Automatica 109 (2019) 108506

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

A fully distributed traffic allocation algorithm for nonconcave utility maximization in connectionless communication networks ☆

Jingyao Wang [a,b,*], Mahmoud Ashour [c], Constantino M. Lagoa [c], Necdet Serhat Aybat [d], Hao Che [e]

[a] The Department of Automation, The Xiamen University, China
[b] The Shenzhen Research Institute of Xiamen University, China
[c] The Department of Electrical Engineering and Computer Science, The Pennsylvania State University, United States of America
[d] The Department of Industrial Engineering, The Pennsylvania State University, United States of America
[e] The Department of Computer Science and Engineering, The University of Texas at Arlington, United States of America

ARTICLE INFO

Article history:
Received 1 September 2017
Received in revised form 21 February 2019
Accepted 8 July 2019
Available online 1 August 2019

Keywords:
Traffic engineering
Network utility maximization problem
Nonconcave utility function
Fully distributed traffic allocation algorithm
Connectionless communication networks

ABSTRACT

As IP video services have emerged to be the predominant Internet application, how to optimize the Internet resource allocation, while satisfying the quality of experience (QoE) for users of video services and other Internet applications becomes a challenge. This is because the QoE perceived by a user of video services can be characterized by a staircase function of the data rate, which is nonconcave and hence it is "hard" to find the optimal operating point. The work in this paper aims at tackling this challenge. It considers the packet routing problem among multiple end points in packet switching networks based on a connectionless, hop-by-hop forwarding paradigm. We model this traffic allocation problem using a fluid flow model and let the link bandwidth be the only resource to be shared. To maximize the utilization of resources and avoid congestion, we formulate the problem as a network utility maximization problem. More precisely, the objective of this paper is to design a Fully Distributed Traffic Allocation Algorithm (FDTAA) that is applicable to a large class of nonconcave utility functions. Moreover, FDTAA runs in a fully distributed way: it enables each router to independently address and route each data unit using immediate local information in parallel, without referring to any global information of the communication network. FDTAA requires minimum computation workload, since the routing decision made at each router is solely based on the destination information carried in each unit. In addition, the network utility values corresponding to the FDTAA iterate sequence converge to the optimal network utility value at the rate of $\mathcal{O}(1/K)$, where K is the iteration counter. These theoretical results are exemplified by the simulation performed on an example communication network.

1. Introduction

The growing popularity of IP video services in recent years puts an ever increasing strain on the network resources at global scale. For example, according to Cisco Inc. annual Visual Networking Index (VNI) forecast report, dated June, 1st, 2016 (Cisco VNI Forecast and Methodology, 2015-2020), IP video traffic will be 82% of all consumer Internet traffic by 2020, up from 70% in 2015. Moreover, the report also forecasts that the IP video traffic will grow threefold from 2015 to 2020. Consequently, the success of the future Internet largely hinges upon its ability to cope with the ever growing resource demands for video services, while being able to provide desired Quality-of-Experience (QoE) to users of video and other Internet services. However, the existing Internet transport-layer flow control solutions (e.g., the window-based transport control protocol (TCP) congestion control mechanism (Allman, Paxson, & Blanton, 2009)) and the internetworking layer in-network traffic engineering (TE) solutions (e.g., the flow-based equal-cost-multipath (ECMP) load balancing mechanism (Hopps, 2000)) cannot provide adequate support of such capabilities for the following reasons.

First, for video services, the user QoE, or equivalently, the user utility, as a function of the flow rate, generally exhibits staircase behaviors. For example, the most popular video streaming is the HTTP-based adaptive video streaming, widely known as, dynamic

☆ The material in this paper was partially presented at the 2017 American Control Conference, May 24–26, 2017, Seattle, WA, USA. This paper was recommended for publication in revised form by Associate Editor Giancarlo Ferrari-Trecate under the direction of Editor Ian R. Petersen.

* Corresponding author at: The Department of Automation, The Xiamen University, China.

E-mail addresses: wangjingyao1@xmu.edu.cn (J. Wang), m.ashour@psu.edu (M. Ashour), lagoa@engr.psu.edu (C.M. Lagoa), nsa10@psu.edu (N.S. Aybat), hche@cse.uta.edu (H. Che).

https://doi.org/10.1016/j.automatica.2019.108506

adaptive streaming over HTTP (DASH) (Sodagar, 2011). The video content for DASH may be encoded at different encoding rates or with different numbers of encoded layers (Sanchez et al., 2012), resulting in a finite set of possible video representations (Juluri, Tamarapalli, & Medhi, 2016). The higher the encoding rate or the more encoded layers, the better the user perceived QoE or the higher the user utility will be; hence, exhibiting staircase behaviors. Such utilities are nonconcave and hence are notoriously difficult to optimize. Unfortunately, the transport-layer protocol that underlies HTTP is TCP whose flow control mechanism can only provide optimal and fair resource sharing among flows with a special concave user utility (Ye, Wang, Che, & Lagoa, 2011). In fact, all the existing distributed optimization-and-utility-based solutions (see Section 2 for details) can only deal with concave user utilities. In other words, the existing Internet infrastructure does not provide the much needed flow adaptation services for video streaming applications in particular, and any other nonconcave-utility-based services in general.

Second, the existing in-network, IP layer load balancing mechanisms, such as ECMP and its variations, are designed, independent of the transport-layer TCP congestion control mechanism. As a result, such load balancing mechanisms running at the IP layer may adversely interact with its upper-layer TCP. In fact, it has been widely recognized (Acemoglu, Johari, & Ozdaglar, 2007; Alizadeh et al., 2014b; Chen, Chen, Bai, & Alizadeh, 2016; Kelly & Voice, 2005; Liu, Zhang, Gong, & Towsley, 2005; Qiu, Yang, Zhang, & Shenker, 2003) that flow control at different layers may adversely interact with one another, leading to sub-optimal, unexpected performance, or even network instability.

Therefore, for the future Internet, it is imperative to develop fully distributed, user-utility-aware, optimization-based, integrated end-to-end and in-network traffic control mechanisms for applications with both concave and nonconcave user utilities. This paper develops the much needed theoretical underpinning upon which such traffic control mechanisms can be developed.

The results presented in this paper enable each router to distribute the incoming traffic among the next set of given hops with the class of service (CoS) requirements satisfied and in an optimal way. This is in line with the connectionless, hop-by-hop forwarding paradigm of the Internet, where the packet forward decision is made on a hop-by-hop basis. Hereafter, a network that uses this packet forwarding paradigm is called a connectionless network. There are two main contributions. First, we overcome the challenges opposed by the nonconvexity of the optimization problem by proposing a convex relaxation whose optimal solution approximates that of the original nonconvex problem. The convex relaxation provided relies on the results from the moment approach to polynomial optimization and measure theory. Second, we handle the difficulties rendered by the lack of global information by developing an iterative decentralized algorithm that distributes the required computation over the nodes (routers) in the network, where each node solely uses its immediate local information to route the incoming traffic. The algorithm is proven to converge to the optimal network utility at a rate $\mathcal{O}(1/K)$, where K is the iteration counter. Moreover, the algorithm is robust against link failures, which means that it quickly responds to link failures and reroutes the traffic to its new optimal point.

2. Related work

First, we note that the end-to-end window-based TCP congestion control mechanism (Allman et al., 2009) is designed empirically. Nevertheless, it is interesting to note that one of our earlier results (Ye et al., 2011) showed that the TCP congestion control effectively achieves the network utility maximization (NUM) design objective with a special concave flow utility for each TCP flow. We also note that the flow-based ECMP is a fully distributed, non-optimization-based solution. It simply distributes flows equally among multiple next-hop nodes, leading to the same destination. Although some recent solutions enable optimization-based dynamic load balancing (Alizadeh et al., 2014a; Wang & Xu, 2014), they are no longer fully distributed by design and require the knowledge of global information. In what follows, we limit ourselves to the literature on fully distributed, optimization-based solutions only, which are most relevant to the current work.

The general approach taken to address the optimization-based, fully distributed traffic control problem is to adopt a fluid-flow model based formulation taking into account link bandwidth constraints, which leads to a nonlinear programming problem. The aim is to find fully distributed traffic control laws with the property that, by working independently of one another, the nodes, e.g., routers, will drive the network to an operation point where a given utility function of flow rates is maximized for the whole network. One approach (e.g., Elwalid, Jin, Low, and Widjaja (2001), Golestani and Bhattacharyya (1998)) is to incorporate link congestion costs into the overall utility function to convert the constrained problem into an unconstrained problem. The resulting optimization problem can then be solved using a gradient-based first-order algorithm. It is generally used for a network where each path is pinned, such as a label switched path (Elwalid et al., 2001). A second approach (e.g., Han, Shakkottai, Hollot, Srikant, and Towsley (2004), Kelly, Maulloo, and Tan (1998), Kelly and Voice (2005)) first converts the original constrained problem into an unconstrained problem, i.e., the Lagrange dual problem, by introducing dual variables. Then the second approach solves a relaxation of the unconstrained problem after closely approximating it through incorporating a price function into the overall utility function, which makes the second approach different from the first one. Fully distributed control laws were found, which were proven to be locally stable in the presence of variable feedback delays. Working independently at a source, a control law adjusts/balances its flow sending rates to multiple paths based on periodic, cumulative price feedbacks from the destination node. Each component price is collected from the intermediate links along the forwarding path.

A third approach is to solve the original problem directly (e.g., Lagoa and Che (2000), Lagoa, Che, and Movsichoff (2004), Movsichoff, Lagoa, and Che (2005, 2007), Su, Liu, Lagoa, Che, Xu, and Cui (2015), Wang, Palaniswami, and Low (2003)). Using a dual model, an algorithm was provided in Wang et al. (2003). In Lagoa and Che (2000), Lagoa et al. (2004), Movsichoff et al. (2005, 2007), this problem was tackled by using a technique based on the Theory of Sliding Modes. Both end-to-end (Lagoa & Che, 2000; Lagoa et al., 2004; Movsichoff et al., 2007) and hop-by-hop (Movsichoff et al., 2005) optimal control laws were proposed. Moreover, fully distributed multi-domain optimal control laws were also developed in Su et al. (2015). These algorithms allow multipath forwarding and enable multiple CoS, and require minimum information feedback for control.

However, all of the above solutions assume that the user utility functions are concave. There are limited results in the literature on developing algorithms that maximize nonconcave utility functions. Fazel and Chiang (2005) proposed a centralized algorithm for solving a NUM problem with polynomial utilities. Hande, Zhang, and Chiang (2007) considered a special class of nonconcave utility functions, for which the duality gap between primal and dual problems is zero, and solved the optimization problem with a fully distributed sub-gradient algorithm.

