Network Utility Maximization in Two-way Flow Scenario

Fei Ge
Computer Science Department
Central China Normal University
430079, Wuhan, Hubei, P.R. China
Feige@mail.ccnu.edu.cn

Liansheng Tan*
Computer Science Department
Central China Normal University
430079, Wuhan, Hubei, P.R. China
L.Tan@mail.ccnu.edu.cn

ABSTRACT

A communication network usually has data packets and acknowledge (ACK) packets being transmitted in opposite directions. ACK packet flows may affect the performance of data packet flows, which is unfortunately not considered in the usual network utility maximization (NUM) model. This paper presents a NUM model in networks with two-way flows (NUMtw) by adding a routing matrix to cover ACK packet flows. The source rates are obtained by solving the dual model and the relation to the routing matrix of ACK packet flows is disclosed. Furthermore, the source rates in networks with one-way flows by the usual NUM model are compared to those in networks with two-way flows by the NUMtw model.

Categories and Subject Descriptors

C.2.5 [**Computer-Communication Networks**]: Local and Wide-Area Networks—*Internet*

Keywords

Network Utility Maximization, Two-way Flows, Throughput Rates

1. INTRODUCTION

In packet switching networks, a flow, also called a traffic flow, a packet flow or a network flow, is a sequence of packets traversing from a source to a destination [3]. Some unidirectional or one-way protocols, *e.g.* User Datagram Protocol (UDP), only have one flow traversing in the system; while others, *e.g.* Transmission Control Protocol (TCP), have two flows transmitting packets in opposite directions, either through the identical or distinct path. Flow control is the process of managing the rate of data transmission between two nodes to prevent a fast sender from outrunning a slow receiver. If the receiver sends feedback to the sender, flow control is a closed loop. A feedback closed loop system has a feedback mechanism that directly relates the input and output signals. For example, TCP adopts Acknowledge (ACK) packet to feed back the congestion information (packet loss or delay) at the destination to the source.

Kelly's Network Utility Maximization (NUM) theory [8], [9] amazingly casts the issue of flow control and bandwidth allocation into a unified optimization framework, which proposes distributed computations running at sources and at links over the network to solve an optimization problem with the objective to maximize the aggregate source utility subject to capacity constrains. The NUM theory has powerful impacts on the following up developments in high speed TCP protocols and other related areas, to mention a few [4], [9], [10], [11], [12], [13], [14], [15] and [16].

Regarding the flow control model, most expositions simplify the impact of ACK packet flow by ignoring the ACK packet size. The consequence is that all transmission protocols can be approximated by one-way protocols. This approximation is reasonable and may not affect the protocols' performance when the ACK packet size is small enough not to affect the data transmission. For example, in Client/Server (C/S) services with symmetric links, ACK packets are transmitted in the direction from the client to the server, and data packets are in the direction from the server to the client. In this case, ACK packets and data packets are transmitted in opposite directions. The volume of ACK packet flow, compared to the uplink capacity, is very small, thus it does not affect the data flow significantly.

However, networks may actually require that ACK packets and data packets be transmitted in the same direction. Therefore, ACK packet and data packet are mixed in one link, which can result in the ACK compression [7]. Moreover, asymmetric link can affect TCP flow's performance [2]. In these cases, the desired TCP throughput rate can be deteriorated by the ACK packet flow. Some techniques [2, 5], *i.e.* ACK Filtering, ACK congestion control are proposed to address this problem in recent years. Recent analysis is performed to study the source rates of two-way FAST TCP [17] flows with the same or different parameters in [6]. The observation made therein is that two-way FAST TCP flow throughput rate is dependent on network parameters and can be significantly deviated from that of one-way flow in some scenarios. Any TCP connection includes two-way flows. Unfortunately, in the existing NUM models the capacity constrains only consider the data packet flow but no ACK packet flow. Subsequently, they are not appropriate to address the source rates of two-way flows.

The main contribution of the current paper is to generalize the usual NUM model from one-way flow scenario to two-way flow scenario to cover the ACK packet flow. We propose a model called the NUM model for networks with two-way flows (NUMtw), which is different from the usual NUM model that only includes the ACK packet flows. To model a network with two-way flows, the routing matrix covering the routing information of ACK packet flow is introduced firstly and the relation between the sizes of data

*Corresponding to Liansheng Tan (email: l.tan@mail.ccnu.edu.cn).

packet and ACK packet is identified next. The NUMtw model is therefore defined and a solution of the source rate to the model is given by its dual problem. The difference of source rates between one-way flow networks and two-way flow networks are discussed.

2. TWO-WAY NUM MODEL

In this section, we present the NUMtw model of two-way flows in networks with duplex links. Furthermore, the optimization problem will be solved by its dual [12] and the source rate solution will be given.

