

Teknillinen Korkeakoulu
Teknillisen fysiikan koulutusohjelma
Mat-2.108 Sovelletun matematiikan erikoistyöt

ERIKOISTYÖ
19.11.2002

Static Fairness Criteria in Telecommunications

Vesa Timonen, e-mail: vesa.timonen@hut.fi
47494R

Static Fairness Criteria in Telecommunications

Vesa Timonen

19th November 2002

Contents

1	Introduction	1
2	Network model	2
2.1	Network setup 1	3
2.2	Network setup 2	3
3	Fairness criteria	4
4	Basic fairness criteria	4
4.1	Maximum throughput	4
4.2	Max-min fairness	5
4.3	Proportional fairness	6
4.4	Potential delay minimization	8
5	Fairness criteria with parameters	9
5.1	Weighted max-min fairness	10
5.2	Weighted proportional fairness	10
5.3	Weighted potential delay minimization	11
6	Utility approach	12
6.1	General utility function	13
6.2	(p, α) -Proportional fairness	14
7	Conclusions	17

Abstract

The main objective of bandwidth sharing in telecommunications network is to use all available bandwidth without disrupting the constraints and maintain a certain fairness. The achieved fairness depends closely on the used fairness criterion. Different fairness criteria favor or discriminate single sources or whole traffic classes on different basis. In this paper the notion of fairness is introduced, essential static fairness criteria and their relationship is studied. A network model is presented and the behavior of different fairness criteria is examined via two simple network setups.

1 Introduction

Basically any telecommunication network can be considered as a set of links with finite capacities as resources and a finite number of sources as network users. The most simple traffic model is such that a data flow between a sending source and the receiving point occupies a constant amount of link capacity in all the links through which it traverses. When several sessions have to compete for the finite resources or capacity of a network, the rate allocated for sources has to be regulated by some control mechanism to avoid congestion and to reduce packet losses in the network. Because of the finite resources, the bandwidth share or equally the rate allocation is a compromise that should be fair, which leads to the concept of “fairness”.

Notion of fairness has no unique definition. It may depend on several different session priorities and service requirements, e.g. a session can require a minimum guaranteed rate for sending data or has a maximum on allowed network delay. It is generally accepted that traffic with the same priority should be treated equally. The most simplified definition is to allocate the same share to each connection. Different fairness criteria favor or discriminate sources or traffic classes on different basis. The objective can be to use the network capacity as effectively as possible without considering a single source (the throughput maximization), or on the contrary, the goal can be to ensure as equal sharing of the resources as possible (max-min fairness).

As a mathematical notion fairness can alternatively be thought of as an optimization problem, where the objective is to find a rate allocation that minimizes or maximizes a utility function specific for the used fairness criterion. In this approach e.g. a cost for achieved rate allocation can be easily added to the examination. Further, the optimization approach provides a generalization of the concept of fairness.

In this paper our objective is to introduce the notion of fairness from the viewpoint of telecommunications and study essential fairness criteria and their relationship as they are considered in some earlier papers, and further, concretize these results via two different example cases. Some assumptions and simplifications are made concerning the network and the traffic. The network topology, consisting of the links with constant capacities and sources, is fixed and known. All sources send elastic traffic with adjustable constant rate to a single target via a fixed route. Round trip times and processing delays are neglected. Thus, this study is limited to static fairness criteria.

Organization of the paper is as follows: in Section 2 the network model of this study and two different example network setups are presented. The notion

of fairness is briefly discussed in Section 3 and basic fairness criteria are studied in Section 4. Weighted fairness criteria containing some additional parameters are presented in Section 5 and in Section 6 fairness as a optimization problem is studied more closely. Relation between the criteria presented in previous sections and the utility approach is stated. The paper is concluded in Section 7.

2 Network model

In this section the basic network model is presented and two different network setups are introduced.

In the following network model some assumptions are made; we assume that there is a fixed number of flows and routes in the network. Also each sources sending rate x_s is assumed to be constant. In a way, the used network model is statical, sending rates, number of flows or used routes do not vary in time.

A network can be considered as a set of links \mathcal{L} , each link $l \in \mathcal{L}$ has a finite capacity $C_l > 0$. Let \mathcal{S} denote the set of all sources in the network and let x_s denote the sending rate of source $s \in \mathcal{S}$. A route r is a non-empty subset of \mathcal{L} . Denote by \mathcal{R} the set of all routes used by the sources $s \in \mathcal{S}$.¹ Let $l \in r$ denote that link l belongs to route r . A 0 – 1 matrix A is defined as follows:

$$A = (A_{ls}) \tag{1}$$

where $A_{ls} = 1_{l \in r_s}$ ², $l \in \mathcal{L}$ and $s \in \mathcal{S}$.

Finite capacities of the links set some restrictions to the rate allocations that can be carried out in the network. A rate allocation that does not disrupt these limitations is called feasible.

Definition 1 (Feasible allocation). *A rate allocation $\{x_s\}_{s \in \mathcal{S}}$ is feasible if $x_s \geq 0 \forall s \in \mathcal{S}$ and it satisfies the capacity constraints*

$$\sum_{s \in \mathcal{S}} A_{ls} x_s \leq C_l, \quad l \in \mathcal{L}. \tag{2}$$

Let $\mathbf{x} = \{x_s\}_{s \in \mathcal{S}}$ denote a rate allocation vector and $\mathbf{C} = \{C_l\}_{l \in \mathcal{L}}$ a vector of the link capacities. Thus, Definition 1 can be presented in a matrix form as follows:

Definition 2 (Feasible allocation). *A rate allocation \mathbf{x} is feasible if $\mathbf{x} \geq 0$ and it satisfies the capacity constraints*

$$\mathbf{Ax} \leq \mathbf{C} \tag{3}$$

Let \mathcal{F} denote the set of all feasible rate allocations, i.e. $\mathcal{F} \stackrel{\text{def}}{=} \{\mathbf{x} \mid \mathbf{x} \text{ is feasible}\}$. The set \mathcal{F} is convex, which is proved as follows:

Proposition Set \mathcal{F} is convex.

¹Set \mathcal{R} can contain identical elements, i.e. several sources can use exactly the same route.

²More generally, we may assume that $A_{ls} \in [0, 1]$, which corresponds to splitting the traffic over parallel paths. [5, 4]

Proof Let \mathbf{x}_1 and \mathbf{x}_2 be two rate allocations such that $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{F}$. Now,

$$\begin{aligned} A(tx_1 + (1-t)x_2) &= t \underbrace{Ax_1}_{\substack{\leq \mathbf{C} \\ \mathbf{x}_1 \in \mathcal{F}}} + (1-t) \underbrace{Ax_2}_{\substack{\leq \mathbf{C} \\ \mathbf{x}_2 \in \mathcal{F}}} \leq t\mathbf{C} + (1-t)\mathbf{C} = \mathbf{C} \\ \Leftrightarrow A(tx_1 + (1-t)x_2) &\leq \mathbf{C} \quad \forall t \in [0, 1] \\ \Leftrightarrow tx_1 + (1-t)x_2 &\in \mathcal{F} \quad \forall t \in [0, 1] \\ \Leftrightarrow \mathcal{F} \text{ is convex} & \quad [1, \text{Def 2.1.1}]. \quad \square \end{aligned}$$

In Sections 2.1 and 2.2 two simple linear network setups are introduced. These networks are used in the later sections to demonstrate the differences between various fairness criteria.