In a related work in Ashour, Wang, Lagoa, Aybat, and Che (2017), the authors of this paper developed a fully distributed

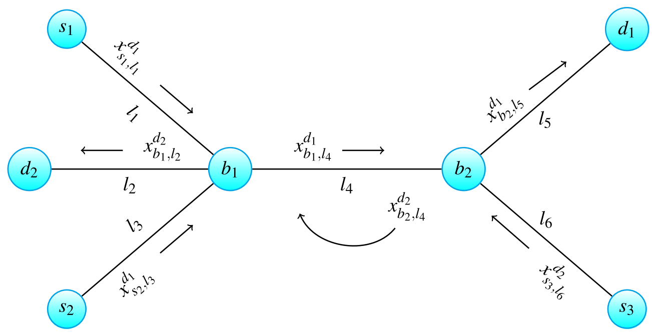


Fig. 1. Notation example.

traffic allocation algorithm that allows end-to-end data adaption and load balancing with minimum information feedback from the network. While the work in Ashour et al. (2017) is useful for providing sophisticated service quality features at the transport or higher layers, the solution in this paper is particularly powerful as it allows similar sophisticated service quality features to be adopted at the inter-networking layer in a connectionless network where routers use hop-by-hop forwarding. A preliminary version of this work has appeared in the 2017 American Control Conference (Wang, Ashour, Lagoa, Aybat, Che, & Duan, 2017); when compared to Wang et al. (2017), the algorithm proposed in this manuscript requires less computation workload as it enables each node to update data rates by per destination. This modification performs better in terms of the computation workload. Moreover, complete proofs of all the results, which were omitted due to space restrictions in Wang et al. (2017), are now provided in this paper.

Finally, we note that there are some interesting application-layer flow control solutions being developed for specific applications, e.g., DASH video streaming (Cofano, De Cicco, Zinner, Nguyen-Ngoc, Tran-Gia, & Mascolo, 2016; Yin, Bartulović, Sekar, & Sinopoli, 2017; Yin, Jindal, Sekar, & Sinopoli, 2015), which may also have a target deployment scenario in mind, e.g., a software-defined-networking environment (Cofano et al., 2016). The current work differs from such solutions in that it is not meant to be used for a specific application, but rather serve as the needed theoretical underpinning for the development of integrated transport-and-IP-layer protocols for the future Internet. To this end, the solution must be distributed by design and in line with the hop-by-hop forwarding paradigm of the Internet to be deployable at global scale.

3. Notation

Let $\mathcal{N}$ denote the set of nodes, and $\mathcal{L} \subset \mathcal{N} \times \mathcal{N}$ denote the set of links connecting particular pairs of nodes. Each link $l \in \mathcal{L}$ has finite capacity $c_l > 0$. Moreover, let $\mathcal{S} \triangleq \{s_1, s_2, \dots, s_n\}$ and $\mathcal{D} \triangleq \{d_1, d_2, \dots, d_{\tilde{n}}\}$ be respectively the set of source and destination nodes contained in $\mathcal{N}$ such that $\mathcal{S} \cap \mathcal{D} = \emptyset$. The set of links connected to source $s \in \mathcal{S}$ is denoted by $\mathcal{L}_s$, whereas the set of intended destinations for flows sent by $s \in \mathcal{S}$ is denoted by $\mathcal{D}_s$. Moreover, let $\mathcal{L}_s^d \subset \mathcal{L}_s$ denote the set of links used to transport flows sent by source $s \in \mathcal{S}$ with an intended destination $d \in \mathcal{D}_s$. Also let $\mathcal{D}_{s,l} \subset \mathcal{D}_s$ be the set of destinations for flows sent by source $s \in \mathcal{S}$ through link $l \in \mathcal{L}_s$. The data rate of a flow sent by source $s \in \mathcal{S}$ with an intended destination $d \in \mathcal{D}_s$ through link $l \in \mathcal{L}_s^d$ is denoted by $x_{s,l}^d$. Furthermore, let $\mathbf{x}_{s,l} = [x_{s,l}^d]_{d \in \mathcal{D}_{s,l}}$, and $\mathbf{x}_s = [\mathbf{x}_{s,l}]_{l \in \mathcal{L}_s}$. Thus, the aggregate data rate of source $s \in \mathcal{S}$ is $r_s = \mathbf{1}^\top \mathbf{x}_s$, where $\mathbf{1}$ denotes a column vector of all ones with appropriate dimension.

Let $\mathcal{B} \triangleq \mathcal{N} \setminus (\mathcal{S} \cup \mathcal{D}) = \{b_1, b_2, \dots, b_m\}$ denote the set of forwarding nodes. Given $b \in \mathcal{B}$, $\mathcal{L}_b$ denotes the set of links directly connected to it and $\mathcal{D}_b$ denotes the set of destinations for flows visiting node b. Let $\mathcal{L}_b^{d,\text{out}} \subseteq \mathcal{L}_b$ be the set of links for outgoing flows from node $b \in \mathcal{B}$ with intended destination $d \in \mathcal{D}_b$. Similarly, let $\mathcal{L}_b^{d,\text{in}} \subseteq \mathcal{L}_b$ be the set of links for flows incoming to node $b \in \mathcal{B}$ with intended destination $d \in \mathcal{D}_b$. The set of destinations for flows departing node $b \in \mathcal{B}$ through link $l \in \mathcal{L}_b$ is denoted by $\mathcal{D}_{b,l}^{\text{out}} \subset \mathcal{D}_b$, and $\mathcal{D}_{b,l}^{\text{in}} \subset \mathcal{D}_b$ denotes the set of destinations for flows received at node $b \in \mathcal{B}$ through link $l \in \mathcal{L}_b$. The data rate of a flow sent by node $b \in \mathcal{B}$ with an intended destination $d \in \mathcal{D}_b$ through link $l \in \mathcal{L}_b^{d,\text{out}}$ is denoted by $x_{b,l}^d$. Also let $\mathbf{x}_{b,l} = [x_{b,l}^d]_{d \in \mathcal{D}_{b,l}^{\text{out}}}$, and $\mathbf{x}_b = [\mathbf{x}_{b,l}]_{l \in \mathcal{L}_b}$. Furthermore, let $e_l(b)$ denote the node connected to node $b \in \mathcal{B}$ via link $l \in \mathcal{L}_b$. Finally, let $\mathbf{x} = \left[[\mathbf{x}_s]_{s\in\mathcal{S}}^\top, [\mathbf{x}_b]_{b\in\mathcal{B}}^\top\right]^\top$, and $\mathbf{r} = [r_s]_{s\in\mathcal{S}}$.

For the convenience of the reader, we summarize the notation in Table 1. Moreover, we present an example network in Fig. 1 with labels using the notation above.

Throughout the paper, $\mathbf{I}_n$ is the $n \times n$ identity matrix, and $\|.\|$ denotes the Euclidean norm. The indicator function of $\mathcal{A}$ is denoted by $\mathbb{1}_{\mathcal{A}}$, i.e., $\mathbb{1}_{\mathcal{A}}(\omega) = 0$ for $\omega \in \mathcal{A}$ and is $+\infty$ otherwise. Let $\Pi_{\mathcal{A}}(\omega) = \text{argmin}\{\|v - \omega\| : v \in \mathcal{A}\}$ be the projection of ω onto $\mathcal{A}$. Let $x^+ = \max(x, 0)$. Given a sequence $\mathbf{t} = \{t_j\}_{j=k}^{k+2h}$ for some nonnegative integer k and positive integer h, $\mathbf{H}(k, k+2h, \mathbf{t}) \in \mathbb{R}^{h+1} \times \mathbb{R}^{h+1}$ is a Hankel matrix of the form

$$\mathbf{H}(k, k+2h, \mathbf{t}) = \begin{bmatrix} t_k & t_{k+1} & \cdots & t_{k+h} \\ t_{k+1} & \iddots & \iddots & t_{k+h+1} \\ \vdots & \iddots & \iddots & \vdots \\ t_{k+h} & \cdots & \cdots & t_{k+2h} \end{bmatrix}. \tag{1}$$

4. Problem statement

Consider a connectionless network consisting of multiple combinations of source/destination pairs. Each source sends data to at least one destination, and the same destination receives data from different sources simultaneously. Traffic flows are modeled using a fluid flow model, and the links' bandwidth is the only shared resource. Next, we list our assumptions on the problem along with their justification.

Assumption 1. The traffic flows for source/destination pairs belong to the same class, i.e., they have the same content and priority.

It is worth emphasizing that although this assumption seems to be restrictive, in fact it is very general. Indeed, we can easily extend the results to multiple-class case: Given flows belonging to different classes with the same destination, we can split the destination into "dummy" destination nodes so that each class will have a unique destination node and then we can perform the proposed control law. In the sequel, we restrict our attention to the single-class case and allow the intermediate nodes to aggregate flows that have the same destination.

Assumption 2. Source nodes are pairwise nonadjacent, i.e., $\mathcal{L}_{s_i} \cap \mathcal{L}_{s_j} = \emptyset$ for $i \neq j$.

In other words, there exists no pair of source nodes that are neighbors. However, this is not assuming that the data of different sources do not share links. Indeed, it is imperative for different source/destination flows to share links. Evidently, our model captures this fact. For instance, in Fig. 1, link l_4 simultaneously carries flows with source/destination pairs (s_1, d_1), (s_2, d_1), and (s_3, d_2).

Assumption 3. Each source $s \in \mathcal{S}$ is aware of the capacities of links $\mathcal{L}_s$. Moreover, each node $b \in \mathcal{B}$ knows the capacities of the links $\mathcal{L}_d^{d,\text{out}}$ for all $d \in \mathcal{D}_b$.

Table 1
List of notation.

Notation	Description
$\mathcal{N}$	Set of nodes in the network.
$\mathcal{D}$	Set of destination nodes.
$\mathcal{S}$	Set of source nodes.
$\mathcal{B}$	Set of intermediate nodes.
$\mathcal{L}$	Set of links in the network.
$e_l(b)$	Node connected to node b through link l.
$\mathcal{L}_s$	Set of links connected to source s.
$\mathcal{L}_b$	Set of links connected to node b.
$\mathcal{D}_s$	Set of destinations for flows sent by source s.
$\mathcal{D}_b$	Set of destinations for flows visiting node b.
$\mathcal{L}_s^d$	Set of links for flows sent by source s with destination d.
$\mathcal{L}_b^{d,\text{out}}$ ($\mathcal{L}_b^{d,\text{in}}$)	Set of links for flows sent (received) by node b with destination d.
$\mathcal{D}_{s,l}$	Set of destinations for flows departing source s using link l.
$\mathcal{D}_{b,l}^{\text{out}}$ ($\mathcal{D}_{b,l}^{\text{in}}$)	Set of destinations for flows sent (received) by node b through link l.
$x_{s,l}^d$	Data rate of flow sent by source s through link l with destination d.
$x_{b,l}^d$	Data rate of flow sent by node b through link l with destination d.
r_s	Aggregate data rate of source s.

Assumption 4. Each source $s \in \mathcal{S}$ has a monotonically nondecreasing possibly nonconcave utility function of its aggregate data rate, $U_s(r_s)$.

The local utility function $U_s(r_s)$ for source $s \in \mathcal{S}$ considered herein belongs to a class of nonconcave polynomial-like functions of the form

$$U_s(r_s) = \sum_{j=0}^{\alpha} p_{s,j} r_s^{j/\ell}, \tag{2}$$

where α and ℓ are some given positive integers, and we consider the case that $\alpha \le \ell$. The main motivating reason for using this large class of utility functions is that the particular form given in (2) is so flexible that it can be used to approximate a wide variety of utility functions arising in real-world applications such as monotonic piecewise affine functions in the case of video streaming (Juluri et al., 2016). Available efficient approximation techniques range from the simple linear regression to more versatile approximation schemes, e.g., Sum of Squares (SoS) (Henrion, Lasserre, & Savorgnan, 2009) and Chebyshev Approximation (Press, Teykolsky, Vetterling, & Flannery, 2007). Fig. 2 shows an example of approximating a monotonic piecewise affine utility function with a polynomial-like function using SoS. As a result of the approximation, the optimum of the approximation function may not be the optimum of the original utility function, especially non-smooth and nonconcave functions, which renders conservativeness to the results in this paper. However, choosing large enough parameters α and ℓ is an effective way to reduce the conservativeness, since the larger both the parameters, the better the approximation result (see also Proposition 1). Of course, available techniques to reduce the conservativeness include but not limited to the way mentioned above. Hence, this type of utility functions can fit various inelastic applications.

