Optimal Resource Allocation for Cross-layer Utility Maximization in Ad Hoc Networks

Guangyou Tu  
Center of Information and Network, Chongqing University, Chongqing, 40044, China  
Email: lefour@cqu.edu.cn

Songtao Guo  
School of Electrical and Information Engineering, Southwest University, Chongqing, 40044, China  
Email: stguocqu@gmail.com

Hongyu Huang  
College of Computer Science, Chongqing University, Chongqing, 40044, China  
Email: hyhuang@cqu.edu.cn

Yuanyuan Yang  
Department of Electrical & Computer Engineering, Stony Brook University, 11794, NY, USA  
Email: yang@ece.sunysb.edu

Abstract—This paper adopt the generalized network utility maximization (GNUM) approach and propose a cross-layer optimized congestion, contention and power control algorithm for ad hoc networks. The goal is to find optimal end-to-end source rates at the transport layer, per-link persistence probabilities at the medium access control layer and transmitting power at the physical layer to maximize the aggregate source utility. Despite the inherent difficulties of non-convexity and non-separability of variables in the original optimization problem, we obtain a decoupled and dual-decomposable convex formulation by applying an appropriate transformation and introducing some new variables. The three decomposed sub-optimization problems are coordinated through the congestion prices. The convergence properties of the three sub-algorithms are also proved. Simulation results further verify the effectiveness and the convergence of our proposed algorithm.

Index Terms—Congestion Control, Power Control, Resource Allocation, Cross-Layer Design, Ad Hoc Networks

I. INTRODUCTION

Unlike cellular wireless networks, ad hoc networks have no preexisting infrastructures or centralized administration such as base stations due to the random access of wireless links. One of the fundamental tasks performed frequently by ad hoc networks is congestion control, whose main objective is to regulate the allowed source rates so that the total traffic load on any link does not exceed its available capacity. At the same time, the data rates that can be achieved for wireless links depend on the interference levels, which in turn are determined by the power control. It is well known that ad hoc nodes are powered by batteries with a limited lifetime, so, how to solve the scarcity of powers is a very serious challenge and how to mitigate congestion of links while maintaining a higher network utility. This is one of motivations of our work.

Concentration control in wireline networks is implemented at the transport layer and is often designed separately from functions of other layers. Since wired links have fixed capacities and are independent, this methodology is well justified. However, these results do not apply directly to ad hoc networks because unlike the wireline counterparts, capacities of wireless links are “elastic”, which depend on transmission powers and channel conditions as well as the specific medium access control (MAC) protocol used. Since the wireless channel is a shared and thus interference-limited medium, concurrent link transmissions in a local area can complicate congestion control due to the fact that transport layer flows can compete if they are located sufficiently close in space. For this reason, congestion control at the transport layer should be optimized and designed jointly with contention control at the MAC layer and power control at the physical layer to ensure efficient utilization and fair sharing of network resources while reduce the power consumption, which further motivates our work.

Network Utility Maximization (NUM) has been extensively used to study congestion control and much of existing research on NUM in wireless networks focuses on the jointly optimal transport layer congestion control and media access control (MAC) layer contention control and channel assignment. In [1-2], a generalized network utility maximization (GNUM) was used to study the interconnection among the layers and cross-layer designs
in wireless networks. In this framework, each source has a utility function and the sum of source utilities is regarded as the objective function to be maximized while other protocol parameters are related by constraints. A utility function can be interpreted as the satisfactions which a user can accept or a function of resource allocation. In [3-5], the joint design of congestion and contention controls has been studied. In [3], resource allocation in ad hoc wireless networks was formulated as a utility maximization problem with constraints that arise from contention for channel access. In [4-5], a joint congestion control and MAC problem was formulated by using the NUM framework over a multihop wireless ad hoc network, whose goal is to find optimal end-to-end source rates at the transport layer and per-link persistence probabilities at the MAC layer to maximize the aggregate source utility. In [6], a novel cross-layer congestion control strategy for WSNs was presented to perform the bandwidth and delay estimation in MAC layer for the current link, and feedbacks them to the transport layer through the cross-layer interaction mechanism. In [7], the problem of congestion control and channel assignment in multi-radio, multi-channel, wireless mesh networks was addressed. To maximize the benefit of shared relaying, in [22] the resource allocation and the scheduling of users among adjacent cell sectors need to be optimized jointly. A heuristic but efficient scheduling and resource allocation algorithm is proposed accordingly. However, these results do not consider the impact of transmission powers on congestion and contention controls.