2.1 Network setup 1

This network setup (Figure 1) is a widely used example to illustrate the complexity related to fairness (see e.g. [2]). There are L links and $L + 1$ traffic classes $\mathcal{R}_0, \dots, \mathcal{R}_L$, each containing n_r flows, $r = 0, \dots, L$. Flows of class \mathcal{R}_0 travel through all links and flows of class \mathcal{R}_l only through link l , $l = 1, \dots, L$.

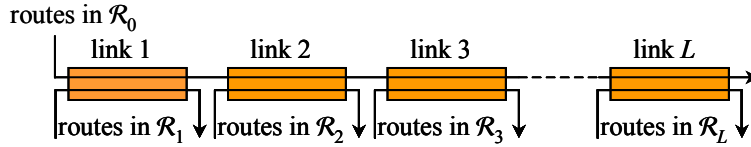


Figure 1: The linear network.

In the following sections *Example 1* refers to this network setup with following parameter values: $L = 2$, $C_l \equiv 1$ and $n_1 \leq n_2$.

2.2 Network setup 2

This network setup (Figure 2) is also known as the “parking lot” scenario. There are L links and L traffic classes $\mathcal{R}_1, \dots, \mathcal{R}_L$ each containing n_r flows, $r = 1, \dots, L$. Flows of class \mathcal{R}_l travel through links $l, l+1, \dots, L$, $l = 1, \dots, L$. In link l there are flows of classes $\mathcal{R}_1, \dots, \mathcal{R}_l$ competing of the links finite capacity.

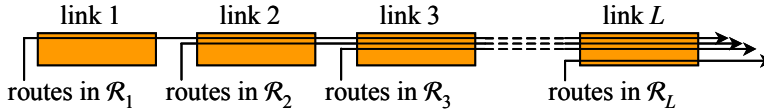


Figure 2: The linear network (the “parking lot” scenario).

In the following sections *Example 2* refers to this network setup with following parameter values: $L = 2$, $C_1 = 1$ and $C_2 = c \geq 1$.

3 Fairness criteria

The main objective of bandwidth sharing is to use all available bandwidth without disrupting the constraints and maintain a certain fairness. The achieved fairness depends closely on the used fairness criterion. Different fairness criteria favor or discriminate single sources or whole traffic classes on different basis. The objective can be to use the network capacity as effectively as possible without considering a single source (the throughput maximization), or on the contrary, the goal can be to ensure as equal sharing of the resources as possible (max-min fairness).

In the following sections various definitions of fairness criteria are introduced. It is assumed that all the sources use a fixed route and send with a constant rate. Also the whole network, its resources and topology is known. Examinations of fairness are restricted to the unicast case, i.e. flows are traversing from one source to a single destination. Thus, the criteria are considered as static, varying of any component in time is excluded from this study.

4 Basic fairness criteria

The most common fairness criteria have no additional parameters. As basic fairness criteria are concerned throughput maximization, max-min fairness, proportional fairness and potential delay minimization.

4.1 Maximum throughput

A straightforward objective for bandwidth allocation is to find a feasible rate allocation that maximizes the total throughput [2, 7, 9, 4]. In a way, it would be the most efficient way to use the network resources. Formulated as an optimization problem:

$$\begin{aligned} \text{maximize } u(\mathbf{x}) &= \sum_{s \in \mathcal{S}} x_s & (4) \\ \text{subject to } \mathbf{x} &\in \mathcal{F} \end{aligned}$$

Example 1 Let us denote $n_i x_i = \sum_{s \in \mathcal{R}_i} x_s$, that is, x_i is the mean rate allocation for flows in the class \mathcal{R}_i , $i = 0, 1, 2$. In a link l the allocated rate for class \mathcal{R}_l has to satisfy equation $n_0 x_0 + n_l x_l \leq 1$. In fact, we clearly have $n_l x_l = 1 - n_0 x_0$. The total throughput of the network is then $\sum_{s \in \mathcal{S}} x_s = \sum_{l=0}^2 n_l x_l = 2 - 2n_0 x_0$, which attains its maximum 2 when $x_0 = 0$. Thus, we have $x_0 = 0$ and $x_l = \frac{1}{n_l}$, $l = 1, 2$.

Example 2 As before, all flows in the same class \mathcal{R}_r have the same mean rate x_r . The maximum of the total throughput is $c = C_2$ and it is obtained with all rate allocations $\mathbf{x} \in \mathcal{F}$ that fulfill the condition $n_1 x_1 + n_2 x_2 = c$ such that $n_1 x_1 \leq 1$. This gives $0 \leq x_1 \leq \frac{1}{n_1}$ and $x_2 = \frac{1}{n_2}(c - n_1 x_1)$. Thus, there is an infinite number of feasible rate allocations that maximize the throughput, e.g.

rate allocation

$$\mathbf{x}_1 = \left\{ \underbrace{0, \dots, 0}_{n_1 \text{ elements}}, \underbrace{c/n_2, \dots, c/n_2}_{n_2 \text{ elements}} \right\}$$

or

$$\mathbf{x}_2 = \left\{ \underbrace{1/n_1, \dots, 1/n_1}_{n_1 \text{ elements}}, \underbrace{(c-1)/n_2, \dots, (c-1)/n_2}_{n_2 \text{ elements}} \right\}.$$

Still, it can be questioned whether the rate allocation \mathbf{x}_1 is as “fair” as \mathbf{x}_2 .

4.2 Max-min fairness

Max-min fairness [2, 4, 9, 7] is the most common definition for the concept of fairness. Its objective is to maximize the minimum of the given bandwidths, i.e. the rate of any source can not be increased without decreasing the rate of some other source that already has a smaller rate.

Definition 3 (Max-min fairness). *A rate allocation \mathbf{x} is max-min fair if $\mathbf{x} \in \mathcal{F}$ and*

$$\forall s \in \mathcal{S} \exists l \in s \text{ such that } \sum_{i \ni l} x_i = C_l \text{ and } x_s = \max_{i \ni l} x_i \quad (5)$$

The bandwidth allocation fulfilling the Definition 3 can be proved to be unique [9, 7].

An alternative definition of max-min fairness is that every source has a “bottleneck link” [2, 4].³ A bottleneck link is defined as follows:

Definition 4 (Bottleneck link). *A link $l \in \mathcal{L}$ is called a bottleneck link for source $s \in \mathcal{S}$ if link l is saturated, i.e.*

$$\sum_{s \in \mathcal{S}} A_{ls} x_s = C_l, \quad l \in \mathcal{L} \quad (6)$$

and source s that sends through link l has the maximum rate among all the other sources using the link l , i.e.

$$x_s \geq x_i \quad \forall x_i \text{ such that } A_{li} = 1, \quad i \in \mathcal{S} \quad (7)$$

where x_s denotes the rate allocation of source $s \in \mathcal{S}$.

