Thesis for Master Degree

Multi-Timescale Cross-Layer Designs for Wireless Multihop Networks

Pham Quoc Viet

Department of Information and Communications System

Graduate School, Inje University

Advisor: Prof. Won-Joo Hwang

Multi-Timescale Cross-Layer Designs for Wireless Multihop Networks

Pham Quoc Viet

Department of Information and Communications System

Graduate School, Inje University

A thesis submitted to the Graduate School of
Inje University in partial fulfillment of the requirements
for the Degree of Master of Science

Advisor: Prof. Won-Joo Hwang

June. 2015

Approved by Committee of the Graduate School of Inje University in partial fulfillment of the requirements for the Degree of Master of Science

Chairman of Committee: Hee-Cheol Kim (signature)

Committee: WON-JOO HWAHG (signature)

Committee: HyunChul Kim (signature)

Graduate School, Inje University

June. 2015

To my parents, my nephew, and my brother.

Acknowledgements

Almost two years ago I packed my bags, boarded an airplane and moved from Vietnam to Korea. I started my M.S. studies in Information and Communications at Inje University; the last two years have been a long journey and this thesis is a final goal.

Foremost, I would like to thank prof. Won-Joo Hwang. I have really enjoyed working with you, and learning from you what research is all about and how to do research successfully. In addition, I'm very grateful for your financial support and academic guidance. I feel deep gratitude to senior members in Computer Networks Lab.: Huy-Thong, Hoang-Hiep, Hoang-Linh, and Minh-Thuyen, for their constructive discussions of and insightful comments that helped me to enhance the quality of this thesis.

Fortunately, the past two years in Korea have not been only research. I actually had the luck of meeting new people and finding so many new friends. I cannot avoid sending heartfelt thanks to Vietnamese friends, especially Vinh Nguyen, Long Truong, Hoa Vo, at Inje University for the great time we spent together between funny talkings and dinners. I would like to thank Mai who was always besides me in the last two years. My grateful thanks go to you for the support and motivations you showed me.

Special thanks to my lovely family, this thesis would never have been completed if I were not received the constant love, encouragement, and continuous support from my family. Finally, this thesis is dedicated to my grandfather who passed away this year.

ABSTRACT

Multi-Timescale Cross-Layer Designs

for Wireless Multihop Networks

Quoc-Viet Pham

(Advisor: Prof. Won-Joo Hwang)

Department of Information and Communication System

Graduate School, Inje University

Since the publication point of the Kelly's paper, a numerous number of researches have been devoted to resource allocation and cross-layer designs in wired networks as well as wireless networks. Most of cross-layer algorithms are established through the concepts of network optimization, especially, convex optimization, which has been the accumulated results of many researches and many years. In the context of cross-layer optimization, we propose two cross-layer designs in fast-fading lossy delay-constrained wireless multihop networks.

The first cross-layer problem we study is to increase the overall utility and decrease the link delay and power consumption subject to constraints on link rate outage probability, link congestion control, and flow rate conservation, in mobile ad hoc networks. As opposed to previous work, the rate outage probability in this work is based on exactly-closed form; therefore, the proposed method can guarantee the globally optimal solution to the underlying problem. The non-convex formulated problem is transformed into a convex one, which is solved by exploiting the duality

technique. Numerical simulations verify that our proposal can achieve considerable benefit over the existing method.

Conventionally, cross-layer designs with same-timescale updates can work well; however, there is a difference in layers' timescales and each layer normally operates at its corresponding timescale when implemented in real systems. Respecting this issue and realizing the same problem as in the first work, the second cross-layer design we propose takes into account the timescale difference among layers. By using the primal decomposition technique, the network optimization problem is decomposed into subproblems at various layers, from which the proposed algorithm can be implemented in a distributed manner and adheres to the natural timescale difference among layers. Our simulation results show that the proposed design yields higher effective rates, consumes less power and suffers less delay in comparison with the current alternative frameworks. In addition, the design adheres to the natural timescale separation, then improves the convergence speed over the corresponding same-timescale method.

Keywords: Rate Control, Link Delay, Power Allocation, Lossy Links, Rayleigh-fading channels, Multi-Timescale, Cross-Layer Optimization.

Contents

List of Figures

Chapter 1

Introduction

1.1 Background and Motivation

In the traditional layered architecture, the networking framework is divided into seven layers (application layer, presentation layer, session layer, transport layer, network layer, and physical layer) according to the Open Systems Interconnection (OSI) reference model and into four layers (application layer, transport layer, network layer, and network access layer) with the Transmission Control Protocol/Internet Protocol (TCP/IP) model, to implement distinct protocols and services [Kuros and Ross, 2007]. Regarding these models, there exist a limited number of control message between adjacent layers and communication among non-adjacent layers are not permitted. However, cross-layer designs, which can be viewed as a coordination model among five layers, allow to share information and configuration with the other layers in order to increase the network performance, transmission reliability, and reduce latency, bit error rate. In [Fu et al., 2014b], the authors showed that cross-layer optimization can be classified via two different approaches. On the one hand, cross-layer designs can be categorized as non-manager and manger methods according to how to share information in one node. On another perspective, based on how to share cross-layer information among all nodes in a network, cross-layer proposals are composed of centralized and distributed algorithms. These classifications are illustrated pictorially as in Fig. 1.1. So why cross-layer designs are preferred to consider instead of hierarchical layers and what are the advantages and disadvantages of cross-layer designs?

