

# Information Theoretic Aspects of Fairness Criteria in Network Resource Allocation Problems

Masato Uchida

Network Design Research Center, Kyushu Institute of Technology  
3-8-1 Asano, Kokura-Kita-ku, Kitakyusyu, Fukuoka 802-0001, JAPAN  
m.uchida@ndrc.kyutech.ac.jp

## ABSTRACT

The present paper provides a novel characterization of fairness criteria in network resource allocation problems based on information theory. Specifically, the optimization problems that motivate fairness criteria for multi-dimensional resource are characterized using information divergence measures that were originally used in information theory. The characteristics of the fairness criteria clarified herein are summarized as follows: (i) The proportional fairness criterion can be derived through the minimization of the Kullback-Leibler divergence. (ii) The  $(\mathbf{p}, \alpha)$ -proportional fairness criterion, which is a generalization of the proportional fairness criterion, can be derived through the minimization of the  $\alpha$ -divergence and the power-divergence. In addition, the optimization of the fairness criterion is closely related to the Tsallis entropy maximization principle. (iii) The above relationships can be generalized using Csiszár's  $f$ -divergence and Bregman's divergence. The information theoretic approach is then applied to a typical example in a practical network resource allocation problem. This example provides a glimpse into the inherent connection between resource allocation problems and information theory.

## Categories and Subject Descriptors

C.2.0 [Computer-Communication Networks]: General—Data communications; C.2.3 [Computer-Communication Networks]: Network Operations—Network management

## General Terms

Algorithms, Theory

## Keywords

Resource Allocation, Fairness, Information Theory

## 1. INTRODUCTION

The present paper considers an abstract network resource allocation problem, where the network has a resource with

finite capacity and supports a finite number of users. When users have to compete for a resource, the share allocated to each user has to be regulated by some control mechanism in order to avoid network congestion and performance degradation. As a result of the finite capacity of the resource, a compromise of resource allocation leads to the concept of fairness.

The present paper examines the mathematical concept of fairness that is formulated as an optimization problem, where the objective is to find a feasible resource allocation that minimizes or maximizes a utility function specific for the used fairness criterion. Examples of such fairness criteria are throughput maximization [6], max-min fairness [6], proportional fairness [19], and potential delay minimization criteria [24]. These examples and their weighted versions are classified as  $(\mathbf{p}, \alpha)$ -proportional fairness criteria [25]. Specifically, the mathematical formulation of the  $(\mathbf{p}, \alpha)$ -proportional fairness realizes a number of fairness criteria by varying the values of the parameters  $\mathbf{p}$  and  $\alpha$ . For example, the weighted and non-weighted versions are defined when  $\mathbf{p} \neq \mathbf{1}$  and  $\mathbf{p} = \mathbf{1}$ , respectively. If  $\mathbf{p} = \mathbf{1}$ , the throughput maximization criterion is obtained when  $\alpha \rightarrow 0$ , the proportional fairness criterion is obtained when  $\alpha \rightarrow 1$ , the potential delay minimization criterion is obtained when  $\alpha \rightarrow 2$ , and the max-min fairness criterion is obtained when  $\alpha \rightarrow \infty$ . A more general concept is the utility fairness criterion [1, 2, 9, 14, 15, 19, 20], which includes the  $(\mathbf{p}, \alpha)$ -proportional fairness as a special case. The utility fairness criterion is defined with a utility function, and fair rate allocation is determined by the solution of the optimization problem specific for the utility fairness criterion.

After their proposal, a number of studies examined the application, extension, and evaluation of the abovementioned fairness criteria. For example, static/dynamic or packet/flow level performance of the fairness criteria in various network models have been examined extensively [1, 2, 7, 8, 17, 20, 23, 26, 31, 32]. However, the origin of the fairness criteria has not yet been clarified and remains unknown despite active research in this field. Thus, the present paper provides an insight into a new cognitive process for establishing a valid background for the abovementioned fairness criteria. To this end, the present paper employs information theory and presents a new aspect of the abovementioned fairness criteria. The result shows that there exist an interesting relationship between the abovementioned fairness criteria and information divergence measures that play an important role in information theory.

Information theory was created by Shannon [27], who also

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*GameComm '07*, October 22, 2007, Nantes, France  
Copyright 2007 ICST 978-963-9799-00-4.

introduced the information entropy measure. A formula similar to Shannon entropy in the sense of statistical physics may be traced back to Boltzmann. In addition, a well-known information divergence measure was introduced by Kullback and Leibler [22]. The information divergence is now referred to as Kullback-Leibler divergence (KL-divergence, in short). The KL-divergence plays an important role in information theory. In addition, the minimization of the KL-divergence is closely related to a principle of statistical inference [21] and has an important relationship with information geometry theory [3, 4]. Information-geometric interpretation of information divergence measures leads to an extension of KL-divergence called  $\alpha$ -divergence [3, 4], which is closely related to Tsallis entropy [29], which is a generalization of Shannon entropy. An important class of information divergences introduced by Csiszár is called  $f$ -divergence [12]. The KL-divergence and  $\alpha$ -divergence are important members of Csiszár's  $f$ -divergence. Another important class of information divergence was introduced by Bregman [10]. An important member of Bregman's divergence is the power divergence introduced by Basu et al. [5]. The power divergence includes the KL-divergence as a special case and has a number of statistical applications [5, 13]. The relationship between the  $\alpha$ -divergence and the power divergence has been clearly presented as a genealogy from Csiszár's  $f$ -divergence [30]. This genealogy enables us to apply the information theory to the phase problem of materials science [28].

The present paper presents the following relationships between resource allocation problems and information theory. (i) The proportional fairness criterion can be derived through minimization of the KL-divergence. (ii) The  $(\mathbf{p}, \alpha)$ -proportional fairness criterion can be derived through minimization of the  $\alpha$ -divergence and the power-divergence. In addition, the optimization of the fairness criterion is closely related to the Tsallis entropy maximization principle. (iii) The above relationships can be generalized using Csiszár's  $f$ -divergence and Bregman's divergence.

In addition, the present paper provides a typical example to which the abovementioned formulation can be applied. This example is a generalization of [18], through which a glimpse of the inherent connection between practical resource allocation problems and information theory is obtained.

The present paper is organized as follows. Sections 2 and 3 briefly review the mathematical formulation of the fairness criteria and the information divergence measures, respectively. Section 4 presents a characterization of  $(\mathbf{p}, \alpha)$ -proportional fairness using  $\alpha$ -divergence and power divergence. Referring to the characterization, Section 5 introduces a possible method for inducing information divergence measures from the utility function specific for the fairness criterion in question. Section 6 generalizes the result presented in Section 4 using the induced information divergence measures. Section 7 presents an example that shows the inherent connection between resource allocation problems and information theory. Section 8 concludes the present paper.

## 2. BRIEF REVIEW OF FAIRNESS CRITERIA

### 2.1 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  $x_n$  be the share of resource that is allocated to user  $n \in \mathcal{N}$ . The resource allocation vector is denoted by  $\mathbf{x} = \{x_1, \dots, x_N\}$ , where  $0 < x_n < \infty$ ,  $n \in \mathcal{N}$ . Let  $\mathcal{C}$  be the feasible region of resource allocation. That is, the resource allocation vector  $\mathbf{x}$  is feasible if  $\mathbf{x} \in \mathcal{C}$ . Suppose that the set  $\mathcal{C}$  is convex and compact. One of the representative resource measures considered in the resource allocation problem is bandwidth.