Consider a network with a set of nodes and L duplex links. One transmission direction of a link is numbered by an integer, *e.g.* l, and the other by an integer $L+l$ [1]. There are S connections. Each connection includes a data flow and a corresponding ACK flow. The direction of a data flow is from its source node to its destination node; while the direction of the ACK flow is oppositive. A data flow is numbered by an integer, *e.g.* s, and its corresponding ACK flow is numbered by $S+s$ [2]. Let c_l, c_{L+l} be the capacities, $x_s, x_s > 0, x_{S+s}, x_{S+s} \geq 0$, be the flow rates. x_{S+s} is related to x_s, *i.e.* $x_{S+s} = f_s(x_s)$, where $f(\bullet)$ is a non-decrease function. The effective flow for a connection is the data flow. A flow is related to a utility function $U_s(x_s)$, which describes the degree of user satisfaction when allocated by a certain amount of bandwidth. Been defined as a curve mapping the amount of bandwidth received by the application to the performance as perceived by the end user, the utility function is assumed to be increasing and continuously differentiable. Like in the usual NUM model the utility function is assumed to be concave in our approach though a non-concave utility function can be extended from our model.

To determine the source rates, the following NUMtw model is proposed:

$$max \sum_{s=1}^{S} U_s(x_s) \tag{1}$$

$$\text{s.t. } B \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \leq \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \tag{2}$$

$$x_{S+s} = f(x_s) \quad 1 \leq s \leq S \tag{3}$$

where

$$B = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

is the routing matrix with $2L \times 2S$ dimensions. The matrices A_{11}, A_{12}, A_{21} and A_{22} are all with $L \times S$ dimensions. $C_1 = \{c_l\}$, $C_2 = \{c_{L+l}\}$, $X_1 = \{x_s\}$, $X_2 = \{x_{S+s}\}$ are the column vectors. $A_{11} = \{b_{ls}\}$ denotes routing information of flow s through link l, $b_{ls} = 1$ if flow s transits through link l; otherwise $b_{ls} = 0$. $A_{12} = \{b_{l(S+s)}\}$ denotes routing information of flow $S+s$ through link l, $b_{l(S+s)} = 1$ if flow $S+s$ transits through link l; otherwise $b_{l(S+s)} = 0$. $A_{21} = \{b_{(L+l)s}\}$ denotes routing information of flow s through link $L+l$, $b_{(L+l)s} = 1$ if flow s transits through link $L+l$; otherwise $b_{(L+l)s} = 0$. $A_{22} = \{b_{(L+l)(S+s)}\}$ denotes routing information of flow $S+s$ through link $L+l$, $b_{(L+l)(S+s)} = 1$ if flow $S+s$ transits through link $L+l$; otherwise $b_{(L+l)(S+s)} = 0$.

[1] l lies in the set $\{l \mid 1 \leq l \leq L\}$ if no further explanation.
[2] s lies in the set $\{s \mid 1 \leq s \leq S\}$ if no further explanation.

Let $f(\bullet)$ be a linear function defined by $f(x_{S+s}) = \theta_s x_s$, $0 \leq \theta_s \leq 1$. By substituting it into (1) and (2), the NUMtw model becomes

$$\text{maximize } \sum_{s=1}^{S} U_s(x_s) \tag{4}$$

$$\text{s.t. } \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} X_1 \\ \theta X_1 \end{bmatrix} \leq \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \tag{5}$$

where $\theta = diag\{s_1, \ldots, s_S\}$ is a diagonal matrix.

The source rates in NUMtw model can be solved by the dual model. Let $\Phi_1 = \{\mu_l, \mu_l \geq 0\}$, $\Phi_2 = \{\mu_{L+l}, \mu_{L+l} \geq 0\}$, the Lagrangian form is defined as

$$L(X_1, \Phi_1, \Phi_2)$$

$$= \sum_{s=1}^{S} U_s(x_s) + \sum_{l=1}^{2L} \mu_l (c_l - \sum_{s=1}^{S} x_s (b_{ls} + b_{l(S+s)}\theta_s))$$

$$= \sum_{s=1}^{S} (U_s(x_s) - x_s \sum_{l=1}^{2L} \mu_l (b_{ls} + b_{l(S+s)}\theta_s)) + \sum_{l=1}^{2L} \mu_l c_l$$

The terms in $L(X_1, \Phi_1, \Phi_2)$ are not coupled with x_s, then

$$maxL(X_1, \Phi_1, \Phi_2)$$

$$= \sum_{s=1}^{S} max \left(U_s(x_s) - x_s \sum_{l=1}^{2L} \mu_l (b_{ls} + b_{l(S+s)}\theta_s) \right) + \sum_{l=1}^{2L} \mu_l c_l$$