Example 1 Let \mathbf{x} be a max-min fair rate allocation. All flows in the same class \mathcal{R}_l have the same rate x_l , $l = 0, 1, 2$. By definition the sources using routes of class \mathcal{R}_0 have a bottleneck link. The bottleneck link is the most crowded link, that is, for the bottleneck link k it holds $k = \arg \max_{l \geq 1} n_l = 2$. Thus, the following equation holds: $n_0 x_0 + n_2 x_2 = 1$. Since link 2 is also the bottleneck link of class \mathcal{R}_2 , we have $x_0 = x_2$, from which it follows that $x_0 = \frac{1}{n_0 + n_2}$. In

³The equality of these definitions is proved e.g. in [2, 4].

the other link 1 the excess capacity is used by the flows of class \mathcal{R}_1 . Thus, the max-min fair rate allocation is

$$x_r = \begin{cases} \frac{1}{n_0+n_2} & \text{for } r \in \mathcal{R}_0, \\ \frac{1}{n_l} \frac{n_2}{n_0+n_2} & \text{for } r \in \mathcal{R}_l, l \in \{1, 2\}. \end{cases}$$

In the case $n_r \equiv 1$ we have $x_r \equiv \frac{1}{2}$ and the total throughput is $\frac{3}{2}$. Max-min fairness gives a smaller total throughput than achieved with throughput maximization, but guarantees that all the classes are given an equal proportion of the bandwidth available. Thus, max-min fairness can be considered as a more “fair” bandwidth sharing objective than pure maximizing of the total throughput.

Example 2 Let \mathbf{x} be the max-min fair rate allocation. All flows in the same class \mathcal{R}_l have the same rate x_l , $l = 1, 2$. By definition the sources using the routes of class \mathcal{R}_2 have a bottleneck link 2. Thus, the following equation holds: $n_1x_1 + n_2x_2 = c$. If the link 2 is also the bottleneck link for flows of class \mathcal{R}_1 , then $x_1 = x_2$. Thus, $x_1 = x_2 = \frac{c}{n_1+n_2}$. Now, if $n_1x_1 > 1$, then the bottleneck link for flows of class \mathcal{R}_1 is link 1. That is, if $\frac{1}{n_1} < \frac{c}{n_1+n_2}$ and further, if $n_1 > \frac{n_2}{c-1}$, then $x_1 = \frac{1}{n_1}$ and $x_2 = \frac{c-1}{n_2}$.

In the case $n_r \equiv 1$ we have $x_1 = x_2 = \frac{c}{2}$ if $n_1 < \frac{n_2}{c-1}$, that is, if $c < 2$. Otherwise $x_1 = 1$ and $x_2 = c - 1$. In both cases max-min fairness gives the maximum total throughput c , but as a contrast to throughput maximization, the rate allocation is unambiguous.

4.3 Proportional fairness

Proportional fairness was proposed in [6]. In proportional fairness deviation from the fair allocation causes a negative average change.

Definition 5 (Proportional fairness). A rate allocation \mathbf{x} is proportionally fair if $\mathbf{x} \in \mathcal{F}$ and

$$\sum_{s \in \mathcal{S}} \frac{x'_s - x_s}{x_s} \leq 0 \quad \forall \mathbf{x}' \in \mathcal{F} \quad (8)$$

The bandwidth allocation fulfilling the Definition 5 can be proved to be unique [9, 7].

Example 1 Let us assume that the proportionally fair allocation is

$$x_r = \begin{cases} \frac{1}{n_0+n_1+n_2} & \text{for } r \in \mathcal{R}_0, \\ \frac{1}{n_l} \frac{n_1+n_2}{n_0+n_1+n_2} & \text{for } r \in \mathcal{R}_l, l \in \{1, 2\}. \end{cases} \quad (9)$$

which is proved as follows:

Proposition The rate allocation \mathbf{x} proposed in (9) is proportionally fair.

Proof Let $\mathbf{y} \in \mathcal{F}$. In links 1 and 2 it holds that $n_0 y_0 + n_l y_l \leq 1$.

$$\begin{aligned}
\sum_{s \in \mathcal{S}} \frac{y_s - x_s}{x_s} &= \sum_{l=0}^2 n_l \frac{y_l - x_l}{x_l} = n_0 \frac{y_0 - x_0}{x_0} + n_1 \frac{y_1 - x_1}{x_1} + n_2 \frac{y_2 - x_2}{x_2} \\
&= \frac{n_0 y_0 - n_0 x_0}{x_0} + \frac{\overbrace{n_1 y_1}^{\leq 1 - n_0 y_0} - \overbrace{n_1 x_1}^{= 1 - n_0 x_0}}{x_1} + \frac{\overbrace{n_2 y_2}^{\leq 1 - n_0 y_0} - \overbrace{n_2 x_2}^{= 1 - n_0 x_0}}{x_2} \\
&\leq n_0 (y_0 - x_0) \left(\frac{1}{x_0} - \frac{1}{x_1} - \frac{1}{x_2} \right) \\
&\stackrel{(9)}{=} n_0 (y_0 - x_0) \left(n_0 + n_1 + n_2 - \frac{n_1}{n_1 + n_2} (n_0 + n_1 + n_2) \right. \\
&\quad \left. - \frac{n_2}{n_1 + n_2} (n_0 + n_1 + n_2) \right) \\
&= 0 \quad \forall \mathbf{y} \in \mathcal{F}.
\end{aligned}$$

Thus, based on Definition 5 rate allocation (9) is proportionally fair. \square

In the case that $n_r \equiv 1$ we have $x_0 = \frac{1}{3}$, $x_1 = x_2 = \frac{2}{3}$ and the total throughput is $1\frac{2}{3} \approx 1.667$. Proportional fairness gives a smaller total throughput than achieved with throughput maximization, but greater than achieved with max-min fairness. Proportional fairness seems to penalize long routes more than max-min fairness with tendency to achieve greater total throughput.

Example 2 Let us assume that the proportionally fair allocation is

$$x_1 = \begin{cases} \frac{c}{n_1 + n_2} & \text{if } n_1 < \frac{n_2}{c-1}, \\ \frac{1}{n_1} & \text{otherwise} \end{cases}$$

and

$$x_2 = \begin{cases} \frac{c}{n_1 + n_2} & \text{if } n_1 < \frac{n_2}{c-1}, \\ \frac{c-1}{n_2} & \text{otherwise} \end{cases}, \tag{10}$$

which is proved as follows:

Proposition The rate allocation \mathbf{x} proposed in (10) is proportionally fair.