Note that the network model introduced in [19] is an example that satisfies the abovementioned definition. The network model is defined as follows: We consider a set of sources (users)  $n = 1, \dots, N$  and links  $l = 1, \dots, L$ . Let  $A_{l,n}$  be the fraction of traffic of user  $n$  which flows on link  $l$ ,  $x_n$  be the sending rate of user  $n$ , and  $c_l$  be the capacity of link  $l$ . Thus, a feasible allocation of rates  $x_n$  is defined by  $\sum_{n=1}^N A_{l,n} x_n \leq c_l$  for all  $l$ . In this case, the feasible region of resource allocation is defined as  $\mathcal{C} = \{\mathbf{x} \mid \sum_{n=1}^N A_{l,n} x_n \leq c_l\}$ . Note that the set  $\mathcal{C}$  is convex and compact.

### 2.2 $(\mathbf{p}, \alpha)$ -Proportional Fairness

A general fairness criterion called  $(\mathbf{p}, \alpha)$ -proportional fairness was proposed in [25]. A resource allocation vector  $\mathbf{x}^* = \{x_1^*, \dots, x_N^*\}$  is  $(\mathbf{p}, \alpha)$ -proportionally fair if it is feasible, that is,  $\mathbf{x}^* \in \mathcal{C}$ , and if for any other feasible vector  $\mathbf{x}$ , the following condition is satisfied:

$$\sum_{n=1}^N p_n \frac{x_n - x_n^*}{x_n^{\alpha}} \leq 0,$$

where  $\mathbf{p} = \{p_1, \dots, p_N\}$  is a weight vector with positive elements.

The  $(\mathbf{p}, \alpha)$ -proportional fairness can be motivated in another way. Consider the following optimization problem:

$$\text{maximize} \quad \sum_{n=1}^N p_n \phi_{\alpha}(x_n), \quad (1)$$

$$\text{subject to} \quad \mathbf{x} \in \mathcal{C}, \quad (2)$$

where  $\phi_{\alpha}(\cdot)$ ,  $\alpha > 0$  is an increasing, strictly concave, and continuously differentiable function on the open interval  $(0, \infty)$ , as follows:

$$\phi_{\alpha}(x) = \begin{cases} \log x, & \text{if } \alpha = 1. \\ \frac{x^{1-\alpha}}{1-\alpha}, & \text{otherwise.} \end{cases} \quad (3)$$

Since the objective function given in Eq. (1) is strictly concave and the feasible region given in Eq. (2) is supposed to be convex and compact, the optimal solution of the above problem exists and is unique.

Note that the  $(\mathbf{p}, \alpha)$ -proportional fairness criterion reduces to several well-known criteria, when the weight vector is set to  $\mathbf{p} = \mathbf{1} = \{1, \dots, 1\}$ . That is, the  $(\mathbf{1}, \alpha)$ -proportional fairness criterion corresponds to the maximum throughput criterion [6] when  $\alpha \rightarrow 0$ , to the proportional fairness criterion [19] when  $\alpha \rightarrow 1$ , to the potential delay minimization criterion [24] when  $\alpha \rightarrow 2$ , and to the max-min fairness criterion [6] when  $\alpha \rightarrow \infty$ .

### 2.3 Utility Fairness

A more general resource allocation criterion called utility fairness was proposed in [1, 2, 9, 14, 15, 19, 20]. A resource allocation vector  $\mathbf{x}^* = \{x_1^*, \dots, x_N^*\}$  is utility fair if it is

feasible, that is,  $\mathbf{x}^* \in \mathcal{C}$ , and if for any other feasible vector  $\mathbf{x}$ , the following condition is satisfied:

$$\sum_{n=1}^N \frac{\partial \phi_n}{\partial x_n}(x_n^*)(x_n - x_n^*) \leq 0,$$

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

Similar to the  $(\mathbf{p}, \alpha)$ -proportional fairness approach, the utility fairness approach is translated into the following optimization problem:

$$\text{maximize } \sum_{n=1}^N \phi_n(x_n), \quad (4)$$

$$\text{subject to } \mathbf{x} \in \mathcal{C}. \quad (5)$$

Since the objective function given in Eq. (4) is strictly concave and the feasible region given in Eq. (5) is supposed to be convex and compact, the optimal solution of the above problem exists and is unique.

Note that the utility fairness is equivalent to the  $(\mathbf{p}, \alpha)$ -proportional fairness when

$$\phi_n(x) = p_n \phi_\alpha(x).$$

### 3. BRIEF REVIEW OF INFORMATION DIVERGENCE MEASURES

Two classes of information divergence measures, the Csiszár's  $f$ -divergence [12] and Bregman's divergence [10], are introduced in this section. These information divergence measures define a kind of distance between vectors  $\mathbf{u} = \{u_1, \dots, u_I\}$  and  $\mathbf{v} = \{v_1, \dots, v_I\}$ , where the components of the vectors are positive (i.e.,  $u_i, v_i > 0, i = 1, \dots, I$ ). Although the original definition is only for the vector that satisfies  $\sum_{i=1}^I u_i = 1$  and  $\sum_{i=1}^I v_i = 1$  (i.e., probability measure), the general definition introduced in this section is applicable for the vector that satisfies  $\sum_{i=1}^I u_i < \infty$  and  $\sum_{i=1}^I v_i < \infty$  (i.e., positive finite measure),

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

$$D_f(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^I u_i f\left(\frac{v_i}{u_i}\right).$$

In addition, Bregman's divergence is defined by a strictly convex, continuously differentiable function  $g(\cdot)$  on the open interval  $(0, \infty)$  as follows:

$$B_g(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^I \{g(u_i) - g(v_i) - g'(v_i)(u_i - v_i)\}.$$

Note that

$$\begin{aligned} D_f(\mathbf{u}, \mathbf{v}) &\geq 0, \\ B_g(\mathbf{u}, \mathbf{v}) &\geq 0, \end{aligned}$$

hold, where the equality holds if and only if  $\mathbf{u} \equiv \mathbf{v}$ . Thus,  $D_f(\cdot, \cdot)$  and  $B_g(\cdot, \cdot)$  can be regarded as a kind of distance between  $\mathbf{u}$  and  $\mathbf{v}$ . However,  $D_f(\mathbf{u}, \mathbf{v}) = D_f(\mathbf{v}, \mathbf{u})$  and  $B_g(\mathbf{u}, \mathbf{v}) = B_g(\mathbf{v}, \mathbf{u})$  (i.e., symmetric property) do not necessary hold

in general. In addition,

$$\begin{aligned} D_{\varepsilon f}(\mathbf{u}, \mathbf{v}) &= \varepsilon D_f(\mathbf{u}, \mathbf{v}), \\ B_{\varepsilon g}(\mathbf{u}, \mathbf{v}) &= \varepsilon B_g(\mathbf{u}, \mathbf{v}) \end{aligned} \quad (6)$$

hold for any positive value  $\varepsilon > 0$ .