More recently, the jointly optimal congestion control and per-link power control have also attracted much attention of researchers [1, 8-9]. In [1], Chiang presented a distributed power control algorithm that couples with TCP congestion control algorithms to increase end-to-end throughput and energy efficiency of multi-hop transmissions in wireless networks by iteratively solving the NUM problem. In [8], they considered the problem of joint congestion control and resource allocation in spatial-TDMA wireless networks. Based on a utility maximization problem subject to link rate constraints which involve both transmission scheduling and power allocation, a transparent and distributed protocol was developed. In [9], a joint opportunistic power scheduling and end-to-end rate control scheme was presented for wireless ad hoc networks. In [10, 11], the problem of joint power and rate control for secondary users in a cellular cognitive radio network was considered by using non-cooperative game theory. Ding et al. [12] proposed a cross-layer opportunistic spectrum access and dynamic routing algorithm for cognitive radio networks to maximize the network throughput by performing joint routing, dynamic spectrum allocation, scheduling, and transmit power control. By considering joint average interference-power and peak transmit-power constraints, Asghari and Aissa [13] obtained the variable rate and power adaptation policy for maximizing the achievable capacity of secondary user over fading channels. Furthermore, the fairness is considered when allocating resource to multiple users. A joint power and subcarrier/end-to-end rate control algorithm for cognitive radio networks was proposed in [14] by restricting the interference to licensed users while fairly maintaining a satisfied data rate or ensuring the interference produced to the primary user within a given limit. In [15, 16], a distributed algorithm was developed to maximize the aggregate source utility and increase end-to-end throughput, by integrating together congestion control, power control and spectrum allocation. In [17], Qu et al. proposed a cross-layer distributed power control and scheduling protocol for delay-constrained applications over mobile CDMA-based ad hoc wireless networks. In [18], Shen et. al. studied the fair resource allocation problem based on Nash bargaining solution (NBS) over wireless amplify-and-forward (AF) relay networks, and proposed a distributed algorithm, including relay selection, relay power allocation, and rate adaptation by the dual decomposition method. Although these results are effective to tackle the joint congestion and power control problems, they do not take into account the effect of MAC contention on congestion control. To the best of our knowledge, this work is the first one jointly considering congestion control, contention control and power control to maximize the network utility.

Therefore, in this paper, to overcome the shortcomings of existing approaches, effectively reduce network congestion and maximize the utilization of network resource, we propose a distributed cross-layer coordinative framework by integrating together congestion control at the transport layer, contention control at the MAC layer and power control at the physical layer, aiming to find optimal end-to-end source rates and per-link persistence probabilities while allocating optimal transmission power for each user. Based on elastic link capacity that is regarded as a function of transmission power, we construct a joint optimization formulation of congestion control, contention control and power control to maximize the network utility, which is a non-convex and non-separable optimization problem. By introducing new variables, the original joint optimization problem becomes decoupled and dual-decomposable. That is, it can be vertically decomposed to the congestion control subproblem, the contention control subproblem and the power control subproblem. The first subproblem is a primal algorithm where the congestion prices are generated based on local aggregate source rate. The last two is dual subgradient algorithms where the transmission power and the persistence probability of each node are updated by the congestion prices. Simulations have been conducted to verify the convergence and effectiveness of the proposed distributed algorithm.

The rest of this paper is organized as follows: Section 2 describes system model. Section 3 formulates the utility-maximization problem of joint congestion control, contention control and power control with elastic link capacities. In Section 4, we propose the distributed congestion control, contention control and power control sub-algorithms based on dual-decomposable and subgradient approach and prove their convergence.
Several numerical results are given in Section 5 to demonstrate the convergence and effectiveness of our proposed algorithm. The conclusions are drawn in Section 6.

II. SYSTEM MODEL

Consider an ad hoc network represented by a directed graph \( G = (N, L) \), where \( N \) denotes the set of nodes and \( L \) is the set of logic links. We assume connectivity to be symmetric, i.e., link \( (j, i) \in L \) from node \( j \) to node \( i \) if and only if \( (i, j) \in L \). We assume \( S \) denotes the set of sources. Each source \( s \in S \) attains a utility function \( U_s(x_s) \), which is an increasing, strictly concave, and continuously differential function of its end-to-end data rate \( x_s \). A utility can be interpreted as the level of satisfaction attained by a user and as a function of resource allocation. It is a quality measure as well. In other words, the higher the data rate, the better the quality measure of source \( s \). Let \( L(s) \subseteq L \) denote the set of links that source \( s \) uses to transmit data, and \( S(l) \subseteq S \) indicate the set of sources that use link \( l \). Define \( L_{\text{out}}(n) \) as the set of outgoing links from node \( n \), \( L_{\text{in}}(n) \) the set of incoming links to node \( n \). Each link \( l \in L \) has an "elastic" capacity \( c_l(w_l) > 0 \), which is a function of transmission power \( w_l \) of link \( l \).

Half-duplex operation is assumed to prevent self-interference, i.e., one transceiver can only transmit or receive at one time. Any two transmissions with a common intended receiver are not allowed to be made simultaneously since collisions will corrupt the packet receptions. Since the transmission range of each node is limited, the contention among links for shared media is relative on the locations of nodes. We define \( N^d_l \) as the set of nodes whose transmissions cause interference to the receiver of link \( l \), excluding the transmitter node of link \( l \), and \( L_{\text{out}}(n) \) as the set of links whose transmissions get interfered from the transmission of node \( n \), excluding outgoing links from node \( n \). Hence, the transmitter of link \( l \) and a node in set \( N^d_l \) cannot transmit data simultaneously; otherwise, the transmission of link \( l \) fails. Similarly, if node \( n \) and the transmitter of a link in set \( L_{\text{out}}(n) \) transmit simultaneously, the transmission of link \( l \) also fails.