The motivations of cross-layer designs are from the unique, opportunistic, and novel properties of wireless networks. The OSI and TCP/IP are originally designed for wireline communication systems [Srivastava and Motani, 2005a]. Wireless networks, however, is quite different characteristics from wireline networks, for example, noise, link impairment, multipath effects, interference and mobility. Consequently, problems created by wireless networks and solved by the layered architecture may lead to unsatisfied results and poor performance. In addition, due to time-varying wireless links and new issues in wireless networks, cross layer designs need to be considered in order to address these properties by exploiting the dependencies and interactions among layers.

A cross-layer design should realize at least one of four specific problems: security, QoS, mobility, and wireless link adaptation, through the concept of four coordination planes [Carneiro et al., 2004; Foukalas et al., 2008; Fu et al., 2014b].

- Security plane: They security plane coordinates the encryption protocols across layers with the objective of reducing repeated encryption functionality, then decreasing power consumption, processing capability, and transmission delay, and enhancing network performance. Based on the network scenarios and security requirements, cross-layer techniques choose which layer should perform the security protocol and which encryption method, e.g., SSH,

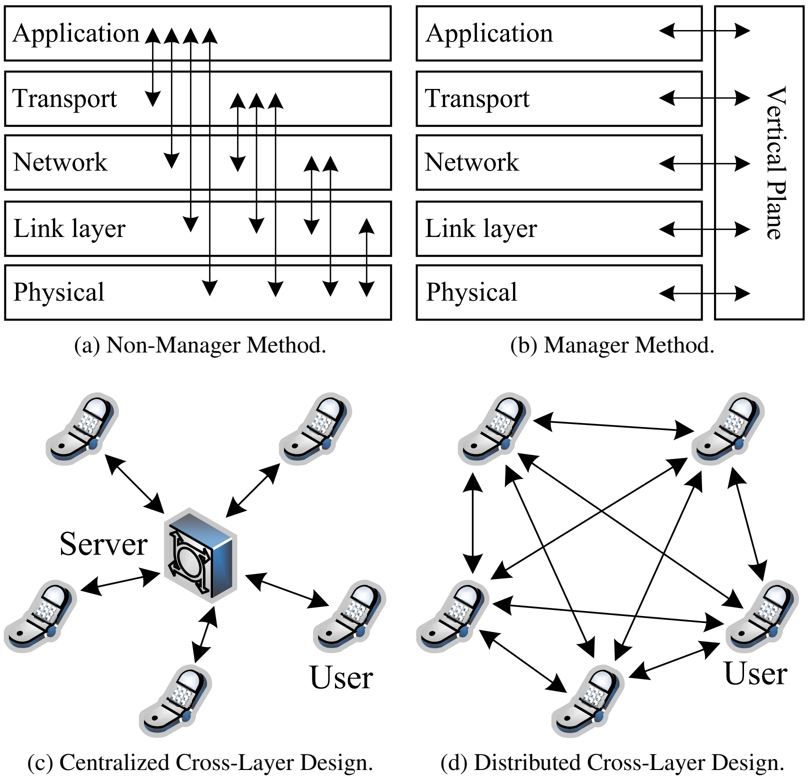


Figure 1.1: Cross-layer design classifications [Foukalas et al., 2008; Fu et al., 2014b; Srivastava and Motani, 2005a] - (a) the non-manager method allows each one of the five layers to directly communicate with each other and this method is attractive to the case of existing a small amount of exchange information, (b) the non-manager method introduces a new abstract layer as a vertical plane or shared data base that can be accessed by all layers and is suitable for vertical calibrations, (c) the centralized cross-layer design uses a central station to interact with exchange information and appeals to cellular networks, (d) there is no central node in the distribute cross-layer design, which is suited to, for instance, sensor networks, and wireless ad hoc networks.

SSL, PGP at the transport layer and application layer, IPSec at the network layer, and Wi-Fi Protected Access (WPA) at the link layer, should be deployed.

- QoS plane: The QoS plane deals with the problem of improving the quality of different services. In the traditional layered architecture, several QoS protocols have been proposed, such as, application/TCP, application/RTP/UDP, and IPQoS [Kuros and Ross, 2007]. In addition, setup QoS requirements of upper layers is not passed to the lower layers and channel information of the lower layers is forbidden to transmit to the upper layers. Due to the time-varying characteristic of wireless links, however, there needs to communicate the channel information state of the link layer and physical layer to the upper layers in order to enhance quality of services in wireless communications.