The  $\alpha$ -divergence [4], which is a member of Csiszár's  $f$ -divergence, is defined using  $f_\alpha(x) = \frac{1}{\alpha(1-\alpha)}\{1 - x^\alpha + \alpha(x - 1)\}$  as follows:

$$D_{f_\alpha}(\mathbf{u}, \mathbf{v}) = \frac{1}{\alpha(1-\alpha)} \sum_{i=1}^I \{(1-\alpha)u_i + \alpha v_i - u_i^{1-\alpha} v_i^\alpha\},$$

where  $\alpha \in \mathfrak{R}$ . The  $\alpha$ -divergence has a duality as follows:

$$D_{f_\alpha}(\mathbf{u}, \mathbf{v}) = D_{f_{1-\alpha}}(\mathbf{v}, \mathbf{u}), \quad (7)$$

where  $D_{f_{1-\alpha}}(\cdot, \cdot)$  is called the dual  $\alpha$ -divergence with respect to  $D_{f_\alpha}(\cdot, \cdot)$ .

In addition, the power divergence [5], which is a member of Bregman's divergence, is defined using  $g_\gamma(x) = \frac{x^{1+\gamma}-1}{\gamma(1+\gamma)}$  as follows:

$$B_{g_\gamma}(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^I \left\{ \frac{1}{\gamma} u_i (u_i^\gamma - v_i^\gamma) - \frac{1}{1+\gamma} (u_i^{1+\gamma} - v_i^{1+\gamma}) \right\},$$

where  $\gamma \in \mathfrak{R}$ . The dualistic property of the power divergence is discussed in [30]. The  $\alpha$ -divergence and the power divergence are related as follows:

$$D_{f_{\gamma/(1+\gamma)}}(\mathbf{u}^{1+\gamma}, \mathbf{v}^{1+\gamma}) = B_{g_\gamma}(\mathbf{u}, \mathbf{v})$$

where  $\mathbf{u}^{1+\gamma} = \{u_1^{1+\gamma}, \dots, u_I^{1+\gamma}\}$  and  $\mathbf{v}^{1+\gamma} = \{v_1^{1+\gamma}, \dots, v_I^{1+\gamma}\}$ . The above relationship has been clearly presented as a genealogy from the  $f$ -divergence [30].

The KL-divergence, which is a common member of the  $\alpha$ -divergence and the power divergence, is defined as follows:

$$D_{f_0}(\mathbf{u}, \mathbf{v}) = B_{g_0}(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^I \left\{ u_i \log \frac{u_i}{v_i} - u_i + v_i \right\},$$

where  $f_0(u) = \lim_{\alpha \rightarrow 0} f_\alpha(u) = -\log u + u - 1$  and  $g_0(u) = \lim_{\alpha \rightarrow 0} g_\alpha(u) = u \log u - u + 1$ .

The dual KL-divergence, which is a member of the  $\alpha$ -divergence is defined using  $f_1(x) = \lim_{\alpha \rightarrow -1} f_\alpha(x) = x \log x - x + 1$  as follows:

$$D_{f_1}(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^I \left\{ v_i \log \frac{v_i}{u_i} - v_i + u_i \right\}.$$

The quadratic divergence, which is a member of the power divergence is defined using  $g_1(x) = \lim_{\alpha \rightarrow 1} g_\alpha(x) = \frac{x^2-1}{2}$  as follows:

$$B_{g_1}(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^I \frac{(u_i - v_i)^2}{2}.$$

Note that the quadratic divergence is a symmetric information divergence.

Neyman's chi-square divergence [11], which is a member of the  $\alpha$ -divergence is defined using  $f_{-1}(x) = \lim_{\alpha \rightarrow -1} f_\alpha(x) = \frac{(x-1)^2}{2x}$  as follows:

$$D_{f_{-1}}(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^I \frac{(u_i - v_i)^2}{2v_i}.$$

The Itakura-Saito divergence [16], which is a member of the power divergence is defined using  $g_{-1}(x) = \lim_{\alpha \rightarrow -1} g_\alpha(x) = -\log x$  as follows:

$$B_{g_{-1}}(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^I \left\{ \frac{u_i}{v_i} - 1 - \log \frac{u_i}{v_i} \right\}.$$

#### 4. FAIRNESS CRITERIA CHARACTERIZED BY $\alpha$ -DIVERGENCE AND POWER DIVERGENCE

A characterization of the  $(\mathbf{p}, \alpha)$ -proportional fairness based on the  $\alpha$ -divergence and the power divergence is given in this section. The characterization provides a reasonable means of understanding the inherent meaning of the  $(\mathbf{p}, \alpha)$ -proportional fairness. Specifically, the result of this section reveals that the objective of  $(\mathbf{p}, \alpha)$ -proportional fairness can be regarded as the maximization of the utility  $\phi_\alpha(\cdot)$  of a pseudo resource allocation vector that is defined as an aggregation of the resource share allocated to the users. Note that the aggregated resource share is well-defined using the weighted sum of the  $\alpha$ -divergence. In addition, a similar characterization is possible for the power divergence.

##### 4.1 Network Model for Multi-Dimensional Resources

Although the original network model introduced in Section 2.1 is defined using a one-dimensional resource measure, the remainder of the present paper uses a network model that is defined using a multi-dimensional resource measure. That is, the resource is defined as an  $I$ -dimensional vector as follows: Let  $\mathcal{N}$  be a set of users who have to compete for the resources, where  $N$  denotes the cardinality of  $\mathcal{N}$ . Let  $x_{n,i}$  be the share of resource  $i = 1, \dots, I$  that is allocated to user  $n \in \mathcal{N}$ . In this way, the resource allocation vector of user  $n \in \mathcal{N}$  is denoted by the  $I$ -dimensional vector  $\mathbf{x}_n = \{x_{n,1}, \dots, x_{n,I}\}$ , where  $0 < x_{n,i} < \infty$ ,  $n = 1, \dots, N$ ,  $i = 1, \dots, I$ . The resource allocation vector of all users is denoted by  $\mathbf{x} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ . In addition, let  $\mathcal{C}_I$  be the feasible region of resource allocation. That is, the resource allocation vector  $\mathbf{x}$  is feasible if  $\mathbf{x} \in \mathcal{C}_I$ . Suppose that the set  $\mathcal{C}_I$  is convex and compact. The above defined network model corresponds to that defined in Section 2.1 when  $I = 1$ .

##### 4.2 $(\mathbf{p}, \alpha)$ -Proportional Fairness for a Multi-Dimensional Resource

The present paper extends the optimization problem defined by Eqs. (1) and (2) to the that for a multi-dimensional resource. The extended optimization problem is given as follows:

$$\text{maximize} \quad \sum_{n=1}^N p_n \phi_\alpha(\mathbf{x}_n), \quad (8)$$

$$\text{subject to} \quad \mathbf{x} \in \mathcal{C}_I, \quad (9)$$

where  $\phi_\alpha(\mathbf{x}_n)$  is defined by

$$\phi_\alpha(\mathbf{x}_n) = \sum_{i=1}^I \phi_\alpha(x_{n,i}). \quad (10)$$

Equation (10) means that utilities are assumed to be additive, so that the total utility of  $\mathbf{x}_n$  is  $\sum_{i=1}^I \phi_\alpha(x_{n,i})$ . The

above mentioned optimization problem equivalent to that defined in Section 2.2 when  $I = 1$ .

#### 4.3 Result

Now, a key result of the present paper can be obtained.

