This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2009 proceedings.

An Information-Theoretic Characterization of Weighted α-Proportional Fairness

Masato Uchida
Network Design Research Center
Kyushu Institute of Technology
Chiyoda-ku, Tokyo 100–0011, Japan
Email: m.uchida@ndrc.kyutech.ac.jp

Jim Kurose
Department of Computer Science
University of Massachusetts
Amherst, MA 01003, USA
Email: kurose@cse.umass.edu

***Abstract*—This paper provides a novel characterization of fairness concepts in network resource allocation problems from the viewpoint of information theory. The fundamental idea adopted in this paper is to characterize the utility functions used in optimization problems, which motivate fairness concepts, based on a trade-off between user and system satisfaction. Here, user satisfaction is evaluated using information divergence measures that were originally used in information theory to evaluate the difference between two probability distributions. In this paper, information divergence measures are applied to evaluate the difference between the implemented resource allocation and a requested resource allocation. The requested resource allocation is assumed to be ideal in some sense from the user's point of view. Also, system satisfaction is evaluated based on the efficiency of the implemented resource utilization, which is defined as the total amount of resources allocated to each user. The results discussed in this paper indicate that the well-known fairness concept called weighted α-proportional fairness can be characterized using the α-divergence measure, which is a general class of information divergence measures, as an equilibrium of the trade-off described above. In the process of obtaining these results, we also obtained a new utility function that has a parameter to control the trade-off. This new function is then applied to typical examples to solve resource allocation problems in simple network models such as those for two-link networks and wireless LANs.**

I. Introduction

This paper considers a network resource allocation problem, where the network has a resource with a finite capacity and supports a finite number of users. When users have to compete for the resource, the share allocated to each must be regulated by some control mechanism to avoid network congestion and degradation in performance. As a result of the finite capacity of the resource, any compromise in resource allocation will lead to the concept of fairness.

The mathematical concept of fairness is formulated as an optimization problem, where the objective is to find a feasible resource allocation that maximizes a utility function specific to the fairness concept used. Examples of such fairness concepts are *throughput maximization* [1], *max-min fairness* [1], *proportional fairness* [2], and *potential delay minimization* [3]. A general class of these examples is *weighted α-proportional fairness* [4]. The mathematical formulation of weighted α-proportional fairness embodies a number of fairness concepts including the above examples by varying the values of parameter α and the weight parameter. A more general class is *utility fairness* [2], [5]–[7], which includes weighted α-proportional fairness as a special case. Utility fairness is defined with a utility function, and resource allocation is determined by solving the optimization problem, where the objective is to find a feasible resource allocation that will maximize the utility function. After these fairness concepts were proposed, a number of studies examined their application, extension, and evaluation. For example, static/dynamic or packet/flow level performance in various network models has been extensively examined [5]–[15].

In this paper, we consider that utility fairness, and α-proportional fairness in particular, should have properties of resource allocation that solves a combined optimization problem of (i) minimizing the "distance" between a requested resource allocation (which may not be feasible) and an implemented resource allocation associated with the overall user utility and (ii) maximizing an overall resource-usage utility in the implemented resource allocation. The fundamental idea behind in this paper is to characterize the fairness concept through insights into the trade-off between user and system satisfaction. That is, we consider that the optimization of user satisfaction should be expressed through the minimization in (i) and the optimization of system satisfaction should be expressed through the maximization in (ii). To realize this goal, we propose a new utility function that expresses the trade-off between user and system satisfaction in the implemented resource allocation. The proposed utility function has a parameter to control the trade-off. We demonstrate that existing utility functions, which motivate fairness concepts such as weighted α-proportional fairness, appear to be an equilibrium in the trade-off expressed by the proposed utility function.

User satisfaction in the implemented resource allocation, which is associated with the overall user utility, is evaluated using information divergence measures that play an important role in information theory. Information divergence is a distance defined on a pair of probability measures (or a pair of positive finite measures), which is positive except in the case of agreement between two probability measures (or two positive finite measures). We use information divergence to evaluate user satisfaction by comparing the implemented resource allocation and the requested resource allocation. The requested resource allocation is assumed to be ideal in some sense from the

978-1-4244-3513-5/09/$25.00

user's point of view. If it is assumed that uniform resource allocation among users is ideal from the users' point of view, weighted α-proportional fairness is obtained from the fundamental idea mentioned above using α-divergence [16], which is an extension of well-known information divergence called Kullback-Libler divergence [17]. Although information divergence measures were also used in [18] to characterize the fairness concept, its approach was different from this paper because they were not used to evaluate the trade-off between user and system satisfaction.

We use the overall resource-usage utility to evaluate system satisfaction in the implemented resource allocation. That is, we assume that efficient utilization of network resources is preferable from the system's point of view.

This paper is organized as follows. Section II briefly reviews the mathematical formulation of fairness concepts and information divergence measures. Section III presents the characterization of weighted α-proportional fairness using α-divergence. This characterization leads to a generalized concept of weighted α-proportional fairness. Section IV applies the generalized concept to resource-allocation problems in simple network models such as those for two-link networks and wireless LANs. Section V concludes the paper.

II. Brief Review of Fairness Concepts and Information Divergence Measures

A. Network Model

Consider a network with a resource. Let $\mathcal{N}$ be a set of users who must compete for the resource, where n denotes the cardinality of $\mathcal{N}$. Let x_i be the share of the resource that is allocated to user $i \in \mathcal{N}$. The resource allocation vector is denoted by $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, where $0 < x_i < \infty$, $i \in \mathcal{N}$. Let $\mathcal{C}$ be a feasible region for allocating the resource. Resource allocation vector $\boldsymbol{x}$ is feasible if $\boldsymbol{x} \in \mathcal{C}$. Bandwidth is a representative resource considered in the resource-allocation problem.

B. Utility Fairness

A general fairness concept called utility fairness has been proposed [2], [5]–[7]. A resource allocation vector, $\boldsymbol{x}^* = (x_1^*, x_2^*, \ldots, x_n^*)$, is utility fair if it is feasible, i.e., $\boldsymbol{x}^* \in \mathcal{C}$, and if for any other feasible vector $\boldsymbol{x}$, the following condition is satisfied:

$$\sum_{i=1}^{n} \frac{\partial \varphi_i}{\partial x_i}(x_i^*)(x_i - x_i^*) \leq 0,$$

where $\varphi_i(\cdot)$ is an increasing, strictly concave, and continuously differentiable function on open interval $(0, \infty)$ for all $i \in \mathcal{N}$.

Utility fairness can be motivated another way. Consider the following optimization problem:

$$\text{maximize} \quad \sum_{i=1}^{n} \varphi_i(x_i), \tag{1}$$

$$\text{subject to} \quad \boldsymbol{x} \in \mathcal{C}. \tag{2}$$

If the feasible region given in Eq. (2) is convex and compact, an optimal solution to the above problem exists and is unique since the objective function given in Eq. (1) is strictly concave. The objective function is referred to as the utility function of utility fairness.