- Mobility plane: The mobility plane helps to guarantee continuous transmissions due to changes in channel quality since, for example, in wireless ad hoc networks, (mobile) nodes usually move around their current positions. A typical example of mobility is handover and a survey on this problem can be found in [Xenakis et al., 2014]. Handover may lead to, such as, disrupted connection, transmission delay, and link failures, the mobility plane, therefore, should well adapt application services to the underlying wireless channels.

- Wireless link adaptation plane: In comparison to wireline networks, transmissions in wireless networks often suffer higher BER, delay, and packet loss ratio due to channel fading and interference. For example, because of channel fading, TCP window-size can reduce to the minimum value and recovering time may be significant. In these context, cross-layer designs should be aware of physical events so that the network performance is not degraded.

Cross-layer designs gain a great number of advantages; however, there exist some limitations and challenges needed to be further studied and tackled, e.g., cross-layer overhead, coexistence of cross-layer solutions, universal reference, and destruction of the layered architecture [Fu et al., 2014b; Srivastava and Motani, 2005a].

- Cross-layer overhead: Control signaling and message passing exchanged among layers or among transceivers are important to adapt to network dynamics. A large amount of such cross-layer information, however, can occupy much bandwidth and becomes a burden of the network performance. It is also valuable to develop cross-layer designs balancing the trade-off between network performance and cross-layer signaling.

- Coexistence of multiple cross-layer proposals: The consideration of this issue is exposed when a numerous number of cross-layer designs for a same problem, e.g., congestion control and power allocation, have been investigated. Until now, there is no answer on, how can different cross-layer designs for congestion and power control be implemented independently and integrated into a unique framework.

- Universal reference: How to find a cross-layer standard for different applications poses difficult challenges. This issue can addressed by first classifying and finding the common viewpoints of existing proposals and then investigating the cross-layer standard for these applications [Fu et al., 2014b].

- Destruction of the layered architecture: Since objectives of cross-layer designs are to improve the network performance by allowing information exchange among layers, cross-layer frameworks can break the traditional layered architecture. Actually, a small change in a layer may lead to a series changes at the other layers. As a result, we should pay great attention of this issue when designing any cross-layer design.

In this thesis, we study issues involved in the cross-layer optimization designs of fast-fading lossy wireless multihop networks that employ congestion control. We aim at formulating the optimization problems, whose objective are to maximize the system effective utility while minimizing the link delay and total consumed power, over multiple layers (transport layer, link layer, and physical layer) and incorporating wireless resources, for example, transmit power, link rate-outage probability, and flow rate, into the problem as constraints. We first transform the non-convex original optimization problem into convex one by log-transformation. By using the Lagrangian and either the primal decomposition technique or dual decomposition technique, the network optimization problem can be decomposed into subproblems at various layers, which can be efficiently and effectively solved by the duality method and can adhere to the timescale of the corresponding layers. Hence, two distributed cross-layer algorithms are achieved with the helps of message passing.

1.2 Contributions

In this thesis, we focus on developing separate cross-layer designs for fast-fading lossy wireless multihop networks. These designs are to maximize the aggregate effective utility while minimizing the average link delay and total power consumption. Our contributions are, as follows:

- To guarantee the globally optimal solutions to the cross-layer problem of congestion, link delay, and power control, we realize the exactly-closed form of rate-outage probability. This design is called novel RENUM (nRENUM). Our simulation results illustrate that nRENUM can achieve greater performance in terms of injection rate, effective rate, total consumed power, flow delay, and link delay compared to the existing alternative framework, RENUM [Guo et al., 2014]. This work was published in Journal of Korea Multimedia Society in Feb.

2015 [Pham et al., 2015a].

- We introduce a cross-layer design adhering to the timescale separation among layers for wireless ad hoc networks, called Multi-Timescale RENUM (MTSRENUM). MTSRENUM can guarantee the globally optimal solutions, which is similar to nRENUM, while considering the timescale difference. The subproblems at the transport layer, link layer, and physical layer are obtained by the primal decomposition technique and operated at their corresponding timescale, e.g., long-timescale, mid-timescale, and short-timescale. Through the simulations and comparison with current frameworks, such as, [Guo et al., 2014], [Gao et al., 2009], [Wang et al., 2013], [Pham et al., 2015a], we find that MTSRENUM not only yields higher injection rates, effective rates, suffers less delay and power consumption, but also proves, the multi-timescale algorithms often converge faster than the corresponding same-timescale algorithms. This work has been under revision of Elsevier, Computer Networks in June 2015 [Pham et al., 2015b].

1.3 Thesis Outline

The thesis is divided into four chapters. In chapter 2, we develop a cross-layer design of congestion control, average link delay, and power control, that can guarantee the globally optimal solutions to the underlying problem. The following chapter presents another aspect of cross-layer design, which not only guarantees the globally optimal solutions, but also adheres to the timescale difference among layers. We finally conclude in the chapter 4.