The dual model of NUMtw is

$$minD(\Phi_1, \Phi_2) \tag{6}$$

where the objective function is

$$D(\Phi_1, \Phi_2) = maxL(X_1, \Phi_1, \Phi_2) = \sum_{s=1}^{S} Y_s(p_s) + \sum_{l=1}^{2L} \mu_l c_l$$

$$Y_s(p_s) = max(U_s(x_s) - x_s p_s) \tag{7}$$

p_s is the entry of the vector P

$$P \ = \begin{bmatrix} A_{11}^T + \theta A_{12}^T, & A_{21}^T + \theta A_{22}^T \end{bmatrix} \begin{bmatrix} \Phi_1 \\ \Phi_2 \end{bmatrix} \tag{8}$$

The Lagrangian multiplier, $\mu_l, 1 \leq l \leq 2L$, is interpreted as the price per unit bandwidth, and p_s is the aggregated link prices observed by the source.

The solution to model NUMtw can be induced to solve maximization (7) by bandwidth charging. For each p_s, a unique maximizer, $x_s(p_s)$, exists. Because $U_s(x_s)$ is increasing, strictly concave and continuously differentiable function of x_s in its argument, $U_s^{'}$ exists and is decreasing. By the Kuhn-Tuker theorem [1], it can be obtained that

$$p_s = U_s^{'}(x_s) \tag{9}$$

The inverse of $U_s^{'}$ is $U_s^{'-1}$, and

$$x_s(\mu) = [U_s^{'-1}(\sum_{l}^{2L} \mu_l (b_{ls} + b_{l(S+s)}\theta_s))]^+ \tag{10}$$

Thus, for given μ_l of all links and the matrix $[A_{11}^T + \theta A_{12}^T,\ A_{21}^T + \theta A_{22}^T]$, an individual connection can solve its rate.

The NUMtw model considers the relation between a connection's data packet flow and its ACK flow, which is not the direct application of the usual NUM model in networks where ACK packet flows are included. If an ACK packet flow is bound to an utility function, the NUM model is

$$max\sum_{s=1}^{2S}U_s(x_s) \qquad \text{s.t. } B\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \le \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \tag{11}$$

In this model, the ACK flows are included, modeled to be isolated flows, and not related to their corresponding data flows. Therefore, the object function is the the sum of $2S$ utility functions. Moreover, the rate of a data flow is not related to that of the corresponding ACK flow.

In the usual NUM theory, the rule is that the transpose of a routing matrix is the price matrix of the connections. The routing matrix of the NUM model described by (11) is the same with that of NUMtw model. It is expected that B^T is the routing matrix in both models. However, it will be shown in the following that it is not the case.

The dual problem of model (11) can be obtained in the same way as that of NUMtw

$$minD(\Phi_1,\Phi_2) \tag{12}$$

$$D(\Phi_1,\Phi_2) = maxL(X_1,X_2,\Phi_1,\Phi_2) = \sum_{s=1}^{2S}Y_s(p_s) + \sum_{l=1}^{2L}\mu_l c_l$$

$$Y_s(p_s) = max(U_s(x_s) - x_s p_s) \tag{13}$$

$$P \ = B^T\begin{bmatrix} \Phi_1 \\ \Phi_2 \end{bmatrix} \tag{14}$$

It can be observed that this dual model is not identified to that of NUMtw. Firstly, there are $2S$ prices in this dual model, while S prices in the dual model of NUMtw. Furthermore, the rate of each flow in this model is

$$x_s(\mu) = [U_s^{'-1}(\sum_{l}^{2L}\mu_l b_{ls})]^+ \quad 1\le s\le 2S \tag{15}$$

which is not equivalent to that in NUMtw model. Lastly, the price matrix of this dual model differs that of NUMtw. That is, the routing matrix's transpose in this model is its price matrix, while that in NUMtw is not. Therefore, the price matrix in networks with two-way flows can not be obtain from the transpose of the routing matrix B.

The routing matrix and the price matrix in networks modeled by NUMtw model should be rediscovered. To achieve this, in NUMtw model, let the matrices K, H

$$K = [A_{11}^T,\ A_{21}^T]^T \quad H = [A_{12}^T,\ A_{22}^T]^T$$

be the routing matrices of the data packet flows and the ACK packet flows, respectively. Recall that $\theta_s, 0\le\theta_s\le 1$, bridges the throughput rates of data packet flow and ACK packet flow of a connection, let the matrix G be the combination of K and H

$$G = [K + H\theta]$$

From the data packet flow's perspective, the matrix G can be called the **extended routing matrix**. This matrix includes not only the routing information of all flows, but also the relation between the throughput rates between a connection's data packet flow and its ACK packet flow. Therefore, the entry of G is not just 0 and 1, but may be θ_s or $1+\theta_s$. Thus the constraint (5) in NUMtw model becomes