A well-known example of the utility fairness is called weighted α-proportional fairness [4]. The formulation of α-proportional fairness is given by substituting $w_i \varphi_\alpha(\cdot)$ into $\varphi_i(\cdot)$ in Eq. (1), where $\varphi_\alpha(\cdot)$, $\alpha > 0$ is an increasing, strictly concave, and continuously differentiable function on open interval $(0, \infty)$, as follows:

$$\varphi_\alpha(x) = \begin{cases} \log x, & \text{if } \alpha = 1. \\ \frac{x^{1-\alpha}}{1-\alpha}, & \text{if } \alpha > 0, \alpha \neq 1. \end{cases} \tag{3}$$

That is, weighted α-proportional fairness can be motivated by the following optimization problem:

$$\text{maximize} \quad \sum_{i=1}^{n} w_i \varphi_\alpha(x_i), \tag{4}$$

$$\text{subject to} \quad \boldsymbol{x} \in \mathcal{C}, \tag{5}$$

where $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$ is a weight vector with positive elements. If $w_i = w_j, \forall i, j \in \mathcal{N}$, then weighted α-proportional fairness is simply called α-proportional fairness. Note that α-proportional fairness reduces to several well-known fairness concepts. For example, *maximum throughput* [1] is obtained when $\alpha \to 0$, *proportional fairness* [2] is obtained when $\alpha \to 1$, *potential delay minimization* [3] is obtained when $\alpha = 2$, and *max-min fairness* [1] is obtained when $\alpha \to \infty$.

C. Csiszár's f-divergence

Information divergence is a measure that defines a distance between vectors $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$, where the components of the vectors are positive (i.e., $x_i,\ y_i > 0$, $i = 1, \ldots, n$). Although the original definition is only for the vector that satisfies $\sum_{i=1}^{n} x_i = 1$ and $\sum_{i=1}^{n} y_i = 1$ (i.e., the probability measure), the general definition introduced in this section can be applied to the vector that satisfies $\sum_{i=1}^{n} x_i < \infty$ and $\sum_{i=1}^{n} y_i < \infty$ (i.e., the positive finite measure) [19].

The most well-known information divergence measure was introduced by Kullback and Leibler [17]. It is now referred to as Kullback-Leibler divergence (KL-divergence, for short). A geometric interpretation of information divergence measures led to an extension of KL-divergence called α-divergence [16]. α-divergence is closely related to Rényi entropy [20] and Tsallis entropy [21], which are generalizations of Shannon entropy. KL-divergence and α-divergence are members of an important class of information divergences called f-divergence that was introduced by Csiszár [22].

Csiszár's f-divergence is defined by a strictly convex and continuously differentiable function $f(\cdot)$ on open interval $(0, \infty)$, satisfying $f(1) = f'(1) = 0$ as:

$$D_f(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} x_i f\left(\frac{y_i}{x_i}\right).$$

Note that

$$D_f(\boldsymbol{x}, \boldsymbol{y}) \geq 0 \tag{6}$$

holds, where the equality holds if and only if $\boldsymbol{x} = \boldsymbol{y}$. Thus, $D_f(\boldsymbol{x}, \boldsymbol{y})$ can be regarded as a kind of distance between $\boldsymbol{x}$ and $\boldsymbol{y}$, even though $D_f(\boldsymbol{x}, \boldsymbol{y}) = D_f(\boldsymbol{y}, \boldsymbol{x})$ (i.e., a symmetric property) does not necessary hold in general. In addition, it is easy to see that

$$\varepsilon D_f(\boldsymbol{x}, \boldsymbol{y}) = D_{\varepsilon f}(\boldsymbol{x}, \boldsymbol{y}), \tag{7}$$

holds for any positive constant value $\varepsilon > 0$.

A member of Csiszár's f-divergence called α-divergence [16] is defined as:

$$D_{f_\alpha}(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{\alpha(1-\alpha)} \sum_{i=1}^{n} \{(1-\alpha)x_i + \alpha y_i - x_i^{1-\alpha} y_i^{\alpha}\},$$

where

$$f_\alpha(u) = \begin{cases} \frac{1}{\alpha(1-\alpha)}\{1 - u^\alpha + \alpha(u-1)\}, & \text{if } \alpha \in \Re \backslash \{0, 1\}. \\ -\log u + u - 1, & \text{if } \alpha = 0. \\ u \log u - u + 1, & \text{if } \alpha = 1. \end{cases}$$

The α-divergence reduces to several well-known information divergences. For example, KL-divergence is obtained when $\alpha \to 0$, Hellinger divergence is obtained when $\alpha = 1/2$, dual KL-divergence is obtained when $\alpha \to 1$, (dual) Neyman's chi-square divergence [23] is obtained when $\alpha = 2$.

III. Characterization of Weighted α-Proportional Fairness

This section first introduces a fundamental idea to characterize fairness concepts in allocating network resources through insights into the trade-off between user and system satisfaction. We then propose a utility function to characterize weighted α-proportional fairness based on this fundamental idea. The proposed utility function is closely related to α-divergence. We demonstrate that the utility function of weighted α-proportional fairness defined by Eq. (4) appears to be an equilibrium of the trade-off expressed in the proposed utility function.

A. Fundamental Idea

First, let vector $\boldsymbol{y}$ represent a requested resource allocation vector that is assumed to be ideal in some sense from the user's point of view. That is, vector $\boldsymbol{y}$ best satisfies a kind of user demand.

For example, let us consider the case where vector $\boldsymbol{y}$ represents resource allocation that makes use of resources as much as possible under the constraint of uniform resource sharing, i.e., $\boldsymbol{y} = \boldsymbol{c} = (c, c, \cdots, c) \in \mathcal{C}$, where $\boldsymbol{c} = \arg\max_{\boldsymbol{z}=(z,z,\ldots,z)\in\mathcal{C}} z$. Vector $\boldsymbol{c}$ can be regarded as an example of a resource allocation that takes into consideration the viewpoint of user fairness.

As another example, let us consider a case when vector $\boldsymbol{y}$ represents an allocation of resources requested by users such that $\boldsymbol{y} = \boldsymbol{r} = (r_1, r_2, \ldots, r_n)$, where r_i is an allowable maximum allocation of resources for user $i \in \mathcal{N}$. Note that condition $\boldsymbol{r} \in \mathcal{C}$ is not necessarily satisfied because vector $\boldsymbol{r}$ is just a request by users. That is, vector $\boldsymbol{r}$ does not have to be feasible. Vector $\boldsymbol{r}$ can be regarded as an example of a resource allocation that takes into consideration the viewpoint of user requirements.

Then, let vector $\boldsymbol{x} \in \mathcal{C}$ represent an implemented allocation of resources that are actually supplied to users. This paper evaluates user satisfaction in such an implemented resource allocation $\boldsymbol{x}$ by using the requested resource allocation $\boldsymbol{y}$ as a basis for comparison. That is, the best resource allocation vector from the user's point of view is obtained from the optimization problem defined by minimizing

$$aD(\boldsymbol{x}, \boldsymbol{y}), \tag{8}$$

subject to $\boldsymbol{x} \in \mathcal{C}$, where $a > 0$ is a constant value and $D(\boldsymbol{x}, \boldsymbol{y})$ is a measure that evaluates the difference between the implemented resource allocation vector $\boldsymbol{x}$ and the requested resource allocation vector $\boldsymbol{y}$. Vector $\boldsymbol{x} \in \mathcal{C}$, which minimizes Eq. (8), should correspond to the requested resource allocation vector $\boldsymbol{y}$ if it is feasible (i.e., $\boldsymbol{y} \in \mathcal{C}$). The difference measure $D(\boldsymbol{x}, \boldsymbol{y})$ reflects the difference in user utility between the implemented resource allocation vector $\boldsymbol{x}$ and the requested resource allocation vector $\boldsymbol{y}$ experienced by users.