Proof Let $\mathbf{y} \in \mathcal{F}$. It holds that $n_1 y_1 \leq 1$ and $n_1 y_1 + n_2 y_2 \leq c$.

$$\begin{aligned}
\sum_{s \in \mathcal{S}} \frac{y_s - x_s}{x_s} &= \sum_{l=1}^2 n_l \frac{y_l - x_l}{x_l} = n_1 \frac{y_1 - x_1}{x_1} + n_2 \frac{y_2 - x_2}{x_2} \\
&= \frac{n_1 y_1 - n_1 x_1}{x_1} + \frac{\overbrace{n_2 y_2}^{\leq c - n_1 y_1} - \overbrace{n_2 x_2}^{= c - n_1 x_1}}{x_2} \\
&\leq (n_1 y_1 - n_1 x_1) \left(\frac{1}{x_1} - \frac{1}{x_2} \right)
\end{aligned}$$

Now, if $n_1 < \frac{n_2}{c-1}$, it follows from (10) that $x_1 = x_2$ and further $\frac{1}{x_1} - \frac{1}{x_2} \equiv 0$. Otherwise $x_1 = \frac{1}{n_1}$, from which it follows that

$$\underbrace{(n_1 y_1)}_{\leq 1} - \underbrace{(n_1 x_1)}_{=1} \leq 0 \text{ and } \frac{1}{x_1} - \frac{1}{x_2} = \underbrace{\frac{n_1}{c-1}}_{\geq \frac{n_2}{c-1}} - \frac{n_2}{c-1} \geq 0,$$

i.e.

$$\sum_{s \in \mathcal{S}} \frac{y_s - x_s}{x_s} \leq 0 \quad \forall \mathbf{y} \in \mathcal{F}.$$

Thus, based on Definition 5 rate allocation (10) is proportionally fair. \square

In the case that $n_r \equiv 1$ we have $x_1 = x_2 = \frac{c}{2}$ if $n_1 < \frac{n_2}{c-1}$, that is, if $c < 2$. Otherwise $x_1 = 1$ and $x_2 = c - 1$. In both cases the total throughput is c , that is, same as in all the previous cases.

4.4 Potential delay minimization

Objective in potential delay minimization is to minimize the time delay needed to complete transfers [7]. The time delay is thought to be inversely proportional to the sending rate of the source. Thus, a fair rate allocation $\mathbf{x} \in \mathcal{F}$ can be defined as solution to following optimization problem:

$$\begin{aligned} \text{minimize } u(\mathbf{x}) &= \sum_{s \in \mathcal{S}} \frac{1}{x_s} & (11) \\ \text{subject to } \mathbf{x} &\in \mathcal{F} \end{aligned}$$

Example 1 Routes inside the same class \mathcal{R}_r get equal allocation of the bandwidth x_r . It is required that $n_0 x_0 + n_l x_l = 1$, from which it follows that $x_l = \frac{1-n_0 x_0}{n_l}$. A fair rate allocation minimizes the utility function $u(\mathbf{x})$, where

$$u(\mathbf{x}) = \sum_{s \in \mathcal{S}} \frac{1}{x_s} = \frac{n_0}{x_0} + \frac{n_1}{x_1} + \frac{n_2}{x_2} = \frac{n_0}{x_0} + \frac{n_1^2 + n_2^2}{1 - n_0 x_0}$$

Differentiating gives

$$\begin{aligned} \frac{\partial u}{\partial x_0} &= -\frac{n_0}{x_0^2} + n_0 \frac{n_1^2 + n_2^2}{(1 - n_0 x_0)^2} = 0 \Leftrightarrow (1 - n_0 x_0)^2 = x_0^2 (n_1^2 + n_2^2) \\ \Rightarrow x_0 &= \frac{1}{n_0 + \sqrt{n_1^2 + n_2^2}} \text{ and } x_l = \frac{1}{n_l} \frac{\sqrt{n_1^2 + n_2^2}}{n_0 + \sqrt{n_1^2 + n_2^2}} \end{aligned}$$

Thus

$$x_r = \begin{cases} \frac{1}{n_0 + \sqrt{n_1^2 + n_2^2}} & \text{for } r \in \mathcal{R}_0, \\ \frac{1}{n_l} \frac{\sqrt{n_1^2 + n_2^2}}{n_0 + \sqrt{n_1^2 + n_2^2}} & \text{for } r \in \mathcal{R}_l, l \in \{1, 2\}. \end{cases}$$

In the case that $n_r \equiv 1$ the rate allocation is $x_0 = \frac{1}{\sqrt{2+1}}$, $x_r = \frac{\sqrt{2}}{\sqrt{2+1}}$ for $r \geq 1$ and the total throughput is $3 - \sqrt{2} \approx 1.586$. Thus, the bandwidth allocation given by this criterion is in between when compared to max-min and proportional fairness.

Example 2 Routes inside the same class \mathcal{R}_r get equal allocation of the bandwidth x_r . It is required that $n_1x_1 + n_2x_2 = c$, from which follows $x_2 = \frac{c-n_1x_1}{n_2}$. A fair rate allocation minimizes the utility function $u(\mathbf{x})$, where

$$u(\mathbf{x}) = \sum_{s \in \mathcal{S}} \frac{1}{x_s} = \frac{n_1}{x_1} + \frac{n_2}{x_2} = \frac{n_1}{x_1} + \frac{n_2^2}{c - n_1x_1}$$

Differentiating gives

$$\begin{aligned} \frac{\partial u}{\partial x_1} &= -\frac{n_1}{x_1^2} + n_1 \frac{n_2^2}{(c - n_1x_1)^2} = 0 \Leftrightarrow (c - n_1x_1)^2 = x_1^2 n_2^2 \\ \Rightarrow x_1 &= \frac{c}{n_1 + n_2} \text{ and } x_2 = \frac{c - n_1 \frac{c}{n_1 + n_2}}{n_2} = \frac{c}{n_1 + n_2} \end{aligned}$$

if the link 1 is not the bottleneck, i.e. $n_1 < \frac{n_2}{c-1}$. Otherwise we have

$$x_1 = \frac{1}{n_1} \text{ and } x_2 = \frac{c-1}{n_2}.$$

That is, the solution reduces to same as in the previous cases

$$x_1 = \begin{cases} \frac{c}{n_1 + n_2} & \text{if } n_1 < \frac{n_2}{c-1}, \\ \frac{1}{n_1} & \text{if otherwise} \end{cases}$$

and

$$x_2 = \begin{cases} \frac{c}{n_1 + n_2} & \text{if } n_1 < \frac{n_2}{c-1}, \\ \frac{c-1}{n_2} & \text{if otherwise} \end{cases}.$$

In the case that $n_r \equiv 1$ the total throughput is c .

5 Fairness criteria with parameters

Concept of weighted shares provides generalizations for max-min and proportional fairness and potential delay minimization. A weighting factor ϕ_s is associated with source $s \in \mathcal{S}$. Increase in weight leads to increase in the received share [7, 9]. Thus, weights can be used e.g. to give a higher rate allocation to a single source in a a specific flow class, or further, to all sources of a flow class.

In the following examples it is assumed that in *Example 1* the weight $\phi_0 = 1$ for flows of class \mathcal{R}_0 , ϕ_1 and ϕ_2 are the weights for flows of classes \mathcal{R}_1 and \mathcal{R}_2 , respectively. It is also demanded that $\phi_1 n_1 \leq \phi_2 n_2$. In *Example 2* the weight $\phi_1 = 1$ for flows of class \mathcal{R}_1 and ϕ_2 is the weight for flows of class \mathcal{R}_2 .