$$GX_1 \le \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \tag{16}$$

Then NUMtw is, from the prospective of the formula form, the same with the usual NUM model in networks with one-way flows. According to the usual NUM theory, the transpose of G, i.e. G^T, is the price matrix. Notice that G^T may not be the transpose of matrices B, K or H, it can be called the **extended price matrix** [3]. Similarly, the entry of the extended price matrix may not be 0 or 1. The price of the data packet sources in NUMtw model, from (8), is

$$P \ = G^T\begin{bmatrix} \Phi_1 \\ \Phi_2 \end{bmatrix} \tag{17}$$

3. MAIN RESULTS ON FLOW RATES

It makes difference to model flows to be one-way flows and two-way flows. Previously, flows are always modeled to be one-way flows in networks with simplex links, while two-way flows exist in real networks. In the same network, the flow rates of one-way flows and those of two-way flows are possibly not the same. In this section, the flow rates in the following two scenarios will be compared. In Scenario S1A, only one transmission direction of a duplex link and one-way flows are considered. In Scenario S1B, another transmission direction of each duplex link and ACK flow are considered, and the routing path of an ACK packet flow is the same as that of the corresponding data packet flow, then it can be find that in the following theorem the parameter θ_s may result in the less rates in S1B.

THEOREM 1. *If $c_{l+L}/c_l \ge \theta_{MAX}$, the flow rates in S1B are identical to those in S1A; while when $c_{L+l}/c_l \le \theta_{MIN}$, the flow rates in S1B are not larger than those in S1A, where $\theta_s > 0, \theta_{MAX} = max\{\theta_s\}, \theta_{MIN} = min\{\theta_s\}$.*

PROOF. The constraint of the usual NUM model in S1A is $A_{11}X_1 \le C_1$. The feasible set $F1A$ is $\{\{x_s\}, \sum_{s=1}^{S} b_{ls}x_s \le c_l\}$. In S1B, $A_{11} = A_{22}, A_{12} = A_{21} = 0$ because the routing path of an ACK flow is the same as that of the corresponding data flow. The constraint of the NUMtw model in S1B is

$$\begin{bmatrix} A_{11} & 0 \\ 0 & A_{11} \end{bmatrix}\begin{bmatrix} X_1 \\ \theta X_1 \end{bmatrix} \le \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$$

Note that the feasible set $F1B$ is $\{\{x_s\}, \sum_{s=1}^{S} b_{ls}x_s \le c_l, \sum_{s=1}^{S} b_{ls}\theta_s x_s \le c_{L+l}\}$. Now it is an immediate result that if $c_{L+l}/c_l \ge \theta_{MAX}$, $F1A = F1B$. Therefore, the flow rates in S1A are identical to those in S1B. In the same way, we can prove that if $c_{L+l}/c_l \le \theta_{MIN}$, $F1A \supseteq F1B$. Because the object function is strictly concave and increasing, the flow rates in S1A are not less than those in S1B.

Now it will be proved that

1. if $c_{L+l}/c_l \ge \theta_{MAX}$, $F1A = F1B$;

2. if $c_{L+l}/c_l \le \theta_{MIN}$, $F1A \supseteq F1B$.

[3]Specially, when $H = 0$, $B = K$, the extended routing matrix G^T is the routing matrix B^T.

The feasible set $F1B$ can be described as the intersection of the two sets $F1B_1 = \{\{x_s\}, \sum_{s=1}^{S} b_{ls}x_s \leq c_l\}$, $F1B_2 = \{\{x_s\}, \sum_{s=1}^{S} b_{ls}\theta_s x_s \leq c_{L+l}\}$.

1. If $c_{L+l}/c_l \geq \theta_{MAX}$, it holds

$$F1B_1 \subseteq \{\{x_s\}, \sum_{s=1}^{S} b_{ls}x_s \leq c_{L+l}/\theta_{MAX}\} = \{\{x_s\}, \sum_{s=1}^{S} b_{ls}\theta_{MAX}x_s \leq c_{L+l}\}$$

   Because θ_{MAX} is not less than θ_s, it follows

$$\{x_s, \sum_{s=1}^{S} b_{ls}\theta_{MAX}x_s \leq c_{L+l}\} \subseteq F1B_2$$

   Therefore, $F1B$ is actually equal to $F1B_1$. Note that $F1A = F1B_1$. As a result, if

$$c_{L+l}/c_l \geq \theta_{MAX}$$

   it holds

$$F1A = F1B$$

2. If $c_{L+l}/c_l \leq \theta_{MIN}$, it follows

$$F1B_2 \subseteq \{\{x_s\}, \sum_{s=1}^{S} b_{ls}\theta_s x_s \leq \theta_{MIN}c_l\} = \{\{x_s\}, \sum_{s=1}^{S} b_{ls}\theta_s/\theta_{MIN}x_s \leq c_l\}$$

