UNIQUENESS OF SOLUTIONS FOR OPTIMAL STATIC ROUTING IN MULTI-CLASS OPEN NETWORKS

Eitan Altman* and Hisao Kameda†

October 27, 1998

ISE-TR-98-156

*INRIA, BP93
2004, Route des Lucioles,
06902 Sophia Antipolis Cedex, France
Phone: +33-4-92-38-77-86, Fax: +33-4-93-65-78-58, Email: Eitan.Altman@sophia.inria.fr

†Institute of Information Sciences and Electronics
University of Tsukuba
Tsukuba Science City, Ibaraki 305-8573, Japan
Phone: +81-298-53-5539, Fax: +81-298-53-5206, Email: kameda@is.tsukuba.ac.jp

This work was supported in part by the Japan Society for the Promotion of Science and in part by the Telecommunications Advancement Organization of Japan.

UNIQUENESS OF SOLUTIONS FOR OPTIMAL STATIC ROUTING IN MULTI-CLASS OPEN NETWORKS

Eitan Altman
INRIA, BP93
2004, Route des Lucioles,
06902 Sophia Antipolis Cedex
France

Hisao Kameda
Institute of Information Sciences
and Electronics
University of Tsukuba
Tsukuba Science City, Ibaraki 305, Japan

Abstract

We study in this paper optimal static routing problems in open multiclass networks with state-independent arrival and service rates. They include static routing problems in communication networks, optimal static load balancing problems in distributed computer systems, and road transportation networks. The goal of the paper is to study the uniqueness of optimal routing under different scenarios. We consider first the overall optimal policy that is the routing policy whereby the overall mean cost (which may represent the average delay) of a job is minimized. We then consider an individually optimal policy whereby jobs are routed so that each job may feel that its own expected cost is minimized if it knows the mean cost for each path. This is thus related to the Wardrop equilibrium concept in a multiclass framework. We finally study the case of class optimization, in which each of several class of jobs tries to minimize the averaged cost per job within that class; this is related to the Nash equilibrium concept. For all three settings, we show that the routing decisions at optimum need not be unique, but that the utilizations in some large class of links are uniquely determined.

Keywords: Routing, Networks, Traffic control (communication), Game theory

1 Introduction

We consider in this paper the problem of optimally routing in networks. This is a typical problem in communication networks, in distributed computer systems and in transportation networks.

Much previous work has been devoted to the routing problem in which at each node one may take new routing decisions. We consider a more general framework in which the sources have to decide how to route their traffic between different existing paths. (These two problems coincide in the case where the set of paths equals to the set of all possible sequences of consecutive directed links which originate at the source and end at the destination.)

In ATM environment, this problem arises when we wish to decide on how to route traffic on a given existing set of virtual paths or virtual connections. Our framework thus allows to handle routing both in a packet switching as well as in a circuit switching environment.

We consider three different frameworks.

The first one is the global optimization criterion, where a single controller makes the routing decisions. This framework is typical for some telecommunication problems (e.g. telephony). Extensive literature exists for this approach, both in telecommunication applications as well as in load balancing for distributed computer systems [3, 5, 8, 9, 10, 12, 13, 22, 23].

The second approach to optimization is that of individual optimality, in which each routed individual (an individual could be an information packet, a job, or a car in a road traffic context) chooses its own path so as to minimize its own cost. This framework has been extensively investigated in transportation science, see [4, 19, 6], and was also considered in the context of telecommunication [11] and in distributed computing [9, 10, 11]. The suitable optimization concept for this setting is of Wardrop equilibrium [24]; it is defined as a set of routing decisions for all individuals such that a path is followed by an individual if and only if it has the lowest cost for that individual.

The third approach is that of class optimization. A class may correspond to a service provider in a telecommunication context, or to a transportation company in a road traffic context. Each class wishes to minimize the average cost per individual, averaged over all individuals within that class; there is thus a single entity per each class which takes the routing decisions for all individuals of that class. The suitable optimization concept for this approach is that of Nash (or Nash-Cournot) equilibrium [6]; it is defined as a set of routing decisions for the different classes such that no class can decrease its own cos by unilaterally deviating from its decision. This approach was used in telecommunication applications in [1, 20, 14, 15, 16, 18], in load balancing problems in distributed computer systems [17, 9, 10] and in transportation science in [6].

An optimization problem does not necessarily have a unique solution. For example, it is possible that the combinations of different values of data flow rates for all paths in a telecommunication network may result in the same minimum overall mean cost (e.g. delay). If they are not unique, it is necessary to make clear the range and characteristics of the solutions, in particular, when we calculate numerically the optimal solutions and when we intend to analyze the effects of the system parameters on the optimal solutions.

Kameda and Zhang [11] already studied the first two approaches (global and individual optimization) in [11] and characterized the uniqueness for a particular cost structure, that of open BCMP queueing networks [2] with state-independent arrival and service rates. We extend here these results to a fairly general cost function. We also extend some results obtained in [20] for the uniqueness of class optimization.

In Section 2 we give some definitions used in this paper and formulate a mathematical model of the network. In Section 3 we obtain the overall optimal solution, and discuss the uniqueness of the overall optimal solution. In Section 4 we show similar results on the uniqueness of the individually optimal solution. Some results on uniqueness for class optimization are finally presented in Section 5. A counter-example for the uniqueness for the individual as well as the class optimal problem is given in Section 6.

2 Notation and assumptions

We consider an open network model that consists of a set $\mathcal{L}$ containing M links. We assume that in the network there are pairs of origin and destination points. We call the pair of one origin and one destination points an *O-D pair*. The unit entity (like a message or packet) that is routed through the network is called a *job*. Each job arrives at one of the origin points and departs from one of the destination points. The origin and destination points of a job are determined when the job arrives in the network.

Jobs are classified into R different classes. For the sake of simplicity, we assume that jobs do not change their class while passing through the network. A class k job may have one of different origin-destination pairs. Such a class may represent all the users of a given service provider in a context of telecommunications. In the context of road traffic it may represent the set of vehicles of a given type, such as busses, or trucks, or bicycle, etc. With this in mind, it is natural to expect that jobs of different classes are faced with different types of routing decisions.

A class k job with the O-D pair (o_1, d_1) originates at node o_1 and destinates for node d_1 through a series of links, which we refer to as a *path*, and then leaves the system.

In many previous papers [1, 20, 14, 15, 16, 18], routing could be done at each node. In this paper we follow the more general approach of [11] in which a job of class k with O-D pair (o_1, d_1) has to choose one of a given finite set of paths. We call this set the *path-classes* of job class k O-D pair (o_1, d_1).

We assume that we can choose the job flow rate of each *path class* in order to achieve a performance objective. A path class may be a given sequence of links that relate the origin and destination nodes. In that case, it may correspond to a virtual connection in ATM. We allow, however, a path class to be some more general object. It may contain a number of sub-paths; we assume however that once the job flow rate of a path class is given, the job flow rate of each sub-path in the path class is fully determined (and is not the object of a control decision). That is, the relative flow rate of each sub-path in the same path-class is governed by some fixed transfer proportions (or probabilities) between the links.

As an example, one may consider path classes that include noisy links in which case lost packets have to be retransmitted locally over the link. Thus, some given proportion of the traffic in this path class use a direct route (no losses) whereas other have to loop (this models losses and retransmissions). Another example are switches that might route arriving traffic in some fixed proportions between outgoing links, in a way that is not controlled by the the entity that takes routing decision for that class.

The solution of a routing problem is characterized by the chosen values of job flow rates of all path classes.

We will use the following notation regarding the network:

$D^{(k)}$ = Set of O-D pairs for class k jobs.

$\Pi_d^{(k)}$ = Set of path classes that class k jobs of O-D pair d flow through.

$\Pi^{(k)}$ = Set of all path classes for class k jobs, i.e., $\Pi^{(k)} = \bigcup_{d \in D^{(k)}} \Pi_d^{(k)}$.

$$\gamma_{pd}^{kk'} = \begin{cases} 1 & \text{if } p \in \Pi_d^{(k')} \text{ and } k = k', \\ 0 & \text{otherwise.} \end{cases}$$

We will use the following notation regarding arrivals to the network and flow rates:

$\phi_d^{(k)}$ = Rate at which class k jobs join O-D pair $d \in D^{(k)}$.