The transmission time is slotted in intervals of equal unit length and the \( i \)-th time slot refers to the time interval \([i, i+1)\), where \( i = 1, 2, \ldots \), i.e., the data transmission attempts of each node occur at discrete time instance \( i \). We assume a MAC protocol based on random access with probabilistic transmissions. At the beginning of a slot, each node \( n \) transmits data with a probability \( P_n \). When it makes a decision to transmit data, it selects one of its outgoing links \( l \in L_{\text{out}}(n) \) with a probability \( q_l \), such that \( \sum_{l \in L_{\text{out}}(n)} q_l = 1 \), and transmits data only on the chosen link. Hence, there is no collision among links that have the same transmitter node. Consequently, link \( l \in L_{\text{out}}(n) \) transmits data with a probability \( p_l = P_n q_l \), such that \( \sum_{l \in L_{\text{out}}(n)} p_l = P_n \), \( \forall n \in N \). \( p_l \) and \( q_l \) are referred to as the persistence probability and conditional persistence probability of link \( l \), respectively. Each link depends on a random access algorithm, which will be studied in the next section, to adjust its persistence probability.

III. PROBLEM FORMULATION

This paper aims to choose optimal source rates \( x = \{x_s | s \in S\} \), transmission powers \( w = \{w_l | l \in L\} \) and link persistence probabilities \( p = \{p_l | l \in L_{\text{out}}(n)\} \) so as to maximize the aggregate source utility of all users in the network. The utility maximization problem with constraints at the transport, MAC and physical layers, in which optimal variables are end-to-end rates \( x \) controlled by TCP, link persistence probability \( p \) controlled by contention-based MAC and transmission power \( w \) at the physical layer, can be formulated as the following optimization (P1):

\[
\begin{align*}
\text{max} & \sum_{s \in S} U_s(x_s) \\
\text{s.t.} & \sum_{x \in L(l)} x_s \leq \eta_s := c_l(w_l) p_l \prod_{k \in N^d_l} \left(1 - \sum_{m \in L_{\text{out}}(k)} p_m\right), \forall l, \\
& \sum_{m \in L_{\text{out}}(k)} p_m \leq 1, \forall n, \\
& w_{l_{\text{min}}} \leq w_l \leq w_{l_{\text{max}}}, \forall l, \\
& x_{s_{\text{min}}} \leq x_s \leq x_{s_{\text{max}}}, \forall s, \\
& 0 \leq p_l \leq 1, \forall l,
\end{align*}
\]

where \( x_{s_{\text{min}}} \) and \( x_{s_{\text{max}}} \) are the minimum and maximum data rates of source \( s \), respectively; \( w_{l_{\text{min}}} \) and \( w_{l_{\text{max}}} \) are the minimum and maximum transmission power of link \( l \), respectively. Since the term \( p_l \prod_{k \in N^d_l} \left(1 - \sum_{m \in L_{\text{out}}(k)} p_m\right) \) indicates the probability that a packet is transmitted over link \( l \) and successfully received by its receiver, \( \eta = \{\eta_l | l \in L\} \) are link throughputs given \( p \).

The problem (P1) entails congestion control at the transport layer (finding \( x \)) and contention control (finding \( p \)) at the MAC layer. The two layers are coupled through the first constraint in (1), which states that the aggregate source rate \( \sum_{s \in L(l)} x_s \) for each link \( l \) does not exceed the link throughput; the second constraint demonstrates that the persistence probability \( P_n \) of each node \( n \) is no more than 1; the third constraint implies that the transmission power over link \( l \) is bounded between \( w_{l_{\text{min}}} \) and \( w_{l_{\text{max}}} \) so as to save the limited energy of the transmitter and satisfy the quality of services; the last one denotes the source rate is...
constrained by $x^\text{min}_s$ and $x^\text{max}_s$ so as to keep the fairness of sources and improve the effectiveness of wireless links. The transport layer source rates, the MAC layer transmission probabilities and the physical layer power control should be jointly optimized to maximize the aggregate source utility.

In interference-limited ad hoc networks, the physical layer link capacity can be expressed by

$$c_j(w_j) = B_j \log(1 + \gamma_j(w_j)) \quad (2)$$

where $B_j$ is the bandwidth of link $l$ and $\gamma_j(w_j)$ denotes the signal-to-interference-plus-noise ratio (SINR) of link $l$. High SINR can always be maintained since CSMA/CA based MAC layer prevents the nearby (i.e., 2-hop) nodes to operate at the same time. High SINR allows us to approximate $c_j(w_j)$ as below

$$c_j(w_j) = B_j \log(\gamma_j(w_j)) \quad (3)$$

Let $\sigma_j$ be the thermal noise power at the receiver of link $l$, and $G_{lm}$ be the direct link gain between the transmitter of link $m$ and the receiver of link $l$. Then SINR of link $l$ can be defined as

$$\gamma_j(w_j) = \frac{G_{lm}w_l}{\sum_{m=1}^{M}G_{lm}w_m + \sigma_j} \quad (4)$$

It is manifest from (3) and (4) that each link capacity $c_j(w_j)$ of link $l$ is “elastic”, and is a function of the transmission power $w_l$ and the channel conditions (such as link gain $G_{lm}$ and thermal noise $\sigma_j$), which depend on path loss, shadowing and multipath fading.

Due to the first constraint, the problem (1) is in general a non-convex and non-separable optimization problem, which is difficult to solve in a distributed manner. For any solution algorithm, convexity is the key for its global optimality and separability for its distributed nature. Therefore, as we will discuss in the next section, the problem (1) will be transformed to a problem which is both convex and dual-decomposable by using appropriate variable transformation and under readily-verifiable conditions. Furthermore, a distributed algorithm will be developed to solve for the globally optimal transport layer rates, MAC layer persistence probabilities and physical layer powers.

