]

where \( T_0(\mu) = 1/(\mu - 1) \) denotes the mean response time of each local-class jobs for given \( \mu \) when the system has no communication means. We have

\[
\Delta(\mu, m, n) = \frac{m - 1}{4mn(\mu - 1)}. \quad (69)
\]

**Remark 4.1** From the above we see that, no forwarding of jobs occurs in the overall, individual, and global-class optima and in case (i) of the local-class optimum. That is, in those optima, jobs arriving at each node are processed only by the node. Thereby the system has no performance improvement or degradation due to adding the communication means (which is not used).

On the other hand, in case (ii) of the local-class optimum, each local class forwards a part of its jobs through the communication means to other nodes for remote processing, and thereby has degradation in its mean response time. The ratio of such degradation increases without bound as the total arrival rate approaches the processing capacity of each node, i.e., \( \mu = 1 \). Consider Example 1 with a fixed number of nodes, \( m \). We see that as \( n \) (and consequently the number of local classes, \( m \times n \)) increases to infinity, the ratio of degradation \( \Delta(\mu, m, n) \) tends to zero. This is in accordance with what we state in the middle of the introduction, i.e., a class optimum (C) approaches an individual optimum (B) when the number of players becomes infinitely many (\( N \to \infty \)).

We note that in Example 1, we may obtain an upper bound of 2 on the ratio of performance degradation by considering the limit as \( m \) tends to infinity. \( \square \)
4.2 Example 2

We have
\[ \tilde{g}(x) = \frac{\theta - (m - 1)(m - 1/n)x}{n(\theta - m(m - 1)x)^2}, \]
and
\[ \mu^{-2}D'(\rho) - n^2\tilde{g}(0) = 1/(\mu - 1)^2 - n/\theta. \]

This is the same as in the above example if we take \( t = \theta^{-1} \). Therefore, we have the following:

(i) If \( \theta^{-1} > 1/[n(\mu - 1)^2] \), we obtain the same solution as in Example 1, i.e., \( \tilde{x}_{-i(i)} = 0 \) and \( T_{ik}(\tilde{x}) = 1/(\mu - 1) \).

(ii) If \( \theta^{-1} \leq 1/[n(\mu - 1)^2] \), the solution \( \tilde{x} \) is given by
\[ \tilde{x}_{ijk} = \tilde{x}, \quad \tilde{x}_{iik} = 1 - (m - 1)\tilde{x}, \quad \text{for all } i, j(\neq i), k, \quad (70) \]
where \( \tilde{x} \) satisfies
\[ \frac{1}{n}(1 - m\tilde{x}) \frac{1}{\mu - 1} = \frac{\theta - (m - 1)(m - 1/n)\tilde{x}}{n(\theta - m(m - 1)\tilde{x})^2}. \quad (71) \]

Remark 4.2 In this example also, we see that, no forwarding of jobs occurs in the overall, individual, and global-class optima and in case (i) of the local-class optimum. That is, in those optima, jobs arriving at each node are processed only by the node. Thereby the system has no performance improvement or degradation due to adding the communication means (which is not used). On the other hand, in case (ii) of the local-class optimum, each local class forwards a part of its jobs through the communication means to other nodes for remote processing, and thereby has degradation in its mean response time. \( \square \)

5 Numerical Examples

We examine Example 1 with \( m = 5 \), i.e., the system with five nodes, and consider the case: \( \mu = 1.01 \). The mean response time is \( T_0(\mu) = 1/(\mu - 1) = 100 \) in the overall optimum, in the individual optimum, and in the case of no communication line and no forwarding of jobs.

First, we consider the case where \( n = 1 \), i.e., the total number of local classes \( R_{i\lambda} \) is 5. \( T = T_{i\lambda} \) takes its maximum value
\[ T(1.01, 5, 1) = 2100 \quad (\text{see (67)}), \]
\( i.e., \) the mean response time increases to 2100 from 100 in the case of no communication line, and the worst ratio of the performance degradation \( \Delta(\mu, m, n) \) in the paradox is
\[ \Delta(1.01, 5, 1) = 20 \quad (\text{i.e., 2000\% degradation}) \quad (\text{see (68)}), \]
when \( t = 1/[2(\mu - 1)^2] = 5000 \) (see (66)). Then
\[ \tilde{x}_{ijk} = (1/5)[1 - t(\mu - 1)^2] = 1/10 \quad \text{for all } i, j(\neq i), k \) (see (64)).
That is, if lines giving expected job forwarding time 5000 are added to the system, 10% of jobs that arrive at each node are forwarded to another node with remaining 60% are processed at the arrival node, thereby the mean response time is 2000% worse than that of the case of no communication line.