We design a traffic allocation algorithm to maximize nonconcave utility functions of the connectionless network, i.e., the sum of individual polynomial-like utility functions, while ensuring that the network resource constraints are satisfied, which include link capacity constraints, Minimum Rate Guaranteed and Upper Bounded Rate Service (MRGUBRS) requirements, and flow conservation through nodes.

Referring to the definitions of the above constraints given in Movsichoff et al. (2007) and the list of notation presented in Table 1, we can formulate the problem of optimal traffic allocation as follows:

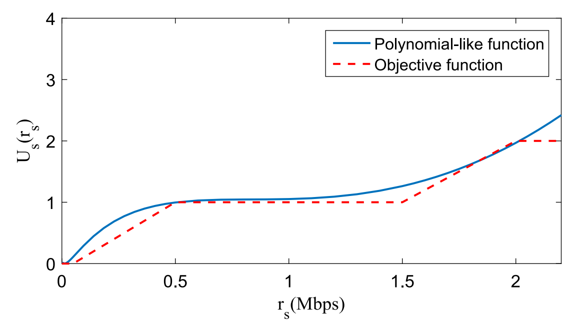


Fig. 2. Utility function approximation for $\alpha = \ell = 20$.

$$\begin{aligned}
\max_{\mathbf{x},\mathbf{r}}.\quad & \sum_{s\in\mathcal{S}}\sum_{j=0}^{\alpha} p_{s,j} r_s^{j/\ell} \\
\text{s.t.}\quad & \mathbf{1}^\top \mathbf{x}_{s,l} \le c_l, \quad l \in \mathcal{L}_s,\ s \in \mathcal{S} \\
& \mathbf{1}^\top \mathbf{x}_{b,l} + \mathbf{1}^\top \mathbf{x}_{e_l(b),l} \le c_l, \quad l \in \mathcal{L}_b,\ b \in \mathcal{B} \\
& \mathbf{B}\mathbf{x} = 0 \\
& x_{b,l}^d \ge 0, \quad l \in \mathcal{L}_b^{d,\text{out}},\ d \in \mathcal{D}_b,\ b \in \mathcal{B} \\
& (\mathbf{x}_s,\ r_s) \in \mathcal{X}_s,\ s \in \mathcal{S},
\end{aligned} \tag{3}$$

where $\mathbf{B}\mathbf{x} = 0$ encodes the flow conservation constraints $\sum_{l\in\mathcal{L}_b^{d,\text{out}}} x_{b,l}^d = \sum_{l\in\mathcal{L}_b^{d,\text{in}}} x_{e_l(b),l}^d$ enforced for all $d \in \mathcal{D}_b$ at each node $b \in \mathcal{B}$. The matrix $\mathbf{B}$ is constructed in a similar fashion to the edge-node incidence matrix of a graph. More precisely, each row of $\mathbf{B}$ is determined by a destination–intermediate node pair (b, d) for $b \in \mathcal{B}$ and $d \in \mathcal{D}_b$, while each column is determined by a node–link–destination triple $(n, l, \tilde{d})$ for $n \in \mathcal{S} \cup \mathcal{B}$, $l \in \mathcal{L}_n$, and $\tilde{d} \in \mathcal{D}_{n,l}$ if $n \in \mathcal{S}$ or $\tilde{d} \in \mathcal{D}_{n,l}^{\text{out}}$ if $n \in \mathcal{B}$. The entry $\mathbf{B}_{(b,d),(n,l,\tilde{d})}$ equals 1 if $d = \tilde{d}$, $n \ne b$, and $x_{n,l}^d$ is received at node b, it equals -1 if $d = \tilde{d}$, $n = b$, and $x_{n,l}^d$ is transmitted by node b, and is 0 otherwise. The set $\mathcal{X}_s$ is defined as:

$$\mathcal{X}_s \triangleq \{(\mathbf{x}_s, r_s) \in \mathbb{R}_+^{\sum_{l\in\mathcal{L}_s}|\mathcal{D}_{s,l}|} \times \mathbb{R}_+ : \xi_s \le r_s \le \zeta_s, r_s = \mathbf{1}^\top \mathbf{x}_s\}.$$

The resulting NUM problem (3) is nonconvex. Nonconvex problems are hard to solve even with centralized algorithms. Moreover, it is not reasonable to assume the availability of global network information at any node when the scale of the network is large. This stimulates the necessity of developing a distributed algorithm that leverages local information at each node.

5. Main results

This section presents the proposed method for overcoming the two main challenges associated with solving (3), namely, the nonconvexity of the problem and the lack of global network information.

5.1. Convex relaxation

The following semidefinite program provides a computationally tractable convex relaxation to (3) inspired by the moment approach to polynomial optimization (Lasserre, 2001).

$$\begin{aligned}
\max_{(\mathbf{m},\mathbf{x},\mathbf{r})} \quad & \sum_{s\in\mathcal{S}} \mathbf{p}_s^\top \mathbf{m}_s \\
\text{subject to} \quad & m_{s,0} = 1, \quad s\in\mathcal{S}, \\
& \mathbf{M}_s \succeq 0,\ \bar{\mathbf{M}}_s \succeq 0, \quad s\in\mathcal{S}, \\
& m_{s,j} \le r_s^{j/\ell}, \quad j\in\{1,\dots,\alpha\},\ s\in\mathcal{S}, \\
& \mathbf{1}^\top \mathbf{x}_{s,l} \le c_l, \quad l\in\mathcal{L}_s,\ s\in\mathcal{S}, \\
& \mathbf{1}^\top \mathbf{x}_{b,l} + \mathbf{1}^\top \mathbf{x}_{e_l(b),l} \le c_l, \quad l\in\mathcal{L}_b,\ b\in\mathcal{B}, \\
& \mathbf{B}\mathbf{x} = 0 \\
& \mathbf{x}_{b,l} \succeq 0, \quad l\in\mathcal{L}_b,\ b\in\mathcal{B}, \\
& (\mathbf{x}_s,\ r_s) \in \mathcal{X}_s, \quad s\in\mathcal{S}.
\end{aligned} \tag{4}$$

The objective function is the sum of local linear functions of moment variables $\mathbf{m}_s = [m_{s,j}]_{j\in\{0,\dots,\alpha\}}$, where $\mathbf{p}_s = [p_{s,j}]_{j\in\{0,\dots,\alpha\}}$. For each $s\in\mathcal{S}$, the matrices $\mathbf{M}_s$ and $\bar{\mathbf{M}}_s$ are constructed using $\mathbf{H}$ in (1) as follows.

- If α is even, then

$$\mathbf{M}_s = \mathbf{H}(0,\alpha,\mathbf{m}_s) \tag{5}$$

$$\bar{\mathbf{M}}_s = \zeta_s^2 \mathbf{H}(0,\alpha-2,\mathbf{m}_s) - \mathbf{H}(2,\alpha,\mathbf{m}_s). \tag{6}$$

- However, if α is odd, then

$$\mathbf{M}_s = \zeta_s \mathbf{H}(0,\alpha-1,\mathbf{m}_s) - \mathbf{H}(1,\alpha,\mathbf{m}_s) \tag{7}$$

$$\bar{\mathbf{M}}_s = \mathbf{H}(1,\alpha,\mathbf{m}_s) + \zeta_s \mathbf{H}(0,\alpha-1,\mathbf{m}_s). \tag{8}$$

Proposition 1. *Given $\ell,\alpha\in\mathbb{Z}_+$, let $(\mathbf{m}^*(\alpha),\mathbf{x}^*(\alpha),\mathbf{r}^*(\alpha))$ be an optimal solution to (4). Then, $(\mathbf{x}^*(\alpha),\mathbf{r}^*(\alpha))$ converges to $(\mathbf{x}^*,\mathbf{r}^*)$ as $\alpha\to\infty$, where $(\mathbf{x}^*,\mathbf{r}^*)$ is an optimal solution to the nonconvex NUM problem in (3) with user utility functions as in (2).*

Proof. See Appendix A.

According to Proposition 1, given the model parameter ℓ, through letting $\alpha \le \ell$, we obtain a convex relaxation of the problem (3), which can be solved efficiently. Hereafter, we set α equal to ℓ.

Problem (4) maximizes the sum of linear functions subject to convex constraints, i.e., (4) is a convex problem. Hence, it can be readily solved by centralized algorithms if global network information is available. However, aggregating global information at a centralized coordinator in a large-scale network renders the network vulnerable to attacks, and is prohibitively expensive from communications and computations perspectives.

5.2. FDTAA

The constraint set of problem (4) consists of not only local constraints, e.g., the MRGUBRS, but also some global constraints, e.g., the flow conservation constraints through nodes. Without assuming global information available, the existence of global constraints renders problem (4) difficult to solve in a fully distributed manner. However, the primal–dual method proposed in Aybat and Yazdandoost Hamedani (2016) provides a way for solving problem (4) using local information at each node, communication among immediate neighbors only. Based on the primal–dual algorithm in Aybat and Yazdandoost Hamedani (2016), we develop a fully distributed method to solve (4), FDTAA summarized in Algorithm 1. The derivation of FDTAA is presented in Appendix B.

Algorithm 1: FDTAA$\{\tau_s\}_{s\in\mathcal{S}},\ \{\tau_{b,l}^d\}_{d\in\mathcal{D}_{b,l},l\in\mathcal{L}_b,b\in\mathcal{B}},\ \{\kappa_{b,l}\}_{l\in\mathcal{L}_b,b\in\mathcal{B}},\gamma$

1: Initialization: $\mathbf{m}^0,\mathbf{x}^0,\mathbf{r}^0,\lambda^0,\mathbf{z}_s^0 \leftarrow \mathbf{x}_s^0,\ s\in\mathcal{S}$, $\mathbf{z}_{b,l}^0 \leftarrow \mathbf{x}_{b,l}^0, l\in\mathcal{L}_b,\ b\in\mathcal{B}$.
2: **for** $k\ge 0$ **do**
3: $(\mathbf{m}_s,\mathbf{x}_s,r_s)^{k+1} \leftarrow \Pi_{\mathcal{A}_s}(\mathbf{m}_s^k + \tau_s\mathbf{p}_s, \mathbf{x}_s^k - \gamma\tau_s\mathbf{u}_s^k, r_s^k), s\in\mathcal{S}$
4: $\mathbf{x}_{b,l}^{k+1} \leftarrow \left(\mathbf{x}_{b,l}^k - \lambda_{b,l}^k[\tau_{b,l}^d + \gamma\tau_{b,l}^d(u_{b,l}^d)^k]_{d\in\mathcal{D}_{b,l}^{\text{out}}}\right)^+, l\in\mathcal{L}_b, b\in\mathcal{B}$
5: Each node $s\in\mathcal{S}$ communicates $\mathbf{x}_s^{k+1}$ to $\mathcal{L}_s$.
6: Each node $b\in\mathcal{B}$ communicates $\mathbf{x}_{b,l}^{k+1}$ to $l\in\mathcal{L}_b$.
7: $\lambda_{b,l}^{k+1} \leftarrow (\lambda_{b,l}^k + \kappa_{b,l}(\mathbf{1}^\top(2\mathbf{x}_{b,l}^{k+1} - \mathbf{x}_{b,l}^k) + \mathbf{1}^\top(2\mathbf{x}_{e_l(b),l}^{k+1} - \mathbf{x}_{e_l(b),l}^k) - c_l))^+, l\in\mathcal{L}_b, b\in\mathcal{B}$
8: $\mathbf{z}_s^{k+1} \leftarrow \mathbf{z}_s^k - \mathbf{x}_s^k + 2\mathbf{x}_s^{k+1}, s\in\mathcal{S}$
9: $\mathbf{z}_{b,l}^{k+1} \leftarrow \mathbf{z}_{b,l}^k - \mathbf{x}_{b,l}^k + 2\mathbf{x}_{b,l}^{k+1}, l\in\mathcal{L}_b, b\in\mathcal{B}$
10: Each node $s\in\mathcal{S}$ communicates $\mathbf{z}_s^{k+1}$ to $\mathcal{L}_s$
11: Each node $b\in\mathcal{B}$ communicates $\mathbf{z}_{b,l}^{k+1}$ to $l\in\mathcal{L}_b$