5.1 Weighted max-min fairness

Max-min fairness defined with weights [7, 9]:

Definition 6. A bandwidth allocation \mathbf{x} is max-min fair with weights ϕ_s if $\mathbf{x} \in \mathcal{F}$ and

$$\forall s \in \mathcal{S} \exists l \in s \text{ such that } \sum_{i \ni l} x_i = C_l \text{ and } \frac{x_s}{\phi_s} = \max_{i \ni l} \frac{x_i}{\phi_i}$$

The definition leads to a rate allocation, in which a single source s is considered as ϕ_s sources. This is verified with following examples.

Example 1 All flows in the same class \mathcal{R}_l have the same rate x_l , $l = 0, 1, 2$. Because $\phi_1 n_1 \leq \phi_2 n_2$, the link 2 is the bottleneck link. By Definition 6 in the bottleneck link $x_0 = \frac{x_2}{\phi_2}$ and $n_0 x_0 + n_2 x_2 = 1$, from which it follows that $x_0 = \frac{1}{n_0 + \phi_2 n_2}$ and $x_2 = \phi_2 x_0 = \frac{\phi_2}{n_0 + \phi_2 n_2}$. In link 1 the excess capacity is used by the flows of class \mathcal{R}_1 . Thus, the max-min fair rate allocation is

$$x_r = \begin{cases} \frac{1}{n_0 + \phi_2 n_2} & \text{for } r \in \mathcal{R}_0, \\ \frac{1}{n_1} \frac{\phi_2 n_2}{n_0 + \phi_2 n_2} & \text{for } r \in \mathcal{R}_l, l \in \{1, 2\}. \end{cases}$$

Example 2 All flows in the same class \mathcal{R}_l have the same rate x_l , $l = 1, 2$. Flows of class \mathcal{R}_2 have a bottleneck link 2. Thus, the following equation holds: $n_1 x_1 + n_2 x_2 = c$. If the link 2 is also the bottleneck link for flows of class \mathcal{R}_1 , then by Definition 6 we have $x_1 = \frac{x_2}{\phi_2}$. Thus, $x_1 = \frac{c}{n_1 + \phi_2 n_2}$ and $x_2 = \phi_2 x_1 = \frac{\phi_2 c}{n_1 + \phi_2 n_2}$. Now, if $n_1 x_1 > 1$, then the bottleneck link for flows of class \mathcal{R}_1 is link 1. That is, if $\frac{1}{n_1} < \frac{c}{n_1 + \phi_2 n_2}$ and further, if $n_1 > \frac{\phi_2 n_2}{c-1}$, then $x_1 = \frac{1}{n_1}$ and $x_2 = \frac{c-1}{n_2}$. In summary,

$$x_1 = \begin{cases} \frac{c}{n_1 + \phi_2 n_2} & \text{if } n_1 < \frac{\phi_2 n_2}{c-1}, \\ \frac{1}{n_1} & \text{otherwise} \end{cases} \text{ and } x_2 = \begin{cases} \frac{\phi_2 c}{n_1 + \phi_2 n_2} & \text{if } n_1 < \frac{\phi_2 n_2}{c-1}, \\ \frac{c-1}{n_2} & \text{otherwise} \end{cases}.$$

5.2 Weighted proportional fairness

Weighted proportional fairness [7, 6] is a generalization of proportional fairness.

Definition 7 (Weighted proportional fairness). A bandwidth allocation \mathbf{x} is proportionally fair with weights ϕ_s if $\mathbf{x} \in \mathcal{F}$ and

$$\sum_{s \in \mathcal{S}} \phi_s \frac{x'_s - x_s}{x_s} \leq 0 \quad \forall \mathbf{x}' \in \mathcal{F} \quad (12)$$

Similarly as in the case of the weighted max-min fairness, the definition leads to a rate allocation, in which a single source s is considered as ϕ_s sources. This is verified with following results.

Example 1 The weighted proportionally fair allocation is

$$x_r = \begin{cases} \frac{1}{n_0 + \phi_1 n_1 + n_2 \phi_2} & \text{for } r \in \mathcal{R}_0, \\ \frac{1}{n_1} \frac{\phi_1 n_1 + \phi_2 n_2}{n_0 + \phi_1 n_1 + \phi_2 n_2} & \text{for } r \in \mathcal{R}_l, l \in \{1, 2\}. \end{cases} \quad (13)$$

Derivation of this result is presented in Section 6.2.

Example 2 The weighted proportionally fair allocation is

$$x_1 = \begin{cases} \frac{c}{n_1 + \phi_2 n_2} & \text{if } n_1 < \frac{\phi_2 n_2}{c-1}, \\ \frac{1}{n_1} & \text{otherwise} \end{cases}$$

and

$$x_2 = \begin{cases} \frac{\phi_2 c}{n_1 + \phi_2 n_2} & \text{if } n_1 < \frac{\phi_2 n_2}{c-1}, \\ \frac{c-1}{n_2} & \text{otherwise} \end{cases}. \quad (14)$$

Derivation of this result is presented in Section 6.2.

5.3 Weighted potential delay minimization

Potential delay minimization defined with weights: a fair rate allocation $\mathbf{x} \in \mathcal{F}$ can be defined as solution to following optimization problem:

$$\begin{aligned} \text{minimize } u(\mathbf{x}) &= \sum_{s \in \mathcal{S}} \frac{\phi_s}{x_s} \\ \text{subject to } \mathbf{x} &\in \mathcal{F} \end{aligned} \quad (15)$$

Example 1 All flows in the same class \mathcal{R}_r get equal allocation of the bandwidth x_r . It is required that $n_0 x_0 + n_l x_l = 1$, from which follows $x_l = \frac{1 - n_0 x_0}{n_l}$. A fair rate allocation minimizes the utility function (15), where

$$u(\mathbf{x}) = \sum_{s \in \mathcal{S}} \frac{\phi_s}{x_s} = \frac{n_0}{x_0} + \frac{\phi_1 n_1}{x_1} + \frac{\phi_2 n_2}{x_2} = \frac{n_0}{x_0} + \frac{\phi_1 n_1^2 + \phi_2 n_2^2}{1 - n_0 x_0}$$

Differentiating gives

$$\begin{aligned} \frac{\partial u}{\partial x_0} &= -\frac{n_0}{x_0^2} + n_0 \frac{\phi_1 n_1^2 + \phi_2 n_2^2}{(1 - n_0 x_0)^2} = 0 \Leftrightarrow (1 - n_0 x_0)^2 = x_0^2 (\phi_1 n_1^2 + \phi_2 n_2^2) \\ \Rightarrow x_0 &= \frac{1}{n_0 + \sqrt{\phi_1 n_1^2 + \phi_2 n_2^2}} \text{ and } x_l = \frac{1}{n_l} \frac{\sqrt{\phi_1 n_1^2 + \phi_2 n_2^2}}{n_0 + \sqrt{\phi_1 n_1^2 + \phi_2 n_2^2}} \end{aligned}$$