Let us consider a vector  $\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)} = \{\bar{v}_1^{(\mathbf{p}, \alpha)}, \dots, \bar{v}_I^{(\mathbf{p}, \alpha)}\}$  that is defined as follows:

$$\bar{v}_i^{(\mathbf{p}, \alpha)} = \left( \sum_{n=1}^N \frac{p_n}{\sum_{m=1}^N p_m} v_{n,i}^{1-\alpha} \right)^{\frac{1}{1-\alpha}}, \quad (11)$$

where  $\mathbf{v}_n = \{v_{n,1}, \dots, v_{n,I}\}$ ,  $0 < v_{n,i} < \infty$ ,  $n = 1, \dots, N$ ,  $i = 1, \dots, I$ . Here, it can be shown that the vector  $\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)}$  satisfies both

$$\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)} = \arg \min_{\substack{0 < u_i < \infty \\ i=1, \dots, I}} \sum_{n=1}^N p_n D_{f_{1-\alpha}}(\mathbf{u}, \mathbf{v}_n), \quad (12)$$

and

$$\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)} = \arg \min_{\substack{0 < u_i < \infty \\ i=1, \dots, I}} \sum_{n=1}^N p_n B_{g_{1-\alpha}}(\mathbf{u}, \mathbf{v}_n). \quad (13)$$

In addition,

$$\sum_{n=1}^N p_n \phi_\alpha(\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)}) = \sum_{n=1}^N p_n \phi_\alpha(\mathbf{v}_n), \quad (14)$$

and equivalently

$$\phi_\alpha(\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)}) = \sum_{n=1}^N \frac{p_n}{\sum_{m=1}^N p_m} \phi_\alpha(\mathbf{v}_n), \quad (15)$$

hold. The proof of the above result is given in the appendix.

This result provides an interesting viewpoint for understanding the inherent meaning of the optimization problem defined by Eqs. (8) and (9).

First, Eq. (14) shows that the utility of  $\mathbf{v}_n$ ,  $n \in \mathcal{N}$  is equivalent to that of  $\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)}$  that is defined as an aggregation of the vectors  $\mathbf{v}_n$ ,  $n \in \mathcal{N}$  in Eq. (11). Thus, the original optimization problem defined by Eqs. (8) and (9) can be translated to

$$\begin{aligned} & \text{maximize} \quad \sum_{n=1}^N p_n \phi_\alpha(\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)}), \\ & \text{subject to} \quad \mathbf{v} \in \mathcal{C}_I. \end{aligned}$$

In the sense of this characterization, the vector  $\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)}$  can be regarded as a kind of pseudo resource allocation vector.

Second, Eqs. (12) and (13) mean that the pseudo resource allocation vector  $\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)}$  can be regarded as the vector that minimizes the weighted sum of  $D_{f_{1-\alpha}}(\cdot, \cdot)$  or  $B_{g_{1-\alpha}}(\cdot, \cdot)$ , where the weight is given by the parameter  $\mathbf{p}$ . Therefore, the pseudo resource allocation vector is a kind of centroid of  $\mathbf{v}_n$ ,  $n \in \mathcal{N}$  with respect to  $D_{f_{1-\alpha}}(\cdot, \cdot)$  or  $B_{g_{1-\alpha}}(\cdot, \cdot)$ .

Finally, in addition to the above observation, note that the definition of  $\phi_\alpha(\cdot)$  is closely related to the definition of Tsallis entropy. The Tsallis entropy is defined for  $\alpha > 0$  as follows:

$$H_\alpha(\mathbf{u}) = \frac{1}{1-\alpha} \left( \sum_{i=1}^I u_i^\alpha - 1 \right).$$

Note that  $\lim_{\alpha \rightarrow 0} H_\alpha(\mathbf{u})$  reduces to Shannon entropy. Here, the following relationship is obtained using the definition of Tsallis entropy:

$$\sum_{i=1}^I \phi_\alpha(v_{n,i}) = \frac{\alpha}{1-\alpha} H_{1-\alpha}(\mathbf{v}_n).$$

Thus, the objective of the  $(\mathbf{p}, \alpha)$ -proportional fairness that is given by Eq. (14) can be regarded as the maximization of the (weighted) Tsallis entropy.

#### 4.4 Examples

Three examples are given in the following. The examples are for maximum throughput criterion (i.e.,  $\alpha \rightarrow 0$ ), proportional fairness criterion (i.e.,  $\alpha \rightarrow 1$ ), and potential delay minimization criterion (i.e.,  $\alpha \rightarrow 2$ ).

##### Example 1: $(\mathbf{p}, 0)$ -proportional fairness

When  $\alpha \rightarrow 0$ , the result given by Eqs. (11), (12), (13) and (15) reduces to the following form that corresponds to the maximum throughput criterion:

$$\begin{aligned} \bar{v}_i^{(\mathbf{p},0)} &= \sum_{n=1}^N \frac{p_n}{\sum_{m=1}^N p_m} v_{n,i}, \\ \bar{\mathbf{v}}^{(\mathbf{p},0)} &= \arg \min_{\substack{0 < u_i < \infty \\ i=1, \dots, I}} \sum_{n=1}^N p_n D_{f_1}(\mathbf{u}, \mathbf{v}_n), \\ \bar{\mathbf{v}}^{(\mathbf{p},0)} &= \arg \min_{\substack{0 < u_i < \infty \\ i=1, \dots, I}} \sum_{n=1}^N p_n B_{g_1}(\mathbf{u}, \mathbf{v}_n), \\ \phi_0(\bar{\mathbf{v}}^{(\mathbf{p},0)}) &= \sum_{n=1}^N \frac{p_n}{\sum_{m=1}^N p_m} \left\{ \sum_{i=1}^I v_{n,i} \right\}. \end{aligned}$$

As described above, the pseudo resource allocation vector  $\bar{\mathbf{v}}^{(\mathbf{p},0)}$  is defined as the centroid of  $\mathbf{v}_n$ ,  $n \in \mathcal{N}$  with respect to the dual KL-divergence  $D_{f_1}(\cdot, \cdot)$  and the quadratic divergence  $D_{g_1}(\cdot, \cdot)$ .

##### Example 2: $(\mathbf{p}, 1)$ -proportional fairness

When  $\alpha \rightarrow 1$ , the result given by Eqs. (11), (12), (13) and (15) reduces to the following form, which corresponds to the proportional fairness criterion:

$$\begin{aligned} \bar{v}_i^{(\mathbf{p},1)} &= \prod_{n=1}^N v_{n,i}^{\frac{p_n}{\sum_{m=1}^N p_m}}, \\ \bar{\mathbf{v}}^{(\mathbf{p},1)} &= \arg \min_{\substack{0 < u_i < \infty \\ i=1, \dots, I}} \sum_{n=1}^N p_n D_{f_0}(\mathbf{u}, \mathbf{v}_n), \\ \bar{\mathbf{v}}^{(\mathbf{p},1)} &= \arg \min_{\substack{0 < u_i < \infty \\ i=1, \dots, I}} \sum_{n=1}^N p_n B_{g_0}(\mathbf{u}, \mathbf{v}_n), \\ \phi_1(\bar{\mathbf{v}}^{(\mathbf{p},1)}) &= \sum_{n=1}^N \frac{p_n}{\sum_{m=1}^N p_m} \left\{ \sum_{i=1}^I \log(v_{n,i}) \right\}. \end{aligned}$$

As described above, the pseudo resource allocation vector  $\bar{\mathbf{v}}^{(\mathbf{p},1)}$  is defined as the centroid of  $\mathbf{v}_n$ ,  $n \in \mathcal{N}$  with respect to the KL-divergence  $D_{f_0}(\cdot, \cdot)$  and  $D_{g_0}(\cdot, \cdot)$ .