For each $s\in\mathcal{S}$, let the set of local constraints $\mathcal{A}_s$ be defined as

$$\begin{aligned}
\mathcal{A}_s = \{ & (\mathbf{m}_s,\mathbf{x}_s,r_s) \in \mathbf{R}^{\ell+1} \times \mathbf{R}_+^{\sum_{l\in\mathcal{L}_s}|\mathcal{D}_{s,l}|} \times \mathbf{R}_+ : m_{s,0} = 1, \\
& \mathbf{M}_s \succeq 0,\ \bar{\mathbf{M}}_s \succeq 0,\ m_{s,j} \le r_s^{j/\ell},\ j\in\{1,\dots,\ell\}, \\
& \mathbf{1}^\top \mathbf{x}_{s,l} \le c_l,\ l\in\mathcal{L}_s,\ (\mathbf{x}_s,\ r_s)\in\mathcal{X}_s\}.
\end{aligned} \tag{9}$$

Note that $\mathcal{A}_s$ is fully characterized by the local information available to node s, e.g., the lower and upper bounds imposed on its own data rates.

In FDTAA, for any given stepsize parameter $\gamma > 0$, the local stepsize $\tau_s > 0$ is determined at each source node $s\in\mathcal{S}$. Each $s\in\mathcal{S}$ keeps the variables $\mathbf{m}_s$, r_s and the parameters $\mathbf{p}_s$ as private information, and need not share them with any other network entity. In addition, source s introduces auxiliary variables $\mathbf{z}_s$. Introducing these auxiliary variables is one of the key steps to make the traffic allocation algorithm fully distributed. Similarly, at each forwarding node $b\in\mathcal{B}$, the stepsize parameters $[\kappa_{b,l}]_{l\in\mathcal{L}_b} \succeq 0$ and $[\tau_{b,l}^d]_{d\in\mathcal{D}_{b,l},l\in\mathcal{L}_b} \succeq 0$ are determined using local information. Moreover, given $b\in\mathcal{B}$ and $l\in\mathcal{L}_b$, $\lambda_{b,l}$ is a scalar nonnegative variable that denotes the price associated with link l. Each forwarding node $b\in\mathcal{B}$ also introduces auxiliary variables $\mathbf{z}_{b,l}$ for all $l\in\mathcal{L}_b$. Hereafter, the superscript k indicates the iteration index. The k-th iteration of FDTAA consists of the following:

- In step 3, all source nodes update their desired sending data rates in parallel and locally by solving a simple semi-definite program, where the vectors $\mathbf{u}_s$ consist of the relative sending data-rate information among neighboring nodes and are computed as shown in (B.9) for each $s\in\mathcal{S}$. Simultaneously, all forwarding nodes locally and in parallel update their desired sending data rates as shown in step 4, where $u_{b,l}^d$, $d\in\mathcal{D}_{b,l}^{\text{out}}$, is the relative sending data-rate information among neighboring nodes and is computed as shown in (B.9) for each $l\in\mathcal{L}_b$ and $b\in\mathcal{B}$.
- In step 7, for each $b\in\mathcal{B}$, each link $l\in\mathcal{L}_b$ updates its link price $\lambda_{b,l}$. This step can be evaluated at one or both of the nodes the link connects. If the communication workload is taken into consideration, the link price can be updated at both nodes simultaneously. Otherwise, the update can be performed at one of the nodes and then sent to the other.

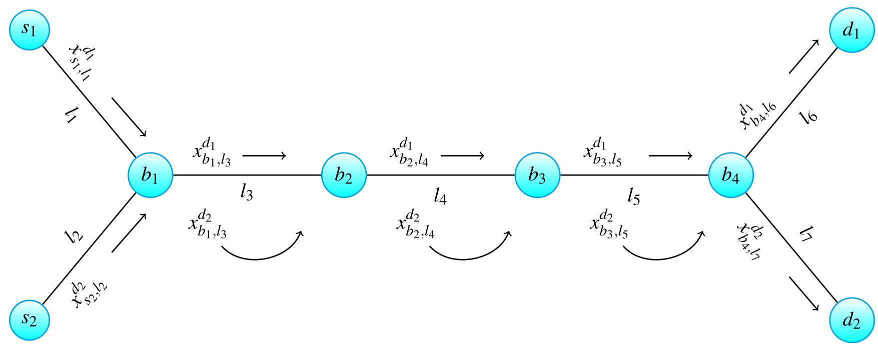


Fig. 3. Routing example.

Overall, FDTAA is a distributed traffic allocation algorithm, where all computations are performed locally at each node, and communication occurs between immediate neighbors only.

At the end of this subsection, it is worth discussing the generality of algorithm FDTAA. This algorithm is proposed to solve the convex relaxation problem of the nonconvex optimization problem (3). Here the convex relaxation problem is actually an approximation of the problem (3) for the following two reasons: (1) when α goes to ∞, the solution of (4) converges to that of (3), since the constraint set of (4) becomes more and more tight; (2) when $\alpha \leq \ell$, the relaxation problem (4) is convex. It follows from the above reasons that the approximation is grounded in well-defined concepts. And FDTAA is general enough to approximately solve problem (3) with high accuracy since it can produce decent data rate allocation results (shown in the simulation section). Moreover, FDTAA is fully distributed, since it just requires immediate local information available to it. To the best of our knowledge that there are no other distributed algorithms that can even approximate the complex problem in this paper.

5.3. Feasibility of the FDTAA iterate sequence

Among the advantages of FDTAA is that if the proposed problem (4) has at least one primal–dual solution, then the iterate sequence converges to some primal–dual solution; moreover, the suboptimality and infeasibility of the ergodic average sequence converges to 0 with $\mathcal{O}(1/K)$ rate, where K is the number of iterations. This conclusion is based on results from Aybat and Yazdandoost Hamedani (2016), which shows that the primal–dual method proposed in Chambolle and Pock (2016) can be used to solve multi-agent consensus optimization problems with local conic constraints over static or time-varying communication networks. Here, in Theorem 1, we discuss the convergence properties of the iterate sequence generated by FDTAA for solving the convex relaxation problem (4).

Theorem 1. *Given the communication network and the convex optimization problem (4). Let $\{\mathbf{m}^k, \mathbf{x}^k, \mathbf{r}^k\}_{k\in\mathbb{N}}$ be the iterate sequence generated by FDTAA, initialized by an arbitrary $(\mathbf{m}^0, \mathbf{x}^0, \mathbf{r}^0)$. Let the primal–dual stepsizes $\{\tau_s\}$, $\{\tau^d_{b,l}\}$, $\{\kappa_{b,l}\}$ and γ be positive constants satisfying the following inequalities*

$$\frac{1}{\tau_s} - \gamma(4+\upsilon_s) > 0, \quad \forall\, s \in \mathcal{S} \tag{10}$$

$$\frac{1}{\kappa_{b,\ l}}\left(\frac{1}{\tau^d_{b,l}} - \gamma(4+\upsilon^d_{b,l})\right) > m_l + 1,\ \forall d \in \mathcal{D}^{\rm out}_{b,l}, l \in \mathcal{L}_b, b \in \mathcal{B}, \tag{11}$$

where m_l is the total number of destinations for flows going through link $l \in \mathcal{L}_b$; $\{\upsilon_s\}$ and $\{\upsilon^d_{b,l}\}$ are positive constants satisfying inequalities (C.1) and (C.2). Then, the iterate sequence $\{\mathbf{m}^k, \mathbf{x}^k, \mathbf{r}^k\}_{k\in\mathbb{N}}$ converges to a maximizer of the utility function of problem (4) subject to the resource constraints.

Proof. See Appendix C.

5.4. Routing as a function of destination and next hop

In FDTAA, each router makes routing decision just with the destination information carried in each data unit. That is to say, if several call types (flows having different pairs of edge nodes or equivalently source/destination nodes) arrive at a forwarding node, the node distinguishes those flows solely according to their destinations, without referring to the specific information of edge node pairs. Consider a communication network with N source nodes and M destination nodes. Suppose that each source node sends data to all destination nodes and thus there exist $N \times M$ types of calls. By using FDTAA, each forwarding node just needs to update M distinct data rate variables, although there are as many as $N \times M$ types of calls received. This obviously reduces the computation burden compared with algorithms that update data rates by call type.

Moreover, the computation burden can be further reduced by routing as a function of destination node and next hop. We illustrate this idea with Fig. 3, which is used to route two types of flows to their respective destination nodes d_1 and d_2. Forwarding node b_1 is aware that the incoming flows have the same next hop b_2 and share the same resource, i.e., the link bandwidth of l_3. Let the next set of hop $\mathcal{H}_{b_1} \triangleq \{b_2\}$ replace the destination set $\mathcal{D}^{\rm out}_{b_1,l_3} \triangleq \{d_1, d_2\}$. Similarly, let per hop data rate variable $x^{b_2}_{b_1,l_3}$ replace $x^{d_1}_{b_1,l_3}$ and $x^{d_2}_{b_1,l_3}$. Then, node b_1 just needs to update the per hop data rate variable $x^{b_2}_{b_1,l_3}$, i.e., per destination information no longer needs to be maintained. Similarly, nodes b_2, b_3 and b_4 can also update their per hop data rate variables. Through this way, we can almost save half of the computation efforts.

6. Simulation results

In this section, we perform simulations to exemplify the behavior of the proposed algorithm FDTAA.

6.1. Simulation setup

We consider the network model shown in Fig. 4, where on each link the corresponding bandwidth is displayed. The network model allows for multiple paths available for flows belonging to each source node. We consider a total of 8 different combinations of source/destination nodes, i.e., s_1/d_3, s_2/d_2, s_3/d_3, s_4/d_2, s_5/d_5, s_6/d_5, s_7/d_7 and s_8/d_3. Moreover, we list the prescribed next hops for all forwarding nodes b_i, $i = 1, \ldots, 8$, in Table 2.