$\phi^{(k)}$ = Total job arrival rate of class k jobs, i.e., $\phi^{(k)} = \sum_{d \in D^{(k)}} \phi_d^{(k)}$.

Φ = System-wide total job arrival rate, i.e., $\Phi = \sum_{k=1}^{R} \phi^{(k)}$.

$x_p^{(k)}$ = Rate at which class k jobs flow through path-class p.

δ_{lp} = Rate at which jobs that flow through path-class p pass through link l assuming that $x_p^{(k)} = 1$, $p \in \Pi^{(k)}$.

$\lambda_l^{(k)}$ = Rate at which class k jobs visit link l, $\lambda_l^{(k)} = \sum_{p \in \Pi^{(k)}} \delta_{lp} x_p^{(k)}$.

We will use the following notation regarding the service and performance values in the open network:

$\mu_l^{(k)}$ = Service rate of class k jobs at link l.

$\rho_l^{(k)}$ = $\lambda_l^{(k)} / \mu_l^{(k)}$. Utilization of link l for class k jobs.

ρ_l = $\sum_{k=1}^{R} \rho_l^{(k)}$. Total utilization of link l.

$\hat{T}_l^{(k)}$ =Mean cost of class k jobs at link l.

$T_l(\rho_l)$ = Weighted cost per unit flow in link l.

$T_p^{(k)}$ =Average class k cost of path-class p, $p \in \Pi^{(k)}$, $k = 1, 2, \ldots, R$.

Δ =Overall mean cost of a job (averaged over all classes).

$\Delta^{(k)}$ =Overall mean cost of a job of class k.

$t_p^{(k)}$ $= \partial(\Phi\Delta)/\partial x_p^{(k)}$, i.e., class k marginal cost of path-class p, $p \in \Pi^{(k)}, k = 1, 2, \ldots, R$.

$\tilde{t}_p^{(k)}$ $= \partial(\Phi\Delta^{(k)})/\partial x_p^{(k)}$, i.e., class k marginal class-cost of path-class p, $p \in \Pi^{(k)}, k = 1, 2, \ldots, R$.

We will use the following notation regarding vectors and matrices:

$\boldsymbol{\rho} = [\rho_1, \rho_2, \ldots, \rho_M]^{\mathrm{T}}$ where T means 'transpose'. We call this the utilization vector.

$\boldsymbol{\lambda} = [\lambda_1^{(1)}, \lambda_2^{(1)}, \ldots, \lambda_M^{(1)}, \ldots, \lambda_1^{(k)}, \lambda_2^{(k)}, \ldots, \lambda_M^{(k)}, \ldots,]^{\mathrm{T}}$, i.e, the vector of total flows over all links.

$\boldsymbol{\phi}$ $=[\phi_1^{(1)}, \phi_2^{(1)}, \ldots, \phi_1^{(2)}, \phi_2^{(2)}, \ldots]^{\mathrm{T}}$, i.e., the arrival rate vector.

$\mathbf{x}$ $=[x_1^{(1)}, x_2^{(1)}, \ldots, x_1^{(2)}, x_2^{(2)}, \ldots]^{\mathrm{T}}$, i.e., the path-class flow rate vector.

$\boldsymbol{\alpha}$ $=[\alpha_1^{(1)}, \alpha_2^{(1)}, \ldots, \alpha_1^{(2)}, \alpha_2^{(2)}, \ldots]^{\mathrm{T}}$, i.e., the vector whose elements are $\alpha_d^{(k)}$, $d \in D^{(k)}$, $k = 1, 2, \ldots, R$; the elements $\alpha_d^{(k)}$ will correspond to some Lagrange multipliers.

$\boldsymbol{\xi}$ $=[\xi_1^{(1)}, \xi_2^{(1)}, \ldots, \xi_1^{(2)}, \xi_2^{(2)}, \ldots]^{\mathrm{T}}$, i.e., the vector whose elements are $\xi_l^{(k)}$, $l \in \mathcal{L}$, $k = 1, 2, \ldots, R$; the elements $\xi_l^{(k)}$ will correspond to some Lagrange multipliers.

$\mathbf{A}$ $=[A_1^{(1)}, A_2^{(1)}, \ldots, A_1^{(2)}, A_2^{(2)}, \ldots]^{\mathrm{T}}$, i.e., the vector whose elements are $A_d^{(k)}$, $d \in D^{(k)}$, $k = 1, 2, \ldots, R$.

$\mathbf{t}$ $=[t_1^{(1)}, t_2^{(1)}, \ldots, t_1^{(2)}, t_2^{(2)}, \ldots]^{\mathrm{T}}$, i.e., the vector whose elements are $t_p^{(k)}, p \in \Pi^{(k)}, k = 1, 2, \ldots, R$.

$\tilde{\mathbf{t}}$ $=[\tilde{t}_1^{(1)}, \tilde{t}_2^{(1)}, \ldots, \tilde{t}_1^{(2)}, \tilde{t}_2^{(2)}, \ldots]^{\mathrm{T}}$, i.e., the vector whose elements are $\tilde{t}_p^{(k)}, p \in \Pi^{(k)}, k = 1, 2, \ldots, R$.

$\mathbf{T}$ $=[T_1^{(1)}, T_2^{(1)}, \ldots, T_1^{(2)}, T_2^{(2)}, \ldots]^{\mathrm{T}}$, i.e., the vector whose elements are $T_p^{(k)}, p \in \Pi^{(k)}$, $k = 1, 2, \ldots, R$.

$\mathbf{x}^{(k)}$ $=[x_1^{(k)}, x_2^{(k)}, \ldots]^{\mathrm{T}}$, i.e., the path-class flow rate vector for class k jobs.

$\mathbf{x}^{-k}$ $=[x_1^{(1)}, x_2^{(1)}, \ldots, x_1^{(k-1)}, \ldots, x_1^{(k+1)}, \ldots]^{\mathrm{T}}$, i.e., the path-class flow rate vector for jobs of the classes other than class k.

$\boldsymbol{\phi}^{(k)}$, $\boldsymbol{\alpha}^{(k)}$, $\mathbf{A}^{(k)}$, $\mathbf{t}^{(k)}$, $\tilde{\mathbf{t}}^{(k)}$, and $\mathbf{T}^{(k)}$ are defined similarly.

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_{11}^{11} & \gamma_{12}^{11} & \cdots & \gamma_{11}^{12} & \gamma_{12}^{12} & \cdots \\ \gamma_{21}^{11} & \gamma_{22}^{11} & \cdots & \gamma_{21}^{12} & \gamma_{22}^{12} & \cdots \\ \vdots & & \ddots & \vdots & & \ddots \\ \gamma_{11}^{21} & \gamma_{12}^{21} & \cdots & \gamma_{11}^{22} & \gamma_{12}^{22} & \cdots \\ \gamma_{21}^{21} & \gamma_{22}^{21} & \cdots & \gamma_{21}^{22} & \gamma_{22}^{22} & \cdots \\ \vdots & & \ddots & \vdots & & \ddots \end{bmatrix}$$

i.e., the incident matrix whose (i,j) element is $\gamma_{pd}^{kk'}$, $p \in \Pi^{(k)}$, $k = 1,2,\ldots,R$, $d \in D^{(k')}$, $k' = 1,2,\ldots,R$, where $i = p + \sum_{\kappa=1}^{k-1} |\Pi^{(\kappa)}|$ and $j = d + \sum_{\kappa=1}^{k'-1} |D^{(\kappa)}|$.

$\mathbf{\Gamma}^{(k)}$ = Incident matrix whose (i,j) element is γ_{pd}^{kk}, $p, d \in D^k$, $k = 1,2,\ldots,R$, where $i = p + \sum_{\kappa=1}^{k-1} |\Pi^{(\kappa)}|$ and $j = d + \sum_{\kappa=1}^{k-1} |D^{(\kappa)}|$.

$\mathbf{x} \cdot \mathbf{y} = \sum_i x_i y_i$, i.e., the inner product of vectors $\mathbf{x} = [x_1, x_2, \ldots]^{\mathrm{T}}$ and $\mathbf{y} = [y_1, y_2, \ldots]^{\mathrm{T}}$.