##### Example 3: $(\mathbf{p}, 2)$ -proportional fairness

When  $\alpha \rightarrow 2$ , the result given by Eqs. (11), (12), (13) and (15) reduces to the following form, which corresponds to the

potential delay minimization criterion:

$$\begin{aligned} \bar{v}_i^{(\mathbf{p},2)} &= \left( \sum_{n=1}^N \frac{p_n}{\sum_{m=1}^N p_m} v_{n,i}^{-1} \right)^{-1}, \\ \bar{\mathbf{v}}^{(\mathbf{p},2)} &= \arg \min_{\substack{0 < u_i < \infty \\ i=1, \dots, I}} \sum_{n=1}^N p_n D_{f_{-1}}(\mathbf{u}, \mathbf{v}_n), \\ \bar{\mathbf{v}}^{(\mathbf{p},2)} &= \arg \min_{\substack{0 < u_i < \infty \\ i=1, \dots, I}} \sum_{n=1}^N p_n B_{g_{-1}}(\mathbf{u}, \mathbf{v}_n), \\ \phi_2(\bar{\mathbf{v}}^{(\mathbf{p},2)}) &= \sum_{n=1}^N \frac{p_n}{\sum_{m=1}^N p_m} v_{n,i}^{-1}. \end{aligned}$$

As described above, the pseudo resource allocation vector  $\bar{\mathbf{v}}^{(\mathbf{p},2)}$  is defined as the centroid of  $\mathbf{v}_n$ ,  $n \in \mathcal{N}$  with respect to Neyman's chi-square divergence  $D_{f_{-1}}(\cdot, \cdot)$  and the Itakura-Saito divergence  $D_{g_{-1}}(\cdot, \cdot)$ .

## 5. INFORMATION DIVERGENCE MEASURES INDUCED FROM THE UTILITY FUNCTION

The results presented in Section 4 indicate that there exists a more general relationship between the concave function that defines fairness criterion and the convex function that defines information divergence. Thus, one possible method for inducing  $f_{1-\alpha}(\cdot)$  and  $g_{1-\alpha}(\cdot)$  from  $\phi_\alpha(\cdot)$  is presented in this section. In addition, this section generalizes the method by which to induce Csiszár's  $f$ -divergence and Bregman's divergence from the utility function.

### 5.1 Induction to Csiszar's $f$ -Divergence

Let us define a new function  $\tilde{\phi}_\alpha(x)$  using the utility function  $\phi_\alpha(\cdot)$ ,  $\alpha > 0$  that is defined in Eq. (3) as follows:

$$\begin{aligned} \tilde{\phi}_\alpha(x) &= -(\phi_\alpha(x) - \phi_\alpha(1)) + \phi'_\alpha(1)(x - 1) \\ &= \frac{1}{1-\alpha} \left\{ (1 - x^{1-\alpha}) - (1-\alpha)(1-x) \right\}. \end{aligned}$$

It is easy to confirm that the function  $\tilde{\phi}_\alpha(\cdot)$  is strictly convex, continuously differentiable on the open interval  $(0, \infty)$ , and satisfies  $\tilde{\phi}_\alpha(1) = \tilde{\phi}'_\alpha(1) = 0$ . Thus, the  $f$ -divergence can be defined using  $\tilde{\phi}_\alpha(\cdot)$  as follows:

$$D_{\tilde{\phi}_\alpha}(\mathbf{u}, \mathbf{v}) = \frac{1}{1-\alpha} \sum_{i=1}^I \{ \alpha u_i + (1-\alpha)v_i - u_i^\alpha v_i^{1-\alpha} \},$$

where  $\alpha > 0$ . Note that since  $\tilde{\phi}_\alpha(\cdot) = \alpha f_{1-\alpha}(\cdot)$  holds, the following is obtained:

$$D_{\tilde{\phi}_\alpha}(\mathbf{u}, \mathbf{v}) = \alpha D_{f_{1-\alpha}}(\mathbf{u}, \mathbf{v}), \quad (16)$$

for  $\alpha > 0$  (see Eq. (6)). This means that the  $f$ -divergence induced from the utility function  $\phi_\alpha(\cdot)$  is equivalent to the dual  $\alpha$ -divergence without the positive constant scaling factor.

Next, let us generalize the above discussion. Let  $\phi(\cdot)$  be a utility function that is an increasing, strictly concave, and continuously differentiable utility function on  $(0, \infty)$ . Using the utility function  $\phi(\cdot)$ , a new function  $\tilde{\phi}(\cdot)$  is introduced as follows:

$$\tilde{\phi}(x) = -(\phi(x) - \phi(1)) + \phi'(1)(x - 1). \quad (17)$$

It is also easy to confirm that the function  $\tilde{\phi}(\cdot)$  is strictly convex, continuously differentiable on the open interval  $(0, \infty)$ , and satisfies  $\tilde{\phi}(1) = \tilde{\phi}'(1) = 0$ . Thus, the  $f$ -divergence can be defined using the function  $\tilde{\phi}(\cdot)$  as follows:

$$D_{\tilde{\phi}}(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^I u_i \tilde{\phi}\left(\frac{v_i}{u_i}\right).$$

In the present paper,  $D_{\tilde{\phi}}(\cdot, \cdot)$  is called the  $f$ -divergence induced from the utility function  $\phi(\cdot)$ .

## 5.2 Induction to Bregman's Divergence

Let us define a new function  $\Phi_\alpha(x)$  using the utility function  $\phi_\alpha(\cdot)$ ,  $\alpha > 0$  that is defined in Eq. (3) as follows:

$$\begin{aligned} \Phi_\alpha(x) &= \int_1^x \phi_\alpha(y) dy \\ &= \frac{x^{2-\alpha} - 1}{(1-\alpha)(2-\alpha)}. \end{aligned}$$

Note that the function  $\Phi_\alpha(\cdot)$  is strictly convex and continuously differentiable on the open interval  $(0, \infty)$ . Thus, Bregman's divergence can be defined using  $\Phi_\alpha(\cdot)$  as follows:

$$\begin{aligned} B_{\Phi_\alpha}(\mathbf{u}, \mathbf{v}) &= \sum_{i=1}^I \left\{ \frac{1}{1-\alpha} u_i (u_i^{1-\alpha} - v_i^{1-\alpha}) - \frac{1}{2-\alpha} (u_i^{2-\alpha} - v_i^{2-\alpha}) \right\}, \end{aligned}$$

where  $\alpha > 0$ . Note that  $\Phi_\alpha(\cdot) = g_{1-\alpha}(\cdot)$  holds, and so the following is obtained:

$$B_{\Phi_\alpha}(\mathbf{u}, \mathbf{v}) = B_{g_{1-\alpha}}(\mathbf{u}, \mathbf{v}), \quad (18)$$

for  $\alpha > 0$ . This means that Bregman's divergence induced from the utility function  $\phi_\alpha(\cdot)$  is equivalent to the power divergence.