In addition, this paper evaluates system satisfaction in the implemented resource allocation $\boldsymbol{x}$ based on the efficiency of resource-usage. The efficiency of resource-usage is defined as the total amount of resources allocated to each user. That is, the best resource allocation vectors from the system's point of view are obtained from the optimization problem defined by maximizing

$$b \sum_{i=1}^{n} x_i, \tag{9}$$

subject to $\boldsymbol{x} \in \mathcal{C}$, where $b > 0$ is a constant value.

Generally, the utility functions defined by Eqs. (8) and (9) are not necessarily compatible. Therefore, it is reasonable to consider the optimization problem given by maximizing

$$-\gamma a D(\boldsymbol{x}, \boldsymbol{y}) + (1-\gamma) b \sum_{i=1}^{n} x_i, \tag{10}$$

subject to $\boldsymbol{x} \in \mathcal{C}$, where $0 < \gamma < 1$ is a constant value. Equation (10) is an important new utility function introduced in this paper. The parameter γ controls the trade-off between user and system satisfaction. For example, Eq. (10) is equivalent to the negative of Eq. (8) when $\gamma \to 1$ and is equivalent to Eq. (9) when $\gamma \to 0$. In this paper, we show that the utility function given by Eq. (10) reduces to those of existing fairness concepts such as weighted α-proportional fairness when $\gamma = 1/2$. This means that existing fairness concepts can be regarded as an equilibrium in the trade-off between user and system satisfaction.

Figure 1 shows the concept underlying Eqs. (8), (9), and (10). The curved surface represents the feasible region $\mathcal{C}$. Since the requested vector $\boldsymbol{y}$ does not lie in $\mathcal{C}$ in this example,

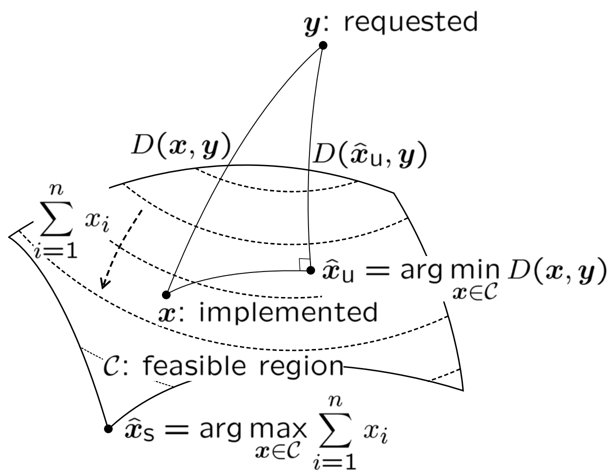

Fig. 1. Concept of Eq. (10)

the best resource allocation vector from the user's point of view is given by the vector $\hat{\boldsymbol{x}}_{\mathrm{u}}$, which is the projection of requested resource allocation vector $\boldsymbol{y}$ onto feasible region $\mathcal{C}$, i.e., $\hat{\boldsymbol{x}}_{\mathrm{u}} = \arg\min_{\boldsymbol{x}\in\mathcal{C}} D(\boldsymbol{x}, \boldsymbol{y})$. Also, the dashed curved lines represent contours of the value of $\sum_{i=1}^{n} x_i$. The value of $\sum_{i=1}^{n} x_i$ increases in the direction of the dashed arrow. Therefore, the best resource allocation vector from the system's point of view is given by vector $\hat{\boldsymbol{x}}_{\mathrm{s}}$, which is in the bottom-left corner of the curved surface, i.e., $\hat{\boldsymbol{x}}_{\mathrm{s}} = \arg\max_{\boldsymbol{x}\in\mathcal{C}} \sum_{i=1}^{n} x_i$. Here, note that the vector $\hat{\boldsymbol{x}}_{\mathrm{s}}$ does not necessary increase user satisfaction because there is a possibility that resources will be exhausted by certain users as a result of maximizing the value of $\sum_{i=1}^{n} x_i$. From this observation, we can see that the feasible resource allocation vector that maximizes the utility function given by Eq. (10) should be located between $\hat{\boldsymbol{x}}_{\mathrm{u}}$ and $\hat{\boldsymbol{x}}_{\mathrm{s}}$. Parameter γ controls the importance between $\hat{\boldsymbol{x}}_{\mathrm{u}}$ (i.e., the user's point of view) and $\hat{\boldsymbol{x}}_{\mathrm{s}}$ (i.e., the system's point of view).

It is important to note that the values of the first and second terms in Eq. (10) (i.e., Eqs. (8) and (9)) should have the same order of asymptotic variation with respect to x_i, $i = 1, 2, \ldots, n$, which is referred to as "scale" in this paper. This will then avoid a mismatch in the scale that causes one term to be dominated by another term and thus lose much of the meaning of weight parameter γ. Specifically, the scale of function $h(x)$ with respect to x is evaluated by $\lim_{x\to\infty} h(x)/x$ in this paper. If this value converges to one, we consider that the scale of $h(x)$ is same as that of x because $h(x)$ varies with x asymptotically. Parameters a and b in Eq. (10) are used to adjust the scale of values.

It is also important to note that the values of the first and second terms on the right hand side of Eq. (10) should have the same physical dimensions. This is because a mismatch in the physical dimensions loses the physical meaning of the value of Eq. (10). In other words, the divergence measure used in the first term should have the same physical dimensions as the second term (i.e., the dimensions of x_i, $i = 1, 2, \ldots, n$). There is a detailed discussion regarding the scale and the physical dimensions of Eq. (10) in the next section.

B. Main Results

In view of the ideas in the previous section, let us then consider a case where $D(\boldsymbol{x}, \boldsymbol{y}) = D_{f_\alpha}(\boldsymbol{x}, \boldsymbol{y})$ in Eq. (10). Here, Eq. (10) reduces to

$$-\gamma a D_{f_\alpha}(\boldsymbol{x}, \boldsymbol{y}) + (1-\gamma) b \sum_{i=1}^{n} x_i. \quad (11)$$

The following demonstrates that Eq. (11) reduces to the utility function of weighted α-proportional fairness when $\gamma = 1/2$ and $\boldsymbol{y}$ is the uniform resource allocation if the scale and the physical dimensions are adjusted appropriately. This means that weighted α-proportional fairness can be regarded as an equilibrium in the trade-off between user and system satisfaction assuming that the uniformity of resource allocation is a request from users.