   Due to $\theta_s/\theta_{MIN} \geq 1$, it holds

$$\{\{x_s\}, \sum_{s=1}^{S} b_{ls}\theta_s/\theta_{MIN}x_s \leq c_l\} \subseteq F1B_1$$

   Therefore, in this situation, $F1B$ is actually equal to $F1B_2$. Further, $F1A \supseteq F1B$ if $c_{L+l}/c_l \leq \theta_{MIN}$.

This concludes the proof. □

If for all sources, $\theta_s = \theta_0$, and the price per unit bandwidth of the two transmission direction of a link is identical, we have the following corollary directly from Theorem 1.

COROLLARY 1. *If $c_{L+l}/c_l \geq \theta_0$, the flow rates in S1B are identical to those in S1A; while $c_{L+l}/c_l \leq \theta_0$, the flow rates in S1B are not larger than those in S1A.*

Next compare the flow rates between the following scenarios S2A and S2B. In Scenario S2A, both transmission directions of a duplex link are used by different data packet flows. The characteristic of S2A is that the routing matrixes A_{11}, A_{21} are not zero and A_{21}, A_{22} are zero matrices. In Scenario S2B, ACK packet flows are considered together with their data packet flows, and the routing path of an ACK packet flow is the same as that of the corresponding data packet flow. Then it has the following result.

THEOREM 2. *The flow rates in S2B are at most equal to those in S2A, and equal to those in S2A if $\theta_s = 0$.*

PROOF. Consider the NUM model in S2A, it is seen that the feasible set $F2A$ is

$$\{\{x_s\}, \sum_{s=1}^{S} b_{ls}x_s \leq c_l, l = 1, \cdots, 2L\}$$

In S2B, $A_{11} = A_{22}, A_{12} = A_{21}$, the feasible set $F2B$ of the NUMtw model in S2B is

$$\{\{x_s\}, \sum_{s=1}^{S} b_{ls}x_s \leq c_l, l = 1, \cdots, L\}$$

The constraints in the NUMtw model in S2B are

$$\begin{bmatrix} A_{11} & A_{21} \\ A_{21} & A_{11} \end{bmatrix} \begin{bmatrix} X_1 \\ \theta X_1 \end{bmatrix} \leq \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$$

The feasible set $F2B$ is

$$\{\{x_s\}, \sum_{s=1}^{S}(b_{ls} + b_{(L+l)s}\theta_s)x_s \leq c_l, \sum_{s=1}^{S}(b_{(L+l)s} + b_{ls}\theta_s)x_s \leq c_{L+l}, l = 1, \cdots, L\}$$

Now it is an immediate result that $F2A \supseteq F2B$. Therefore, the flow rates in S2B are at most equal to those in S2A. If $\theta_s = 0$, $F2A = F2B$, the flow rates in S2B are equal to those in S2A.

This concludes the proof. □

In the two-way scenario, when the connections share the same path and $U_s(x_s) = U(x_s)$, the fairness of two flow rates is described in Theorems 3 and 4. When the connections sharing the same path have logarithmic utilities, it is well-known that the connection rates are proportionally fair in the one-way scenario; while the fairness is related to θ_s in two-way scenario.

THEOREM 3. *For all connections with positive flow rates in the two-way scenario, if the data flows of the connections i and j share the same path, and their ACK flows share the same path, when $\theta_i > \theta_j$, $x_i \leq x_j$; when $\theta_i < \theta_j$, $x_i \geq x_j$; when $\theta_i = \theta_j$, $x_i = x_j$.*

PROOF. Due to $b_{li} = b_{lj}, b_{l(S+i)} = b_{l(S+j)}$, when $\theta_i > \theta_j$, it follows

$$\sum_{l}^{2L} \mu_l(b_{li} + b_{l(S+i)}\theta_i) \geq \sum_{l}^{2L} \mu_l(b_{lj} + b_{l(S+j)}\theta_j)$$

Because $U^{'-1}(\bullet)$ is a strictly decreasing function, thus

$$U^{'-1}(\sum_{l}^{2L} \mu_l(b_{li} + b_{l(S+i)}\theta_i)) \leq U^{'-1}(\sum_{l}^{2L} \mu_l(b_{lj} + b_{l(S+j)}\theta_j))$$

From (10), it follows the result $x_i \leq x_j$

In the same way, it can be proved that when $\theta_i < \theta_j$, $x_i \geq x_j$ and when $\theta_i = \theta_j$, $x_i = x_j$.