Next, let us generalize the above discussion. Let  $\phi(\cdot)$  be a utility function that is an increasing, strictly concave, and continuously differentiable utility function on  $(0, \infty)$ . Using the utility function  $\phi(\cdot)$ , a new function  $\Phi(\cdot)$  is introduced as follows:

$$\Phi(x) = \int_1^x \phi(y) dy \quad (19)$$

It is easy to confirm that the function  $\Phi(\cdot)$  is strictly convex and continuously differentiable on the open interval  $(0, \infty)$ . Thus, Bregman's divergence can be defined using  $\Phi(\cdot)$  as follows:

$$\begin{aligned} B_\Phi(\mathbf{u}, \mathbf{v}) &= \sum_{i=1}^I \{ \Phi(u_i) - \Phi(v_i) - \Phi'(v_i)(u_i - v_i) \} \\ &= \sum_{i=1}^I \{ \Phi(u_i) - \Phi(v_i) - \phi(v_i)(u_i - v_i) \}. \end{aligned}$$

In the present paper,  $B_\Phi(\cdot, \cdot)$  is called Bregman's divergence induced from the utility function  $\phi(\cdot)$ .

## 6. GENERALIZATION OF FAIRNESS CONCEPT USING INFORMATION DIVERGENCE MEASURES

In this section, the characterization of the  $(\mathbf{p}, \alpha)$ -proportional fairness given in Section 4 is generalized using the result

in Section 5. That is, the generalization is achieved using Csiszár's  $f$ -divergence and Bregman's divergence that are induced from the utility function specific for the used fairness criterion.

### 6.1 Generalization of the Fairness Concept using Csiszar's $f$ -Divergence

Using the relationship given in Eq. (16), the definition of  $\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)}$  given in Eq. (12) is equivalent to the following definition:

$$\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)} = \arg \min_{\substack{0 < u_i < \infty \\ i=1, \dots, I}} \sum_{n=1}^N p_n D_{\tilde{\phi}_\alpha}(\mathbf{u}, \mathbf{v}_n),$$

Thus, the natural generalization of the above definition would be given as follows:

$$\bar{\mathbf{v}}^{(\mathbf{p}, \tilde{\phi}_{\mathcal{N}})} = \arg \min_{\substack{0 < u_i < \infty \\ i=1, \dots, I}} \sum_{n=1}^N D_{\tilde{\phi}_n}(\mathbf{u}, \mathbf{v}_n),$$

where  $\tilde{\phi}_n(\cdot)$  is defined by the utility function  $\phi_n(\cdot)$  using Eq. (17). Note that the vector  $\bar{\mathbf{v}}^{(\mathbf{p}, \tilde{\phi}_{\mathcal{N}})}$  exists and is unique because the function  $\tilde{\phi}_n(\cdot)$ ,  $n \in \mathcal{N}$  is strictly concave and the feasible region  $\mathcal{C}_I$  is supposed to be convex and compact.

Summarizing the above, the generalized fairness concept can be defined based on Csiszár's  $f$ -divergence as follows:

$$\text{maximize} \quad \sum_{n=1}^N \phi_n(\bar{\mathbf{x}}^{(\mathbf{p}, \tilde{\phi}_{\mathcal{N}})}), \quad (20)$$

$$\text{subject to} \quad \mathbf{x} \in \mathcal{C}_I. \quad (21)$$

Although the above generalization does not lead to the optimization problem motivated by the utility fairness approach introduced in Section 2.3, it conversely indicates another possibility for the generalization of fairness criteria in the resource allocation problem.

### 6.2 Generalization of Fairness Concept using Bregman's Divergence

Using the relationship presented in Eq. (18), the definition of  $\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)}$  given in Eq. (13) is equivalent to the following definition:

$$\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)} = \arg \min_{\substack{0 < u_i < \infty \\ i=1, \dots, I}} \sum_{n=1}^N p_n B_{\Phi_\alpha}(\mathbf{u}, \mathbf{v}_n),$$

Thus, the natural generalization of the above definition would be given as follows:

$$\bar{\mathbf{v}}^{(\mathbf{p}, \Phi_{\mathcal{N}})} = \arg \min_{\substack{0 < u_i < \infty \\ i=1, \dots, I}} \sum_{n=1}^N B_{\Phi_n}(\mathbf{u}, \mathbf{v}_n),$$

where  $\Phi_n(\cdot)$  is defined by the utility function  $\phi_n(\cdot)$  using Eq. (19). Note that the vector  $\bar{\mathbf{v}}^{(\mathbf{p}, \Phi_{\mathcal{N}})}$  exists and is unique because the function  $\Phi_n(\cdot)$  is strictly concave and the feasible region  $\mathcal{C}_I$  is supposed to be convex and compact.

Summarizing the above, the generalized fairness concept can be defined based on Bregman's divergence as follows:

$$\text{maximize} \quad \sum_{n=1}^N \phi_n(\bar{\mathbf{v}}^{(\mathbf{p}, \Phi_{\mathcal{N}})}), \quad (22)$$

$$\text{subject to} \quad \mathbf{v} \in \mathcal{C}_I. \quad (23)$$

In the following, let us consider the relationship between the optimization problem motivated from the utility fairness approach introduced in Section 2.3 and the optimization problem defined by Eqs. (22) and (23). Here, the present paper extends the optimization problem defined by Eqs. (4) and (5) to that for the multi dimensional resource. The extension is given by the following optimization problem:

$$\text{maximize} \quad \sum_{n=1}^N \phi_n(\mathbf{x}_n), \quad (24)$$

$$\text{subject to} \quad \mathbf{x} \in \mathcal{C}_I, \quad (25)$$

where  $\phi_n(\mathbf{x}_n)$  is defined by

$$\phi_n(\mathbf{x}_n) = \sum_{i=1}^I \phi_n(x_{n,i}).$$

The optimization problem defined by Eqs. (24) and (25) equivalent to that defined in Section 2.3 when  $I = 1$ .

Now, the following relationship can be obtained. The proof is given in the appendix.

$$\sum_{n=1}^N \phi_n(\bar{\mathbf{v}}^{(\mathbf{p}, \Phi_{\mathcal{N}})}) = \sum_{n=1}^N \phi_n(\mathbf{v}_n) \quad (26)$$

This means that the objective function of the original optimization problem can be regarded as the utility of the pseudo resource allocation vector  $\bar{\mathbf{v}}^{(\mathbf{p}, \Phi_{\mathcal{N}})}$  that is defined as the centroid in terms of Bregman's divergence  $B_{\Phi_n}(\cdot, \cdot)$  that is induced from the utility function  $\phi_n(\cdot)$ . Note that Eq. (26) is equivalent to Eq. (14) when  $\phi_n(\cdot) = p_n \phi_\alpha(\cdot)$ .

## 7. RELATED EXAMPLE

In this section, a typical example is provided that gives a glimpse into the inherent connection between resource allocation problems and information theory. In the following, the practical resource allocation problem introduced in [18] is used as an example and its result is generalized. The goal of this section is to provide a deeper understanding of practical resource allocation problems based on information theory.