Before turning to the discussion about the relationship, let us consider the scale and the physical dimensions of Eq. (11). First, let us determine the values of a and b by taking the scale of the values of the first and second terms in Eq. (11) into consideration. We rewrite Eq. (11) as

$$\sum_{i=1}^{n} \left\{ -\gamma a d_{f_\alpha}(x_i, y_i) + (1-\gamma) b x_i \right\},$$

where the function, $d_{f_\alpha}(\cdot, \cdot)$, is defined as

$$\begin{aligned} d_{f_\alpha}(x_i, y_i) &= x_i f_\alpha\left(\frac{y_i}{x_i}\right) \\ &= \frac{1}{\alpha(1-\alpha)}\{(1-\alpha)x_i + \alpha y_i - x_i^{1-\alpha} y_i^{\alpha}\}. \end{aligned}$$

Here, we can see that the values of the first and second terms in Eq. (11) have the same scale with respect to x_i if the value of $d_{f_\alpha}(x_i, y_i)$ has the same scale as x_i, $i = 1, 2, \ldots, n$. The scale of the value of $d_{f_\alpha}(x_i, y_i)$ with respect to x_i is evaluated as

$$\lim_{x_i\to\infty} \frac{d_{f_\alpha}(x_i, y_i)}{x_i} = \begin{cases} \frac{1}{\alpha} & \text{if } \alpha > 0. \\ \infty & \text{if } \alpha < 0. \end{cases}$$

In addition,

$$f_\alpha(0) \stackrel{\text{def}}{=} \lim_{u\to 0} f_\alpha(u) = \lim_{x_i\to\infty} \frac{d_{f_\alpha}(x_i, y_i)}{x_i}$$

holds. Therefore, the value of a is set to $a = 1/f_\alpha(0)$ and that of b is set to $b = 1$ in this paper. By using these settings, the values of the first and second terms in Eq. (11) can have the same scale. As a result, Eq. (11) is rewritten as

$$-\gamma D_{\bar{f}_\alpha}(\boldsymbol{x}, \boldsymbol{y}) + (1-\gamma) \sum_{i=1}^{n} x_i, \quad (12)$$

where $\bar{f}_\alpha(u) \stackrel{\text{def}}{=} f_\alpha(u)/f_\alpha(0)$ and we use the relationship given in Eq. (7). Note that $\bar{f}_\alpha(u)$ satisfies the conditions for constructing f-divergence only when $\alpha > 0$, i.e., $\bar{f}_\alpha(u)$ is strictly convex and a continuously differentiable function on open interval $(0, \infty)$, satisfying $f(1) = f'(1) = 0$ only when $\alpha > 0$. Thus, from now on, we will restrict this to $\alpha > 0$.

Then, note that the physical dimensions of $D_{\bar{f}_\alpha}(\boldsymbol{x}, \boldsymbol{y})$ are the same as those for x_i, $i = 1, 2, \ldots, n$. This is because $D_{\bar{f}_\alpha}(\boldsymbol{x}, \boldsymbol{y})$ is defined as

$$D_{\bar{f}_\alpha}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} x_i \bar{f}_\alpha\left(\frac{y_i}{x_i}\right)$$

and the value of y_i/x_i, i.e., the value of $\bar{f}_\alpha(y_i/x_i)$, is dimensionless since x_i and y_i have the same dimensions.

The above discussion shows that the objective function given by Eq. (12) is well-defined from the viewpoint of scale and physical dimensions. In the following, we demonstrate that the utility function defined by Eq. (12) reduces to that of weighted α-proportional fairness when $\gamma = 1/2$.

Now, let us look closely at the properties of Eq. (12). To find the relationship with weighted α-proportional fairness, we replace x_i with $x'_i = w_i x_i$ and y_i with $y'_i = w_i y_i$ ($i = 1, 2, \ldots, n$), where $\boldsymbol{w} = (w_1, w_2, \ldots, w_i)$ is a weight vector with positive elements. Here, Eq. (12) reduces to

$$\begin{aligned}
&-\gamma D_{\bar{f}_\alpha}(\boldsymbol{x}', \boldsymbol{y}') + (1-\gamma)\sum_{i=1}^{n} x'_i \\
&= -\gamma \frac{1}{1-\alpha}\sum_{i=1}^{n}\{(1-\alpha)w_i x_i + \alpha w_i y_i - (w_i x_i)^{1-\alpha}(w_i y_i)^{\alpha}\} \\
&\quad + (1-\gamma)\sum_{i=1}^{n} w_i x_i \\
&= \gamma \sum_{i=1}^{n} w_i \frac{x_i^{1-\alpha} y_i^{\alpha}}{1-\alpha} + (1-2\gamma)\sum_{i=1}^{n} w_i x_i - \gamma \sum_{i=1}^{n} \frac{\alpha}{1-\alpha} w_i y_i.
\end{aligned}$$

In addition,

$$\begin{aligned}
&\lim_{\alpha \to 1}\left\{-\gamma D_{\bar{f}_\alpha}(\boldsymbol{x}', \boldsymbol{y}') + (1-\gamma)\sum_{i=1}^{n} x'_i\right\} \\
&= \gamma \sum_{i=1}^{n}\left\{-w_i y_i \log \frac{y_i}{x_i}\right\} + (1-2\gamma)\sum_{i=1}^{n} w_i x_i + \gamma \sum_{i=1}^{n} w_i y_i
\end{aligned}$$

holds. Therefore, the optimization problem, which uses the utility function given by Eq. (12) as an objective function, is defined as

$$\text{maximize} \quad C_{\bar{f}_\alpha,\gamma}(\boldsymbol{x}', \boldsymbol{y}'), \tag{13}$$

$$\text{subject to} \quad \boldsymbol{x} \in \mathcal{C}, \tag{14}$$

where $C_{\bar{f}_\alpha,\gamma}(\boldsymbol{x}', \boldsymbol{y}')$ is defined as:

$$C_{\bar{f}_\alpha,\gamma}(\boldsymbol{x}', \boldsymbol{y}') = \begin{cases} \gamma \sum_{i=1}^{n} w_i y_i \log x_i + (1-2\gamma)\sum_{i=1}^{n} w_i x_i & \text{if } \alpha = 1. \\ \gamma \sum_{i=1}^{n} w_i \frac{x_i^{1-\alpha} y_i^{\alpha}}{1-\alpha} + (1-2\gamma)\sum_{i=1}^{n} w_i x_i & \text{if } \alpha > 0, \alpha \neq 1. \end{cases}$$

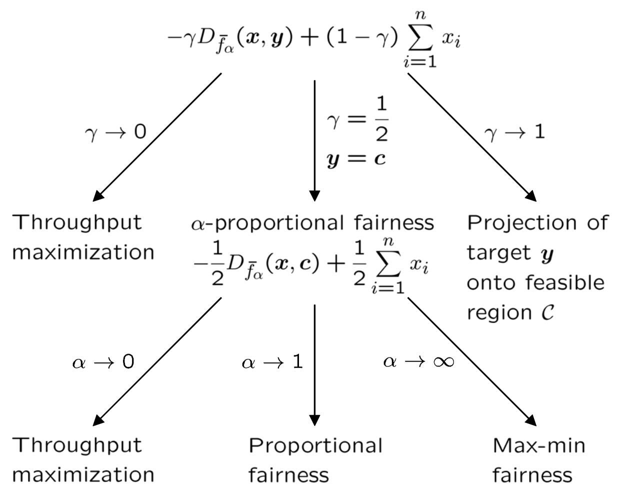

Fig. 2. Characteristics of Eq. (12)

For example, if $\gamma = 1/2$, the utility function reduces to

$$C_{\bar{f}_\alpha,1/2}(\boldsymbol{x}', \boldsymbol{y}') \propto \begin{cases} \sum_{i=1}^{n} w_i y_i \log x_i & \text{if } \alpha = 1. \\ \sum_{i=1}^{n} w_i \frac{x_i^{1-\alpha} y_i^{\alpha}}{1-\alpha} & \text{if } \alpha > 0, \alpha \neq 1. \end{cases} \tag{15}$$