In this case, as $t$ increases from 0 to 10000 ($= 1/(\mu - 1)^2$), the job forwarding ratio $\bar{x}_{ijk} (= \bar{x}_k)$ decreases from 1/5 down to 0, and for $t > 10000$, no forwarding of jobs occurs.

Consider the case where the communication delay is $t = 5000$. In this particular example, it is surprising that each local class keeps to forward a part of its jobs equally to the other nodes, even though the communication delay for forwarding, $t = 5000$, itself is much greater than the processing delay at the node, $T_{ik}(\mu) = 100$, at which its jobs arrive.

Then we consider the case where $n = 100$, i.e., the total number of local classes is 500. $T = T_{ik}$ takes its maximum value

$$\bar{T}(1.01, 5, 100) = 120 \text{ (see (67))},$$

and the worst ratio of the performance degradation $\Delta(\mu, m, n)$ in the paradox is

$$\Delta(1.01, 5, 100) = 0.2 \text{ (i.e., 20% degradation) \ (see (68))},$$

when $t = 1/[2n(\mu - 1)^2] = 50 \text{ (see (66))}$. Then

$$\bar{x}_{ijk} = (1/5)[1/n - t(\mu - 1)^2] = 1/1000, \quad \text{for all } i, j(\neq i), k \text{ (see (64))}.$$

In this case, as $t$ increases from 0 to 100 ($= 1/(\mu - 1)^2$), the job forwarding ratio $\bar{x}_{ijk} (i \neq j)$ decrease from 1/500 down to 0, and for $t > 100$, no forwarding of jobs occurs.

Thus we see that the chances of paradoxes and the magnitude in the performance degradation in the paradox are greatly reduced compared to the case $n = 1$.

We also consider other values of $\mu$ with $n = 1$.

For $\mu = 1.001$, $\Delta(1.001, 5, 1) = 200 \text{ (i.e., 20000% degradation)}$, and

for $\mu = 1.00001$, $\Delta(1.00001, 2, 1) = 20000 \text{ (i.e., 2000000% degradation)}$, etc.

In this way, we see that the worst ratio of the performance degradation $\Delta(\mu, m, n)$ in the paradox increases without bound as $\mu$ approaches 1 with $n = 1$.

6 Concluding Remarks

In this paper, we have examined the model of a system consisting of homogeneous nodes with identical arrival processes. In such a system, forwarding of jobs to the other nodes through communication means with nonzero delays may clearly lead to the degradation of performance. We have confirmed that in the overall optimization and in the individual optimization such forwarding never occurs. We have shown that, for some parameter setting of the local-class optimization (Nash equilibrium for one set of decision makers), there is mutual forwarding of jobs for remote processing, leading to an increase in the degree of performance degradation
which may increase without bound, as observed in Example 1. We have also shown that in the
global-class optimization (Nash equilibrium for another set of decision makers) such forwarding
never occurs. That is, such a paradoxical behavior may occur only in the local-class optimum
and does never occur for the overall, individual, and global-class optima, in the same setting of
this symmetrical node model.

We have obtained unique solutions of local-class and global-class optima on the basis of
some assumptions on communication means (i.e., dedicated lines and bus-type connections).
It has been quite hard to extend the proofs to the case of more general assumptions. It is not
certain whether in some cases of the communication means the optima may still be unique. It
has been also difficult for us to analyze asymmetrical models. These are open future problems.

In the system of symmetrical nodes, adding means of job forwarding looks apparently inef-
tective. We thus feel it counter-intuitive that in some class optimum, adding the means causes
mutual job forwarding among nodes and brings about the paradox, in such a symmetrical sys-
tem. Moreover, we think that the phenomenon described in this paper is an example of “local
decision making which does not lead to a globally best decision,” which a very large number of
people still do not seem to recognize.