IV. JOINT CONGESTION, CONTENTION AND POWER CONTROL ALGORITHM

A. Non-convex Transformation and Dual Problem

As stated in the above section, the first constraint induces that (1) is in general a non-convex and non-separable optimization problem. Under certain conditions, however, it can be transformed into a convex one by introducing the auxiliary variables and the variable transformation. We first take the logarithm on both sides of the first constraint and replace the variables by their logarithmic counterparts, i.e., $x'_s = \log x_s$, $x_s^\text{min} = \log x^\text{min}_s$ , $x_s^\text{max} = \log x^\text{max}_s$, $w'_l = \log w_l$, $w_l^\text{min} = \log w^\text{min}_l$, $w_l^\text{max} = \log w^\text{max}_l$, $U'_j(x_j) = U_j(e^{x_j})$ and $c'_j(w'_j) = c_j(e^{w'_j})$, the first constraint of the problem (1) can be rewritten as

$$\log \sum_{l \in S_l} e^{c'_j(w'_j)} - \log c_j(w_j) - \log p_l = 0, \quad \forall l \quad (5)$$

It can be verified that the term $\log \sum_{l \in S_l} e^{c'_j(w'_j)}$ is convex with respect to $x'$ through the second derivative test. However, the difficulty of solving (1) in a distributed way arises due to the non-separability of the term $\log \sum_{l \in S_l} e^{c'_j(w'_j)}$, although it is a convex function. Accordingly, to overcome this challenge, we introduce a set of new variables

$$\alpha_s = \left\{ \alpha_s \mid 0 \leq \alpha_s \leq 1, \sum_{s \in S(l)} \alpha_s = l \in L, s \in S \right\} \quad (6)$$

where each $\alpha_s$ can be interpreted as the fraction of the traffic from source $s \in S(l)$ over the overall traffic on link $l$. We observe that the first constraint in (1) for each link $l$ is equivalent to a number $|S(l)|$ of separable constraints:

$$x_s \leq \alpha_s c_j(w_j) p_l \prod_{k=1}^{N_S(l)} \left(1 - \sum_{s \in S(l)} \alpha_s \right), \quad \forall s \in S(l), \quad (7)$$

where $|S(l)|$ represents the cardinality of a set.

Similarly, we take logarithm on both sides of (7) and obtain the following optimization problem (P2):

$$\text{max} \sum_{l \in L} U'_j(x'_l) \quad \text{s.t.} \quad x'_s - \log \alpha_s - \log c_j(w'_j) - \log p_l = 0, \quad \forall l \in l, \forall s \in S(l), \quad (8)$$

$$\sum_{s \in S(l)} \alpha_s = 1, \quad \forall l \in L, \forall s \in S(l), \quad 0 \leq \alpha_s \leq 1, \quad \forall l \in L, \forall s \in S(l), \quad \text{and } \sum_{s \in S(l)} \alpha_s = l, \quad \forall l \in L, \forall s \in S(l), \quad (9)$$

for problem (P2) to be a convex optimization problem, we need to check the convexity of the constraint set and the concavity of the objective function, which may not be true for any $U'_j(x'_l)$ even when $U_j(x_j)$ is concave. We define

$$g_s(x_s) = \frac{dU_s(x_s)}{dx_s} x_s + \frac{dU'_j(x'_l)}{dx'} x'_s \quad (10)$$

Then we have the following theorem:
Theorem 1. The constraint set in (8) is convex and if \( g_s(x_s) < 0 \), then the objective function \( U'(x_s') \) is a strictly convex function with respect to \( x_s' \).

Proof. It can be obtained from (3) and (4) that

\[
c_i(w_i) = B \left[ \log G_i + \log w_i - \log \left( \sum_{m \in S(i)} G_m w_m + \sigma_i \right) \right] \quad (10)
\]

Substituting \( w_i' = \log w_i \) into (10), (10) can be rewritten as

\[
c_i(w_i') = B \left[ \log G_i + w_i' - \log \left( \sum_{m \in S(i)} G_m e^{w_i'} + \sigma_i \right) \right] \quad (11)
\]

We can verify that the term \(-\log \left( \sum_{m \in S(i)} G_m e^{w_i'} + \sigma_i \right)\) is a concave function with respect to \( w_i' \) through the second derivative test. Accordingly, we can obtain that \( c_i(w_i') \) is convex with respect to \( w_i' \) while we can verify that \( \sum_{s \in S(i)} \log(\cdot) \) is concave with respect to \( p_m \). Thus the first constraint of (8) is convex. Additionally, because all constraints of (8) are linear and bounded, the constraint set of (8) is convex.

Similar to the proof in [17], since \( x_i = e^{w_i} \),

\[
d^2 U'(x_s') \over dx_i'^2 = d^2 U'_i(x_s') \left( \frac{dx_i}{dx_i'} \right)^2 + dU'_i(x_s') \frac{d^2 x_i}{dx_i'^2} - e^{w_i'} \left( \frac{d^2 U'_i(x_s')}{dx_i'^2} + \frac{dU'_i(x_s')}{dx_i'} \right) + e^{w_i'} g_i(x_s')
\]

Therefore, if \( g_i(x_s) < 0 \), \( U'(x_s') \) is a strictly convex function with respect to \( x_s' \). This completes the proof.