We make the following assumptions on the cost:

B1: Associated with each link $l \in \mathcal{L}$ there is a cost $T_l(\rho_l)$ per flow unit, which is further weighted by the class dependent factor $\mu_l^{(k)}$. Thus, the cost per unit flow of class k on link l is given by $\hat{T}_l^{(k)} = T_l/\mu_l^{(k)}$. Thus the average cost per unit flow of class k job that passes through path-class $p \in \Pi^{(k)}$ is

$$T_p^{(k)} = \sum_{l=1}^{M} \delta_{lp} \hat{T}_l^{(k)} = \sum_{l \in \mathcal{L}} \frac{\delta_{lp}}{\mu_l^{(k)}} T_l(\rho_l) \tag{1}$$

B2: $T_l : [0, \infty) \to [0, \infty]$, and $T_l(0)$ is finite.

B3: The set $\mathcal{L}$ of links contains two types of links:
(i) the type N for which $T_l(\rho_l)$ are convex and strictly increasing,
(ii) the type I for which $T_l(\rho_l) = T_l$ are constant (and do not depend on ρ_l).

B4: $T_l(\rho_l)$ are continuous. Moreover, they are continuously differentiable whenever they are finite.

Denote

$\boldsymbol{\rho}_U = \boldsymbol{\rho}|_{\rho_l=0, l \in I}$. This is the same as $\boldsymbol{\rho}$ except that $\rho_l = 0$ for all $l \in I$.

The overall mean cost of a job, Δ, can be written as

$$\Delta = \sum_{k=1}^{R} \sum_{p \in \Pi^{(k)}} \frac{x_p^{(k)}}{\Phi} T_p^{(k)} = \frac{1}{\Phi} \sum_{l \in \mathcal{L}} \rho_l T_l(\rho_l).$$

The mean cost of a job of class k, $\Delta^{(k)}$, can be written as

$$\Delta^{(k)} = \sum_{p \in \Pi^{(k)}} \frac{x_p^{(k)}}{\phi^{(k)}} T_p^{(k)} = \frac{1}{\phi^{(k)}} \sum_{l \in \mathcal{L}} \rho_l^{(k)} T_l(\rho_l).$$

Note that the following conditions should be satisfied.

$$\sum_{p\in\Pi_d^{(k)}} x_p^{(k)} = \phi_d^{(k)}, \quad d \in D^{(k)}, \quad k = 1, 2, \ldots, R, \tag{2}$$

$$x_p^{(k)} \geq 0, \; p \in \Pi^{(k)}, k = 1, 2, \ldots, R. \tag{3}$$

We can express (2) as

$$\sum_{k'=1}^{R} \sum_{p\in\Pi^{(k')}} \gamma_{pd}^{k'k} x_p^{(k')} = \phi_d^{(k)}, \; d \in D^{(k)}, \; k = 1, 2, \ldots, R,$$

or, equivalently,

$$\mathbf{\Gamma}^{\mathrm{T}}\mathbf{x} = \boldsymbol{\phi}. \tag{4}$$

Remark We easily see that our model includes those discussed for the static routing problems of communications networks [3, 4, 7]. We also see that our model includes those of the load balancing problems of distributed computer systems such as given by Tantawi and Towsley [8, 9]. From condition B3 it is easy to see that $T_l^{(k)}$ is a convex function of $\lambda_l^{(k)}, l \in N, \; k = 1, 2, \ldots, R$. It follows that $T_l^{(k)}$ is also convex with respect to $\mathbf{x}$.

3 Uniqueness of the overall optimal solution

By the overall optimal policy we mean the policy whereby routing is determined so as to minimize the overall mean cost of a job. The problem of minimizing the overall mean cost is stated as follows:

$$\text{minimize:} \quad \Delta \;=\; \frac{1}{\Phi}\sum_{l\in\mathcal{L}} \rho_l T_l(\rho_l) \tag{5}$$

with respect to $\mathbf{x}$ subject to

$$\mathbf{\Gamma}^{\mathrm{T}}\mathbf{x} \;=\; \boldsymbol{\phi}, \tag{6}$$

$$\mathbf{x} \;\geq\; 0, \tag{7}$$

where $\rho_l = \sum_{k=1}^{R} \lambda_l^{(k)}/\mu_l^{(k)}$ and $\lambda_l^{(k)} = \sum_{p\in\Pi^{(k)}} \delta_{lp} x_p^{(k)}$. Note that (6) and (7) are the same as (2) and (3), respectively. We call the above problem the *overall optimization problem*, and its solution the *overall optimal solution.*

Lemma 3.1 $\mathbf{x}$ *is an optimal solution of the problem* (5) *if and only if* $\mathbf{x}$ *satisfies the following conditions. There exist Lagrange multipliers* $\boldsymbol{\alpha}$ *such that*

$$[\mathbf{t}(\mathbf{x}) - \mathbf{\Gamma}\boldsymbol{\alpha}] \cdot \mathbf{x} \;=\; 0, \tag{8}$$

$$\mathbf{t}(\mathbf{x}) - \mathbf{\Gamma}\boldsymbol{\alpha} \;\geq\; 0, \tag{9}$$

$$\mathbf{\Gamma}^{\mathrm{T}}\mathbf{x} - \boldsymbol{\phi} \;=\; 0, \tag{10}$$

$$\mathbf{x} \;\geq\; 0. \tag{11}$$

Proof. Since the objective function (5) is convex and the feasible region of its constraints is a convex set, any local solution of the problem is a global solution point. To obtain the optimal solution, we construct the Lagrangian function

$$L(\mathbf{x},\boldsymbol{\alpha}) \;=\; \Phi\Delta + \boldsymbol{\alpha}\cdot(\boldsymbol{\phi} - \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{x}) \tag{12}$$

for (5). By the Kuhn-Tucker theorem, $\mathbf{x}$ is an optimal solution if and only if there exist Lagrange multipliers $\boldsymbol{\alpha}$ such that

$$\frac{\partial L}{\partial \mathbf{x}} \;=\; \mathbf{t}(\mathbf{x}) - \boldsymbol{\Gamma}\boldsymbol{\alpha} \geq 0, \tag{13}$$

$$\frac{\partial L}{\partial \mathbf{x}}\cdot\mathbf{x} \;=\; [\mathbf{t}(\mathbf{x}) - \boldsymbol{\Gamma}\boldsymbol{\alpha}]\cdot\mathbf{x} = 0, \tag{14}$$

$$\frac{\partial L}{\partial \boldsymbol{\alpha}} \;=\; \boldsymbol{\phi} - \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{x} = 0, \tag{15}$$

$$\mathbf{x} \;\geq\; 0, \tag{16}$$

where $(\partial L)/(\partial \mathbf{x})$ denotes the vector whose elements are $(\partial L)/(\partial x_p^{(k)})$, $p \in \Pi^{(k)}$, $k = 1, 2, \ldots, R$. The relations (13) - (16) are exactly the same as the relations (8) - (11). □

Lemma 3.2 $\bar{\mathbf{x}}$ *is an optimal solution of the problem* (5) *if and only if*

$$\mathbf{t}(\bar{\mathbf{x}})\cdot(\mathbf{x} - \bar{\mathbf{x}}) \;\geq\; 0, \quad \text{for all } \mathbf{x} \tag{17}$$

$$\text{such that } \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{x} \;=\; \boldsymbol{\phi} \text{ and } \mathbf{x} \geq 0.$$

Proof. (17) holds for some $\bar{\mathbf{x}}$ if and only if $\bar{\mathbf{x}}$ is the solution of the following linear program (where the decision variables are $\mathbf{x}$):

$$\begin{array}{ll} \text{minimize:} & \mathbf{t}(\bar{\mathbf{x}})\cdot\mathbf{x} \\ \text{subject to} & \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{x} = \boldsymbol{\phi}, \quad \mathbf{x} \geq 0 \end{array}$$

with $\bar{\mathbf{x}}$ fixed. $\mathbf{x}$ is an optimal solution of the linear program if and only if $\mathbf{x}$ satisfies the Kuhn-Tucker conditions (see e.g. pp. 158-165 in [21]) for the Lagrangian

$$L(\mathbf{x},\boldsymbol{\alpha}^*) \;=\; \mathbf{t}(\bar{\mathbf{x}})\cdot\mathbf{x} + \boldsymbol{\alpha}^*\cdot(\boldsymbol{\phi} - \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{x}). \tag{18}$$