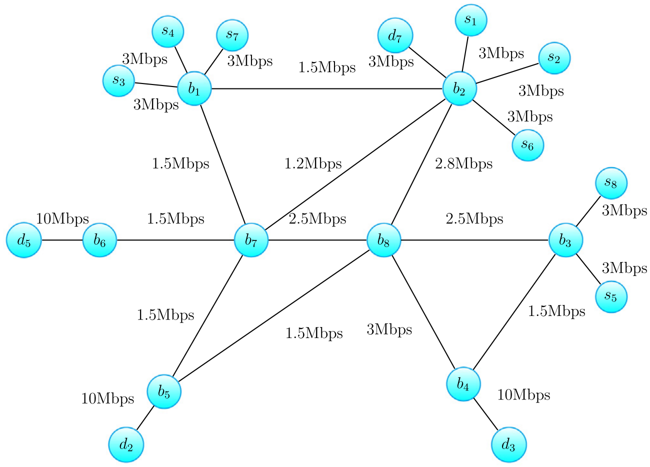


Fig. 4. Topology of the communication network.

Table 2
Routing decisions by destination nodes.

	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8
d_2	b_2, b_7	b_7, b_8	–	–	d_2	–	b_5	b_5, b_7
d_3	b_2, b_7	b_7, b_8	b_4	d_3	–	–	b_8	b_3, b_4
d_5	b_7	b_1, b_7, b_8	b_4, b_8	b_8	b_7	d_5	b_6	b_5, b_7
d_7	b_2, b_7	d_7	–	–	–	–	b_2, b_8	b_2

For example, according to Table 2, node b_1 forwards the data with destination d_2 to nodes b_2 and b_7.

The objective throughout the simulation is to maximize the sum utility of source nodes, where each source has the utility function given by

$$U_{s_i}(r_{s_i}) = \begin{cases} 0, & \text{if } 0 \le r_{s_i} \le 1, \\ 1, & \text{if } 1 \le r_{s_i} \le 2, \quad i = 1, \dots, 8, \\ 2, & \text{if } 2 \le r_{s_i} \le 3. \end{cases} \tag{12}$$

$U_{s_i}(r_{s_i}),\ i = 1, \dots, 8$, are monotonic piecewise affine nonconcave utility functions. The reason we consider monotonic piecewise affine functions is because this class of functions is appropriate for describing the video quality perceived by a user in a video streaming application (Nekouei, Nair, & Alpcan, 2015). Note that it is easy to form the resource constraints with the information provided in Fig. 4 and Table 2. Throughout the simulations, we impose the lower and upper bounds on the aggregate data rate of each user as $\xi_{s_i} = 0.1$ and $\zeta_{s_i} = 3$, i=1,…,8, respectively.

6.2. Performance of FDTAA

One advantage of the proposed algorithm is that efficient approximation techniques exist for approximating general utility functions with polynomial-like functions, i.e., functions of the form (2). Obviously, the utility functions (12) are not in the polynomial-like form. We use polynomial-like functions as in (2) to approximate the monotonic piecewise affine functions in (12). This way we obtain the coefficient vectors $\mathbf{p}_{s_i}$ for $i = 1, \dots, 8$ that make (2) a close approximation of (12) according to some defined metric. We choose to show the approximation result for $\alpha = \ell = 6$ and do the following simulations for this approximation function. We refrain from providing the details of the approximation techniques for that purpose since it is not the main focus of this paper and due to space limitation. Interested readers can refer to Henrion et al. (2009), which presents theoretical results and tutorial examples on the approximation technique.

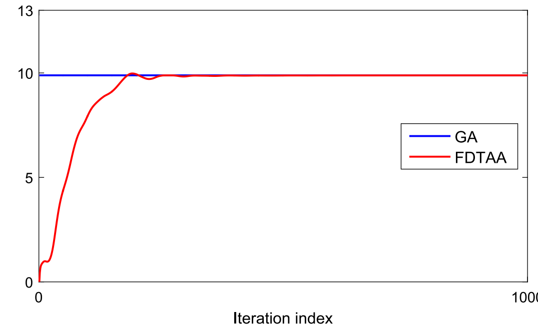


Fig. 5. Utility function value for GA and FDTAA iterates.

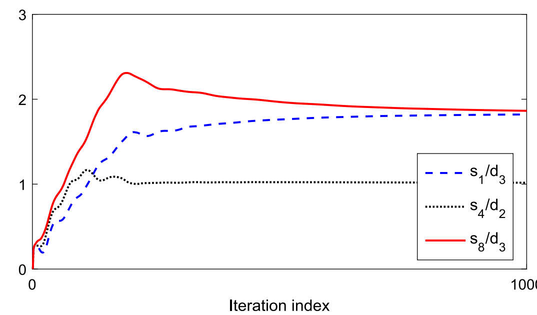


Fig. 6. Data rates in response to MRGUBRS requirements on the flows.

We focus on approximating a global optimal solution to problem in (3) having an objective that is the sum of polynomial-like utility functions as in (2); therefore, any global optimal solution to (3) serves as a benchmark solution with which we can evaluate the performance of the proposed algorithm FDTAA. Since (3) is nonconvex, in this simulation, the global optimal solution is estimated using a global optimization technique, in particular the genetic algorithm (GA). However, we emphasize that GA is a centralized solution that requires global network information, and is prohibitively expensive to be implemented in practice.

Fig. 5 shows the performance of FDTAA for stepsize parameters chosen to satisfy the convergence condition set forth by Theorem 1. It can be seen that the utility function values converge to the optimal one, which is obtained by using the genetic algorithm. Although all the computations of FDTAA are performed locally at each node, it attains almost the same network utility value obtained with a centralized optimization algorithm, which requires the availability of global information.

Fig. 6 shows the representative data rate trajectories for MRGUBRS flows transporting through source/destination nodes s_1/d_3, s_4/d_2 and s_8/d_3. All data rate sequences are generated by FDTAA. Fig. 6 shows that the MRGUBRS requirements are satisfied.

6.3. Robustness of FDTAA

In this section, we show that FDTAA is robust to link failures. When a link failure is detected, the algorithm reroutes the traffic such that flows avoid going through that link. More precisely, when a link failure happens, only the nodes connected to that link are aware of the link failure and change the corresponding link capacity. Other nodes are oblivious to the failure. We show that FDTAA excels at this task. At iteration 130, we shutdown the link connecting nodes $b_7 - b_8$. We choose to test the response of FDTAA to the failure of this particular link since it is used by all the source nodes to transport their flows to their destinations, which

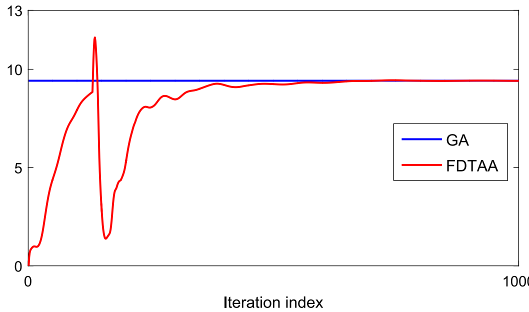


Fig. 7. Utility function value in response to a link failure.

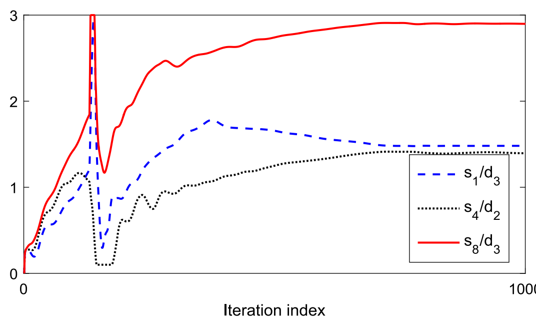


Fig. 8. Data rates in response to a link failure.

follows from Table 2. Moreover, we do not need to update all the stepsize parameters, since they still satisfy Theorem 1 for the new communication topology. Fig. 7 shows that FDTAA responds to the link failure immediately and the utility function converges to its new optimal value, which is obtained by using genetic algorithm and it requires the availability of global information. Fig. 8 shows the representative data rate trajectories for flows transporting among source/destination nodes s_1/d_3, s_4/d_2 and s_8/d_3. It also shows that FDTAA rapidly redistributes the traffic and in this way steers the network towards its new optimal point of operation.

7. Discussion

This paper formulates the traffic allocation problem in connectionless networks as a NUM problem which maximizes a nonconcave utility function subject to the resource constraints. The resulting NUM problem is nonconvex and is hard to analyze and solve even with centralized algorithms. We proposed a convex relaxation whose optimal value approximates that of the given nonconvex NUM problem. Furthermore, we developed FDTAA which is a fully distributed algorithm and its iterate sequence converges to an optimal solution of the convex relaxation. FDTAA has low computational workload since it updates data rates as per destination node.

Simulation results show that the behavior of FDTAA mimics the optimal traffic allocation. In particular, the simulation studies demonstrate that FDTAA can effectively reroute the traffic to a new optimal state when link failures occur.

Although the robustness of FDTAA against link failures has been tested, there are many other issues that need further consideration, like the delay of information transmission. Future research directions include studying these detailed implementation issues and testing the implementation in large-scale network settings.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (61803319), Shenzhen Science and Technology Projects, China (JCYJ20180306172720364), the Fundamental Research Funds for the Central Universities of China (20720190015), Natural Science Foundation of Fujian Province of China (No. 2019J05021), National Institutes of Health R01 HL142732 (United States), National Science Foundation #1808266, XPS-1629625 (United States), SHF-1704504, CMMI-1400217, CMMI-1635106, ARO grant W911NF-17-1-0298 (United States), and the China Scholarship Council.

Appendix A. Proof of Proposition 1

In this Appendix, we first recall some results which are fundamental for the development of Proposition 1, and then present the main steps of the proof.

Lemma 1. *Let f be a real-valued polynomial, $\mathcal{F} \subset \mathbb{R}^n$ be a semialgebraic set, and $\mathcal{M}(\mathcal{F})$ be the set of finite positive Borel measures supported on $\mathcal{F}$. Then, the problems*

$$\min \{f(\mathbf{x}) : \mathbf{x} \in \mathcal{F}\}, \tag{A.1a}$$

$$\min \left\{ \int_{\mathcal{F}} f d\mu : \mu \in \mathcal{M}(\mathcal{F}),\ \mu(\mathcal{F}) = 1 \right\}, \tag{A.1b}$$

are equivalent in the following sense

(a) the optimal values of (A.1a) and *(A.1b) are the same;*
(b) if $\mathbf{x}^*$ is an optimal solution to (A.1a), then the Dirac measure centered at $\mathbf{x}^*$ is optimal to (A.1b);
(c) if μ^* is an optimal solution to (A.1b), then any point in the support of μ^* is optimal to (A.1a).

Proof. (a) follows from Lasserre (2010, Theorem 5.1). (b) and (c) easily follow from (a).

Lemma 2. *Let μ be a Borel probability measure supported on the closed and bounded interval* $[0, 1]$. *Then the moment sequence $\mathbf{t} \triangleq \{t_j\}_{j\in\mathbb{N}}$ of measure μ satisfies $0 \le t_j \le 1$ for all $j \in \mathbb{N}$.*

Proof. Since μ is a probability measure with support contained on $[0, 1]$, we have $0 \le t_j = \mathbb{E}_\mu[y^j] = \int_{\mathcal{Y}} y^j d\mu(y) \le \int_{\mathcal{Y}} y\, d\mu(y) \le 1$ for each $j \in \mathbb{N}$.

Lemma 3. *Let $\mathbf{t} = \{t_j\}_{j\in\mathbb{N}}$ be a real sequence. If there exists a constant $c > 0$ such that $\mathbf{H}(0, \alpha, \mathbf{t}) \succeq 0$ and $t_j \le c$ for all $\alpha \in \mathbb{N}$ and $j \in \mathbb{N}$, then there exists a representing measure μ with support on* $[0, 1]$.