The utility function defined by Eq. (15) represents an equilibrium in the trade-off between user and system satisfaction since $\gamma = 1/2$. In this case, if the requested resource allocation is given by any uniform resource allocation vector, i.e., $\boldsymbol{y} = \boldsymbol{c} = (c, c, \ldots, c) \in \mathcal{C}$, the utility function reduces to

$$C_{\bar{f}_\alpha,1/2}(\boldsymbol{x}', \boldsymbol{c}') \propto \begin{cases} \sum_{i=1}^{n} w_i \log x_i & \text{if } \alpha = 1, \\ \sum_{i=1}^{n} w_i \frac{x_i^{1-\alpha}}{1-\alpha} & \text{if } \alpha > 0, \alpha \neq 1, \end{cases} \tag{16}$$

where $\boldsymbol{c}' = (w_1 c, w_2 c, \ldots, w_n c)$. The utility function defined by Eq. (16) is equivalent to Eq. (4). This means that the utility function of weighted α-proportional fairness can be characterized as an equilibrium in the trade-off between user and system satisfaction assuming that the uniformity of resource allocation is a request from users. It is worth noting that the utility function defined by Eq. (16) is common for any uniform requested resource allocation. That is, uniformity is a sufficient condition for the requested resource allocation to obtain Eq. (16). The utility function, $C_{\bar{f}_\alpha,\gamma}(\boldsymbol{x}', \boldsymbol{y}')$, does not specifically depend on the value of c if $\gamma = 1/2$ as long as $\boldsymbol{y} = \boldsymbol{c} = (c, c, \ldots, c)$ holds, whereas it depends on the value of c if $\gamma \neq 1/2$ even when $\boldsymbol{y} = \boldsymbol{c} = (c, c, \ldots, c)$ holds.

As has been shown, the optimization problem defined by Eqs. (13) and (14) can be regarded as a generalization of the optimization problem of the weighted α-proportional fairness defined by Eqs. (4) and (5). The characteristics of Eq. (12) discussed in the above are summarized in Fig. 2.

It is important to note that the characterization of weighted α-proportional fairness obtained in this section can be regarded as a necessary result when we use α-divergence as a measure to evaluate the difference between the implemented resource allocation vector and the requested resource allocation vector. That is, α-divergence derives weighted α-proportional fairness as its natural corresponding fairness concept. In other words, the physical meaning of parameter α in the definition of α-divergence is naturally given by the connection with weighted α-proportional fairness.

Here, one might suspect that $C_{\bar{f}_\alpha,1/2}(\boldsymbol{x}', \boldsymbol{y}')$ is equivalent to the utility function of weighted α-proportional fairness with weight parameter $w_i := w_i y_i^\alpha$, and therefore the definition of the utility function in Eq. (12), which employs the notion of requested resource allocation vector $\boldsymbol{y}$, is redundant. Although this interpretation is reasonable to some extent, it is important to remember that the notion of requested resource allocation vector $\boldsymbol{y}$ is vital to characterize α-proportional fairness as an equilibrium in the trade-off between user and system satisfaction.

It is also important to note that the function, $C_{\bar{f}_\alpha,\gamma}(\boldsymbol{x}', \boldsymbol{y}')$, satisfies the common properties of the utility function given in Section II-B only when $0 < \gamma \leq 1/2$. That is, function $C_{\bar{f}_\alpha,\gamma}(\boldsymbol{x}', \boldsymbol{y}')$ is increasing and strictly concave with respect to $\boldsymbol{x} > \boldsymbol{0}$ only when $0 < \gamma \leq 1/2$. This means that it is not valid to use $1/2 < \gamma < 1$ to attach much more importance to the closeness of the implemented resource allocation to the requested resource allocation. This is because a part of the resource might be left unused in such cases even when the unused resource can be completely expended without having to sacrifice any user's resource allocation.

Summarizing the above observations, we can characterize function $C_{\bar{f}_\alpha,1/2}(\boldsymbol{x}', \boldsymbol{y}')$ from two points of view. That is, (i) function $C_{\bar{f}_\alpha,1/2}(\boldsymbol{x}', \boldsymbol{y}')$ applies a maximum weight on user satisfaction (i.e., closeness to requested resource allocation) under the constraint that it makes sense as a utility function, (ii) while it can also be regarded as an equilibrium in the trade-off between user and system satisfaction as described in the main discussion in this section.

Here, it is worth noting that the maximization of Eq. (11) can be regarded as the maximization of Rényi entropy or the maximization of Tsallis entropy with respect to $\boldsymbol{x} \in \mathcal{C}$ under certain constraints. Since the proof for the above is easy to derive, the details have been omitted due to space limitations.

A generalization of the results in this section is given in the Appendix. In the generalization, we first propose a utility function that uses Csiszár's f-divergence in place of α-divergence in Eq. (11). Then, we show that an overall utility, which is defined with Csiszár's f-divergence, appears to be an equilibrium of the trade-off expressed in the proposed utility function. The above generalization means that the definition of a divergence measure, which is used as a distance measure to evaluate the difference between an implemented resource allocation vector and a requested resource allocation vector, is equivalent to the definition of its corresponding fairness concept. Therefore, if the divergence measure is defined to reflect various kinds of properties of subjective or objective evaluations concerning the distance from the ideal quality of

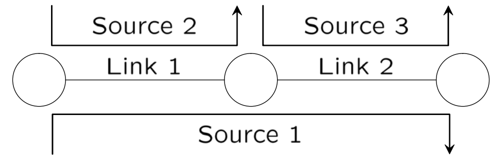

Fig. 3. Network model

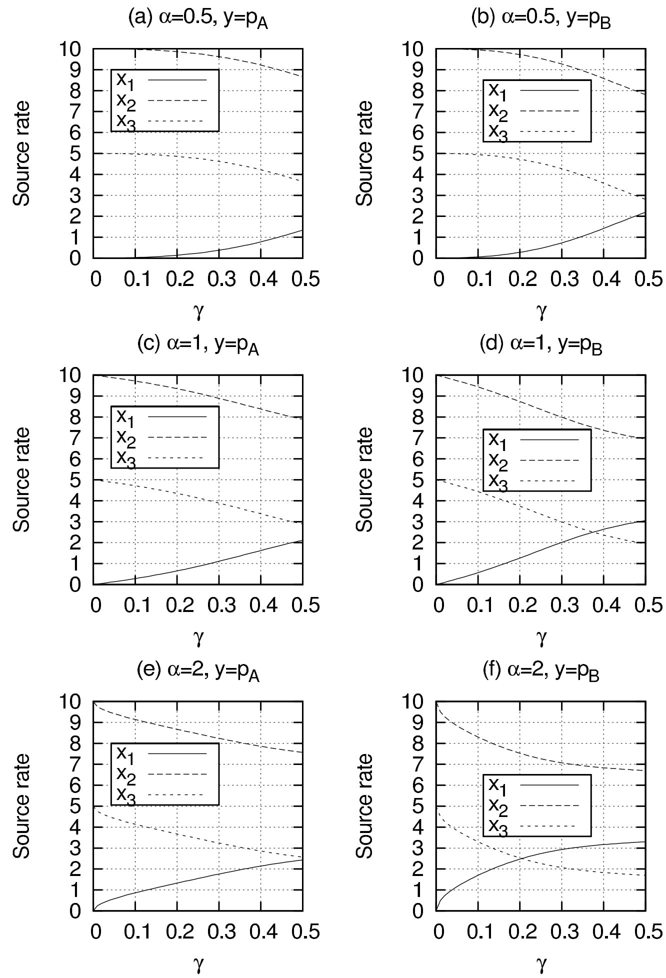

Fig. 4. Solutions of the optimization problem defined by Eqs. (13) and (14)

communication services, the definition enables us to derive its natural corresponding fairness concept with a utility function.