It is obvious that given that Theorem 1 is satisfied, problem (8) is a convex optimization problem, and all log rates are decoupled, which enables the dual decomposition. This indicates that the minimal solution to dual problem is equal to the maximal solution to (8), i.e., there is no duality gap. Thus, it can be efficiently tackled by using modern convex programming schemes [19]. In order to develop a primal-dual iteration which solves the optimization problem P2, we need Lagrangian function with the well-known Lagrangian Multipliers, which can be given by

\[
L(\lambda_s, x_s, p, w, \alpha) = \sum_{s \in S} U'(x_s') - \sum_{i \in S(i)} \bar{\lambda}_i \left[ x_i - \log \alpha_i - \log c_i(w_i') \right] - \log p_i \left[ x_l + \log \left( 1 - \sum_{i \in S(l)} p_i \right) \right]
\]

\[
= \sum_{s \in S} \left( U'(x_s') - \lambda_s x_s' \right) + \sum_{i \in S(i)} \bar{\lambda}_i \log \alpha_i + \sum_{i \in S(i)} \lambda_i \log p_i
\]

\[
+ \sum_{i \in S(i)} \lambda_i \left( 1 - \sum_{s \in S(i)} \alpha_s \right)
\]

where \( \lambda_i := \sum_{i \in S(i)} \lambda_i, \bar{\lambda}_i := \lambda_i | s \in S, l \in S(l) \)

The Lagrangian dual function can be defined as

\[
D(\lambda) = \max L(\lambda, p, w, \alpha) \quad \text{s.t.} \quad 0 \leq \alpha_s \leq 1, \sum_{s \in S(l)} \alpha_s = 1, \forall l \in S(l),
\]

\[
\sum_{m \in L(a)(e)} p_m \leq 1, \forall m, \quad p_m \geq 0, \forall l \in S(l)
\]

Accordingly, the dual problem to (8) is

\[
\min_{\lambda} D(\lambda)
\]

Note that the only complexity involved with the constraint reformulation (7) arises from the fact that each link has to store individual price information per flow going through it, which incurs only a linearly increasing memory usage. However, the communication complexity remains identical to that with a common link price, since the source only requires the sum-price information.

B. Distributed Algorithms

For a given \( \lambda_s \), the maximization of Lagrangian dual function (13) can be decomposed into three subalgorithms: one is congestion control subalgorithm at each source; other two are contention control and power control at each node respectively.

(i) Congestion control problem at each source is

\[
\max \left( U'(x_s') - \lambda_s x_s' \right), \forall s \in S
\]

\[
\text{s.t.} \quad x_s'^{\text{max}} \leq x_s' \leq x_s'^{\text{min}}, \forall s.
\]

The Lagrange multiplier \( \lambda_s \) is interpreted as the price of congestion control per unit of log bandwidth that link \( l \) charges to source \( s \). The strategy of congestion control at source \( s \) is to maximize its net benefit \( U'(x_s') - \lambda_s x_s' \), since \( \lambda_s x_s' \) is just the sum bandwidth cost charged to all links on its path if source \( s \) transmits at log rate \( x_s' \). Since \( U'(x_s') \) is strictly concave over \( x_s' \), a unique maximizer exists. The source rates in congestion control algorithm are adjusted as follows:

\[
x_s'(t+1) = \left[ x_s'(t) + \beta(t) \left( \frac{dU'(x_s')}{dx_s'} - \lambda_s \right) \right]_{x_s'^{\text{min}} \leq x_s' \leq x_s'^{\text{max}}}
\]

where \( \alpha = \min(\alpha, b, c), \quad \beta(t) > 0 \) is a step size.

(ii) The contention control subproblem at the MAC layer and the power control at the physical layer at each node \( n \) for every outgoing link \( l \in L_{n}(e) \) are, respectively

\[
\max \sum_{l \in L_{n}(e)} \lambda_i \log p_i + \sum_{m \in L_{n}(e)} \lambda_i \log \left( 1 - \sum_{l \in L_{n}(e)} p_m \right)
\]

\[
\text{s.t.} \quad \sum_{m \in L_{n}(e)} p_m \leq 1, \forall m, \quad 0 \leq p_i \leq 1, \forall l.
\]

and
Taking the partial derivative of (25) with respect to \( w'_l \), we can obtain the subgradient of the objective function (25) over \( w'_l \) as follows

\[
V_j(w') = B \left[ \lambda_j - w \sum_{l \in S_j} \frac{\lambda_j G_j e^{w'_l}}{G_j w_s + \sigma_j} \right]
\]

(26)

Thus by using the subgradient projection method, the power \( w'_l \) can be updated by

\[
w'_l(t + 1) = \left[ w'_l(t) + \beta(t) V_j(w') \right]^{-}
\]

(27)

It is worth noting that \( \alpha_n \) represents the fraction of the overall traffic on link \( l \) contributed by source \( s \in S(l) \). The solution (21) indicates that \( \alpha_n \) is equal to the normalized price that source \( s \) pays for link \( l \). The higher the price that the source \( s \) is willing to pay at, the more traffic the link allows to pass through. We assume \( \mu(n) \neq 0 \) for the solution in (20). If a packet is successfully transmitted over link \( l \in L_{out}(n) \), this link contributes a reward \( \lambda_l \) to the system in local area. However, all other links \( l' \in L_{in}(n) \cup L_{out}(n) \), \( l' \neq l \) must remain silent during the transmission of link \( l \). If those links can transmit simultaneously without collisions, then the total reward would be \( \sum_{l' \in L_{out}(n)} \lambda_{l'} + \sum_{l' \in L_{in}(n)} \lambda_{l'} \).