Chapter 2

Globally Optimal Solutions for Cross-Layer Design in Fast-Fading Lossy Delay-Constrained MANETs

2.1 Introduction

Traditionally, protocols for wireless networks are relied on the strictly-layered structure and implemented isolatedly. Followed from the seminal work on resource allocation in wired networks proposed by Kelly [Kelly et al., 1998], a lot of researches have been studied and devoted to showing the significant benefits of cross-layer designs (CLDs). Unlike wired networks, resource allocation in wireless networks is critical due to, e.g., scare resource, interference and environment disturbance. Network Utility maximization (NUM) framework has been seen as the efficient and versatile tool to deal with CLD problems, for example, routing at the network layer, congestion control at the transport layer and power control at the physical layer. By jointly using NUM and CLD, the problem can be decoupled and the algorithm can be implemented in a distributed manner.

The lossy feature was first taken into consideration in the effective network utility maximization (ENUM) framework [Gao et al., 2009], where the transmission rate at the source is called the injection rate and the correctly received data rate at the destination is called the effective rate. Nevertheless, since ENUM does not consider the eects of the transmission power and take it into account of the optimization objective, transmit powers are not adjusted dynamically according to the channel conditions. Followed from [Gao et al., 2009], in [Wang et al., 2013], the authors considered the problem of congestion control in interference-limited wireless networks with power control, namely ENUM with power control (ENUMP). ENUMP, however, still does not integrate the power control and consider the link delay in the optimization objective. Guo et al. [2014] examined the effects of lossy features on the power control and link delay as well, namely rate-effective NUM (RENUM), with constraints on rate outage probability, data rate reduction and delay-constrained traffics, by taking them into consideration of the objective function. In [Guo et al., 2014], the rate outage probability is, however, based on approximated form; therefore, RENUM may produce suboptimal solutions to the problem [Pham et al., 2015a].

In this chapter, we propose a novel RENUM (nRENUM) which can generate the globally optimal solutions to RENUM, in fast-fading lossy delay-constrained mobile ad hoc networks (MANETs). Summarily, Our main contributions and the considerable differences of this chapter can be listed, as follows:

- In section 3, we show the network model and then formulate a joint optimization problem of congestion control, link delay and power allocation with constraints on rate outage probability, link delay and lossy rate. As opposed to RENUM, in nRENUM, the rate outage probability is based on exactly closed form; therefore, nRENUM can guarantee the globally optimal solutions to the underlying problem. nRENUM is solved by the duality techniques in section 4.

- In section 5, we investigate the numerical simulation to further verify the outperformance of nRENUM compared to RENUM.

2.2 Related Work

Recent researches whose principles focus on designing optimal CLD policies have been proposed [Bui et al., 2008; Chiang, 2005; Fu et al., 2014a; Gao et al., 2009; Guo et al., 2014; Papandriopoulos et al., 2008; Pham and Hwang, 2014; Tran et al., 2013; Wang et al., 2013], ranging from wired networks to wireless networks. The seminal work on network resource allocation was first proposed by Kelly et al. in [Kelly et al., 1998]. In [Chiang, 2005], the authors analyzed a joint design of optimal congestion and power control and made use of the high-SIR regime to transform the non-convex underlying problem to a convex one. Due to assuming that link transmissions are orthogonal, this design is not suitable for interference wireless networks. A survey on and challenges of CLDs in wireless networks can be found in [Fu et al., 2014a]. A framework of congestion control and power allocation with an outage probability in fast-faded wireless channels has been studied in [Tran et al., 2013]. Like [Papandriopoulos et al., 2008], the non-convex underlying problem is transformed into a new convex problem and then solved by the duality method.

In addition, the authors proposed a successive convex approximation method in order to turn the original problem to approximated convex problems and keep the TCP stack.

Gao et al. [2009] first investigated the leaky-pipe ow model, called ENUM, where the transmission rate of a ow decreases along its route. Henceforth, two schemes: ENUM with link outage probability and ENUM with path outage probability have been considered to examine the eects of lossy wireless links. However, power control is not examined and just xed regardless what the current channel quality is. In addition, ENUM is not suitable for interference-limited wireless environments. Followed from [Guo et al., 2014], S. Guo et al. proposed RENUM framework in which the overall transmission power and the link average delay are also used in the optimization objective with a constraint on the link delay requirement. Nonetheless, RENUM may just produce the suboptimal points. Our goal in this chapter is to present a design that can provide the accurately optimal solutions to the RENUM problem.

2.3 System Model

2.3.1 Network Model

We consider a WANET with L logical links and S sources. Let $\Phi_s = \{1, 2, ..., S\}$ and $\Psi_l = \{1, 2, ..., L\}$ denote sets of sources and links, respectively. Let $L(s)$ be the set of links that flow s uses and $S(l)$ be the set of sources using link l.

Each flow is associated with a utility function which is assumed to be strictly concave, nondecreasing, continuously differentiable. We consider a family of utility functions which have been thoroughly discussed in [Mo and Walrand, 2000].