IV. Applications to Resource Allocation Problems in Simple Network Models

To see how the optimization problem defined by Eqs. (13) and (14) works, let us look at two simple examples of applications. The first is an application to the two-link network model, and the second is an application to wireless LANs.

A. Application to Two-Link Network

Consider the simple network in Fig. 3. The network consists of two links and three sources. Source 1's route includes both links 1 and 2, source 2's route consists of link 1, and source 3's route consists of link 2. Suppose the capacity of link 1 is

10 and that of link 2 is 5. Therefore, the feasible region is given by $\mathcal{C} = \{\boldsymbol{x}|x_1 + x_2 \le 10, x_1 + x_3 \le 5, x_1, x_2, x_3 \ge 0\}$, which is convex and compact.

Figure 4 plots the solutions to the optimization problem for different values of α and γ, where we have used $0 < \gamma \le 1/2$ taking into consideration the discussion in Section III-B. We have also used $\boldsymbol{w} = (1, 1, \ldots, 1)$ for simplicity. Figures 4(a), 4(c), and 4(e) are for $\boldsymbol{y} = \boldsymbol{p}_A = \{2.5, 2.5, 2.5\}$, and Figs. 4(b), 4(d), and 4(f) are for $\boldsymbol{y} = \boldsymbol{p}_B = \{5, 2.5, 2.5\}$. The requested vector $\boldsymbol{p}_B$ is not feasible (i.e., $\boldsymbol{p}_B \notin \mathcal{C}$).

As seen in these figures, the rate allocated to source 1 (the rates allocated to sources 2 and 3) becomes larger (smaller) when $\boldsymbol{p}_B$ is used as a requested vector, compared with when $\boldsymbol{p}_A$ is used as a requested vector. This is because the requested value of source 1 (i.e., the value of y_1) of $\boldsymbol{p}_B$ is larger than that of $\boldsymbol{p}_A$. In addition, we can see that the solution reacts more sensitively to the value of γ as the value of α increases. This is because the gap in the order with respect to x_i, $i = 1, 2, 3$, between the first and the second terms in Eq. (12) increases as the value of α rises. In other words, the impact of the first term increases as the value of α rises.

B. Application to Wireless LANs

Here, we apply the optimization problem defined by Eqs. (13) and (14) to a resource allocation problem in wireless LANs (WLANs) [24].

1) Problem Formulation: Let t_i be the total amount of air time used by wireless station (WS) $i \in \mathcal{N}$ measured over a very long period. The fraction of air time used by WS i is then

$$s_i = \frac{t_i}{\sum_{j=1}^{N} t_j}.$$

In addition, let r_i be the transmission rate of WS $i \in \mathcal{N}$. The throughput of WS i is then

$$x_i = s_i r_i. \tag{17}$$

Note that a suitable transmission rate r_i, $i \in \mathcal{N}$, is predetermined by the distances of WSs from the access point and the channel conditions, or simply by the IEEE standards (802.11b or 802.11g) used by their wireless cards. For example, 802.11b stations have a maximum data rate of 11 Mbps, while 802.11g stations have a maximum data rate of 54 Mbps. Thus, r_i, $i \in \mathcal{N}$ is constant for the optimization problem.

Referring to Eqs. (13) and (14), let us define the following optimization problem:

$$\text{maximize} \quad C_{\bar{f}_\alpha,\gamma}(\boldsymbol{x}', \boldsymbol{y}'), \tag{18}$$

$$\text{subject to} \quad \boldsymbol{s} \in \left\{ \boldsymbol{s} \,\middle|\, \sum_{i=1}^{n} s_i = 1, s_i \ge 0, i \in \mathcal{N} \right\}, \tag{19}$$

where $\boldsymbol{x}' = (w_1 x_1, w_2 x_2, \ldots, w_n x_n)$ and $\boldsymbol{s} = (s_1, \ldots, s_n)$. Note that $\boldsymbol{x}$ is a function of $\boldsymbol{s}$ as defined in Eq. (17).

2) Analysis: To solve the above optimization problem, let us define the Lagrangian as:

$$L_\alpha(\boldsymbol{s}, \mu_\alpha) = C_{\bar{f}_\alpha,\gamma}(\boldsymbol{x}', \boldsymbol{y}') + \mu_\alpha \left(1 - \sum_{i=1}^{n} s_i\right),$$

where μ_α is a Lagrange multiplier. Solution $\boldsymbol{s}^* = (s_1^*, \ldots, s_n^*)$ satisfies

$$\frac{\partial L_\alpha}{\partial s_i}(\boldsymbol{s}^*, \mu_\alpha) = \gamma w_i s_i^{*-\alpha} r_i^{1-\alpha} y_i^\alpha + (1-2\gamma) w_i r_i - \mu_\alpha = 0,$$

for all $n \in \mathcal{N}$. Thus, the solution is obtained as:

$$s_i^* = \left\{ \frac{\gamma w_i r_i^{1-\alpha} y_i^\alpha}{\mu_\alpha - (1-2\gamma) w_i r_i} \right\}^{\frac{1}{\alpha}}, \tag{20}$$

$$x_i^* = \left\{ \frac{\gamma w_i r_i y_i^\alpha}{\mu_\alpha - (1-2\gamma) w_i r_i} \right\}^{\frac{1}{\alpha}}, \tag{21}$$

where $x_i^* = s_i^* r_i$ and μ_α is determined to satisfy the condition given by Eq. (19). A closed-form solution can be obtained when $\gamma = 1/2$ as:

$$s_i^* = \frac{w_i^{\frac{1}{\alpha}} r_i^{\frac{1-\alpha}{\alpha}} y_i}{\sum_{j=1}^{n} w_j^{\frac{1}{\alpha}} r_j^{\frac{1-\alpha}{\alpha}} y_j},$$

$$x_i^* = \frac{w_i^{\frac{1}{\alpha}} r_i^{\frac{1}{\alpha}} y_i}{\sum_{j=1}^{n} w_j^{\frac{1}{\alpha}} r_j^{\frac{1-\alpha}{\alpha}} y_j},$$

where $x_i^* = s_i^* r_i$.

Two kinds of examples for Eqs. (20) and (21) are presented in Table I. The first example is for $\gamma = 1/2$, $\boldsymbol{w} = (1, 1, \ldots, 1)$ and $\boldsymbol{y} = \boldsymbol{c} = (c, c, \cdots, c) \in \mathcal{C}$. Vector $\boldsymbol{c}$ represents a uniform allocation of resources. As discussed in Section III-B, this case corresponds to α-proportional fairness. The second example is for $\gamma = 1/2$, $\boldsymbol{w} = (1, 1, \ldots, 1)$, and $\boldsymbol{y} = \boldsymbol{r}$, where $\boldsymbol{r}$ is the transmission rate of WS $i \in \mathcal{N}$. The results in this table indicate that the optimization problem defined by Eqs. (13) and (14) is richly expressive in network resource allocation.