This concludes the proof. □

Further, it can be obtained that, for $l = 1, \cdots, L$

1. when $\theta_i > \theta_j$, $x_i < x_j$ if $\mu_l = 0$;
2. when $\theta_i < \theta_j$, $x_i > x_j$ if $\mu_l = 0$;
3. $x_i = x_j$ if $\mu_{L+l} = 0$.

THEOREM 4. *For all connections with positive flow rates in the two-way scenario, if the data flow of connection i and the ACK flow of connection j share the same links in $\{1, \cdots, L\}$, the ACK flow of connection i and the data flow of connection j share the same links in $\{L+1, \cdots, 2L\}$, when c_l is large enough,* i.e. *$\mu_l = 0$, it has $x_j = U'^{-1}(U'(x_i)/\theta_i)$; while if c_{L+l} is large enough,* i.e. *$\mu_{L+l} = 0$, it has $x_i = U'^{-1}(U'(x_j)/\theta_j)$.*

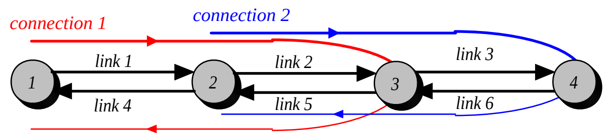

Figure 1: A network with two connections. The thick and thin lines mean the data packet flows and ACK packet flows, respectively.

PROOF. If the data flow and of the connection i and the ACK flow of the connection j share the same path, it has $b_{li} = b_{l(S+j)}$. If the ACK flow of the connection i and the data flow of the connection j share the same path, it has $b_{(L+l)(S+i)} = b_{(L+l)j}$. When $\mu_l = 0$, it has $U'(x_i) = \theta_i U'(x_j)$ from (10). Therefore,

$$x_j = U'^{-1}(U'(x_i)/\theta_i)$$

Similarly, when $\mu_{L+l} = 0$, it can be proved that $x_i = U'^{-1}(U'(x_j)/\theta_j)$.

This concludes the proof. □

In particular, when $U(x_s) = log x_s$, the flow rates of the connections i and j have, by Theorem 4,

1. when $\mu_l = 0$, one has $x_j = x_i\theta_i$;

2. $\mu_{L+l} = 0$, one has $x_i = x_j\theta_j$.

Next the flow rates are accurately provided in two examples. In both examples, suppose $x_s > 0, c_l > 0, 0 < \theta_s < 1$.

Example 1: Consider a network with three duplex links, as in Fig.1. The capacities of the six links are, in order, $100c_2$, c_2, $100c_2$, $100c_5$, c_5 and $100c_5$. The utility functions are $U_1(x_1) = log x_1, U_2(x_2) = 9 log x_2$.

The data packet flows of the connections 1 and 2 are from node 1 to node 3, from node 2 to node 4, respectively. The constraints of the NUM model in S1A are: $x_1 \le 100c_2, x_1 + x_2 \le c_2, x_2 \le 100c_2$. By solving the optimization problem [1], the flow rates in S1A are

$$x_1 = 0.1c_2 \quad x_2 = 0.9c_2 \tag{18}$$

The ACK packet flows of the connections 1 and 2 are from node 3 to node 1, from node 4 to node 2, respectively. Another three constraints are added to the NUMtw model in S1B compared to the NUM model in S1A, which are $\theta_1 x_1 \le 100c_5, \theta_1 x_1 + \theta_2 x_2 \le c_5, \theta_2 x_2 \le 100c_5$. There are two possible solutions: one is

$$x_1 = c_5/(10\theta_1) \quad x_2 = 9c_5/(10\theta_2)$$

under the condition $1/\theta_1 + 9/\theta_2 < 10c_2/c_5$; the other is as (18) under $\theta_1 + 9\theta_2 < 10c_5/c_2$ or $\theta_1 = \theta_2, c_5 = \theta_1 c_2$.

In this example, Theorem 1 can be verified by the accurate flow rates, that is, if $c_5/c_2 > \theta_{MAX}$, the flow rates in S1B are identical to those in S1A; while $c_5/c_2 < \theta_{MIN}$, the flow rates in S1B are less than those in S1A.

Specially, setting the data packet size to 1000 bytes and the ACK packet size 40 bytes for all connections, and $c_2 =$ 500 [Mb/s], $c_5 = 10$ [Mb/s] in Example 1, the flow rates x_1, x_2 in S1A are 50 [Mb/s] and 450 [Mb/s]; while in S1B, they are 25 [Mb/s] and 225 [Mb/s].

Example 2: Add connection 3 to the network in Example 1 and each connection has the same utility function.