Appendix: Glossary of symbols

- \( m \) — the number of nodes (host computers or processors)
- \( n \) — the number of job types
- \( x_{ik} \) — the ratio of jobs that are processed at node \( i \), out of type-\( k \) jobs arriving at node \( i \)
- \( x_{ijk} \) (\( j \neq i \)) — the ratio of jobs that are forwarded upon arrival through the communication
  means to another node \( j \) to be processed there, out of type-\( k \) jobs arriving at node \( i \). Thus
  \( \sum_j x_{ijk} = 1 \) and \( 0 \leq x_{ijk} \leq 1 \), for all \( i, j, k \).
  That is, the rate \( \phi_k x_{ijk} \) of type-\( k \) jobs that arrive at node \( i \) is forwarded through the com-
munication means to node \( j \), while the rate \( \phi_k x_{ik} \) of local-class \( R_{ik} \) jobs is processed at
the arrival node \( i \).
- \( x_{ik} = (x_{i1k}, \cdots, x_{imk}) \) — \( m \)-dimensional vector that shows ‘local-class \( R_{ik} \) strategy.’
- \( x_{k} = (x_{1k}, x_{2k}, \cdots, x_{mk}) \) — \( mn \)-dimensional vector that shows global-class \( J_k \) strategy
- \( x \) — \( mnn \)-dimensional vector that shows the strategy profile, the vector of strategies con-
cerning all local classes,

\[
x = (x_{11}, x_{12}, \cdots, x_{1n}, x_{21}, \cdots, x_{2n}, \cdots, x_{m1}, \cdots, x_{mn}),
\]

or \( x = (x_1, x_2, \cdots, x_n) \).
- \( C \) — the set of \( x \)’s that satisfy the constraints: \( \sum_i x_{ijk} = 1, x_{ijk} \geq 0 \), for all \( i, j, k \)
- \( D(\cdot) \) — a strictly increasing, convex and continuously differentiable function. \( \mu_k^{-1} D(\beta_i) \)
is the expected processing (including queueing) time of a type-\( k \) job that is processed at
node \( i \) given \( \beta_i \).
• \( G_{ijk}(x) \) — the mean communication delay (including queueing delay) or the cost for forwarding type-\( k \) jobs arriving at node \( i \) to node \( j \) \((i \neq j)\), given a strategy profile \( x \). It is a positive, nondecreasing, convex and continuously differentiable function of \( x \). \( G_{ijk}(x) = 0 \).

• \( G(\cdot) \) — a nondecreasing, convex, and differentiable function. \( G(0) = 1 \)

• \( J_k \) — the global class that consists in the collection of local class \( R_{ik} \), \( i.e. \), \( J_k = \bigcup_i R_{ik} \). It thus consists of all jobs of type \( k \).

• \( R_{ik} \) — the local class that consists of type-\( k \) jobs that arrive at node \( i \).

• \( T_{ik} \) — the mean response time of a local-class \( R_{ik} \) job

• \( T_k \) — the mean response time of a global-class \( J_k \) job

• \( T \) — the overall mean response time of a job that arrives at the system

• \( \beta_i \) — load on node \( i \), \( i.e. \), \( \beta_i = \beta_i(x) = \sum_{j,k} \rho_k x_{ijk} \) [eq.(1)].

• \( \beta_i^{(k)} \) — contribution on the load of node \( i \) by type-\( k \) jobs, \( i.e. \),

\[
\beta_i^{(k)}(x) = \sum_j \rho_k x_{ijk} \text{ [eq.(2)]} \quad \text{and} \quad \beta_i = \sum_k \beta_i^{(k)}.
\]

• \( \mu_k \) — service rate of type-\( k \) jobs by each node. \( \mu_k^{-1} \) is the average processing (service) time (without queueing delays) of a type-\( k \) job at any node and is node-independent.

• \( \rho_k = \phi_k / \mu_k \)

• \( \rho = \sum_k \rho_k \).

• \( \sigma_k \) — coefficient of job forwarding rate as to traffic intensity.

• \( u_k = \phi_k / \omega_k \)

• \( \phi_k \) — node-independent rate at which jobs of type \( k \) arrive at each node

• \( \phi \) — the total arrival rate at each node \( (\phi = \sum_k \phi_k) \). We have the time scale whereby \( \phi = 1 \).

• \( \omega_k \) — a constant concerning type-\( k \) jobs. \( \omega_k^{-1} \) can be regarded as the mean communication time (without queueing delays) for forwarding a type-\( k \) job from the arrival node to another processing node under the assumption \( \Pi 3 \).

• \( \Gamma_k = \rho_k^2 u_k^{-1} \)

**Acknowledgments**

The authors thank the editor for careful consideration and anonymous referees for very careful reviewing of the manuscript and very constructive and detailed comments improving this paper. Thanks are also due Eitan Altman for helpful comments and cooperation relating to this problem.
References


27