V. Conclusion

An information theoretic cognitive process for establishing a valid background of fairness concepts in network resource allocation problems has been presented. The fundamental idea behind this paper was to characterize the utility function used in the optimization problem that motivates weighted α-proportional fairness as an equilibrium in the trade-off between user and system satisfaction in an implemented resource allocation. User satisfaction was evaluated based on the difference between the implemented resource allocation and a requested resource allocation using α-divergence. System satisfaction was evaluated based on the efficiency of the implemented resource utilization that is defined as the total amount of resources allocated to each user. We demonstrated that α-divergence derives weighted α-proportional fairness as its natural corresponding fairness concept. This means that the natural corresponding fairness concept would be necessarily derived if the properties of the difference measure between the

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2009 proceedings.

TABLE I
EXAMPLES OF EQS. (20) AND (21)

	$\gamma = \frac{1}{2}$, $\boldsymbol{w} = (1, 1, \ldots, 1)$, $\boldsymbol{y} = \boldsymbol{c}$	$\gamma = \frac{1}{2}$, $\boldsymbol{w} = (1, 1, \ldots, 1)$, $\boldsymbol{y} = \boldsymbol{r}$
$\alpha \to 0$	$s_i^* = \begin{cases} \frac{1}{m}, & \text{if } r_i = \max\{r_1, \ldots, r_n\} \\ 0, & \text{otherwise} \end{cases}$, $x_i^* = \begin{cases} \frac{r_i}{m}, & \text{if } r_i = \max\{r_1, \ldots, r_n\} \\ 0, & \text{otherwise} \end{cases}$, where $m = \lvert\{r_i \mid r_i = \max\{r_1, \ldots, r_N\}\}\rvert$	
$\alpha = \frac{1}{2}$	$s_i^* = \frac{r_i}{\sum_{j=1}^n r_j}$, $x_i^* = \frac{r_i^2}{\sum_{j=1}^n r_j}$	$s_i^* = \frac{r_i^2}{\sum_{j=1}^n r_j^2}$, $x_i^* = \frac{r_i^3}{\sum_{j=1}^n r_j^2}$
$\alpha = 1$	$s_i^* = \frac{1}{n}$, $x_i^* = \frac{r_i}{n}$	$s_i^* = \frac{r_i}{\sum_{j=1}^n r_j}$, $x_i^* = \frac{r_i^2}{\sum_{j=1}^n r_j}$
$\alpha = 2$	$s_i^* = \frac{1/\sqrt{r_i}}{\sum_{j=1}^n 1/\sqrt{r_j}}$, $x_i^* = \frac{\sqrt{r_i}}{\sum_{j=1}^n 1/\sqrt{r_j}}$	$s_i^* = \frac{\sqrt{r_i}}{\sum_{j=1}^N \sqrt{r_j}}$, $x_i^* = \frac{r_i\sqrt{r_i}}{\sum_{j=1}^N \sqrt{r_j}}$
$\alpha \to \infty$	$s_i^* = \frac{r_i^{-1}}{\sum_{j=1}^n r_j^{-1}}$, $x_i^* = \frac{1}{\sum_{j=1}^n r_j^{-1}}$	$s_i^* = \frac{1}{n}$, $x_i^* = \frac{r_i}{n}$

implemented resource allocation and the requested resource allocation could be defined accordingly. There is a detailed discussion on this in the Appendix. Also, we gave two examples of applications to simple network models to demonstrate how the characterization presented in this paper works. We believe that the results in this paper will provide a glimpse into the inherent connection between resource allocation problems and information theory.

ACKNOWLEDGMENTS

This work was supported in part by the Japan Society for the Promotion of Science through a Grant-in-Aid for Scientific Research (S) (18100001). We would like to thank the anonymous reviewers who provided us with invaluable comments and feedback on this paper.

REFERENCES

[1] D. Bertsekas and R. Gallager, *Data Networks*. Prentice Hall, 1987.

[2] F. P. Kelly, "Charging and rate control for elastic traffic," *European Transactions on Telecommunications*, vol. 8, no. 1, pp. 33–37, January 1997.

[3] L. Massoulié and J. Roberts, "Bandwidth sharing: Objectives and algorithms," in *Proceedings of the 18th Annual IEEE Conference on Computer Communications (INFOCOM'99)*, New York, NY, USA, March 1999, pp. 1395–1403.

[4] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Transactions on Networking*, vol. 8, no. 5, pp. 556–567, October 2000.

[5] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, no. 3, pp. 237–252, March 1998.

[6] R. Agrawal and V. Subramanian, "Optimality of certain channel aware scheduling policies," in *Proceedings of the 40th Annual Allerton Conference on Communication, Control and Computing*, Allerton, IL, USA, October 2002, pp. 1532–1541.

[7] J.-Y. L. Boudec, "Rate adaptation, congestion control and fairness: A tutorial," December 2006, Ecole Polytechnique Fédérale de Lausanne (EPFL).

[8] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," in *Proceedings of IEEE Vehicular Technology Conference (VTC'00)*, vol. 3, Boston, MA, USA, September 2000, pp. 1854–1858.

[9] T. Bonald and L. Massoulie, "Impact of fairness on internet performance," *ACM SIGMETRICS Performance Evaluation Review*, vol. 29, no. 1, pp. 82–91, June 2001.

[10] S. Borst and P. Whiting, "Dynamic rate control algorithms for HDR throughput optimization," in *Proceedings of 20th Annual IEEE Conference on Computer Communications (INFOCOM'01)*, Anchorage, AK, USA, April 2001, pp. 976–985.

[11] P. Viswaanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1277–1294, June 2002.

[12] R. Srikant, *The Mathematics of Internet Congestion Control*. Birkhäuser, 2003.

[13] Y. Ohta, M. Tsuru, and Y. Oie, "Fairness property and TCP-level performance of unified scheduling algorithm in HSDPA networks," in *Proceedings of the 3rd European Conference on Universal Multiservice Networks (ECUMN'04)*, Porto, Portugal, October 2004, pp. 185–195 (LNCS 3262).

[14] L. Massoulié, "Structural properties of proportional fairness: Stability and insensitivity," Microsoft Research, Tech. Rep. MSR-TR-2005-102, August 2005.

[15] Y. Zhang, M. Uchida, M. Tsuru, and Y. Oie, "Scheduling algorithms with error rate consideration in HSDPA networks," in *Proceedings of International Wireless Communications and Mobile Computing Conference (IWCMC'06)*, Vancouver, Canada, July 2006, pp. 1241–1246.

[16] S. Amari and H. Nagaoka, *Methods of Information Geometry*. American Mathematical Society, 2000.

[17] S. Kullback, *Information Theory and Statistics*. John Wiley and Sons, 1959.

[18] M. Uchida, "Information theoretic aspects of fairness criteria in network resource allocation problems," in *Proceedings of the First International Workshop on Game Theory for Communication Networks (GameComm'07)*, Nantes, France, October 2007.

[19] M. Uchida and H. Shioya, "A study on an extended formula of divergence measures using invariance," *Electronics and Communications in Japan (Part 3: Fundamental Electronic Science)*, vol. 88, no. 4, pp. 35–42, December 2004.

[20] A. Rényi, "On measures of entropy and information," in *Proceedings of the 4th Berkeley Symposium on Mathematical Satatistics and Probability*, vol. 1, 1961, pp. 547–561.

[21] C. Tsallis, "Possible generalization of Boltzmann-Gibbs statistics," *Journal of Statistical Physics*, vol. 52, no. 1–2, pp. 479–487, July 1988.

[22] L. Csiszár, "Information-type measures of different of probability distributions," *Studia Scientiarum Mathematicarum Hungarica*, vol. 2, pp. 299–318, 1967.