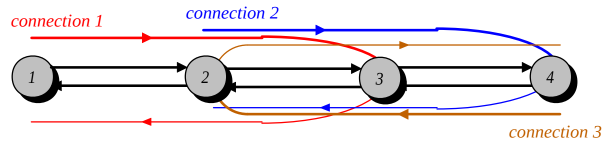

Figure 2: A network with three connections, where the transmission directions of connection 3 are reverse to those of connection 2.

The data packet flow and the ACK packet flow of the connection are from node 4 to node 2 and from node 2 to node 4, respectively, as shown in Fig. 2.

The solution to the NUM model in S2A can be solved by $U_1'(x_1) = U_2'(x_2), x_1 + x_2 = c_2, x_3 = c_5$.

The extended routing matrix G in S2B is

$$G^T = \begin{pmatrix} 1 & 1 & 0 & \theta_1 & \theta_1 & 0 \\ 0 & 1 & 1 & 0 & \theta_2 & \theta_2 \\ 0 & \theta_3 & \theta_3 & 0 & 1 & 1 \end{pmatrix}$$

There are four possible solutions in S2B: one possible solution satisfies

$$\theta_1 U_1'^{-1}(\theta_1 U_3'(x_3)) + \theta_2 U_2'^{-1}(\theta_2 U_3'(x_3)) + x_3 = c_5$$

$$x_1 = U_1'^{-1}(\theta_1 U_3'(x_3))$$

$$x_2 = U_2'^{-1}(\theta_2 U_3'(x_3))$$

$$x_1 + x_2 + \theta_3 x_3 < c_2 \quad \mu_5 > 0$$

the other possible solution satisfies

$$U_1'^{-1}(U_3'(x_3)/\theta_3) + U_2'^{-1}(U_3'(x_3)/\theta_3) + \theta_3 x_3 = c_2$$

$$x_1 = U_1'^{-1}(U_3'(x_3)/\theta_3)$$

$$x_2 = U_2'^{-1}(U_3'(x_3)/\theta_3)$$

$$\theta_1 x_1 + \theta_2 x_2 + x_3 < c_5 \quad \mu_2 > 0$$

the third possible solution satisfies

$$c_2 - \theta_3 c_5 - (1 - \theta_1\theta_3)x_1 = (1 - \theta_2\theta_3)U_2'^{-1}(U_1'(x_1))$$

$$x_1 = U_1'^{-1}(U_2'(x_2))$$

$$x_2 = U_2'^{-1}(U_1'(x_1))$$

$$\mu_2 = \frac{U_1'(x_1) - \theta_1 U_3'(x_3)}{1 - \theta_1\theta_3} > 0$$

$$\mu_5 = \frac{U_3'(x_3) - \theta_3 U_1'(x_1)}{1 - \theta_1\theta_3} > 0$$

the fourth possible solution satisfies

$$U_3'\left(\frac{c_5 - \theta_2 c_2 - (\theta_1 - \theta_2)x_1}{1 - \theta_2\theta_3}\right) =$$

$$\frac{(\theta_2\theta_3 - 1)U_1'(x_1) - (\theta_1\theta_3 - 1)U_2'\left(\frac{c_2 - \theta_3 c_5 - (1-\theta_1\theta_3)x_1}{1-\theta_2\theta_3}\right)}{\theta_2 - \theta_1}$$

$$x_2 = \frac{c_2 - \theta_3 c_5 - (1 - \theta_1\theta_3)x_1}{1 - \theta_2\theta_3}$$

$$x_3 = c_5 - \theta_1 x_1 - \theta_2 x_2$$

$$\mu_5 = \frac{U_1'(x_1) - U_2'(x_2)}{\theta_1 - \theta_2} > 0$$

$$\mu_2 = \frac{\theta_2 U_1'(x_1) - \theta_1 U_2'(x_2)}{\theta_2 - \theta_1} > 0$$

In this example, if $U_s(x_s) = log x_s$, $\theta_s = \theta_0$, the flow rates in S2A are

$$x_1 = x_2 = c_2/2 \quad x_3 = c_5$$

while when the settings satisfy

$$\theta_0 < \frac{3\theta_0}{1+2\theta_0^2} < \frac{c_2}{c_5} < \frac{2+\theta_0^2}{3\theta_0} < \frac{1}{\theta_0}$$

the flow rates in S2B are

$$x_1 = x_2 = \frac{c_2 - \theta_0 c_5}{2(1-\theta_0^2)} \quad x_3 = \frac{c_5 - \theta_0 c_2}{1-\theta_0^2}$$

which are less than those in S2A.

Specially, setting the data packet size to 500 bytes and the ACK packet size 40 bytes for all connections, and $c_2 = c_5 = 100$ [Mb/s] in Example 2, the flow rates x_1, x_2 and x_3 in S1A are 50 [Mb/s], 50 [Mb/s] and 100 [Mb/s]; while in S1B, they are 46.3 [Mb/s], 46.3 [Mb/s] and 92.6 [Mb/s].