### 7.1 Problem Formulation

In wireless LANs (WLANs) and ad hoc networks, fairness and efficiency can have two conflicting objectives. For multi-rate WLANs and ad hoc networks, in which nodes have the choice of transmitting at varying bit rates (e.g., in IEEE 802.11b, data rate could be 1 Mbps, 2 Mbps, 5.5 Mbps, or 11 Mbps), this issue becomes even more pronounced.

In a wireless network, a wireless station (WS) may use an "auto-rate" mechanism to choose a suitable data rate for transmission according to the channel condition experienced. Co-existing WSs could also transmit at different speeds simply because they use different generations of compatible technologies. For example, 802.11g stations have a maximum data rate of 54 Mbps, while 802.11b stations have a maximum data rate of 11 Mbps.

In [18], an interesting property of proportional fairness in multi-rate WLANs has been given. The property is that proportional fairness is achieved when the fractions of air-time usage by the WSs are equal. In the following, the abovementioned result presented in [18] is generalized for the  $(\mathbf{p}, \alpha)$ -proportional fairness.

Let  $t_n$  be the total amount of air-time used by WS  $n \in \mathcal{N}$  measured over a very long period. The fraction of air-time used by WS  $n$  is then

$$s_n = \frac{t_n}{\sum_{m=1}^N t_m}$$

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

$$v_n = s_n r_n$$

Note that  $r_n$ ,  $n \in \mathcal{N}$  is pre-determined by the distances (or channel condition) of the WSs from the access point, or by the standards (802.11b or 802.11g) used by their wireless cards. Thus,  $r_n$ ,  $n \in \mathcal{N}$  is constant for the optimization problem.

Referring to Eqs. (1) and (2), let us define the following optimization problem:

$$\text{maximize} \quad \psi_\alpha(\mathbf{s}) = \sum_{n=1}^N p_n \frac{(s_n r_n)^{1-\alpha}}{1-\alpha}, \quad (27)$$

$$\text{subject to} \quad \mathbf{s} \in \left\{ \mathbf{s} \mid \sum_{n=1}^N s_n = 1, s_n \geq 0, n \in \mathcal{N} \right\}, \quad (28)$$

where  $\mathbf{s} = \{s_1, \dots, s_N\}$ . As shown in Eqs. (12) and (13), the right-hand side of Eq. (27) can be characterized as the utility of the pseudo throughput that is defined as an aggregation of the throughput values allocated to all WSs. In addition, the pseudo throughput is derived as a result of the minimization of the weighted sum of the  $\alpha$ -divergence.

Note that it is possible to introduce a further generalization of the optimization problem defined by Eqs. (27) and (28) considering the multi-dimensional resource case. For reference, the problem formulation that is defined along with Eqs. (8) and (9) is provided. The problem is formulated as follows:

$$\text{maximize} \quad \psi_\alpha(\mathbf{s}) = \sum_{n=1}^N p_n \left\{ \sum_{i=1}^I \frac{(s_{n,i} r_{n,i})^{1-\alpha}}{1-\alpha} \right\},$$

$$\text{subject to} \quad \mathbf{s} \in \left\{ \mathbf{s} \mid \sum_{n=1}^N s_{n,i} = 1, s_{n,i} \geq 0, n \in \mathcal{N}, i = 1, \dots, I \right\},$$

where  $\mathbf{s}_n = \{s_{n,1}, \dots, s_{n,I}\}$  and  $\mathbf{s} = \{\mathbf{s}_1, \dots, \mathbf{s}_I\}$ .

### 7.2 Analysis

Although the optimization problem is generalized for the multi-dimensional resource case, for simplicity, the present paper hereinafter focuses on the one-dimensional resource case (i.e.,  $I = 1$ ). To solve the above optimization problem, let us define the Lagrangian as follows:

$$L(\mathbf{s}, \mu) = \psi_\alpha(\mathbf{s}) + \mu \left( 1 - \sum_{n=1}^N s_n \right),$$

where  $\mu$  is a Lagrange multiplier. Solution  $\mathbf{s}^* = \{s_1^*, \dots, s_N^*\}$  satisfies

$$\frac{\partial L}{\partial s_n}(\mathbf{s}^*, \mu) = p_n s_n^{*-\alpha} r_n^{1-\alpha} - \mu = 0,$$

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

$$s_n^* = \frac{p_n^\alpha r_n^{\frac{1-\alpha}{\alpha}}}{\sum_{m=1}^N p_m^\alpha r_m^{\frac{1-\alpha}{\alpha}}},$$

$$v_n^* = \frac{p_n^\alpha r_n^{\frac{1}{\alpha}}}{\sum_{m=1}^N p_m^\alpha r_m^{\frac{1}{\alpha}}},$$

where  $v_n^* = s_n^* r_n$ . Note that the solution is closely related to the Tsallis entropy maximization principle, as pointed out in Section 4.3, where the constraint is given in Eq. (28).

A few examples are presented in the following. These examples are for  $\mathbf{p} = \mathbf{1}$ . The first example is for  $\alpha = 1$  that corresponds to the proportional fairness criterion. In this case, the following result, which was given in [18], is obtained:

$$s_n^* = \frac{1}{N},$$

$$v_n^* = \frac{r_n}{N}.$$

Note that  $s_n^*$  does not depend on  $r_n$  if and only if  $\alpha = 1$ . Next example is for  $\alpha \rightarrow 0$  that corresponds to the maximum throughput criterion. In this case, the following is obtained:

$$s_n^* = \begin{cases} \frac{1}{M}, & \text{if } r_n = \max\{r_1, \dots, r_N\} \\ 0, & \text{otherwise} \end{cases},$$

$$v_n^* = \begin{cases} \frac{r_n}{M}, & \text{if } r_n = \max\{r_1, \dots, r_N\} \\ 0, & \text{otherwise} \end{cases},$$

where  $M = |\{r_n | r_n = \max\{r_1, \dots, r_N\}\}|$ . When  $\alpha = 2$ , the following result that corresponds to the minimum potential delay criterion is obtained:

$$s_n^* = \frac{1/\sqrt{r_n}}{\sum_{m=1}^N 1/\sqrt{r_m}},$$

$$v_n^* = \frac{\sqrt{r_n}}{\sum_{m=1}^N 1/\sqrt{r_m}}.$$

When  $\alpha \rightarrow \infty$ , the following result that corresponds to the max-min fairness criterion is obtained:

$$s_n^* = \frac{r_n^{-1}}{\sum_{m=1}^N r_m^{-1}},$$

$$v_n^* = \frac{1}{\sum_{m=1}^N r_m^{-1}}.$$

## 8. CONCLUSION

An information theoretic cognitive process for establishing a valid background of various fairness criteria in network resource allocation problems has been presented. This is believed to be the first paper to deal with resource allocation problems in the context of information theory. The discussion is based mainly on two classes of information divergence measures, namely Csiszár's  $f$ -divergence and Bregman's divergence, which define a kind of distance using a strictly convex function between two positive vectors. The key result of the present paper is based on the observation on the interesting characterization of the  $(\mathbf{p}, \alpha)$ -proportional fairness criterion using the  $\alpha$ -divergence and the power divergence, which are members of Csiszár's  $f$ -divergence and Bregman's divergence, respectively. In addition, the characterization was generalized using Csiszár's  $f$ -divergence and

Bregman's divergence, where the relationship between the convex function that defines the information divergence and the concave function that defines fairness criterion was applied. A typical example was then presented to which the above characterization was applied. This example provided a glimpse of the inherent connection between resource allocation problems and information theory.