The instantaneous capacity on link l is modeled by the Shannon capacity

$$C_l(\mathbf{P}) = W \log \left(1 + \zeta \gamma_l(\mathbf{P})\right), \tag{2.1}$$

where $\boldsymbol{P}$ is a vector of transmission powers, W is the baseband bandwidth, ζ is a constant value depending on particular modulation, coding scheme, and bit-error-rate. Here $\gamma_l(\boldsymbol{P})$ is the instantaneous signal-to-interference-plus-noise ratio (SINR) at link l, which is

$$\gamma_l(P) = \frac{p_l G_{ll} F_{ll}}{\sigma_l^2 + \sum_{k \neq l} p_k G_{lk} F_{lk}}, \tag{2.2}$$

where σ_l^2 is the thermal noise power at the receiver on link l, $\sum\limits_{k \neq l} p_k G_{lk} F_{lk}$ is the interference experienced at the receiver on link l. Similar to [Tran et al., 2013], we consider the non-light-of-sight propagation and the average link SINR and capacity by utilizing the statistics of the SINR γ_l. Then,

$$\overline{\gamma}_l(P) = \frac{E\left[p_l G_{ll} F_{ll}\right]}{E\left[\sigma_l^2 + \sum_{k \neq l} p_k G_{lk} F_{lk}\right]} = \frac{p_l G_{ll}}{\sigma_l^2 + \sum_{k \neq l} p_k G_{lk}}, \tag{2.3}$$

where exponentially random variables F_{lk} are assumed to be independent and identically distributed (i.i.d) and $E\left[F_{lk}\right] = 1, \forall k$. Accordingly, $\overline{C}_l(P) = W \log(1 + \zeta \overline{\gamma}_l(\mathbf{P}))$.

2.3.2 Average Delay

We consider the average link delay as a criteria in the optimization problem. Followed from [Gao et al., 2009], each link is modeled as a $M/M/1$ queuing system. Let $\tau(l)$ be the sum of the transmission delay and queuing delay on link l. The average packet delay [Bertsekas et al., 1987;

Guo et al., 2014] on link l is

$$E(\tau(l)) = \frac{K}{(C_l(\mathbf{P}) - \sum_{s\in S(l)} x_s},\tag{2.4}$$

where K (bits) is the mean of exponentially distributed rate of the input Poisson process and x_s is the transmission rate of flow s. For delay-sensitive applications, it is required that the upper bound on the link delay is lower than a threshold, i.e., we have $E(\tau(l)) \leq \upsilon_l$. Therefore,

$$\sum_{s\in S(l)} x_s \leq C_l(\mathbf{P}) - \frac{K}{\upsilon_l}.\tag{2.5}$$

2.3.3 Rate-Outage Probability and Effective Rate

It is required to re-track the instantaneous SINR and compelled to rerun the algorithm to seek the optimal solutions when channel states change. It is not efficient and impractical, especially, for the fast-fading environments. To overcome this issue, we consider the term of outage probability [Goldsmith, 2005], which is Pr$(\gamma_l < \gamma_l^{th})$ where γ_l^{th} is a target minimum SINR that below which performance becomes unacceptable.

We formulate the outage probability, as follows:

$$\text{Pr}(\gamma_l \leq \gamma_l^{th}) = 1 - \phi_l(P),\tag{2.6}$$

where $\phi_l(P)$ can be viewed as reduction of data rate over link l. In the Rayleigh fading model, the exactly closed-from expression [Kandukuri and Boyd, 2002; Papandriopoulos et al., 2008] of

$\phi_l(P)$ is given as

$$\phi_l(P) = \exp\left(-\frac{\sigma_l^2 \gamma_l^{th}}{p_l G_{ll}}\right) \prod_{k \neq l} \left(1 + \gamma_l^{th} \frac{p_k G_{lk}}{p_l G_{ll}}\right)^{-1}. \tag{2.7}$$

Let ϵ_l is the maximal outage probability of link l. Then a combination of (2.6) and (2.7) leads to the following equation

$$\prod_{k \neq l} \left(1 + \gamma_l^{th} \frac{p_k G_{lk}}{p_l G_{ll}}\right) \leq \Omega_l(p_l), \tag{2.8}$$

where $\Omega_l(p_l) = (1 - \epsilon_l)^{-1} \exp\left(-\frac{\sigma_l^2 \gamma_l^{th}}{p_l G_{ll}}\right)$.

We consider the leaky-pipe flow model [Gao et al., 2009], where the transmission rate of each flow changes hop by hop and decrease along its route. The effective rate y_s at the destination is calculated as the multiplication of the outage probability on links that flow s traverses and the injection rate x_s, i.e., $y_s = x_s \prod_{l \in L(s)} [1 - \Pr(l, \gamma_l)]$. We assume that data generated by source s travels H_s hops before reaching the receiver. The data rate at link i of flow s is x_s^i and specified, as follows:

$$x_s^{i+1} = x_s^i \left(1 - \Pr\left(l(s,i)\right)\right), \ i = 1, ..., H_s, \tag{2.9}$$

where $\Pr\left(l(s,i)\right)$ is the outage probability on link i.