The Kuhn-Tucker conditions are

$$\frac{\partial L}{\partial \mathbf{x}} \;=\; \mathbf{t}(\bar{\mathbf{x}}) - \boldsymbol{\Gamma}\boldsymbol{\alpha}^* \geq 0, \tag{19}$$

$$\frac{\partial L}{\partial \mathbf{x}}\cdot\mathbf{x} \;=\; [\mathbf{t}(\bar{\mathbf{x}}) - \boldsymbol{\Gamma}\boldsymbol{\alpha}^*]\cdot\mathbf{x} = 0, \tag{20}$$

$$\frac{\partial L}{\partial \boldsymbol{\alpha}^*} \;=\; \boldsymbol{\phi} - \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{x} = 0, \tag{21}$$

$$\mathbf{x} \;\geq\; 0. \tag{22}$$

That is, the relation (17) (i.e., the statement that $\bar{\mathbf{x}}$ is a solution of the above linear program) is equivalent to the set of relations in Lemma 3.1 for some (finite) Lagrange multiplier $\boldsymbol{\alpha}^*$ (for the finiteness, see Cor. 5.1 p. 165 in [21]). □

From condition B1 we see that Δ depends only on the utilization of each link, ρ_l, which results from the path flow rate matrix. It is possible, therefore, that different values of the path flow rate matrix result in the same utilization of each link and the same minimum mean cost. We are uncertain, however, about whether distinct optimal solutions should have the same utilization of each link or not. We first define the concept of monotonicity of vector-valued functions with vector-valued arguments.

Definition Let $\mathbf{F}(\bullet)$ be a vector-valued function that is defined on a domain $S \subseteq R^n$ and that has values $\mathbf{F}(\mathbf{x})$ in R^n. This function is *monotone* in S if for every pair $\mathbf{x}, \mathbf{y} \in S$

$$(\mathbf{x}-\mathbf{y}) \cdot [\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})] \geq 0.$$

It is *strictly monotone* if, for every pair $\mathbf{x} \in S$ and $\mathbf{y} \in S$ with $\mathbf{x} \neq \mathbf{y}$,

$$(\mathbf{x}-\mathbf{y}) \cdot [\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})] > 0.$$

We need the following property:

Lemma 3.3 *Consider assumptions B1–B4, and let $l \in N$. Then*
(i) $T_l(\rho_l)$ is finite if and only if its derivative $T_l'(\rho_l)$ is finite.
(ii) If $T_l'(\rho_l)$ is infinite then for any $\mathbf{x}$ for which the load on link l is ρ_l, the corresponding cost $\Delta(\mathbf{x})$ is infinite.

Proof. (i) We have

$$T_l(\rho_l) = \int_0^{\rho_l} T_l'(\zeta_l) d\zeta_l \leq \rho_l T_l'(\rho_l).$$

If $T_l(\rho_l) = \infty$ then $\rho_l > 0$, which implies by the latter equation that $T_l'(\rho_l)$ is infinite. To show the converse, assume that $T_l(\rho_l)$ is finite. Then by continuity, there exists some $\epsilon > 0$ such that $T_l(\rho_l + \epsilon)$ is finite. Since T_l is convex,

$$T_l'(\rho_l) \leq \epsilon^{-1}(T_l(\rho + \epsilon) - T_l(\rho_l))$$

and is thus finite as well.

(ii) If $T_l'(\rho_l)$ is infinite then by (i), $T_l(\rho_l)$ is infinite; moreover, $\rho_l > 0$ by assumption B2, so that $\Delta(\rho) = \infty$ (by (5)). □

For the function $\mathbf{t}(\mathbf{x})$ we have the following property.

Lemma 3.4 *Consider assumptions B1–B4. Whenever finite, $\mathbf{t}(\mathbf{x})$ is monotone but is not strictly monotone, i.e., for arbitrary $\mathbf{x}$ and $\mathbf{x}'$ ($\mathbf{x} \neq \mathbf{x}'$), if $\Delta(\mathbf{x})$ or if $\Delta(\mathbf{x}')$ is finite then*

$$(\mathbf{x}-\mathbf{x}') \cdot [\mathbf{t}(\mathbf{x}) - \mathbf{t}(\mathbf{x}')] > 0 \quad \text{if} \quad \boldsymbol{\rho}_U \neq \boldsymbol{\rho}_U', \tag{23}$$

$$= 0 \quad \text{if} \quad \boldsymbol{\rho}_U = \boldsymbol{\rho}_U', \tag{24}$$

where $\boldsymbol{\rho}_U$ and $\boldsymbol{\rho}_U'$ are the utilization vectors that $\mathbf{x}$ and $\mathbf{x}'$ result in, respectively.

Proof. Assume that $\boldsymbol{\rho}_U \neq \boldsymbol{\rho}'_U$. Then

$$
\begin{aligned}
& (\mathbf{x}-\mathbf{x}') \cdot [\mathbf{t}(\mathbf{x}) - \mathbf{t}(\mathbf{x}')] \\
= & \sum_{k=1}^{R} \sum_{p \in \Pi^{(k)}} (x_p^{(k)} - x_p'^{(k)})[t_p^{(k)}(\mathbf{x}) - t_p^{(k)}(\mathbf{x}')] \\
= & \sum_{k=1}^{R} \sum_{p \in \Pi^{(k)}} \sum_{l \in \mathcal{L}} (x_p^{(k)} - x_p'^{(k)}) \frac{\delta_{lp}}{\mu_l^{(k)}} \Big\{ \big[T_l(\rho_l) - T_l(\rho_l') \big] \\
& + \Big[\rho_l \frac{dT_l(\rho_l)}{d\rho_l} - \rho_l' \frac{dT_l(\rho_l')}{d\rho_l'} \Big] \Big\} \\
= & \sum_{l \in N} (\rho_l - \rho_l') \Big\{ \big[T_l(\rho_l) - T_l(\rho_l') \big] + \Big[\rho_l \frac{dT_l(\rho_l)}{d\rho_l} - \rho_l' \frac{dT_l(\rho_l')}{d\rho_l'} \Big] \Big\} > 0 \qquad (25)
\end{aligned}
$$

Due to Lemma 3.3, if $\Delta(\rho)$ is finite then $T_l'(\rho_l)$ is finite for all links $l \in \mathcal{L}$ (and similarly for $\Delta(\rho')$). The last inequality follows since by condition $B3$, $T_l(\rho_l)$ are strictly monotone and $\rho_l dT_l(\rho_l)/d\rho_l$ are increasing for $l \in N$. Therefore we have the relations (23) and (24). □

Theorem 3.5 *Consider assumptions B1–B4 and assume that there exists some finite feasible solution. The overall optimal solution may not be unique. However, the utilization in each link $k \in N$ is uniquely determined and is the same for all overall optimal solutions.*

Proof. The former half of this theorem is clear by noting that T depends only on $\boldsymbol{\rho}$ (see (5)) (thus if $\mathbf{x}$ is globally optimal then any solution $\mathbf{x}'$ that gives rise to the same value of $\boldsymbol{\rho}$ will be optimal as well). In Section 5 of [11] there is an example of the cases where more than one optimal solution exists.