Thus

$$x_r = \begin{cases} \frac{1}{n_0 + \sqrt{\phi_1 n_1^2 + \phi_2 n_2^2}} & \text{for } r \in \mathcal{R}_0, \\ \frac{1}{n_l} \frac{\sqrt{\phi_1 n_1^2 + \phi_2 n_2^2}}{n_0 + \sqrt{\phi_1 n_1^2 + \phi_2 n_2^2}} & \text{for } r \in \mathcal{R}_l, l \in \{1, 2\}. \end{cases}$$

Example 2 All flows in the same class \mathcal{R}_r get equal allocation of the bandwidth x_r . It is required that $n_1 x_1 + n_2 x_2 = c$, from which follows $x_2 = \frac{c - n_1 x_1}{n_2}$. A fair rate allocation minimizes the utility function (15), where

$$u(\mathbf{x}) = \sum_{s \in \mathcal{S}} \frac{\phi_s}{x_s} = \frac{n_1}{x_1} + \frac{\phi_2 n_2}{x_2} = \frac{n_1}{x_1} + \frac{\phi_2 n_2^2}{c - n_1 x_1}$$

Differentiating gives

$$\begin{aligned} \frac{\partial u}{\partial x_1} &= -\frac{n_1}{x_1^2} + n_1 \frac{\phi_2 n_2^2}{(c - n_1 x_1)^2} = 0 \Leftrightarrow (c - n_1 x_1)^2 = \phi_2 n_2^2 x_1^2 \\ \Rightarrow x_1 &= \frac{c}{n_1 + \sqrt{\phi_2} n_2} \text{ and } x_2 = \frac{\sqrt{\phi_2} c}{n_1 + \sqrt{\phi_2} n_2} \end{aligned}$$

if the link 2 is the bottleneck, i.e. $n_1 x_1 = \frac{n_1 c}{n_1 + \sqrt{\phi_2} n_2} < 1$, that is $n_1 < \frac{\sqrt{\phi_2} n_2}{c-1}$. Otherwise we have

$$x_1 = \frac{1}{n_1} \text{ and } x_2 = \frac{c-1}{n_2}.$$

Thus, we have

$$x_1 = \begin{cases} \frac{c}{n_1 + \sqrt{\phi_2} n_2} & \text{if } n_1 < \frac{\sqrt{\phi_2} n_2}{c-1}, \\ \frac{1}{n_1} & \text{otherwise} \end{cases}$$

and

$$x_2 = \begin{cases} \frac{\sqrt{\phi_2} c}{n_1 + \sqrt{\phi_2} n_2} & \text{if } n_1 < \frac{\sqrt{\phi_2} n_2}{c-1}, \\ \frac{c-1}{n_2} & \text{otherwise} \end{cases}.$$

As a difference to the case of weighted max-min fairness and weighted proportional fairness, the effect of the weights is less significant; comparison between previous two examples and the corresponding non-weighted examples shows that with the weighted criterion a single source s is considered as $\sqrt{\phi_s}$ sources.

6 Utility approach

Utility approach is a more general concept of fairness [4]. Every source $s \in \mathcal{S}$ has a utility function u_s , where $u_s(x_s)$ indicates the value to source s of having the rate x_s . Every link $l \in \mathcal{L}$ has a cost function g_l , where $g_l(f_l)$ indicates the cost to the network of supporting an amount of flow f_l on link l . A ‘‘utility fair’’ allocation $\mathbf{x} \in \mathcal{F}$ is defined as solution to following optimization problem:

$$\begin{aligned} \text{maximize } H(\mathbf{x}) &= \sum_{s \in \mathcal{S}} u_s(x_s) - \sum_{l \in \mathcal{L}} g_l(f_l) & (16) \\ \text{subject to } \mathbf{x} &\in \mathcal{F} \end{aligned}$$

where

$$f_l = \sum_{s \in \mathcal{S}} A_{ls} x_s$$

Different fairness criteria can be presented with criterion specific utility and cost functions, and the fair rate allocation is found by this constrained optimization

problem. The fairness criteria presented in sections 4 and 5 share a common cost function g_l , that guarantees the feasibility of a rate allocation:

$$g_l(f_l) = \begin{cases} 0 & \text{for } f_l \leq C_l, \\ \infty & \text{for } f_l > C_l. \end{cases}$$

Disrupting the feasibility constrain causes an infinite cost, and thus forces the rate allocation into the feasible region.

6.1 General utility function

The following form of the objective function was introduced in [3], based on the criterion introduced in [8]:

$$u_\alpha^s(x) = \begin{cases} \phi_s \log x & \text{if } \alpha = 1, \\ \phi_s \frac{x^{1-\alpha}}{1-\alpha} & \text{otherwise.} \end{cases} \quad (17)$$

Local optimal solution of (16) is unique global optimum, if function $H : \mathcal{F} \mapsto \mathbb{R} : H(\mathbf{x})$ is strictly concave for all $\mathbf{x} \in \mathcal{F}$ and set $\mathcal{F} \neq \emptyset$ is convex [1, Def 3.1.1, Th 3.4.2]. Convexity of \mathcal{F} is proved in Section 2. Now, let us proof the concavity of function $H(\mathbf{x}) = \sum_{s \in \mathcal{S}} u_\alpha^s(x_s)$ in two parts.

Proposition Let us define function f_α as follows:

$$f_\alpha(x) \stackrel{\text{def}}{=} \begin{cases} \log x & \text{if } \alpha = 1, \\ \frac{x^{1-\alpha}}{1-\alpha} & \text{otherwise.} \end{cases}$$

Function $f_\alpha(x)$ is strictly concave for all $x > 0$ and $\alpha \geq 0$.

Proof Function f_α is strictly concave, if its second derivative is negative for all $x > 0$ [1, Def 3.1.1, Th 3.3.8]. Let us first examine the case $\alpha = 1$:

$$\frac{\partial^2}{\partial x^2} \log x = -\frac{1}{x^2} < 0 \quad \forall x > 0.$$

In case $\alpha \neq 1$ we have

$$\frac{\partial^2}{\partial x^2} \frac{x^{1-\alpha}}{1-\alpha} = -\alpha \underbrace{x^{-\alpha-1}}_{\geq 0 \quad \forall x \geq 0} < 0 \quad \forall x > 0 \quad \forall \alpha \in \mathbb{R}_+ \setminus \{1\}.$$

Thus, function $f_\alpha(x)$ is strictly concave for all $x > 0$ and $\alpha \geq 0$. \square

We have $u_\alpha^s(x) = \phi_s f_\alpha(x)$, where $\phi_s \geq 0$. Multiplying with a positive constant maintains the concavity, thus also function u_α^s is strictly concave.

Proposition Let $\{f_i\}_{i \in \mathcal{I}}$ be a finite set of strictly concave functions, $f_i : \mathcal{F} \mapsto \mathbb{R}$, where \mathcal{F} is a nonempty convex set. Let us define function g as follows:

$$g : \mathcal{F} \mapsto \mathbb{R} : g(\mathbf{x}) \stackrel{\text{def}}{=} \sum_{i \in \mathcal{I}} f_i(\mathbf{x})$$

Function g is strictly concave.