2.3.4 Optimization Problem

We formulate a joint cross-layer problem with the optimization objective of maximizing the total effective utility and minimizing the transmission consumption and total link delay, as follows:

$$\max_{x_s, p_l, \upsilon_l} \sum_{s \in \Phi_s} \mathrm{U}_s(x_s^{H_s+1}) - \omega_1 \sum_{l \in \Psi_l} p_l - \omega_2 \sum_{l \in \Psi_l} \upsilon_l \tag{2.10}$$

$$\text{s.t.} \quad \chi = \left\{ x_s | x_s^{\min} \leq x_s \leq x_s^{\max} \right\} \ \forall s, \tag{2.11}$$

$$\Gamma = \left\{ p_l | p_l^{\min} \leq p_l \leq p_l^{\max} \right\} \ \forall l, \tag{2.12}$$

$$\prod_{k \neq l} \left(1 + \gamma_l^{th} \frac{p_k G_{lk}}{p_l G_{ll}} \right) \leq \Omega_l(p_l) \ \forall l, \tag{2.13}$$

$$\sum_{s \in S(l)} x_s \leq C_l(\overline{\gamma}_l(\mathbf{P})) - \frac{K}{\upsilon_l} \ \forall l, \tag{2.14}$$

$$x_s^{i+1} \leq x_s^i \left(1 - \Pr\left(l(s,i)\right)\right) \ i = 1, 2, ..., H_s, \forall s, \tag{2.15}$$

where ω_1 and ω_2 are the costs per unit of consumed power and suffered delay, respectively. The above optimization problem can be explained as follows. The first and second constraint are feasible sets of flow rate and transmission power, respectively. The third one is the constraint on the rate outage probability. The fourth constrain is the requirement on the link average delay while the last one is a constraint due to the lossy nature of wireless links.

The problem (2.13) is not a joint convex problem due to the existence of the third and fifth constraints. Therefore, KKT conditions are just necessary, not sufficient for solution optimality [Boyd and Vandenberghe, 2009]. To ease the optimization problem, we transform it into a new one, as follows:

$$\max_{\hat{x}_s,\hat{p}_l,\hat{\upsilon}_l} \sum_{s\in\Phi_s} \mathrm{U}_s(e^{\hat{x}_s^{H_s+1}}) - \omega_1 \sum_{l\in\Psi_l} e^{\hat{p}_l} - \omega_2 \sum_{l\in\Psi_l} e^{\hat{\upsilon}_l} \tag{2.16}$$

$$\text{s.t.} \quad \sum_{k\neq l} \log\left(1+\gamma_l^{th}\frac{e^{\hat{p}_k}G_{lk}}{e^{\hat{p}_l}G_{ll}}\right) \leq \log\Omega_l(e^{\hat{p}_l})\ \forall l, \tag{2.17}$$

$$\sum_{s\in S(l)} e^{\hat{x}_s^l} \leq C_l(\overline{\gamma}_l(e^{\hat{\mathbf{P}}})) - \frac{K}{e^{\hat{\upsilon}_l}}\ \forall l, \tag{2.18}$$

$$\hat{x}_s^{i+1} \leq \hat{x}_s^i - \psi_{l(s,i)}(\hat{\mathbf{P}})\ i=1,2,...,H_s, \forall s, \tag{2.19}$$

where

$$x_s^l = e^{\hat{x}_s^l},\ \hat{\chi} = \left\{\hat{x}_s | \log\left(x_s^{\min}\right) \leq \hat{x}_s \leq \log\left(x_s^{\max}\right)\right\}, \tag{2.20}$$

$$p_l = e^{\hat{p}_l},\ \hat{\Gamma} = \left\{\hat{p}_l | \log\left(p_l^{\min}\right) \leq \hat{p}_l \leq \log\left(p_l^{\max}\right)\right\}, \tag{2.21}$$

$$\upsilon_l = e^{\hat{\upsilon}_l}, \tag{2.22}$$

$$x_s^{i+1} \leq x_s^i \phi\left(l(s,i)\right),\ \psi_{l(s,i)}(\hat{\mathbf{P}}) = -\log\left(\phi\left(l(s,i)\right)\right). \tag{2.23}$$

In order to make the optimization problem (2.16) convex, we assume the following assumption

$$\frac{d^2U_s(x_s)}{dx_s^2} + \frac{dU_s(x_s)}{dx_s} < 0. \tag{2.24}$$

Then, the problem (2.16) is the convex optimization problem. Furthermore, strong duality also holds for this problem.