Proof. Direct application of Jasour, Aybat, and Lagoa (2015, Lemma 2.4).

We proceed with the following theorem that provides necessary and sufficient conditions for the existence of Borel measures whose support is included in bounded and symmetric intervals of the real line.

Theorem 2. *Given a sequence $\mathbf{t} = \{t_j\}_{j=1}^n$, there exists a Borel measure μ with support contained in $\mathcal{X} \doteq [-\epsilon, \epsilon]$ such that $\mu(\mathcal{X}) = 1$ and $t_j = \mathbb{E}_\mu[y^j] = \int_{\mathcal{C}} y^j d\mu(y)$ holds if and only if*

- *when n is odd, i.e., $n = 2k + 1$ for some $k \in \mathbb{Z}_+$, the following holds*

$$\epsilon \mathbf{H}(0, 2k, \mathbf{t}) \succeq \mathbf{H}(1, 2k + 1, \mathbf{t}) \tag{A.2}$$

$$\mathbf{H}(1, 2k + 1, \mathbf{t}) \succeq -\epsilon \mathbf{H}(0, 2k, \mathbf{t}), \tag{A.3}$$

- *when n is even, i.e., $n = 2k$ for some $k \in \mathbb{Z}_+$, the following holds*

$$\mathbf{H}(0, 2k, \mathbf{t}) \succeq 0 \tag{A.4}$$

$$\epsilon^2 \mathbf{H}(0, 2k-2, \mathbf{t}) \succeq \mathbf{H}(2, 2k, \mathbf{t}), \tag{A.5}$$

where $\mathbf{H}(k, k+2h, \mathbf{t}) \in \mathbb{R}^{(h+1)\times(h+1)}$ is the Hankel matrix given in (1) and $t_0 = 1$.

Proof. Direct application of Theorem III.2.3. and Theorem III.2.4. in Krein and Nudelman (1977).

We proceed with the proof of Proposition 1 using the results mentioned above. With the variable change $y_s = r_s^{1/\ell}$, problem (3) can be converted into a polynomial optimization form. The equivalent problem is stated as:

$$\begin{aligned}
\max_{\mathbf{x},\mathbf{y},\mathbf{r}}. \quad & \sum_{s\in\mathcal{S}} \sum_{j=0}^{\alpha} p_{s,j} y_s^j \\
\text{s.t.} \quad & 0 \le y_s \le r_s^{1/\ell}, \quad s \in \mathcal{S} \\
& \mathbf{1}^\top \mathbf{x}_{s,l} \le c_l, \quad l \in \mathcal{L}_s, s \in \mathcal{S} \\
& \mathbf{1}^\top \mathbf{x}_{b,l} + \mathbf{1}^\top \mathbf{x}_{e_l(b),l} \le c_l, \ l \in \mathcal{L}_b, \ b \in \mathcal{B} \\
& \mathbf{B}\mathbf{x} = 0 \\
& \mathbf{x}_{b,l} \succeq 0, \quad l \in \mathcal{L}_b, \ b \in \mathcal{B} \\
& (\mathbf{x}_s, \ r_s) \in \mathcal{X}_s, \quad s \in \mathcal{S}.
\end{aligned} \tag{A.6}$$

Now, we consider transforming problem (A.6) into an optimization problem over the space of probability measures of y_s with support contained in the feasible set of (A.6). Define $\mathcal{Y}_s \triangleq \{y_s \in \mathbb{R} : 0 \le y_s \le r_s^{1/\ell}\}$. Then, Lemma 1 implies that (A.6) is equivalent to:

$$\begin{aligned}
\max_{\mathbf{x},\mathbf{r},\{\mu_s\}_{s\in\mathcal{S}}}. \quad & \int \sum_{s\in\mathcal{S}} \sum_{j=0}^{\alpha} p_{s,j} y_s^j d\mu_s(y) \\
\text{s.t.} \quad & \text{supp}(\mu_s) \subset \mathcal{Y}_s, \ \mu_s(\mathcal{Y}_s) = 1, \quad s \in \mathcal{S} \\
& \mathbf{1}^\top \mathbf{x}_{s,l} \le c_l, \quad l \in \mathcal{L}_s, s \in \mathcal{S} \\
& \mathbf{1}^\top \mathbf{x}_{b,l} + \mathbf{1}^\top \mathbf{x}_{e_l(b),l} \le c_l, \ l \in \mathcal{L}_b, \ b \in \mathcal{B} \\
& \mathbf{B}\mathbf{x} = 0 \\
& \mathbf{x}_{b,l} \succeq 0, \quad l \in \mathcal{L}_b, \ b \in \mathcal{B} \\
& (\mathbf{x}_s, \ r_s) \in \mathcal{X}_s, \quad s \in \mathcal{S},
\end{aligned} \tag{A.7}$$

where the support of μ_s is denoted by $\text{supp}(\mu_s)$.

We now note that:

- The objective functions of (A.7) equals that of (4) following our definition of moment variables, i.e., $m_{s,j}$ is the jth order moment of μ_s; hence, $m_{s,j} = \int y_s^j d\mu_s(y)$.
- Theorem 2 justifies the first three constraints in (4).
- Given $s \in \mathcal{S}$, we can establish the equivalence between the sequence of constraints $\{m_{s,j} \le r_s^{j/\ell}\}_{j\in\mathbb{N}}$ in (4) and the constraint $\text{supp}(\mu_s) \subset \mathcal{Y}_s$. Using dilation, we can reduce the case $y_s \in [0, r_s^{1/\ell}]$ to $\pi_s \triangleq y_s/(r_s^{1/\ell}) \in [0, 1]$. Then, it follows from Lemma 2 that the moments of π_s satisfy the sequence of constraints $\{m_{s,j}/r_s^{j/\ell} \le 1\}_{j\in\mathbb{N}}$. Similarly, the other side of the equivalence can be established by using Lemma 3 together with the second constraint in (4).

Note that $r_s^{j/\ell}$ is a concave function for all $j \in \{1, \dots, \ell\}$. Thus, $m_{s,j} \le r_s^{j/\ell}$ for all $j \in \{1, \dots, \alpha\}$ and $s \in \mathcal{S}$ is a set of convex constraints if $\alpha \le \ell$. Hence, problem (4) is convex provided that $\alpha \le \ell$.

Solving a sequence of problems of the form in (4) with increasing size converges to the solution of (A.7). The convergence can be shown using arguments similar to those in Jasour et al. (2015, Theorem 3.3). To summarize briefly, given $\alpha \in \mathbb{Z}_+$, let an optimal solution to (4) be denoted by $(\mathbf{m}^*(\alpha), \mathbf{x}^*(\alpha), \mathbf{r}^*(\alpha))$. Next, for all $s \in \mathcal{S}$, let $\{\bar{\mathbf{m}}_s^*(\alpha)\}_{\alpha\in\mathbb{Z}_+} \subset \mathbb{R}^{\mathbb{N}}$ be a sequence of infinite real sequences such that $\bar{\mathbf{m}}_s^*(\alpha)$ is obtained by zero-padding $\mathbf{m}_s^*(s)$. Then, for all $s \in \mathcal{S}$, $\{\bar{\mathbf{m}}_s^*(\alpha)\}_{\alpha\in\mathbb{Z}_+}$ weakly converges to an accumulation point such that there exists a representing measure μ_s^* of this point that is optimal to (A.7). Furthermore, for each $s \in \mathcal{S}$, any y_s^* in the support of μ_s^* is optimal to (A.6). It then follows by equivalence of problems (3) and (A.6) that $(\mathbf{x}^*(\alpha), \mathbf{r}^*(\alpha)) \to (\mathbf{x}^*, \mathbf{r}^*)$ as $\alpha \to \infty$.

Appendix B. Derivation of FDTAA

The problem in (4) can be compactly stated as

$$\begin{aligned}
\max_{\mathbf{m},\mathbf{x},\mathbf{r}}. \quad & \sum_{s\in\mathcal{S}} \mathbf{p}_s^\top \mathbf{m}_s \\
\text{s.t.} \quad & \mathbf{1}^\top \mathbf{x}_{b,l} + \mathbf{1}^\top \mathbf{x}_{e_l(b),l} - c_l \le 0, \ l \in \mathcal{L}_b, \ b \in \mathcal{B} \\
& \mathbf{B}\mathbf{x} = 0 \\
& (\mathbf{m}_s, \mathbf{x}_s, r_s) \in \mathcal{A}_s, \ s \in \mathcal{S} \\
& \mathbf{x}_{b,l} \succeq 0, \quad l \in \mathcal{L}_b, \ b \in \mathcal{B},
\end{aligned} \tag{B.1}$$

where $\mathcal{A}_s$ is defined in (9) for each $s \in \mathcal{S}$. The convex–concave saddle-point form of the primal problem (B.1) is given by

$$\min_{\mathbf{m},\mathbf{x},\mathbf{r}} \max_{\lambda, \ \theta} \mathrm{L}(\mathbf{m}, \mathbf{x}, \mathbf{r}, \lambda, \theta), \tag{B.2}$$

where $\mathrm{L}(\mathbf{x}, \ \mathbf{m}, \ \mathbf{r}, \ \lambda, \ \theta)$ is the Lagrangian function, i.e.,

$$\begin{aligned}
\mathrm{L}(\mathbf{m}, \mathbf{x}, \mathbf{r}, \lambda, \theta) = & - \sum_{s\in\mathcal{S}} (\mathbf{p}_s^\top \mathbf{m}_s - \mathbb{1}_{\mathcal{A}_s}(\mathbf{m}_s, \mathbf{x}_s, r_s)) \\
& + \sum_{b\in\mathcal{B}} \sum_{l\in\mathcal{L}_b} (\mathbb{1}_{\mathbb{R}_+^{|\mathcal{D}_{b,l}^{\text{out}}|}}(\mathbf{x}_{b,l}) - \mathbb{1}_{\mathbb{R}_+}(\lambda_{b,l})) + \langle \mathbf{B}\mathbf{x}, \ \theta \rangle \\
& + \sum_{b\in\mathcal{B}} \sum_{l\in\mathcal{L}_b} \left\langle \mathbf{1}^\top \mathbf{x}_{b,l} + \mathbf{1}^\top \mathbf{x}_{e_l(b),l} - c_l, \lambda_{b,l} \right\rangle.
\end{aligned} \tag{B.3}$$

and $\theta \in \mathbb{R}^{\sum_{b\in\mathcal{B}} |\mathcal{D}_b|}$ denotes the dual variables associated with the flow conservation constraint $B\mathbf{x} = 0$. Given $l \in \mathcal{L}_b$ and $b \in \mathcal{B}$, the dual variable $\lambda_{b,l}$ is associated to the capacity inequality constraint $\mathbf{1}^\top \mathbf{x}_{b,l} + \mathbf{1}^\top \mathbf{x}_{e_l(b),l} \le c_l$, and we define $\lambda = [\lambda_{b,l}]_{l\in\mathcal{L}_b, b\in\mathcal{B}}$.