Proof Let $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{F}$. Now,

$$\begin{aligned}
g(t\mathbf{x}_1 + (1-t)\mathbf{x}_2) &= \sum_{i \in \mathcal{I}} \underbrace{f_i(t\mathbf{x}_1 + (1-t)\mathbf{x}_2)}_{\substack{> tf_i(\mathbf{x}_1) + (1-t)f_i(\mathbf{x}_2) \\ f_i \text{ is strictly concave}}} \\
&> \sum_{i \in \mathcal{I}} (tf_i(\mathbf{x}_1) + (1-t)f_i(\mathbf{x}_2)) \\
&= t \sum_{i \in \mathcal{I}} f_i(\mathbf{x}_1) + (1-t) \sum_{i \in \mathcal{I}} f_i(\mathbf{x}_2) \\
&= tg(\mathbf{x}_1) + (1-t)g(\mathbf{x}_2) \quad \forall t \in [0, 1] \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{F} \\
&\Leftrightarrow g \text{ is strictly concave.} \quad \square
\end{aligned}$$

Now, function H is a sum of strictly concave functions and thereby function $H(\mathbf{x})$ is also strictly concave, when $\mathbf{x} \in \mathcal{F}$ and \mathcal{F} is convex. Thus, we have proved that the solution of (16) is unique global optimum. We can define a general utility function as follows:

$$u_{\text{Gen}} : \mathcal{F} \mapsto \mathbb{R} : u_{\text{Gen}}(\mathbf{x}) \stackrel{\text{def}}{=} \sum_{s \in \mathcal{S}} u_{\alpha}^s(x_s). \quad (18)$$

With different values of α and ϕ_s different fairness criteria are achieved [3, 8]. Let us consider the case $\phi_s \equiv 1$. Now, when $\alpha \rightarrow 0$, optimization problem (16) reduces to (4) and the solution is the allocation that maximizes the throughput. When $\alpha \rightarrow 1$ the solution of (16) is proportionally fair [8]. In the case that $\alpha \rightarrow 2$ optimization problem (16) reduces to (11) and the solution minimizes the potential delay. In the case that $\alpha \rightarrow \infty$ the solution of (11) is equal with solution of (5) [8]. Weighted versions of proportional fairness and potential minimization are achieved when $\phi_s \neq 1$.

6.2 (p, α) -Proportional fairness

(p, α) -Proportional fairness (presented in [8]) is a generalization of proportional fairness.

Definition 8 ((p, α) -Proportional fairness). *A bandwidth allocation \mathbf{x} is (p, α) -proportionally fair if $\mathbf{x} \in \mathcal{F}$ and*

$$\sum_{s \in \mathcal{S}} p_s \frac{x'_s - x_s}{x_s^\alpha} \leq 0 \quad \forall \mathbf{x}' \in \mathcal{F} \quad (19)$$

where $\alpha \in \mathbb{R}_+$.

Note that if $p_s = \phi_s$ and $\alpha = 1$, (19) reduces to (12). Further, if $p_s \equiv 1$ and $\alpha = 1$, (19) reduces to (8).

In [8] it is proved that (p, α) -proportionally fair bandwidth allocation is also solution of optimization problem (16), where $H = u_{\text{Gen}}$, as defined in (18). All basic fairness criteria can be derived from (p, α) -proportional fairness and also the weighted potential delay minimization is achieved when $p_s \neq 1$ and $\alpha = 2$. An interesting observation is that when $\alpha \rightarrow \infty$, the (p, α) -proportionally fair

allocation is max-min fair regardless of the values of parameters p_s ⁴. This is confirmed in the following two examples.

Now, let us derive previous results by solving optimization problem (16), where $H(\mathbf{x}) = u_{\text{Gen}}(\mathbf{x})$. Note, that $\frac{\partial u_\alpha}{\partial x_s} = \phi_s x_s^{-\alpha}$ for all α . Thus, it is not required to divide following examinations into two parts.

Example 1 All flows in the same class \mathcal{R}_r get equal allocation of the bandwidth x_r . It is required that $n_0 x_0 + n_l x_l = 1$, from which follows $x_l = \frac{1 - n_0 x_0}{n_l}$. (p, α) -Proportionally fair rate allocation maximizes function $u(\mathbf{x})$, where

$$\begin{aligned} u(\mathbf{x}) &= \sum_{s \in \mathcal{S}} u_\alpha(x_s) = \sum_{i=0}^2 n_i u_\alpha(x_i) \\ &= n_0 u_\alpha(x_0) + n_1 u_\alpha\left(\frac{1 - n_0 x_0}{n_1}\right) + n_2 u_\alpha\left(\frac{1 - n_0 x_0}{n_2}\right) \end{aligned}$$

Differentiating gives

$$\begin{aligned} \frac{\partial u}{\partial x_0} &= n_0 x_0^{-\alpha} + n_1 \phi_1 \left(\frac{1 - n_0 x_0}{n_1}\right)^{-\alpha} \left(-\frac{n_0}{n_1}\right) \\ &\quad + n_2 \phi_2 \left(\frac{1 - n_0 x_0}{n_2}\right)^{-\alpha} \left(-\frac{n_0}{n_2}\right) \\ &= n_0 (x_0^{-\alpha} - (\phi_1 n_1^\alpha + \phi_2 n_2^\alpha) (1 - n_0 x_0)^{-\alpha}) = 0 \\ &\Leftrightarrow (1 - n_0 x_0)^\alpha = (\phi_1 n_1^\alpha + \phi_2 n_2^\alpha) x_0^\alpha \\ &\Leftrightarrow 1 - n_0 x_0 = \sqrt[\alpha]{\phi_1 n_1^\alpha + \phi_2 n_2^\alpha} x_0 \\ \Rightarrow x_0 &= \frac{1}{n_0 + \sqrt[\alpha]{\phi_1 n_1^\alpha + \phi_2 n_2^\alpha}} \text{ and } x_l = \frac{1}{n_l} \left(1 - \frac{n_0}{n_0 + \sqrt[\alpha]{\phi_1 n_1^\alpha + \phi_2 n_2^\alpha}}\right) \end{aligned}$$

Thus

$$x_r = \begin{cases} \frac{1}{n_0 + \sqrt[\alpha]{\phi_1 n_1^\alpha + \phi_2 n_2^\alpha}} & \text{for } r \in \mathcal{R}_0, \\ \frac{1}{n_l} \frac{\sqrt[\alpha]{\phi_1 n_1^\alpha + \phi_2 n_2^\alpha}}{n_0 + \sqrt[\alpha]{\phi_1 n_1^\alpha + \phi_2 n_2^\alpha}} & \text{for } r \in \mathcal{R}_l, l \in \{1, 2\}. \end{cases} \quad (20)$$

Now, different fairness criteria are achieved with different values of parameters α and ϕ_s : When $\alpha \rightarrow 0$ and $1 < \phi_1 + \phi_2 < \infty$, solution (20) maximizes the throughput, i.e. solution reduces to be a solution of (4). If $0 < \phi_1 + \phi_2 < 1$, letting $\alpha \rightarrow 0$ leads to solution in which all the capacity is allocated for class 0 flows. This actually minimizes the total throughput. When $\alpha \rightarrow 1$, solution (20) reduces to (13) giving the weighted proportionally fair allocation. Setting $\phi_s \equiv 1$ gives the non-weighted solution (9). When $\alpha \rightarrow 2$, solution minimizes the potential delay, i.e. solution reduces to be a solution of (15). Setting $\phi_s \equiv 1$ gives the solution of the non-weighted criteria (11). When $\alpha \rightarrow \infty$ and $0 < \phi_s < \infty$, solution (20) is max-min fair.