The latter half is proved as follows. Suppose that the overall optimal policy has two distinct solutions $\hat{\mathbf{x}}$ and $\tilde{\mathbf{x}}$, which result in the utilization vectors $\hat{\boldsymbol{\rho}}_U$ and $\tilde{\boldsymbol{\rho}}_U$, respectively, and $\hat{\boldsymbol{\rho}}_U \neq \tilde{\boldsymbol{\rho}}_U$. Then we have from Lemma 3.2,

$$
\mathbf{t}(\hat{\mathbf{x}}) \cdot (\tilde{\mathbf{x}} - \hat{\mathbf{x}}) \geq 0,
$$
$$
\mathbf{t}(\tilde{\mathbf{x}}) \cdot (\hat{\mathbf{x}} - \tilde{\mathbf{x}}) \geq 0.
$$

Then we have

$$
(\hat{\mathbf{x}} - \tilde{\mathbf{x}}) \cdot \Big[\mathbf{t}(\hat{\mathbf{x}}) - \mathbf{t}(\tilde{\mathbf{x}}) \Big] \leq 0.
$$

From Lemma 3.4 we have

$$
(\hat{\mathbf{x}} - \tilde{\mathbf{x}}) \cdot \Big[\mathbf{t}(\hat{\mathbf{x}}) - \mathbf{t}(\tilde{\mathbf{x}}) \Big] \; > \; 0,
$$

since $\hat{\boldsymbol{\rho}}_U \neq \tilde{\boldsymbol{\rho}}_U$. This leads to a contradiction. That is, if there exist two distinct optimal solutions, the utilization vectors of both the solutions must be the same. Note that the utilization of link $l \in I$ is considered always zero. Naturally, in that case, $\sum_{l \in I} \rho_l$ must be unique but each of ρ_l, $l \in I$, need not be unique. □

Now let us consider the range of the optimal solutions. From the above, we obtain the following relations that characterize the range of the optimal solutions.

$$\sum_{k=1}^{R} \sum_{p \in \Pi^{(k)}} \delta_{lp} \frac{x_p^{(k)}}{\mu_l^{(k)}} = \rho_l, \quad l \in N, \tag{26}$$

$$\sum_{l \in I} \sum_{k=1}^{R} \sum_{p \in \Pi^{(k)}} \delta_{lp} \frac{x_p^{(k)}}{\mu_l^{(k)}} = \sum_{l \in I} \rho_l, \tag{27}$$

$$\sum_{p \in \Pi_d^{(k)}} x_p^{(k)} = \phi_d^{(k)}, \; d \in D^{(k)}, \; k = 1, 2, \ldots, R, \tag{28}$$

$$x_p^{(k)} \geq 0, \; p \in \Pi^{(k)}, k = 1, 2, \ldots, R, \tag{29}$$

where the value of each ρ_l is what an optimal solution $\mathbf{x}$ results in. From the relations (26)–(29) we see that optimal path flow rates belong to a convex polyhedron. Then we have the following proposition about the uniqueness of the optimal solutions.

Corollary 3.6 *The overall optimal solution is unique if and only if the total number of elements in* $\mathbf{x}$ *does not exceed the number of linearly independent equations in the set of linear equations* (26)–(28).

4 Uniqueness of the individually optimal solution

By the individually optimal policy we mean that jobs are scheduled so that each job may feel that its own mean cost is minimum if it knows the mean cost of each path of O-D pair d, $T_p^{(k)}(\mathbf{x})$, $p \in \Pi_d^{(k)}$. By the *individual optimization problem* we mean the problem of obtaining the routing decision that achieves the objective of the individually optimal policy. We call the solution of the individual optimization problem the *individually optimal solution* or the *equilibrium.* In the equilibrium, no user has any incentive to make a unilateral decision to change his route. Wardrop [24] considered the policy for a transportation network. We assume that there is a routing decision and that $\mathbf{x}$ is the path flow rate matrix which results from the routing decision. The individually optimal policy requires that a class k job of O-D pair d should follow through such a path class $\hat{p}$ that satisfies

$$T_{\hat{p}}^{(k)}(\mathbf{x}) = \min_{p \in \Pi_d^{(k)}} T_p^{(k)}(\mathbf{x}) \quad \text{for all } d \in D^{(k)}, \; k = 1, 2, \ldots, R. \tag{30}$$

If a routing decision satisfies the above condition we say the routing decision realizes the individually optimal policy.

Definition The path flow rate vector $\mathbf{x}$ is said to satisfy the equilibrium conditions for a multi-class open network if the following relations are satisfied for all $d \in D^{(k)}$, $k =$

$1, 2, \ldots, R,$

$$
\begin{aligned}
T_p^{(k)}(\mathbf{x}) &\geq A_d^{(k)}, \;\; x_p^{(k)} = 0, & (31)\\
T_p^{(k)}(\mathbf{x}) &= A_d^{(k)}, \;\; x_p^{(k)} > 0, & (32)\\
\sum_{p \in \Pi_d^{(k)}} x_p^{(k)} &= \phi_d^{(k)}, & (33)\\
x_p^{(k)} &\geq 0, \;\; p \in \Pi_d^{(k)}, & (34)
\end{aligned}
$$

where

$$
A_d^{(k)} = \min_{p \in \Pi_d^{(k)}} T_p^{(k)}(\mathbf{x}), \;\; d \in D^{(k)}, \;\; k = 1, 2, \ldots, R. \quad (35)
$$

Note that the set of the relations (31)–(33) is identical with the following set of relations.

$$
\begin{aligned}
[\mathbf{T}(\mathbf{x}) - \mathbf{\Gamma A}] \cdot \mathbf{x} &= 0, & (36)\\
\mathbf{T}(\mathbf{x}) - \mathbf{\Gamma A} &\geq 0, & (37)\\
\mathbf{\Gamma}^{\mathrm{T}} \mathbf{x} - \boldsymbol{\phi} &= 0, & (38)\\
\mathbf{x} &\geq 0. & (39)
\end{aligned}
$$

The above definition is the natural extension of the notion of Wardrop [24] equilibrium to our setting.

Theorem 4.1 *Consider assumptions* $B1$–$B4$*. There exists an individually optimal solution* $\mathbf{x}$ *which satisfies the relations* (36)–(39).

Proof. Define $\tilde{T}(\mathbf{x})$ by

$$
\tilde{T}(\mathbf{x}) = \frac{1}{\Phi}\Big[\sum_{l \in N} \log_e\Big(\frac{1}{T_l(\rho_l)}\Big) + \sum_{l \in I} \rho_l T_l\Big],
$$

where $\rho_l = \sum_{k=1}^{R} \lambda_l^{(k)} / \mu_l^{(k)}$ and $\lambda_l^{(k)} = \sum_{p \in \Pi^{(k)}} \delta_{lp} x_p^{(k)}$. Note that $\tilde{T}(\mathbf{x})$ is a convex function of $\mathbf{x}$. Then we have by noting (1)

$$
T_p^{(k)}(\mathbf{x}) = \frac{\partial}{\partial x_p^{(k)}} (\Phi \tilde{T}(\mathbf{x})).
$$

Let us consider the following convex nonlinear program:

minimize	$\tilde{T}(\mathbf{x})$
with respect to	$\mathbf{x}$
subject to (38) and (39).	

The Kuhn-Tucker conditions are the same as (36) - (39). Therefore, the program should have an optimal solution which must satisfy relations (36) - (39). □

We can express the individually optimal solution in the variational inequality form by using the same way as that for the overall optimal solution as follows.

Corollary 4.2 *Consider assumptions B1–B4.* $\bar{\mathbf{x}}$ *is an individually optimal solution if and only if*

$$\begin{aligned} \mathbf{T}(\bar{\mathbf{x}}) \cdot (\mathbf{x} - \bar{\mathbf{x}}) &\geq 0, \quad \text{for all } \mathbf{x} \\ \text{such that } \mathbf{\Gamma}^{\mathrm{T}}\mathbf{x} &= \phi \text{ and } \mathbf{x} \geq 0. \end{aligned}$$

Proof. Similar to the proof of Lemma 3.2. □

Lemma 4.3 *Consider assumptions B1–B4. Whenever finite, the function* $\mathbf{T}(\mathbf{x})$ *is monotone but is not strictly monotone. That is, for arbitrary* $\mathbf{x}$ *and* $\mathbf{x}'$ $(\mathbf{x} \neq \mathbf{x}')$, *if* $\mathbf{T}(\mathbf{x})$ *are finite or if* $\mathbf{T}(\mathbf{x}')$ *are finite then*

$$(\mathbf{x} - \mathbf{x}') \cdot [\mathbf{T}(\mathbf{x}) - \mathbf{T}(\mathbf{x}')] > 0 \quad \text{if} \quad \boldsymbol{\rho}_U \neq \boldsymbol{\rho}'_U, \tag{40}$$

$$= 0 \quad \text{if} \quad \boldsymbol{\rho}_U = \boldsymbol{\rho}'_U \tag{41}$$

where $\boldsymbol{\rho}_U$ *and* $\boldsymbol{\rho}'_U$ *are the utilization vectors that* $\mathbf{x}$ *and* $\mathbf{x}'$ *result in, respectively.*

Proof. This Lemma can be proved by the same way as that for the Lemma 3.4. Assume that $\boldsymbol{\rho}_U \neq \boldsymbol{\rho}'_U$. Then

$$\begin{aligned} (\mathbf{x} - \mathbf{x}') \cdot [\mathbf{T}(\mathbf{x}) - \mathbf{T}(\mathbf{x}')] &= \sum_{k=1}^{R} \sum_{p \in \Pi^{(k)}} \sum_{l \in N} (x_p^{(k)} - x_p'^{(k)}) \frac{\delta_{lp}}{\mu_l^{(k)}} (T_l(\rho_l) - T_l(\rho'_l)) \\ &= \sum_{l \in N} (\rho_l - \rho'_l)(T_l(\rho_l) - T_l(\rho'_l)) > 0 \end{aligned}$$

The last inequality follows since by condition B3, $T_l(\rho_l)$ are strictly monotone for $l \in N$. Therefore we have the relations (40) and (41). □

Theorem 4.4 *Consider assumptions B1–B4. The equilibrium is unique in the following sense: There is a unique utilization for each link* $l \in N$ *for all equilibria which have finite costs for all routes. But the individually optimal solution may not be unique.*

Proof. We can prove this theorem in the same way as that for Theorem 3.5. □

The range of the individually optimal solutions (related to finite costs) is given by the same set of relations as (26) - (29) but with possibly different values of ρ_l, $l = 1, 2, \ldots, M$.