[23] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 1991.

[24] L. B. Jiang and S. C. Liew, "Proportional fairness in wireless LANs and ad hoc networks," in *Proceedings of IEEE Wireless Communications and*

Networking Conference (WCNC'05), vol. 3, New Orleans, LA, USA, March 2005, pp. 1551–1556.

[25] T. Harks and T. Poschwatta, "Priority pricing in utility fair networks," in *Proceedings of the 13TH IEEE International Conference on Network Protocols (ICNP'05)*, Boston, MA, USA, November 2005, pp. 311–320.

APPENDIX

The results in Section III-B are generalized in this appendix. Let $\varphi(\cdot)$ be an increasing, strictly concave, and continuously differentiable function on the open interval $(0, \infty)$. Using the function $\varphi(\cdot)$, let us define a function $\psi(\cdot)$ as:

$$\psi(u) = -\left\{ u\varphi\left(\frac{1}{u}\right) - \varphi(1) \right\} + \{\varphi(1) - \varphi'(1)\}(u-1).$$

It is easy to see that function $\psi(\cdot)$ is strictly convex and continuously differentiable on open interval $(0, \infty)$, satisfying $\psi(1) = \psi'(1) = 0$. Therefore, the f-divergence can be defined using function $\psi(\cdot)$ as

$$\begin{aligned} &D_\psi(\boldsymbol{x}, \boldsymbol{y}) \\ &= \sum_{i=1}^{n} x_i \psi\left(\frac{y_i}{x_i}\right) \\ &= \sum_{i=1}^{n} \left\{ -y_i \varphi\left(\frac{x_i}{y_i}\right) + \varphi'(1)x_i + (\varphi(1) - \varphi'(1))y_i \right\}. \end{aligned}$$

It is important to note that if $\varphi(u) = \varphi_\alpha(u)$, then $\psi(u) = \bar{f}_\alpha(u)$.

Here, let us consider a case where $D(\boldsymbol{x}, \boldsymbol{y}) = D_\psi(\boldsymbol{x}, \boldsymbol{y})$ in Eq. (10). In this case, the objective function reduces to

$$-\gamma a D_\psi(\boldsymbol{x}, \boldsymbol{y}) + (1-\gamma) b \sum_{i=1}^{n} x_i. \tag{22}$$

Along the same lines of the discussion in Section III-B, let us determine the values of a and b considering the scale of the values of the first and second terms in Eq. (22). Equation (22) is rewritten as

$$\sum_{i=1}^{n} \left\{ -\gamma a d_\psi(x_i, y_i) + (1-\gamma) b x_i \right\},$$

where the function $d_\psi(\cdot, \cdot)$ is defined as

$$\begin{aligned} d_\psi(x_i, y_i) &= x_i \psi\left(\frac{y_i}{x_i}\right) \\ &= -y_i \varphi\left(\frac{x_i}{y_i}\right) + \varphi'(1)x_i + (\varphi(1) - \varphi'(1))y_i. \end{aligned}$$

The scale of the value of $d_\psi(x_i, y_i)$ with respect to x_i is evaluated as

$$\psi(0) \stackrel{\text{def}}{=} \lim_{u \to 0} \psi(u) = \lim_{x_i \to \infty} \frac{d_\psi(x_i, y_i)}{x_i} = \varphi'(1),$$

where $\varphi(u) = o(u)$ holds because the function $\varphi(\cdot)$ is defined as increasing, strictly concave, and continuously differentiable on open interval $(0, \infty)$. Thus, this paper sets the values of a and b to $a = 1/\psi(0) = 1/\varphi'(1)$ and $b = 1$. By using these values, the values of the first and second terms in Eq. (22) can have the same scale. In this case, Eq. (22) is rewritten as

$$-\gamma D_{\bar{\psi}}(\boldsymbol{x}, \boldsymbol{y}) + (1-\gamma) \sum_{i=1}^{n} x_i, \tag{23}$$

where $\bar{\psi}(u) = \psi(u)/\psi(0) = \psi(u)/\varphi'(1)$. Also, for the same reason as that in Section III-B, the first and second terms in Eq. (23) have the same physical dimensions. Therefore, we can see that the objective function given by Eq. (23) is well-defined from the viewpoint of scale and physical dimensions.

If we replace x_i with $x'_i = w_i x_i$ and y_i with $y'_i = w_i y_i$ $(i = 1, 2, \ldots, n)$, where $\boldsymbol{w} = (w_1, w_2, \ldots, w_i)$ is a weight vector with positive elements, Eq. (23) reduces to

$$\begin{aligned} &-\gamma D_{\bar{\psi}}(\boldsymbol{x}', \boldsymbol{y}') + (1-\gamma) \sum_{i=1}^{n} x'_i \\ &= -\gamma \sum_{i=1}^{n} w_i x_i \bar{\psi}\left(\frac{y_i}{x_i}\right) + (1-\gamma) \sum_{i=1}^{n} w_i x_i \\ &= -\gamma \sum_{i=1}^{n} \left\{ -w_i y_i \varphi\left(\frac{x_i}{y_i}\right) \frac{1}{\varphi'(1)} + w_i x_i + \frac{\varphi(1) - \varphi'(1)}{\varphi'(1)} y_i \right\} \\ &\quad + (1-\gamma) \sum_{i=1}^{n} w_i x_i \\ &= \frac{\gamma}{\varphi'(1)} \sum_{i=1}^{n} w_i y_i \varphi\left(\frac{x_i}{y_i}\right) \\ &\quad + (1-2\gamma) \sum_{i=1}^{n} w_i x_i - \frac{\gamma}{\varphi'(1)} (\varphi(1) - \varphi'(1)) y_i. \end{aligned}$$

Therefore, the optimization problem, which uses the objective function defined by Eq. (23), is given by

$$\text{maximize} \quad C_{\bar{\psi},\gamma}(\boldsymbol{x}', \boldsymbol{y}'), \tag{24}$$

$$\text{subject to} \quad \boldsymbol{x} \in \mathcal{C}, \tag{25}$$

where $C_{\bar{\psi},\gamma}(\boldsymbol{x}, \boldsymbol{y})$ is defined as

$$C_{\bar{\psi},\gamma}(\boldsymbol{x}', \boldsymbol{y}') = \frac{\gamma}{\varphi'(1)} \sum_{i=1}^{n} w_i y_i \varphi\left(\frac{x_i}{y_i}\right) + (1-2\gamma) \sum_{i=1}^{n} w_i x_i.$$

For example, if $\gamma = 1/2$, the objective function reduces to

$$C_{\bar{\psi},1/2}(\boldsymbol{x}', \boldsymbol{y}') \propto \sum_{i=1}^{n} \varphi_i(x_i), \tag{26}$$

where $\varphi_i(u) = w_i y_i \varphi(u/y_i)$. The objective function defined by Eq. (26) can be regarded as a special case defined by Eq. (1). The above result means that a sum of utility $\varphi_i(x_i)$, which is associated with Csiszár's f-divergence, appears to be an equilibrium of the trade-off expressed in Eq. (23). Here, for the same reason as in Section III-B, the value of control parameter γ should be restricted to $0 < \gamma \leq 1/2$.

Although we have focused on concave utility functions in this paper, we intend to generalize our results in future studies by taking into consideration other fairness concepts defined with non-concave utility functions as in [25].