But this cannot happen because of interference. Accordingly, the true fraction of the reward in a local area contributed by link \( l \) should be \( \lambda_l \) normalized by the total \( \sum_{l' \in L_{out}(n)} \lambda_{l'} + \sum_{l' \in L_{in}(n)} \lambda_{l'} \), which is equal to \( p_l \) in (20). In other words, the attempt probability of a link is equal to the fraction of the reward it can generate locally with a successful transmission.

Subsequently, we are ready to solve the dual problem (14) by using subgradient project method [20]. At each node \( n \) for \( \forall l \in L_{in}(n) \) and \( \forall s \in S(l) \), the outgoing link prices for sources are adjusted by

\[
\lambda_l(t + 1) = \left[ \lambda_l(t) - \beta(t) \frac{\partial D}{\partial \lambda_l} \left( \lambda_l(t) \right) \right]
\]

(28)

where \( \delta = \max \{0, a\} \). According to (12) and (13), we have

\[
\frac{\partial D}{\partial \lambda_l} = \left[ x'_l - \log \alpha_n - \log c_j(w'_l) \right] - \log p_l \left( 1 - \sum_{l' \in L_{in}(n)} p_{l'} \right)
\]

(29)

Substituting (29) into (28), we obtain the following adjustment rule for link \( \forall l \in L_{in}(n) \) at each node \( n \):

\[
\lambda_l(t + 1) = \left[ \lambda_l(t) - \beta(t) \left( x'_l - \log \alpha_n - \log c_j(w'_l) - \log p_l \left( 1 - \sum_{l' \in L_{in}(n)} p_{l'} \right) \right) \right]
\]
\[ \lambda_n(t+1) = \left[ \lambda_n(t) + \beta(t) \left[ x'_n(t) - \log \alpha_n(t) - \log c_i(w'_i(t)) \right] \right. \\
- \log p_i(t) - \sum_{k \in N'_n(t)} \left( 1 - \sum_{m \in \text{I}_m(k)} p_m(t) \right) \left. \right] \]

where \( x'_n(t) \), \( p_i(t) \), \( \alpha_n(t) \) and \( w'_i(t) \) are the solutions of the problems (15), (17), (21) and (22), respectively. Thus the problems (15), (17), (21) and (22) and (30) can be solved at each node (including source node) in a distributed manner by using local message \( \lambda_n \).

This link price adjustment rule follows the law of supply and demand in a fashion similar to that for congestion control in wireline networks. One apparent difference is that all source \( s \in S(t) \) sharing the same link \( l \) are charged with an individual price, which may be different from others, while in wireline networks all sources using the same link share a common link price.

In the following, we summarize the joint congestion, contention and power control algorithm.

Algorithm 1: The joint congestion, contention and power control algorithm

1. Construct the local interference graph for each node \( n \), and obtain sets \( L_{out}(n) \), \( L_{in}(n) \), \( N'_n \) and \( I'_m \), \( \forall l \in L_{out}(n) \).
2. Set iteration \( t = 0 \) for each node \( n \) (including source \( s \)), let \( \lambda_n(0) = 1 \), \( p_i(0) = \frac{1}{P_{m_{in}}[n]} \), and \( \alpha_n(0) = \frac{1}{\delta} \), and initialize \( x'_n(0) \) and \( w'_i(0) \), \( \forall l \in L_{out}(n) \).
3. At each node \( n \), do
   3.1 Set \( t = t+1 \) and step size \( \beta(t) \);
   3.2 Compute the congestion price \( \lambda_n(t) \) using (30) and inform \( \lambda_n(t) \) all neighbor nodes in \( N'_n \), \( \forall l \in L_{out}(n) \);
   3.3 Compute the congestion control rate \( \lambda_n(t) \) by (16) and computes \( \alpha_n(t) \) using (23);
   3.4 Source \( s \) adjusts its transmission logarithmic rates \( x'_n(t) \) by (17) and computes \( \alpha_n(t) \)
   3.5 Receive the local congestion price \( \lambda_n \) from its neighbor nodes and computes \( \lambda_n(t) = \sum_{l \in L_{out}(n)} \lambda_n(t) \).
   3.6 Compute \( \rho(n,t) = \sum_{l \in L_{out}(n)} \lambda_n(l) + \sum_{l \in L_{in}(n)} \lambda_n(l) \).
   3.7 Update the persistence probability \( p_i(t) \) of its outgoing links using (20);
   3.8 Computes the transmission probability \( P_i(t) = \sum_{l \in L_{out}(n)} p_i(t) \) of node \( n \) and the conditional persistence of each of its outgoing links \( q_i(l) = p_i(t) / P_i(t) \), \( \forall l \in L_{out}(n) \).
   3.9 Node \( n \) decides if it will transmit data with a probability \( P_i(t) \). If it decides to transmit data, it chooses to transmit on one of its outgoing links with a probability \( q_i(t) \), \( l \in L_{out}(n) \);
   3.10 According to the results obtained in 3.6, node \( n \) computes the subgradient \( \nabla_i(w'_i) \) of the objective function (25) over \( w'_i \) using (26);

3.11 Node \( n \) adjusts the transmission logarithmic power \( w'_i(t) \) on the link \( l \in L_{out}(n) \) using (27);
4. Until \( \left| x'_n(t+1) - x'_n(t) \right| < \delta_1 \), \( \left| p_i(t+1) - p_i(t) \right| < \delta_2 \) and \( \left| w'_i(t+1) - w'_i(t) \right| < \delta_3 \), where \( \delta_1 \), \( \delta_2 \) and \( \delta_3 \) are tiny positive real numbers.