⁴Weighted max-min fairness is achieved as a solution of optimization problem (16), where $H(\mathbf{x}) = \sum_{s \in \mathcal{S}} \frac{(\phi_s x_s)^{1-\alpha}}{1-\alpha}$ and $\alpha \rightarrow \infty$

Example 2 All flows in the same class \mathcal{R}_r get equal allocation of the bandwidth x_r . It is required that $n_1x_1 + n_2x_2 = c$, from which follows $x_2 = \frac{c - n_1x_1}{n_2}$. (p, α) -Proportionally fair rate allocation maximizes function $u(\mathbf{x})$, where

$$\begin{aligned} u(\mathbf{x}) &= \sum_{s \in \mathcal{S}} u_\alpha(x_s) = n_1 u_\alpha(x_1) + n_2 u_\alpha(x_2) \\ &= n_1 u_\alpha(x_1) + n_2 u_\alpha\left(\frac{c - n_1x_1}{n_2}\right) \end{aligned}$$

Differentiating gives

$$\begin{aligned} \frac{\partial u}{\partial x_1} &= n_1 x_1^{-\alpha} + n_2 \phi_2 \left(\frac{c - n_1x_1}{n_2}\right)^{-\alpha} \left(-\frac{n_1}{n_2}\right) \\ &= n_1 (x_1^{-\alpha} - \phi_2 n_2^\alpha (c - n_1x_1)^{-\alpha}) = 0 \\ &\Leftrightarrow (c - n_1x_1)^\alpha = \phi_2 n_2^\alpha x_1^\alpha \\ &\Leftrightarrow c - n_1x_1 = \sqrt[\alpha]{\phi_2 n_2} x_1 \\ \Rightarrow x_1 &= \frac{c}{n_1 + \sqrt[\alpha]{\phi_2 n_2}} \text{ and } x_2 = \frac{1}{n_2} \left(c - \frac{n_1 c}{n_1 + \sqrt[\alpha]{\phi_2 n_2}}\right) \end{aligned}$$

if the link 2 is the bottleneck, i.e. $n_1x_1 = \frac{n_1 c}{n_1 + \sqrt[\alpha]{\phi_2 n_2}} < 1$, that is $n_1 < \frac{\sqrt[\alpha]{\phi_2 n_2}}{c-1}$. Otherwise

$$x_1 = \frac{1}{n_1} \text{ and } x_2 = \frac{c-1}{n_2}.$$

Thus,

$$\begin{aligned} x_1 &= \begin{cases} \frac{c}{n_1 + \sqrt[\alpha]{\phi_2 n_2}} & \text{if } n_1 < \frac{\sqrt[\alpha]{\phi_2 n_2}}{c-1}, \\ \frac{1}{n_1} & \text{otherwise} \end{cases} \\ \text{and} & \\ x_2 &= \begin{cases} \frac{\sqrt[\alpha]{\phi_2} c}{n_1 + \sqrt[\alpha]{\phi_2 n_2}} & \text{if } n_1 < \frac{\sqrt[\alpha]{\phi_2 n_2}}{c-1}, \\ \frac{c-1}{n_2} & \text{otherwise} \end{cases}. \end{aligned} \tag{21}$$

Now, different fairness criteria are achieved with different values of parameters α and ϕ_s : When $\alpha \rightarrow 0$, two different solutions are achieved depending on the value of ϕ_2 ; if $1 < \phi_2 < \infty$, all the capacity is given to class 2 flows, whereas case $0 < \phi_2 < 1$ leads to allocation in which class 1 flows are given the maximum amount of capacity. Nevertheless, solution (21) maximizes the throughput, i.e. solution reduces to be a solution of (4). When $\alpha \rightarrow 1$, solution (21) reduces to (14) giving the weighted proportionally fair allocation. Setting $\phi_s \equiv 1$ gives the non-weighted solution (10). When $\alpha \rightarrow 2$, solution minimizes the potential delay, i.e. solution reduces to be a solution of (15). Setting $\phi_s \equiv 1$ gives the solution of the non-weighted criteria (11). When $\alpha \rightarrow \infty$ and $0 < \phi_s < \infty$, solution (21) is max-min fair.

7 Conclusions

In this paper we have presented all the most common static fairness criteria and derived the rate allocations given by these criteria in two simple example network setups. We have also examined the relations between different criteria and presented the equality of some fairness criteria and the utility approach, which generalizes the concept of fairness.

The main objective of bandwidth sharing is to use all available bandwidth without disrupting the constraints and maintain a certain fairness. The achieved fairness depends closely on the used fairness criterion. Different fairness criteria favor or discriminate single sources or whole traffic classes on different basis.

As it was verified by the examples, the throughput maximization gives the most efficient rate allocation at the expense of a single sources or some flow classes. The max-min fair allocation can be considered as the most fair criterion on the grounds that its objective is to provide as equal rate allocation as possible to all sources. This happens at the expense of the total throughput. The proportional fairness penalizes long routes more than max-min fairness with tendency to achieve greater total throughput. The objective of potential delay minimization is to minimize the delay needed to complete transfers and the provided rate allocation is between when compared to max-min and proportional fairness. In the case of a single bottleneck, all basic criteria provide the same allocation.

The study in this paper was strongly bounded and a great number of simplifications was made concerning the network model, traffic and the time scale. Fairness in TCP traffic and the relation between different congestion control schemes and the realized fairness are topical themes concerned in recent research. Thus, this paper can be considered as preliminary study for further work.

References

- [1] Mokhtar S. Bazaraa, Hanif D. Sherali, and C. M. Shetty. *Nonlinear Programming : theory and algorithms*. John Wiley & Sons, second edition, 1993.
- [2] Dimitri Bertsekas and Robert Gallager. *Data Networks*. Prentice Hall, second edition, 1992.
- [3] Thomas Bonald and Laurent Massoulié. Impact of fairness on internet performance. In *SIGMETRICS/Performance*, pages 82–91, 2001.
- [4] Jean-Yves Le Boudec. Selected lecture notes. SSC-D02, October 1998.
- [5] Paul Hurley, Jean-Yves Le Boudec, and Patrick Thiran. A note on the fairness of additive increase and multiplicative decrease. In *ITC-16*, June 1999.
- [6] Frank Kelly. Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, 8:33 – 37, January 1997.
- [7] Laurent Massoulié and James Roberts. Bandwidth sharing: Objectives and algorithms. In *INFOCOM'99 (3)*, pages 1395 – 1403, 1999.
- [8] Jeonghoon Mo and Jean Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on Networking*, 8(5):556 – 567, 2000.
- [9] James Roberts and Laurent Massoulié. Bandwidth sharing and admission control for elastic traffic. In *ITC Specialist Seminar*, October 1998.