As we shall show in Section 6, there may be different equilibria where only one has finite costs; those with infinite costs may have utilizations which are different than those in the case of finite cost in all the links.

5 Uniqueness of the class optimal solution

We shall make the following assumption throughout this section:
G: If not all classes have finite cost then at least one of the classes which has infinite cost can change its own flow to make this cost finite.

By the class optimal policy we mean that jobs are scheduled so that the expected cost of each class may be minimum under the condition that the scheduling decisions on jobs of the other classes are given and fixed. By the *class optimization problem* we mean the problem of obtaining the routing decision $\mathbf{x}$ that achieves the objective of the class optimal policy. We call the solution of the class optimization problem the *class optimal solution* or the *Nash equilibrium.* In the Nash equilibrium, no class has any incentive to make a unilateral decision to change the decision on the routes of the jobs of the class.

Assumption G above implies that in any Nash equilibrium, all classes have finite costs.

We assume that there is a routing decision and that $\mathbf{x}$ is the path flow rate matrix which results from the routing decision. The class optimal policy requires that

$$\Delta^{(k)}(\mathbf{x}^{(k)},\mathbf{x}^{-k}) = \min_{\mathbf{x}'^{(k)}} \Delta^{(k)}(\mathbf{x}'^{(k)},\mathbf{x}^{-k}) \quad \text{for all } k=1,2,\ldots,R. \tag{42}$$

If a routing decision $\mathbf{x}$ satisfies the above condition we say the routing decision $\mathbf{x}$ realizes the class optimal policy. The problem of minimizing the mean cost for jobs of class k is stated as follows:

$$\text{minimize:} \quad \Delta^{(k)} = \frac{1}{\Phi^{(k)}} \sum_{l\in\mathcal{L}} \rho_l^{(k)} T_l(\rho_l) \tag{43}$$

with respect to $\mathbf{x}^{(k)}$ with $\mathbf{x}^{-k}$ being fixed subject to

$$\mathbf{\Gamma}^{\mathrm{T}}\mathbf{x} = \boldsymbol{\phi}, \tag{44}$$

$$\mathbf{x} \geq 0, \tag{45}$$

where $\rho_l^{(k)} = \lambda_l^{(k)}/\mu_l^{(k)}$, $\rho_l = \sum_{k=1}^R \rho_l^{(k)}$ and $\lambda_l^{(k)} = \sum_{p\in\Pi^{(k)}} \delta_{lp} x_p^{(k)}$.

Lemma 5.1 $\mathbf{x}$ *is an optimal solution of the problem* (42) *if and only if* $\mathbf{x}$ *satisfies the following conditions*

$$[\tilde{\mathbf{t}}(\mathbf{x}) - \mathbf{\Gamma}\boldsymbol{\alpha}]\cdot\mathbf{x} = 0, \tag{46}$$

$$\tilde{\mathbf{t}}(\mathbf{x}) - \mathbf{\Gamma}\boldsymbol{\alpha} \geq 0, \tag{47}$$

$$\mathbf{\Gamma}^{\mathrm{T}}\mathbf{x} - \boldsymbol{\phi} = 0, \tag{48}$$

$$\mathbf{x} \geq 0. \tag{49}$$

Proof. Since the objective function (43) is convex and the feasible region of its constraints is a convex set, any local solution of the problem is a global solution point. To obtain the optimal solution, we construct the Lagrangian function for (43), $k=1,2,\ldots,R$,

$$L^{(k)}(\mathbf{x},\boldsymbol{\alpha}^{(k)}) = \Phi^{(k)}\Delta^{(k)} + \boldsymbol{\alpha}^{(k)}\cdot(\boldsymbol{\phi}^{(k)} - \mathbf{\Gamma}^{(k)\mathrm{T}}\mathbf{x}^{(k)}). \tag{50}$$

By the Kuhn-Tucker theorem $\mathbf{x}$ is an optimal solution if and only if there exists some (finite) $\boldsymbol{\alpha}$ and the following relations hold for $k = 1, 2, \ldots, R$

$$\frac{\partial L^{(k)}}{\partial \mathbf{x}^{(k)}} = \tilde{\mathbf{t}}^{(k)}(\mathbf{x}) - \boldsymbol{\Gamma}^{(k)}\boldsymbol{\alpha}^{(k)} \geq 0, \tag{51}$$

$$\frac{\partial L^{(k)}}{\partial \mathbf{x}^{(k)}} \cdot \mathbf{x}^{(k)} = [\tilde{\mathbf{t}}^{(k)}(\mathbf{x}) - \boldsymbol{\Gamma}^{(k)}\boldsymbol{\alpha}^{(k)}] \cdot \mathbf{x}^{(k)} = 0, \tag{52}$$

$$\frac{\partial L^{(k)}}{\partial \boldsymbol{\alpha}^{(k)}} = \boldsymbol{\phi}^{(k)} - \boldsymbol{\Gamma}^{(k)\mathrm{T}}\mathbf{x}^{(k)} = 0, \tag{53}$$

$$\mathbf{x}^{(k)} \geq 0, \tag{54}$$

where $(\partial L^{(k)})/(\partial \mathbf{x})$ denotes the vector whose elements are $(\partial L^{(k)})/(\partial x_p^{(k)})$, $p \in \Pi^{(k)}$, $k = 1, 2, \ldots, R$ (for the finiteness, see [21] Cor. 5.1). We easily see that the relations (51) - (54) for $k = 1, 2, \ldots, R$ are exactly the same as the relations (46) - (49). □

We can express the class optimal solution in the variational inequality form by using the same way as that for the overall optimal solution as follows.

Corollary 5.2 *Consider assumptions B1–B4.* $\bar{\mathbf{x}}$ *is a class optimal solution if and only if*

$$\begin{aligned} \tilde{\mathbf{t}}(\bar{\mathbf{x}}) \cdot (\mathbf{x} - \bar{\mathbf{x}}) &\geq 0, \quad \text{for all } \mathbf{x} \\ \text{such that } \boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{x} &= \boldsymbol{\phi} \text{ and } \mathbf{x} \geq 0. \end{aligned}$$

Proof. Similar to the proof of Lemma 3.2. □

The question of uniqueness for the class optimal solution has only been treated for some special cases [17, 20]. A counter example in [20] shows that different class optimal solutions may exist, with different utilizations.

Our purpose in this section is to extend the known uniqueness results to more general assumptions.

5.1 All positive flows

We make the following assumptions:

- μ_l^i can be represented as $a^{(k)}\mu_l$, and $0 < \mu_l^i$ is finite.
- At each node, each class may re-route all the flow that it sends through that node to any of the out-going links of that node.
- The rate of traffic of class k that enters the network at node v is given by $\phi_v^{(k)}$; if this quantity is negative this means that traffic of class k leaves node v at a rate of $|\phi_v^{(k)}|$. We assume that $\sum_v \phi_v^{(k)} = 0$.

For each node u, denote by $In(u)$ the set of its in-going links, and denote by $Out(u)$ the set of its out-going links.