Note that the above algorithm is conducted at each node \( n \) to calculate \( P_i(t) \), \( p_i(t) \), \( \lambda_n \), \( w'_i(t) \) and \( x'_n(t) \) for its outgoing link \( l \in L_{out}(n) \). Hence, the above algorithm is carried out at the transmitter node of each node in a distributed manner. If we assume that two nodes within interference range can communicate with each other, each node in the above algorithm requires information from nodes within two-hop distance from it.

To compute \( P_i(t) \) and \( p_i(t) \) for its outgoing link \( l \in L_{out}(n) \), node \( n \) needs \( \lambda_n(t) \) from the transmitter of link \( k \in L_{out}(n) \) that is interfered from the transmission of node \( n \). Also, to calculate \( \lambda_n(t) \) for its outgoing link \( l \in L_{out}(n) \), node \( n \) needs \( P_i(t) \) from node \( k \), \( k \in N'_n \), of its transmission outgoing link \( l \in L_{out}(n) \) is interfered.

Remark. The number of message passing required in the above algorithm depends on the network topology. The average numbers of message passing in each iteration, \( M_i \), is obtained as

\[ M_i = H \sum_{n} \left[ |I'_m(n)| + \sum_{l \in L_{out}(n)} |N'_n(l)| \right] \],

where \( H \) is the average number of hops that each message traverses, \( 1 \leq H \leq 2 \).

For the convergence and optimality of the distributed algorithm, we have the following result.

Theorem 2. Let \( \lambda^* \) denote a minimizer of the dual problem (14). If the following condition is satisfied at the optimal dual solution \( \lambda^* \)

\[ \mu^*(n) = \sum_{l \in L_{out}(n) \cap L_{in}(l)} \lambda'(l) + \sum_{l \in L_{out}(n) \cap L_{in}(l)} \lambda'(l) = 0, \forall n \in N \] 

then there are the limited step sizes \( \{ \beta(t) \}_t^\infty \) guarantee that \( \lambda(t) \) converges to the optimal dual solution \( \lambda^* \), i.e., \( \lim_{t \to \infty} \lambda(t) = \lambda^* \). In addition, at \( \lambda^* \), solutions to (15), (17), (21) and (22) denoted as \( \lambda^* \), \( P^* \), \( \alpha^* \) and \( w^* \) optimize problem (8).

Proof. According to Danskin Theorem [20], if

\[ \frac{\partial D(\lambda)}{\partial \lambda_n} = -\left[ \frac{x'_n - \log \alpha_n - \log c_i(w'_i)}{P_{m_{in}}[n]} \right] \]

is a subgradient of the dual problem (14), then (30) is also a subgradient of the dual problem (14). Thus there exists a step size \( \beta(t) \) (e.g. \( \beta(t) = 1/t \)) to guarantee that \( \lambda(t) \) converges to the optimal dual solution \( \lambda^* \). In fact, the fixed step size (e.g. \( \beta(t) = \omega \forall t \)) can more efficiently
track system variations and more practical for implementation than a diminishing step size. In this case, \( \lambda(t) \) converges to the neighborhood of \( \lambda^* \).

Problem (8) is a convex optimization problem. By the assumption in Theorem 2, problems (15), (17), (21) and (22) have a unique solution \( x^* \), \( p^* \), \( \alpha^* \) and \( w^* \) at \( \lambda^* \), respectively. As a result, it follows from Property 6.5 in [21] that \( x^* \), \( p^* \), \( \alpha^* \) and \( w^* \) are also the optimal solutions for problem (8). This completes the proof.

V. NUMERICAL SIMULATIONS

In this section, we provide numerical examples to complement the analysis in the previous sections. We consider a simple ad hoc network shown in Fig. 1. We assume that only if the distance between the transmitter of one link and the receiver of the other link is less than 2\( d \), then transmission of the first link will cause interference strong enough to influence reception of the second link.

![Figure 1](image1.png)

The wireless ad hoc network topology

Suppose that there are three end-to-end flows in Fig. 1: \( f1: A \rightarrow B \rightarrow C \rightarrow D \rightarrow E \), \( f2: F \rightarrow G \) and \( f3: H \rightarrow I \). The sets, \( L_{\text{out}}(n) \), \( L_{\text{in}}(n) \), \( N_{\text{in}}(l) \) and \( L_{\text{out}}(n) \), can thus be obtained from the given network. In all simulations, we use the step size \( \beta(t) = 1/t \). We test utilities \( U_i(x_i) = (1 - \theta)^{-1} x_i^{1-\theta} \) with \( \theta = 2, 10 \) for source 1, both of which leads to strictly concave transformed utilities \( U'_i(x'_i) = (1 - \theta)^{-1} \exp[x'_i(1 - \theta)] \). We initialize the link price \( \lambda_i \) to 1 and link bandwidth \( B_i = 5\text{MHz} \). We let \( w_{\text{max}} = 0.8\text{mW} \), \( w_{\text{max}}' = 28\text{mW} \), \( x_{\text{max}} = 0\text{Mbps} \), \( x_{\text{max}}' = 5\text{Mbps} \), \( \sigma_i = 8\text{dB} \) and \( G_n = d^{-4} \), where \( d \) denotes the distance between the transmitter of link \( l \) and the receiver of link \( k \).