## Acknowledgment

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) and in part by the Ministry of Internal Affairs and Communications, Japan.

## 9. REFERENCES

- [1] R. Agrawal, A. Bedekar, R. J. La, and V. Subramanian. Class and channel condition based weighted proportional fair scheduler. In Proceedings of 17th International Teletraffic Congress (ITC'01), pages 553–565, Salvador da Bahia, Brazil, December 2001.
- [2] 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, pages 1532–1541, Allerton, IL, USA, October 2002.
- [3] S. Amari. Differential-Geometrical Methods in Statistics. Springer-Verlag, 1985.
- [4] S. Amari and H. Nagaoka. Methods of Information Geometry. American Mathematical Society, 2000.
- [5] A. Basu, I. R. Harris, N. L. Hjort, and M. C. Jones. Robust and efficient estimation by minimising a density power divergence. *Biometrika*, 85(3):549–559, September 1998.
- [6] D. Bertsekas and R. Gallager. Data Networks. Prentice Hall, 1987.
- [7] T. Bonald and L. Massoulié. Impact of fairness on internet performance. *ACM SIGMETRICS Performance Evaluation Review*, 29(1):82–91, June 2001.
- [8] 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), pages 976–985, Anchorage, AK, USA, April 2001.
- [9] J.-Y. L. Boudec. Rate adaptation, congestion control and fairness: A tutorial, December 2006. Ecole Polytechnique Fédérale de Lausanne (EPFL).
- [10] L. M. Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *U.S.S.R. Computational Mathematics and Mathematical Physics*, 7(1):200–217, 1967.
- [11] T. M. Cover and J. A. Thomas. Elements of Information Theory. Wiley-Interscience, 1991.
- [12] L. Csiszár. Information-type measures of different of probability distributions. *Studia Scientiarum Mathematicarum Hungarica*, 2:299–318, 1967.
- [13] S. Eguchi and M. Minami. Robust blind source separation by  $\beta$ -divergence. *Neural Computation*, 14:1859–1886, 2002.
- [14] T. Harks. Utility proportional fair bandwidth allocation: An optimization oriented approach. In



Proceedings of QoS in Multiservice IP Networks (QoS-IP'05), pages 61–74 (LNCS 3375), Catania, Italy, February 2005.

- [15] T. Harks and T. Poschwatta. Utility fair congestion control for real-time traffic. In Proceedings of 24th Annual IEEE Conference on Computer Communications (INFOCOM'05), pages 2786–2791, Miami, FL, USA, March 2005.
- [16] F. Itakura and S. Saito. Analysis synthesis telephony based on the maximum likelihood method. In Reports of the 6th International Congress on Acoustics, pages 17–20, Tokyo, Japan, August 1968.
- [17] 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), volume 3, pages 1854–1858, Boston, MA, USA, September 2000.
- [18] 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), volume 3, pages 1551–1556, New Orleans, LA, USA, March 2005.
- [19] F. P. Kelly. Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, 8(1):33–37, January 1997.
- [20] 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*, 49(3):237–252, March 1998.
- [21] S. Kullback. *Information Theory and Statistics*. John Wiley and Sons, 1959.
- [22] S. Kullback and R. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22(1):79–86, March 1951.
- [23] L. Massoulié. Structural properties of proportional fairness: Stability and insensitivity. Technical report, Microsoft Research.
- [24] L. Massoulié and J. Roberts. Bandwidth sharing: Objectives and algorithms. In Proceedings of the 18th Annual IEEE Conference on Computer Communications (INFOCOM'99), pages 1395–1403, New York, NY, USA, March 1999.
- [25] J. Mo and J. Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on Networking*, 8(5):556–567, October 2000.
- [26] 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), pages 185–195 (LNCS 3262), Porto, Portugal, October 2004.
- [27] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423 and 623–656, July and October 1948.
- [28] H. Shioya and K. Gohara. Generalized phase retrieval algorithm based on information measures. *Optics Communications*, 266(1):88–93, October 2006.
- [29] C. Tsallis. Possible generalization of boltzmann-gibbs statistics. *Journal of Statistical Physics*, 52(1–2):479–487, July 1988.

- [30] 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)*, 88(4):35–42, December 2004.
- [31] P. Viswaanath, D. N. C. Tse, and R. Laroia. Opportunistic beamforming using dumb antennas. *IEEE Transactions on Information Theory*, 48(6):1277–1294, June.
- [32] 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), pages 1241–1246, Vancouver, Canada, July 2006.

## APPENDIX

### Derivation of Eqs. (12) and (13)

Consider the vector  $\mathbf{u} = \{u_1, \dots, u_I\}$  that satisfies

$$\begin{aligned} & \frac{\partial}{\partial u_i} \sum_{n=1}^N p_n D_{f_{1-\alpha}}(\mathbf{u}, \mathbf{v}_n) \\ &= \sum_{n=1}^N p_n \{(1-\alpha) - (1-\alpha) u_i^{\alpha-1} v_{n,i}^{1-\alpha}\} \\ &= 0, \end{aligned}$$

for  $n = 1, 2, \dots, N$  and  $i = 1, 2, \dots, I$ . The vector  $\mathbf{u}$  is easily confirmed to be equivalent to  $\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)}$ . The convexity of the  $\alpha$ -divergence concludes the proof of Eq. (12).

Similarly, consider the vector  $\mathbf{u} = \{u_1, \dots, u_I\}$  that satisfies

$$\begin{aligned} & \frac{\partial}{\partial u_i} \sum_{n=1}^N p_n B_{g_{1-\alpha}}(\mathbf{u}, \mathbf{v}_n) \\ &= \sum_{n=1}^N p_n \left[ \frac{1}{1-\alpha} \left\{ (2-\alpha) u_i^{1-\alpha} - v_{n,i}^{1-\alpha} \right\} - u_i^{1-\alpha} \right] \\ &= \sum_{n=1}^N p_n \left[ \frac{1}{1-\alpha} (u_i^{1-\alpha} - v_{n,i}^{1-\alpha}) \right] \\ &= 0, \end{aligned}$$

for  $n = 1, 2, \dots, N$  and  $i = 1, 2, \dots, I$ . The vector  $\mathbf{u}$  is easily confirmed to be equivalent to  $\bar{\mathbf{v}}^{(\mathbf{p}, \alpha)}$ . The convexity of the power divergence concludes the proof of Eq. (13).

### Derivation of Eq. (26)

Consider the vector  $\mathbf{u} = \{u_1, \dots, u_I\}$  that satisfies

$$\begin{aligned} & \frac{\partial}{\partial u_i} \sum_{n=1}^N B_{\Phi_n}(\mathbf{u}, \mathbf{v}_n) \\ &= \sum_{n=1}^N \{\Phi'_n(u_i) - \phi_n(v_{n,i})\} \\ &= \sum_{n=1}^N \{\phi_n(u_i) - \phi_n(v_{n,i})\} \\ &= 0. \end{aligned} \tag{29}$$

Here, the convexity of Bregman's divergence and the definition of the vector  $\bar{\mathbf{v}}^{(\mathbf{p}, \Phi_{\mathcal{N}'})}$  conclude the proof of Eq. (26).