Due to the second assumption, we may work directly with the decision variables $\lambda_l^{(k)}$ instead of working with the path flows. We can then replace the constraint (4) by:

$$\sum_{l \in Out(v)} \lambda_l^{(k)} = \sum_{l \in In(v)} \lambda_l^{(k)} + \phi_v^{(k)}$$

for each node v.

We shall use the Kuhn-Tucker condition. To do so, we define the Lagrangian

$$J^{(k)}(\boldsymbol{\lambda}, \boldsymbol{\xi}^{(k)}) = \sum_{l \in \mathcal{L}} \rho_l^{(k)} T_l - \sum_u \xi_u^{(k)} \left[\sum_{l \in Out(u)} \lambda_l^{(k)} - \sum_{l \in In(u)} \lambda_l^{(k)} - \phi_u^{(k)} \right].$$

Here, $\boldsymbol{\xi}^{(k)} = [\xi_1^{(k)}, \xi_2^{(k)}, \ldots, \xi_M^{(k)}]^{\mathrm{T}}$ is the vector of Lagrange multipliers for class k.

An assignment $\boldsymbol{\lambda}^*$ is class-optimal if and only if the following Kuhn-Tucker conditions hold. There exists some $\boldsymbol{\xi}^{(k)} = [\xi_u^{(k)}]$ such that

$$\frac{\partial J^{(k)}(\boldsymbol{\lambda}^*, \boldsymbol{\xi}^{(k)})}{\partial \lambda_l^{(k)}} \geq 0, \tag{55}$$

$$\frac{\partial J^{(k)}(\boldsymbol{\lambda}^*, \boldsymbol{\xi}^{(k)})}{\partial \lambda_l^{(k)}} = 0 \text{ if }, \lambda_l^{(k)} > 0; \tag{56}$$

$$\lambda_l^{(k)} \geq 0, \quad \sum_{l \in Out(v)} \lambda_l^{(k)} = \sum_{l \in In(v)} \lambda_l^{(k)} + \phi_v^{(k)}.$$

Define

$$K_l^{(k)}(\boldsymbol{\lambda}) := \frac{\partial \rho_l^{(k)} T_l(\rho_l)}{\partial \lambda_l^{(k)}}.$$

Then

$$K_l^{(k)}(\rho_l^{(k)}, \rho_l) = \frac{1}{\mu_l^{(k)}} \left(\rho_l^{(k)} \frac{\partial T_l(\rho_l)}{\partial \rho_l} + T_l(\rho_l) \right)$$

Conditions (55)-(56) can be rewritten as

$$K_l^{(k)}(\rho_l^{(k)}, \rho_l) \geq \xi_u^{(k)} - \xi_v^{(k)}, \text{ with equality if } \lambda_l^{(k)} > 0 \tag{57}$$

Note that condition B3 implies that $K_l^{(k)}(\rho_l^{(k)}, \rho_l)$ is strictly monotonically increasing in both arguments.

Lemma 5.3 *Consider assumptions B1-B4. Assume that $\boldsymbol{\lambda}$ and $\hat{\boldsymbol{\lambda}}$ are two class-optimal solutions with finite costs.*
If $\rho_l = \hat{\rho}_l$ for some link l of type N (see assumption B3) then $\lambda_l^{(k)} = \hat{\lambda}_l^{(k)}$ for all $k = 1, ..., R$.

Proof. Consider some link $l = (u, v)$ of type N. The Lemma clearly holds if $\rho_l = 0$, so assume that $\rho_l > 0$. Let $\boldsymbol{\lambda}$ and $\hat{\boldsymbol{\lambda}}$ be two class-optimal solutions satisfying $\rho_l = \hat{\rho}_l$ and such that

$$\hat{\xi}_u^{(k)} - \hat{\xi}_v^{(k)} > \xi_u^{(k)} - \xi_v^{(k)}. \tag{58}$$

Let k be such that $\rho_l^{(k)} > 0$. Then (57) implies that

$$K_l^{(k)}(\hat{\rho}_l^{(k)}, \hat{\rho}_l) \geq \hat{\xi}_u^{(k)} - \hat{\xi}_v^{(k)} > \xi_u^{(k)} - \xi_v^{(k)} = K_l^{(k)}(\rho_l^{(k)}, \rho_l) = K_l^{(k)}(\rho_l^{(k)}, \hat{\rho}_l) \tag{59}$$

Since $K_l^{(k)}$ is monotone in its first argument, this implies that $\hat{\rho}_l^{(k)} > \rho_l^{(k)}$. Since this holds for all k such that $\rho_l^{(k)} > 0$ (and there is at least one such k), and since for other k's $\hat{\rho}_l^{(k)} \geq \rho_l^{(k)} = 0$, this contradicts the fact that $\hat{\rho}_l = \rho_l$. We thus conclude that (58) holds with equality replacing the strict inequality, and that $\hat{\rho}_l^{(k)} \geq \rho_l^{(k)}$ for all k.

Now that we know that (58) holds with equality, we can interchange in the above argument between $\boldsymbol{\lambda}$ and $\hat{\boldsymbol{\lambda}}$ and conclude that $\hat{\rho}_l^{(k)} \leq \rho_l^{(k)}$ for all k as well, which implies the Lemma.

Theorem 5.4 *Consider assumptions B1–B4. Denote by $\mathcal{L}_1(\boldsymbol{\lambda})$ the sets of links l such that $\lambda_l^{(k)} > 0, \forall k \in N$ for an assignment $\boldsymbol{\lambda}$.*
Assume that $\boldsymbol{\lambda}$ and $\hat{\boldsymbol{\lambda}}$ are two class-optimal solutions with finite costs for all players.
Assume that $\lambda_l^{(k)} = 0, \forall k, \forall l \notin \mathcal{L}_1(\boldsymbol{\lambda})$, $\hat{\lambda}_l^{(k)} = 0, \forall k, \forall l \notin \mathcal{L}_1(\hat{\boldsymbol{\lambda}})$
Then $\lambda_l^{(k)} = \hat{\lambda}_l^{(k)}$ for all $l \in N$ (see assumption B3 for the definition of N).

Remark 5.1 The Theorem and its proof are extensions of [20] who considered the special case where μ_l^i do not depend on l and i, where there is a single source-destination pair which is the same for all users (all paths and all classes), and where $\mathcal{L}_1(\boldsymbol{\lambda}) = \mathcal{L}_1(\hat{\boldsymbol{\lambda}})$.

Proof. Denote $\xi_u = \sum_{k=1}^{R} a^{(k)} \xi_u^{(k)}$, where $a^{(k)}$ is defined in the beginning of Subsection 5.1, and

$$S_l(\rho_l) = \rho_l \frac{\partial T_l(\rho_l)}{\partial \rho_l} + R T_l(\rho_l).$$

Note that the assumption that costs are finite together with Lemma 3.3 imply that $S_l(\rho_l)$ are finite.

Let $\hat{\boldsymbol{\xi}}$ denote the vector of the Lagrange multipliers corresponding to $\hat{\boldsymbol{\lambda}}$. It follows from (57) that

$$\frac{1}{\mu_{uv}} S_{uv}(\rho_{uv}) \geq \xi_u - \xi_v, \text{ with equality for } (u, v) \in \mathcal{L}_1(\boldsymbol{\lambda}). \tag{60}$$

A similar relation holds for $\hat{\boldsymbol{\lambda}}$

We obtain

$$
\begin{aligned}
0 &\leq \sum_{(u,v)\in\mathcal{L}} (\rho_{uv} - \hat{\rho}_{uv})(S_{uv}(\rho_{uv}) - S_{uv}(\hat{\rho}_{uv})) \\
&\leq \sum_{(u,v)\in\mathcal{L}} \mu_{uv}(\rho_{uv} - \hat{\rho}_{uv})\Big((\xi_u - \hat{\xi}_u) - (\xi_v - \hat{\xi}_v)\Big) = 0
\end{aligned} \tag{61}
$$

The first inequality follows from the strict monotonicity and the convexity of $T_l(\rho_l)$ for $l \in N$; for $l \in I$ this relation is trivial. The second inequality holds in fact for each pair u, v (and not just for the sum). Indeed, for $(u,v) \in \mathcal{L}_1(\boldsymbol{\lambda}) \cap \mathcal{L}_1(\hat{\boldsymbol{\lambda}})$ this relation holds with equality due to (60). This is also the case for $(u,v) \notin \mathcal{L}_1(\boldsymbol{\lambda}) \cup \mathcal{L}_1((\hat{\boldsymbol{\lambda}})$, since in that case $\rho_{uv} = \hat{\rho}_{uv} = 0$. Consider next the case $(u,v) \in \mathcal{L}_1(\boldsymbol{\lambda}), (u,v) \notin \mathcal{L}_1(\hat{\boldsymbol{\lambda}})$. Then we have

$$
\begin{aligned}
&(\rho_{uv} - \hat{\rho}_{uv})(S_{uv}(\rho_{uv}) - S_{uv}(\hat{\rho}_{uv})) \\
&= \rho_{uv}(S_{uv}(\rho_{uv}) - S_{uv}(\hat{\rho}_{uv})) \\
&\leq \mu_{uv}\rho_{uv}\Big((\xi_u - \hat{\xi}_u) - (\xi_v - \hat{\xi}_v)\Big).
\end{aligned}
$$

A symmetric argument establishes the case $(u,v) \in \mathcal{L}_1(\hat{\boldsymbol{\lambda}}), (u,v) \notin \mathcal{L}_1(\boldsymbol{\lambda})$.