Table 1 gives the relationships among the iteration number \( t \), the optimal transmission rates \( x_i^* \), the optimal transmission powers \( w_i^* \), and the optimal consistence probability \( p_i^* \) over the different \( \theta \) values. It is shown from Table 1 that the larger the value of \( \theta \) is, the smaller the step size should be, and the faster the convergence speed is. Furthermore, we find that since source 1 traverses more heavily interfered links, compared to other sources, it is allocated to the lowest rate at the optimal rate allocation that maximizes the network utility. However, as the value of \( \theta \) increases, the gap among the sources decreases and the fairness among sources is improved. It is observed that the total power consumed by sources decreases as well and the fairness of link usage is improved with the increase of the value of \( \theta \).

Figures 2 and 3 indicate the evolution of the transmission rates of sources with the numbers of iterations for different \( \theta \). We can observe from figures 2 and 3 that the source rate can converge to the optima after some numbers of iterations and the one traversing the links with better channel conditions (e.g. source 3) converges faster. Meanwhile, it is further verified that the larger the value of \( \theta \) is, and the faster the convergence is.

![Figure 2](image2.png)

The evolution of source rates for \( \theta = 2 \)

![Figure 3](image3.png)

The evolution of source rates for \( \theta = 10 \)

TABLE I. RELATIONSHIPS AMONG \( \theta \) AND \( t, x_i^*, w_i^* \) AND \( p_i^* \)

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>( t )</th>
<th>( x_1^* )</th>
<th>( x_2^* )</th>
<th>( x_3^* )</th>
<th>( w_1^* )</th>
<th>( w_2^* )</th>
<th>( w_3^* )</th>
<th>( p_1^* )</th>
<th>( p_2^* )</th>
<th>( p_3^* )</th>
<th>( p_4^* )</th>
<th>( p_5^* )</th>
<th>( p_6^* )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>116</td>
<td>2.87</td>
<td>3.35</td>
<td>3.66</td>
<td>2.13</td>
<td>6.75</td>
<td>6.71</td>
<td>0.96</td>
<td>0.59</td>
<td>0.46</td>
<td>0.33</td>
<td>0.58</td>
<td>0.52</td>
</tr>
<tr>
<td>10</td>
<td>52</td>
<td>3.01</td>
<td>3.13</td>
<td>3.22</td>
<td>2.29</td>
<td>4.84</td>
<td>4.68</td>
<td>0.95</td>
<td>0.55</td>
<td>0.41</td>
<td>0.37</td>
<td>0.33</td>
<td>0.32</td>
</tr>
</tbody>
</table>
Figures 4 and 5 show the evolutions of the links' persistence probabilities and the sources' transmission powers with the numbers of iterations for $\theta=10$, respectively. The results show that both the links' persistence probabilities and the sources' transmission powers converge to the optima after some numbers of iterations. It is also found that the powers of sources with better channel conditions converge faster and are consumed less. The links with better channel conditions have higher persistence probability and are easier to access.

![Figure 4. The evolutions of the links' persistence probabilities](image)

![Figure 5. The evolutions of the sources' transmission powers](image)

VI. CONCLUSIONS

This paper studied the joint cross-layer design of congestion control at the transport layer, contention control at the MAC layer and power control at the physical layer for ad hoc networks. By taking into account "elastic" link capacity, in this paper, we construct a generalized network utility maximization (GNUM) with the constraints of the link capacity, the persistence probability, the transmission power and the fairness. The original GNUM problem is non-convex and non-separable due to the effect of the persistence probability at the MAC layer on the link capacity at the physical layer. By introducing the auxiliary variables and the variable transformation, we provide a decoupled and dual-decomposable convex formulation and propose subgradient-based cross-layer algorithms to solve the dual problem in a distributed fashion. The proposed algorithms can be decomposed into three layers to implement: the transport layer where sources adjust their end-to-end rates to control congestion, the MAC layer where sources adjust their persistence probability to control contention, and the physical layer where sources control their transmission powers. These three layers interact and are coordinated through link prices to maximize the network utility. We further prove the convergence and optimality of the proposed algorithms. Finally, numerical examples verify the effectiveness and convergence of the algorithms.

REFERENCES

Guangyou Tu was born in 1978. He received his MS degree in Computer Architecture from Chongqing University. He is currently pursuing the Ph.D. degree of Computer Software and Theory at Chongqing University. His research interests include network security, network congestion control.

Songtao Guo was born in 1975. Professor of Southwest University. He received his BS, MS, and Ph.D. degree in Computer Software and Theory from Chongqing University, Chongqing, China, in 1999, 2003, and 2008 respectively. His main research interests include resource allocation in wireless networks, network congestion control, stability analysis and bifurcation control of nonlinear systems with time delays.

Hongyu Huang was born in 1979. He received his Ph.D. degree in Computer Science from Shanghai Jiao Tong University in 2009. He is now an associate professor at Chongqing University. His research interests include vehicular ad hoc networks, mobile ad hoc networks, and wireless sensor networks. He is a member of the IEEE and the IEEE Computer Society.

Yuanyuan Yang is currently a Professor of computer engineering and computer science with Stony Brook University, Stony Brook, NY, and the Director of Communications and Devices Division, New York State Center of Excellence in Wireless and Information Technology.