4. CONCLUSION

In this paper, the NUM problem in networks with two-way flows is investigated. Firstly, the NUMtw model in networks with two-way flows is proposed, and the flow rates are obtained by its dual problem. The model takes into account the routing information of the ACK packet flows. Secondly the flow rates in the two-way flow scenario are studied and compared to those in the one-way flow scenario. It is found that if two-way flows are present, the flow rates may be equal to or less than those of one-way flows.

To the best of our knowledge, this paper is the first attempt to investigate the NUM theory in the two-way flow scenario. The modeling and solution lead to some straightforward discoveries of a critical problem with network bandwidth allocation and performance optimization in the NUM framework to account for the impact of ACK packet flow.

Flow control approaches consist of two components: a source algorithm that dynamically adjusts rate (or window size) in the response to congestion in its path, and a link algorithm that updates, implicitly or explicitly, a congestion measure and sends it back to sources that uses that link. These algorithms, well-known as the primal-dual algorithm, are all resulted from the NUM model. However, these algorithms unfortunately failed to address the impacts of ACK packet flow. With an extension of the usual NUM model being proposed in the current paper, future work would be worthwhile to develop the new primal-dual algorithms for the two-way scenario by following the NUMtw model to take into account the ACK packet flow. Guided by the NUMtw model and theory, new congestion control approaches and protocols would possibly be developed for a general network with data packet flow and ACK packet flow.

5. ACKNOWLEDGMENT

The work described in this paper was supported by grants from the National Natural Science Foundation of China (No. 61070197, No. 61370107) and grants from the self-determined research funds (CCNU13F009, CCNU13A05013).

6. REFERENCES

[1] M. Avriel. *Nonlinear Programming: Analysis and Methods*. Dover Publishing, 2003.

[2] H. Balakrishnan and V. Padmanabhan. How network asymmetry affects TCP. *IEEE Communications Magazine*, 39(4):60–67, 2001.

[3] N. Brownlee, C. Mills, and G. Ruth. RFC 2722 - Traffic Flow Measurement: Architecture, October 1999.

[4] M. Chiang, S. Low, A. Calderbank, and J. Doyle. Layering as optimization decomposition: A mathematical theory of network architectures. *Proceedings of the IEEE*, 95(1):255–312, 2007.

[5] S. Floyd, A. Arcia, D. Ros, and J. Iyengar. RFC 5690 - adding acknowledgement congestion control to TCP, February 2010.

[6] F. Ge, S. Chan, L. L. H. Andrew, F. Li, L. Tan, and M. Zukerman. Performance effects of two-way FAST TCP. *Computer Networks*, 55:2976–2984, 2011.

[7] L. Kalampoukas, A. Varma, and K. K. Ramakrishnan. Two-way TCP traffic over rate controlled channels: effects and analysis. *IEEE/ACM Transactions on Networking*, 6(6):729–743, 1998.

[8] F. P. Kelly. Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, 8(1):33–37, 1997.

[9] 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(9):237–252, 03/14 1998.

[10] F. P. Kelly and T. Voice. Stability of end-to-end algorithms for joint routing and rate control. *Computer Communication Review*, 35(2):5–12, 2005.

[11] S. H. Low. A duality model of TCP and queue management algorithms. *IEEE/ACM Transactions on Networking*, 11(4):525–536, 2003.

[12] S. H. Low and D. E. Lapsley. Optimization flow control I: basic algorithm and convergence. *IEEE/ACM Transactions on Networking*, 7:861–874, 1999.

[13] D. Palomar and M. Chiang. A tutorial on decomposition methods for network utility maximization. *IEEE Journal on Selected Areas in Communications*, 24(8):1439–1451, 2006.

[14] L. Tan, P. Yang, W. Zhang, and F. Ge. On utility-optimised router-level bandwidth allocation. *Transactions on Emerging Telecommunications Technologies*, 24(3):303–316, 2013.

[15] A. Tang, J. Wang, S. H. Low, and M. Chiang. Equilibrium of heterogeneous congestion control: existence and uniqueness. *IEEE/ACM Transactions on Networking*, 15:824–837, 2007.

[16] A. Tang, X. Wei, S. H. Low, and M. Chiang. Equilibrium of heterogeneous congestion control: optimality and stability. *IEEE/ACM Transactions on Networking*, 18:844–857, 2010.

[17] D. X. Wei, C. Jin, S. H. Low, and S. Hegde. FAST TCP: Motivation, architecture, algorithms, performance. *IEEE/ACM Transactions on Networking*, 14(6):1246–1259, 2006.