Now, given the initial iterates $\mathbf{m}^0, \mathbf{x}^0, \mathbf{r}^0, \lambda^0, \theta^0$ and parameters $\gamma > 0$, $\tau_s > 0$ for all $s \in \mathcal{S}$, $\tau_{b,l}^d > 0$, $\kappa_{b,l} > 0$ for all $d \in \mathcal{D}_{b,l}^{\text{out}}$, $l \in \mathcal{L}_b$ and $b \in \mathcal{B}$, applying the basic primal–dual algorithm in Chambolle and Pock (2016) on (B.2), we compute the iterate sequences as follows:

$$\begin{aligned}
(\mathbf{m}, \mathbf{x}, \mathbf{r})^{k+1} = & \operatorname*{argmin}_{\mathbf{m},\mathbf{x}\succeq 0,\mathbf{r}} - \sum_{s\in\mathcal{S}} (\mathbf{p}_s^\top \mathbf{m}_s - \mathbb{1}_{\mathcal{A}_s}(\mathbf{m}_s, \mathbf{x}_s, r_s)) \\
& + \sum_{s\in\mathcal{S}} \frac{1}{2\tau_s} (\|\mathbf{m}_s - \mathbf{m}_s^k\|^2 + \|\mathbf{x}_s - \mathbf{x}_s^k\|^2 + (r_s - r_s^k)^2) \\
& + \left\langle \mathbf{B}\mathbf{x}, \ \theta^k \right\rangle + \sum_{b\in\mathcal{B}} \sum_{l\in\mathcal{L}_b} \Bigg(\left\langle \mathbf{1}^\top \mathbf{x}_{b,l} + \mathbf{1}^\top \mathbf{x}_{e_l(b),l} - c_l, \lambda_{b,l}^k \right\rangle \\
& \qquad + \sum_{d\in\mathcal{D}_{b,l}^{\text{out}}} \frac{1}{2\tau_{b,l}^d} (x_{b,l}^d - (x_{b,l}^d)^k)^2 \Bigg)
\end{aligned} \tag{B.4}$$

$$\begin{aligned}
\lambda_{b,l}^{k+1} = & \operatorname*{argmax}_{\lambda_{b,l}\ge 0} \left\langle \mathbf{1}^\top (2\mathbf{x}_{b,l}^{k+1} - \mathbf{x}_{b,l}^k + 2\mathbf{x}_{e_l(b),l}^{k+1} - \mathbf{x}_{e_l(b),l}^k) - c_l, \lambda_{b,l} \right\rangle \\
& - \frac{1}{2\kappa_{b,l}} (\lambda_{b,l} - \lambda_{b,l}^k)^2, \ l \in \mathcal{L}_b, \ b \in \mathcal{B}
\end{aligned} \tag{B.5}$$

$$\begin{aligned}
\theta^{k+1} = & \operatorname*{argmax}_{\theta} \left\langle \mathbf{B}(2\mathbf{x}^{k+1} - \mathbf{x}^k), \theta \right\rangle - \frac{1}{2\gamma} \|\theta - \theta^k\|^2 \\
= & \ \theta^k + \gamma \mathbf{B}(2\mathbf{x}^{k+1} - \mathbf{x}^k).
\end{aligned} \tag{B.6}$$

These iterations are still not in a distributed form since one needs to know global network information to be able to solve the subproblem for updating the primal variables $\mathbf{x}$. Indeed, this is caused by the presence of the term $\langle \mathbf{Bx},\ \theta\rangle$, which is associated with the flow conservation constraints at nodes. Nevertheless, investigating the structure of the inner product of the two vectors $\mathbf{Bx}$ and θ reveals that it is a summation of local linear functions of the local variables; therefore, it is possible to develop an optimal decentralized traffic allocation algorithm. As in Aybat and Yazdandoost Hamedani (2016), one can use recursion to write θ^{k+1} as a partial summation of primal variables $\mathbf{x}^k$, i.e.,

$$\theta^k = \theta^0 + \gamma \sum_{n=0}^{k-1} B(2\mathbf{x}^{n+1} - \mathbf{x}^n). \tag{B.7}$$

Let $\theta^0 = \gamma \mathbf{B}\mathbf{x}^0$, $\mathbf{z}^0 = \mathbf{x}^0$ and $\mathbf{z}^k \triangleq \mathbf{x}^k + \sum_{n=1}^k \mathbf{x}^n$ for $k \geq 1$. Then, one has

$$\left\langle \mathbf{Bx},\ \theta^k \right\rangle = \gamma \left\langle \mathbf{x}, \mathbf{B}^\top B\mathbf{z}^k \right\rangle = \gamma \sum_{b\in\mathcal{B}} \sum_{l\in\mathcal{L}_b} \mathbf{x}_{b,l}^\top \mathbf{u}_{b,l}^k + \gamma \mathbf{x}_s \mathbf{u}_s^k, \tag{B.8}$$

where the variables $\mathbf{u}_{b,l} \triangleq [u_{b,l}^d]_{d\in\mathcal{D}_{b,l}^{\text{out}}}$ for each $l \in \mathcal{L}_b$ and $b \in \mathcal{B}$, and $\mathbf{u}_s \triangleq [[u_{s,l}^d]_{d\in\mathcal{D}_{s,l}}]_{l\in\mathcal{L}_s}$ for each $s \in \mathcal{S}$ are introduced to denote the relative sending data-rate information among neighbors, and their entries are defined as

$$\begin{aligned} u_{b,l}^d \triangleq{}& z_{b,l}^d + \sum_{\tilde{l}\in\mathcal{L}_b^{d,\text{out}}} z_{b,\tilde{l}}^d - \sum_{\bar{l}\in\mathcal{L}_b^{d,\text{in}}} z_{e_{\bar{l}}(b),\bar{l}}^d - \sum_{\hat{l}\in\mathcal{L}_{e_l(b)}^{d,\text{out}}} z_{e_l(b),\hat{l}}^d \\ &+ \sum_{\check{l}\in\mathcal{L}_{e_l(b)}^{d,\text{in}}} z_{e_{\check{l}}(e_l(b)),\check{l}}^d,\ d \in \mathcal{D}_{b,l}^{\text{out}},\ l \in \mathcal{L}_b,\ b \in \mathcal{B} \end{aligned} \tag{B.9a}$$

$$u_{s,l}^d \triangleq z_{s,l}^d - \sum_{\tilde{l}\in\mathcal{L}_{e_l(s)}^{d,\text{out}}} z_{e_l(s),\tilde{l}}^d,\ d \in \mathcal{D}_{s,l},\ l \in \mathcal{L}_s,\ s \in \mathcal{S}. \tag{B.9b}$$

Next, (B.5) implies that $\lambda_{b,l}^{k+1}$ is simply updated as:

$$\lambda_{b,l}^{k+1} = (\lambda_{b,l}^k + \mathbf{1}^\top(2\mathbf{x}_{b,l}^{k+1} - \mathbf{x}_{b,l}^k + 2\mathbf{x}_{e_l(b),l}^{k+1} - \mathbf{x}_{e_l(b),l}^k) - c_l)^+. \tag{B.10}$$

Substituting (B.8) and (B.9) into (B.4) and (B.6) yields a fully distributed traffic allocation algorithm shown in Algorithm 1.

Appendix C. Proof of Theorem 1

Proof. Let $\{\upsilon_s\}$ and $\{\upsilon_{b,l}^d\}$ be positive constants satisfying the following inequalities

$$\upsilon_s > \max_{d\in\mathcal{D}_s} |\mathcal{L}_s^d| + \sum_{l\in\mathcal{L}_s} (|\mathcal{L}_{e_l(s)}^{d,\text{out}}| + |\mathcal{L}_{e_l(s)}^{d,\text{in}}|) - 3,\ \forall\, s \in \mathcal{S} \tag{C.1}$$

$$\upsilon_{b,l}^d > |\mathcal{L}_b^{d,\text{out}}| + |\mathcal{L}_b^{d,\text{in}}| + \sum_{\tilde{l}\in\mathcal{L}_b^{d,\text{out}}} (|\mathcal{L}_{e_{\tilde{l}}(b)}^{d,\text{out}}| + |\mathcal{L}_{e_{\tilde{l}}(b)}^{d,\text{in}}|) - 4 \tag{C.2}$$

for all $d \in \mathcal{D}_{b,l}^{\text{out}}, l \in \mathcal{L}_b, b \in \mathcal{B}$. Choose the primal–dual step sizes $\{\tau_s\}$, $\{\tau_{b,l}^d\}$, $\{\kappa_{b,l}\}$ and γ to be positive constants such that conditions (10) and (11) are satisfied. Let $\mathbf{1}_n$ ($\mathbf{0}_n$) be a column vector of n ones (zeros), and $\mathbf{1}_{m\times n}$ ($\mathbf{0}_{m\times n}$) be an $m \times n$ matrix of all ones (zeros). Then, it can be shown that the following matrix is positive definite by using Schur complement Lemma and Gershgorin circle Theorem, i.e.,

$$Q(A, A_0) \triangleq \begin{bmatrix} D_\tau & -A^\top & -A_0^\top \\ \star & D_\kappa & \mathbf{0} \\ \star & \star & D_\gamma \end{bmatrix} \succ 0, \tag{C.3}$$

where $A \triangleq \text{diag}([\mathbf{0}_{(\ell+1)|\mathcal{S}|}^\top, \mathbf{1}_{|\mathbf{x}|}^\top, \mathbf{0}_{|\mathcal{S}|}^\top])$, and $D_\kappa \triangleq \text{diag}([\mathbf{0}_{(\sum_{s\in\mathcal{S}}(\ell+1+\sum_{l\in\mathcal{L}_s}|\mathcal{D}_{s,l}|))}^\top, [\kappa_b^\top]_{b\in\mathcal{B}}^\top, \mathbf{0}_{|\mathcal{S}|}^\top])$ for $\kappa_b \triangleq [\frac{1}{\kappa_{b,l}}]_{l\in\mathcal{L}_b}$,; $D_\gamma \triangleq \text{diag}([\mathbf{0}_{(\ell+1)|\mathcal{S}|}^\top, \frac{1}{\gamma}\mathbf{1}_{\sum_{b\in\mathcal{B}}|D_b|}^\top, \mathbf{0}_{|\mathcal{S}|}^\top])$; $A_0 \triangleq \text{diag}([\mathbf{0}_{((\ell+1)|\mathcal{S}|)\times((\ell+1)|\mathcal{S}|)}, \mathbf{B}, \mathbf{0}_{|\mathcal{S}|\times|\mathcal{S}|}])$ is a block-diagonal matrix; $D_\tau \triangleq \text{diag}([[\tilde{\upsilon}_s^\top]_{s\in\mathcal{S}}^\top, [\upsilon_b^\top]_{b\in\mathcal{B}}^\top, [\frac{1}{\tau_s}]_{s\in\mathcal{S}}^\top])$ where $\tilde{\upsilon}_s \triangleq [\frac{1}{\tau_s}\mathbf{1}_{\ell+1+\sum_{l\in\mathcal{L}_s}|\mathcal{D}_{s,l}|}^\top]_{s\in\mathcal{S}}$, $\upsilon_b \triangleq [\upsilon_{b,l}^\top]_{l\in\mathcal{L}_b}$, and $\upsilon_{b,l} \triangleq [\frac{1}{\tau_{b,l}^d}]_{d\in\mathcal{D}_{b,l}^{\text{out}}}$.

By using Theorem 2.2 in Aybat and Yazdandoost Hamedani (2016), we have that the conclusions of Theorem 1 hold.

References

Acemoglu, D., Johari, R., & Ozdaglar, A. (2007). Partially optimal routing. *IEEE Journal on Selected Areas in Communications*, *25*(6).