We finally establish the last equality in (61).

$$
\begin{aligned}
&\sum_{(u,v)\in\mathcal{L}} \mu_{uv}(\rho_{uv} - \hat{\rho}_{uv})\Big((\xi_u - \hat{\xi}_u) - (\xi_v - \hat{\xi}_v)\Big) \\
&= \sum_j (\xi_j - \hat{\xi}_j)\sum_w (\rho_{jw} - \hat{\rho}_{jw})\mu_{jw} - \sum_j (\xi_j - \hat{\xi}_j)\sum_w (\rho_{wj} - \hat{\rho}_{wj})\mu_{wj} \\
&= \sum_j (\xi_j - \hat{\xi}_j)\left(\sum_{l\in Out(j)} (\rho_l - \hat{\rho}_l)\mu_l - \sum_{l\in In(j)} (\rho_l - \hat{\rho}_l)\mu_l\right) \\
&= \sum_{k=1}^{R} a^{(k)}\left[\sum_j (\xi_j - \hat{\xi}_j)\left(\sum_{l\in Out(j)} (\lambda_l^{(k)} - \hat{\lambda}_l^{(k)}) - \sum_{l\in In(j)} (\lambda_l^{(k)} - \hat{\lambda}_l^{(k)})\right)\right] = 0
\end{aligned}
$$

We conclude from (61) that $\rho_l = \hat{\rho}_l$ for all links in N. The proof now follows from Lemma 5.3. □

6 Counter example

In this subsection we show that both in class as well as in individual optimization, utilization are not unique at equilibrium even if all links are of type N. Indeed, in addition to an equilibrium with finite costs there may be several other equilibria with infinite costs.

Consider the following network. There are 4 nodes: $\{1,2,3,4\}$ and 2 symmetric classes. The set of links is $\{(12),(13),(24),(34),(23)\}$. Each class has half a unit of flow to ship

$\phi^{(1)} = \phi^{(2)} = 0.5$ between the source node 1 and the destination node 4. The cost per link is given by

$$T_l(\rho_l) = \frac{1}{1-\rho_l}.$$

The strategy in which both classes send all their flows along the path $(1 \to 2 \to 3 \to 4)$ is both class as well as individually optimal, but it gives rise to infinite cost for both players.

However, there exists another strategy which is optimal for both the individual and the class optimum problem: it is to route half of the flow of each class along the path $1 \to 2 \to 4$ and the other half through the path $1 \to 3 \to 4$. This is the unique equilibrium that has finite cost for all players.

7 Concluding remarks

We studied in this paper multiclass static routing problems. We considered several types of optimization concepts in networks: the global optimization, individual optimization (related to Wardrop equilibrium) and class optimization (related to Nash equilibrium). The routing problem is of the type studied in [11], where one has to determine the assignment of the flow rates among different paths. This setting is more general than the one in which routing decisions may be taken at each node (see [14, 15, 16, 20]); it is of special importance to telecommunications networks (in particular ATM networks) in which the users have to route their traffic through predetermined virtual paths. We established the uniqueness of the optimal solution for the different types of optimization problems.

References

[1] E. Altman, E. Başar, T. Jiménez and N. Shimkin, "Routing in networks: distributed game-theoretic algorithms" *Proceedings of the 8th International Symposium on Dynamic Games and Applications*, July, Maastricht, The Netherlands, 1998.

[2] F. Baskett, K.M. Chandy, R.R. Muntz and F. Palacios, "Open, closed, and mixed networks of queues with different classes of customers", *J. ACM* **22**, No. 2, 248–260, 1975.

[3] D.G. Cantor and M. Gerla, "Optimal routing in a packet-switched computer network," *IEEE Trans. Computers* **C-23**, No. 10, 1062–1069, 1973.

[4] S. Dafermos, "The traffic assignment problem for multiclass-user transportation networks", *Transportation Sci.* **6**, 73–87, 1972.

[5] L. Fratta, M. Gerla and L. Kleinrock, "The flow deviation method—an approach to the store-and-forward communication network design", *Networks* **3**, 97–133, 1973.

[6] A. Haurie and P. Marcott, "On the relationship between Nash-Cournot and Wardrop equilibria", *Networks* **15**, pp. 295-308, 1985.

[7] M. D. Intriligator, *Mathematical Optimization and Economic Theory*, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1971.

[8] H. Kameda and A. Hazeyama, "Individual vs. overall optimization for static load balancing in distributed computer systems", *Computer Science Report*, The University of Electro-Communications, 1988.

[9] H. Kameda, T. Kozawa and J. Li, "Anomalous relations among various performance objectives in distributed computer systems", *Proc. World Congress on Systems Simulation*, pp. 459-465, Sept, 1997.

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

[11] H. Kameda and Y. Zhang, "Uniqueness of the solution for optimal static routing in open BCMP queueing networks", *Math. Comput. Modelling* **22**, No. 10-12, pp. 119-130, 1995.

[12] C. Kim and H. Kameda, "An algorithm for optimal static load balancing in distributed computer systems", *IEEE Trans. Computers* **41**, No. 3, 381–384, 1992.

[13] L. Kleinrock, *Queueing Systems, Vol 2: Computer Applications*, Wiley - Interscience, New York, 1976.

[14] Korilis, Y. A., Lazar, A.A. and Orda, A., "Architecting noncooperative networks", *Journal on Selected Areas in Communications*, Volume 13, No. 7, pp. 1241-1251, September 1995.

[15] Korilis, Y. A., Lazar, A.A. and Orda, A., "Achieving network optima using Stackelberg routing games", *IEEE Transactions on Networking*, Vol. 5, No. 1, pp. 161-173, February 1997.

[16] Korilis, Y. A., Lazar, A.A. and Orda, A., "Capacity allocation under non-cooperative routing", *IEEE Transactions on Automatic Control*, Vol. 42, No. 3, pp.309-325, March 1997.

[17] H. Kameda, E. Altman and Kozawa, "Braess-like paradoxes of Nash equilibria for load balancing in distributed systems", submitted.

[18] R. J. La and V. Anantharam, "Optimal routing control: game theoretic approach", *Proc. of the 36th IEEE Conference on Decision and Control*,San Diego, California, Dec. 1997.

[19] T.L. Magnanti, "Models and algorithms for predicting urban traffic equilibria", *Transportation Planning Models*, (M. Florian, ed.) Elsevier Science Publishers B.V., North-Holland, 153–185, 1984.

[20] A. Orda, R. Rom and N. Shimkin, "Competitive routing in multi-user environments", *IEEE/ACM Transactions on Networking*, vol. 1, pp. 510-521, October 1993.

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

[22] A. N. Tantawi and D. Towsley, "A general model for optimal static load balancing in star network configurations", *Proc. of PERFORMANCE' 84*, North-Holland, New York, 277–291, 1984.

[23] A. N. Tantawi and D. Towsley, "Optimal static load balancing in distributed computer systems", *J. ACM* **32**, No. 2, 445–465, 1985.

[24] J.G. Wardrop, "Some theoretical aspects of road traffic research", *Proc. Inst. Civ. Eng.* Part 2, 1, 325–378, 1952.