2.4 nRENUM Distributed Algorithm

We will apply the Lagrangian dual method to solve the problem (2.16). Here, let λ, μ, ν be the Lagrangian multipliers that associate, respectively, with the constraints (2.17), (2.18), and (2.19). Then the Lagrangian function is defined, as follows:

$$\begin{aligned} L(\hat{x}, \hat{\upsilon}, \hat{\mathbf{P}}, \lambda, \mu, \nu) &= \sum_{s\in\Phi_s} \mathrm{U_s}(e^{\widehat{x}_s^{H_s+1}}) - \omega_1 \sum_{l\in\Psi_l} e^{\widehat{p}_l} - \omega_2 \sum_{l\in\Psi_l} e^{\widehat{\upsilon}_l} \\ &+ \sum_l \lambda_l \left(\log \Omega_l \left(e^{\widehat{p}_l}\right) - \sum_{k\neq l} \log \left(1 + \gamma_l^{th} \frac{e^{\widehat{p}_k} G_{lk}}{e^{\widehat{p}_l} G_{ll}}\right)\right) \\ &+ \sum_l \mu_l \left(C_l(e^{\widehat{\mathbf{P}}}) - \sum_{s\in S(l)} e^{\widehat{x}_s^l} - \frac{K}{e^{\widehat{\upsilon}_l}} \right) + \sum_s \sum_{i=1}^{H_s} \nu_s^i \left(\hat{x}_s^i - \hat{x}_s^{i+1} - \psi_{l(s,i)}(\hat{\mathbf{P}})\right). \end{aligned} \tag{2.25}$$

The Lagrangian dual function is given as

$$g(\lambda, \mu, \nu) = \max_{\hat{x}, \hat{\upsilon}, \hat{\mathbf{P}}} L(\hat{x}, \hat{\upsilon}, \hat{\mathbf{P}}, \lambda, \mu, \nu). \tag{2.26}$$

Accordingly, the dual problem is, as follows:

$$\min_{\lambda, \mu, \nu} g(\lambda, \mu, \nu). \tag{2.27}$$

Due to the separable nature of the Lagrangian function (2.25), we can use the decomposition method to derive separate subfunctions, as follows:

$$L(\hat{x}, \hat{v}, \hat{\mathbf{P}}, \lambda, \mu, \nu) = L(\hat{x}, \mu, \nu) + L(\hat{v}, \nu) + L(\hat{\mathbf{P}}, \lambda, \mu, \nu), \tag{2.28}$$

where

$$L(\hat{x}, \mu, \nu) = \sum_{s \in \Phi_s} \mathrm{U_s}(e^{\widehat{x}_s^{H_s+1}}) - \nu_s^{H_s} \hat{x}_s^{H_s+1} + \sum_s \sum_{i=1}^{H_s} \hat{x}_s^i (\nu_s^i - \nu_s^{i-1}) - \sum_l \mu_l \sum_{s \in S(l)} e^{\hat{x}_s^l}, \tag{2.29}$$

$$L(\hat{v}, \nu) = -\omega_2 \sum_l e^{\hat{v}_l} - \sum_l \mu_l \frac{K}{e^{\hat{v}_l}}, \tag{2.30}$$

$$\begin{aligned} L(\hat{\mathbf{P}}, \lambda, \mu, \nu) = &-\omega_1 \sum_l e^{\hat{p}_l} + \sum_l \mu_l C_l(e^{\hat{\mathbf{P}}}) - \sum_l \nu_l \psi_l(e^{\hat{\mathbf{P}}}) \\ &+ \sum_l \lambda_l \left(\log \Omega_l \left(e^{\widehat{p}_l} \right) - \sum_{k \neq l} \log \left(1 + \gamma_l^{th} \frac{e^{\widehat{p}_k} G_{lk}}{e^{\widehat{p}_l} G_{ll}} \right) \right). \end{aligned} \tag{2.31}$$

The subfunction (2.29) is the congestion control problem which intends to determine flow rates at each link and at the destination. The subfunction (2.30) is the delay-constrained problem which aims at specifying link delays. The last one subfunction (2.31) that is the resource allocation problem which determines the transmission power of each link. Problems are all interacted through the dual variables.

The dual problem (2.27) can be solved by the subgradient method, from which the proposed method can be implemented in a distributed manner. Let $\delta(t)$ be a positive scalar stepsize.

Data rate: updates as

$$\hat{x}^i(t+1) = \left[\hat{x}^i(t) + \delta(t)\nabla L(\hat{x}_s^i)(t)\right]_{\hat{\mathcal{X}}},\ i = 1, 2, ..., H_{s+1}, \tag{2.32}$$

where $\nabla L(\hat{x}_s^i)(t)$ is the subgradient of L with respect to x_s^i. We have

$$\nabla L(\hat{x}_s^i)(t) = \nu_s^i(t) - \nu_s^{i-1}(t) - \mu_{l(s,i)} e^{\hat{x}_s^i(t)}, \tag{2.33}$$

for $i = 1, ..., H_s$. For $i = H_{s+1}$,

$$\nabla L(\hat{x}_s^{H_s+1})(t) = U_s^{'}(e^{\hat{x}_s^{H_s+1}}) e^{\hat{x}_s^{H_s+1}} - \nu_s^{H_s}, \tag{2.34}$$

where $U_s^{'}(.)$ is the first derivative of utility function.