Alizadeh, M., Edsall, T., Dharmapurikar, S., Vaidyanathan, R., Chu, K., Fingerhut, A., et al. (2014a). CONGA: Distributed congestion-aware load balancing for datacenters. In *ACM SIGCOMM computer communication review: Vol. 44(4)* (pp. 503–514). ACM.

Alizadeh, M., Edsall, T., Dharmapurikar, S., Vaidyanathan, R., Chu, K., Fingerhut, A., et al. (2014b). CONGA: Distributed congestion-aware load balancing for datacenters. In *ACM SIGCOMM computer communication review: Vol. 44(4)* (pp. 503–514). ACM.

Allman, M., Paxson, V., & Blanton, E. (2009). TCP congestion control. *Technical report*, IETF RFC.

Ashour, M., Wang, J., Lagoa, C., Aybat, N., & Che, H. (2017). Non-concave network utility maximization: a distributed optimization approach. In *IEEE conference on computer communications* (pp. 1908–1916).

Aybat, N. S., & Yazdandoost Hamedani, E. (2016). A primal-dual method for conic constrained distributed optimization problems. In *Advances in neural information processing systems, Vol. 29* (pp. 5049–5057). Curran Associates, Inc..

Chambolle, A., & Pock, T. (2016). On the ergodic convergence rates of a first-order primal–dual algorithm. *Mathematical Programming*, *159*(1–2), 253–287.

Chen, L., Chen, K., Bai, W., & Alizadeh, M. (2016). Scheduling mix-flows in commodity datacenters with karuna. In *Proceedings of the 2016 ACM SIGCOMM conference* (pp. 174–187). ACM.

(2015–2020). Cisco VNI Forecast and Methodology. http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.pdf.

Cofano, G., De Cicco, L., Zinner, T., Nguyen-Ngoc, A., Tran-Gia, P., & Mascolo, S. (2016). Design and experimental evaluation of network-assisted strategies for HTTP adaptive streaming. In *Proceedings of the 7th international conference on multimedia systems* (p. 3). ACM.

Elwalid, A., Jin, C., Low, S., & Widjaja, I. (2001). MATE: MPLS adaptive traffic engineering. In *Proceedings of the twentieth annual joint conference of the IEEE computer and communications societies: Vol. 3* (pp. 1300–1309). IEEE.

Fazel, M., & Chiang, M. (2005). Network utility maximization with nonconcave utilities using sum-of-squares method. In *Proceedings of the 44th IEEE conference on decision and control* (pp. 1867–1874). IEEE.

Golestani, S., & Bhattacharyya, S. (1998). A class of end-to-end congestion control algorithms for the Internet. In *Proceedings of the sixth international conference on network protocols* (pp. 137–150). IEEE.

Han, H., Shakkottai, S., Hollot, C., Srikant, R., & Towsley, D. (2004). Overlay TCP for multi-path routing and congestion control. In *IMA workshop on measurements and modeling of the internet*.

Hande, P., Zhang, S., & Chiang, M. (2007). Distributed rate allocation for inelastic flows. *IEEE/ACM Transactions on Networking*, *15*(6), 1240–1253.

Henrion, D., Lasserre, J., & Savorgnan, C. (2009). Approximate volume and integration for basic semialgebraic sets. *SIAM Review*, *51*(4), 722–743.

Hopps, C. (2000). Analysis of an equal-cost multi-path algorithm. *RFC Editor*.

Jasour, A., Aybat, N., & Lagoa, C. (2015). Semidefinite programming for chance constrained optimization over semialgebraic sets. *SIAM Journal on Optimization*, *25*(3), 1411–1440.

Juluri, P., Tamarapalli, V., & Medhi, D. (2016). Measurement of quality of experience of video-on-demand services: A survey. *IEEE Communications Surveys & Tutorials*, *18*(1), 401–418.

Kelly, F., Maulloo, A., & Tan, D. (1998). Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research society*, *49*(3), 237–252.

Kelly, F., & Voice, T. (2005). Stability of end-to-end algorithms for joint routing and rate control. *ACM SIGCOMM Computer Communication Review*, *35*(2), 5–12.

Krein, M., & Nudelman, A. (1977). *The markov moment problem and extremal problems*. Providence, Rhode Island: American Mathematical Society.

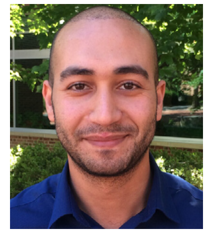
Lagoa, C., & Che, H. (2000). Decentralized optimal traffic engineering in the Internet. *ACM SIGCOMM Computer Communication Review*, *30*(5), 39–47.

Lagoa, C., Che, H., & Movsichoff, B. (2004). Adaptive control algorithms for decentralized optimal traffic engineering in the Internet. *IEEE/ACM Transactions on Networking*, *12*(3), 415–428.

Lasserre, J. B. (2001). Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, *11*(3), 796–817.

Lasserre, Jean-Bernard (2010). Moments, positive polynomials and their applications. *1*, World Scientific.
Liu, Y., Zhang, H., Gong, W., & Towsley, D. (2005). On the interaction between overlay routing and underlay routing. In *INFOCOM 2005. 24th annual joint conference of the IEEE computer and communications societies. Proceedings IEEE: Vol. 4* (pp. 2543–2553). IEEE.
Movsichoff, B., Lagoa, C., & Che, H. (2005). Decentralized optimal traffic engineering in connectionless networks. *IEEE Journal on Selected Areas in Communications*, *23*(2), 293–303.
Movsichoff, B., Lagoa, C., & Che, H. (2007). End-to-end optimal algorithms for integrated QoS, traffic engineering, and failure recovery. *IEEE/ACM Transactions on Networking*, *15*(4), 813–823.
Nekouei, E., Nair, G., & Alpcan, T. (2015). Convergence analysis of quantized primal-dual algorithm in quadratic network utility maximization problems. In *Proceedings of the 54th IEEE conference on decision and control* (pp. 2655–2660). IEEE.
Press, W., Teykolsky, S., Vetterling, W., & Flannery, B. (2007). *Numerical recipes in C++: The art of scientific computing*. New York: Cambridge University Press.
Qiu, L., Yang, Y., Zhang, Y., & Shenker, S. (2003). On selfish routing in internet-like environments. In *Proceedings of the 2003 conference on applications, technologies, architectures, and protocols for computer communications* (pp. 151–162). ACM.
Sanchez, Y., Schierl, T., Hellge, C., Wiegand, T., Hong, D., De Vleeschauwer, D., et al. (2012). Efficient HTTP-based streaming using scalable video coding. *Signal Processing: Image Communication*, *27*(4), 329–342.
Sodagar, I. (2011). The MPEG-DASH standard for multimedia streaming over the internet. *IEEE MultiMedia*, *18*(4), 62–67.
Su, W., Liu, C., Lagoa, C., Che, H., Xu, K., & Cui, Y. (2015). Integrated, distributed traffic control in multidomain networks. *IEEE Transactions on Control Systems Technology*, *23*(4), 1373–1386.
Wang, J., Ashour, M., Lagoa, C., Aybat, N., Che, H., & Duan, Z. (2017). Non-concave network utility maximization in connectionless networks: a fully distributed traffic allocation algorithm. In *Proceedings of the Amercian control conference* (pp. 3980–3985).
Wang, W., Palaniswami, M., & Low, S. (2003). Optimal flow control and routing in multi-path networks. *Performance Evaluation*, *52*(2), 119–132.
Wang, P., & Xu, H. (2014). Expeditus: distributed load balancing with global congestion information in data center networks. In *Proceedings of the 2014 CoNEXT on student workshop* (pp. 1–3). ACM.
Ye, L., Wang, Z., Che, H., & Lagoa, C. (2011). TERSE: A unified end-to-end traffic control mechanism to enable elastic, delay adaptive, and rate adaptive services. *IEEE Journal on Selected Areas in Communications*, *29*(5), 938–950.
Yin, X., Bartulović, M., Sekar, V., & Sinopoli, B. (2017). On the efficiency and fairness of multiplayer HTTP-based adaptive video streaming. In *American control conference* (pp. 4236–4241). IEEE.
Yin, X., Jindal, A., Sekar, V., & Sinopoli, B. (2015). A control-theoretic approach for dynamic adaptive video streaming over HTTP. *ACM SIGCOMM Computer Communication Review*, *45*(4), 325–338.

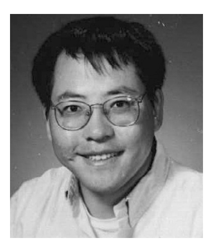
Jingyao Wang received the B. S. degree in mathematics from Inner Mongolia University, Hohhot, China in 2010, and Ph.D. degree in engineering from Peking University, Beijing in 2017. Since July 2017, she has been with the School of Aerospace Engineering at Xiamen University, Xiamen, China, and the Shenzhen Research Institute of Xiamen University, Shenzhen, China, where she currently holds the position of Assistant Professor. Her current research interests include distributed coordination control of multi-agent systems and network resource management.



Mahmoud Ashour received the B.Sc. and M.Sc. degrees in electrical engineering from Cairo University and Nile University, Egypt, in 2010 and 2013, respectively. He was a research assistant with the Computer Science and Engineering Department, Qatar University, Qatar, for one year in 2014. He is currently pursuing the Ph.D. degree in the Electrical Engineering and Computer Science Department, Penn State University, University Park, PA, USA. His research interests lie in the broad area of communication networks with an emphasis on distributed optimization algorithms.

Constantino M. Lagoa received the B.S. and M.S. degrees from the Instituto Superior Tecnico, Techical University of Lisbon, Portugal in 1991 and 1994, respectively, and the Ph.D. degree from the University of Wisconsin at Madison in 1998. He joined the Electrical Engineering Department of Pennsylvania State University, University Park, PA, in August 1998, where he currently holds the position of Professor. He has a wide range of research interests including robust optimization and control, chance constrained optimization, controller design under risk specifications, system identification and control of computer networks. Dr. Lagoa has served as Associate Editor of IEEE Transactions on Automatic Control (2012–2017) and IEEE Transactions on Control systems Technology (2009–2013) and he is currently Associate Editor of Automatica.

Necdet Serhat Aybat received the B.S. and M.S. degrees in industrial engineering from Bogazici University, Istanbul, Turkey, in 2003 and 2005, respectively, and the Ph.D. degree in operations research from Columbia University, New York, NY, USA, in 2011. He joined the Department of Industrial Engineering, Penn State University, State College, PA, USA, in August 2011. His current research interests include developing first-order algorithms for large-scale convex optimization problems from diverse application areas, such as compressed sensing, matrix completion, convex regression, and distributed optimization. Dr. Aybat is a runner-up of the INFORMS Computing Society Student Paper Award in 2010, and received the SIAM Student Paper Prize in 2011.



Hao Che received the B.S. degree from Nanjing University, Nanjing, China; the M.S. degree in physics from the University of Texas at Arlington, Arlington, Texas, USA; and Ph.D. degree in electrical engineering from the University of Texas at Austin, Austin, Texas, USA. He is currently a full professor in the Department of Computer Science and Engineering at the University of Texas at Arlington, Texas. Prior to joining UTA, he was a system architect with Santera Systems, Inc. in Plano (2000–2002) and an assistant professor of Electrical Engineering at the Pennsylvania State University (1998 to 2000). His current research interests include network architecture and Internet traffic control, datacenter resource management and job scheduling, edge computing and IoT. He is senior member of IEEE.