Link delay: the link delay is updated as

$$\upsilon_l(t+1) = \left[\mu_l(t)\frac{K}{\omega_2}\right]^{1/2}. \tag{2.35}$$

Link transmit power:

$$p_l(t+1) = \left[\frac{M_l(t) + (\lambda_l(t) + \nu_l(t))\,\sigma_l^2 m_l^{th}(t)}{\omega_1 + \sum\limits_{k\neq l} G_{kl}(m_k(t) + (\lambda_k(t) + \nu_k(t))\frac{G_{kl} m_k^{th}(t)}{1+G_{kl} m_k^{th}(t) p_l(t)})}\right]_{p_l^{\min}}^{p_l^{\max}}, \tag{2.36}$$

where $[x]_a^b = \max\{\min\{x, b\}, a\}$,

$$M_n(t) = W\mu_n(t)\frac{\zeta\overline{\gamma}_n(t)}{1+\zeta\overline{\gamma}_n(t)},\ m_n(t) = M_n(t)\frac{\overline{\gamma}_n(t)}{G_{nn}p_n(t)},\ m_n^{th}(t) = \frac{\overline{\gamma}_n^{th}(t)}{G_{nn}p_n(t)}. \tag{2.37}$$

Lagrange multipliers:

$$\lambda_l(t+1) = \left[\lambda_l(t) - \delta(t)\nabla L(\lambda_l)(t)\right]^+, \tag{2.38}$$

where $[z]^+ = \max\{z, 0\}$ and $\nabla L(\lambda_l)(t)$ is the subgradient of L with respect to λ_l and given by

$$\nabla L(\lambda_l)(t) = \log \Omega_l\left(e^{\widehat{p}_l(t)}\right) - \sum_{k\neq l}\log\left(1+\gamma_l^{th}\frac{e^{\widehat{p}_k(t)}G_{lk}}{e^{\widehat{p}_l(t)}G_{ll}}\right). \tag{2.39}$$

$$\mu_l(t+1) = \left[\mu_l(t) - \delta(t)\left(C_l(e^{\widehat{\mathbf{P}}(t)}) - \sum_{s\in S(l)} e^{\widehat{x}^l_s(t)} - \frac{K}{e^{\widehat{v}_l(t)}}\right)\right]^+. \tag{2.40}$$

$$\nu_s^i(t+1) = \left[\nu_s^i(t) - \delta(t)\left(\hat{x}_s^i(t) - \hat{x}_s^{i+1}(t) - \psi_{l(s,i)}(\hat{\mathbf{P}}(t))\right)\right]^+. \tag{2.41}$$

From above updates, we propose an iterative algorithm as illustrated in Algorithm 1.

Here, we make several remarks on the nRENUM iterative algorithm.

Remark 1.

- *As stated, the rate outage probability constraint in nRENUM is in rightly and explicitly closed-form; therefore, nRENUM can provide exactly optimal solutions. In a meanwhile, that of RENUM is based on approximated form, so there is no guarantee of the globally optimal solutions.*

Algorithm 1 nRENUM iterative algorithm

Initialization

1) Initialize $t = 0$, $x_s^l = x_s^l(0)$, $\upsilon_l = \upsilon_l(0)$ $p_l = p_l(0)$, $\lambda_l = \lambda(0)$, $\mu_l = \mu_l(0)$, $\nu_s^i = \nu_s^i(0)$, which are required to be non-negative.

Iteration

1) At each link l of flow s, the link transmits with date rate $\hat{x}_s^l$. The transformed link rate and the transformed effective rates are updated via (2.32) with gradients computed by (2.33) and (2.34), respectively. Then, it is come back to the solution space as $x_s^l = \exp(\hat{x}_s^l)$.

2) At each link l, for a given dual price $\mu_l(t)$, the link delay is updated based on (2.35).

3) At each link l, the link transmitter uses power $p_l(t)$ and for given (λ, μ, ν), the transmission power is updated by Eq. (2.36).

4) The Lagrange multipliers are updated via (2.38), (2.40), and (2.41).

5) Set t increase by 1, $t = t + 1$, and go to the next step. The iteration stops when satisfying the termination condition.

- *The proposal can be implemented in a distributed manner through message passing among links. At each link message-passing components (e.g., $m_k^{th}(t)$, $m_k(t)$, $\lambda_k(t)$ and $\nu_s^k(t)$) are computed based and then broadcast to the other links. The transmitter on link l receives broadcast message passing from all the other links, estimates the channel gain G and then updates its transmission power via* (2.36).

Theorem 1. ***Convergence of the nRENUM algorithm****: Let $x(0)$, $\upsilon(0)$, $\boldsymbol{P}(0)$, $\lambda(0)$, $\mu(0)$, and $\nu(0)$ are initial values. Here, $x(t)$, $\upsilon(t)$, $\boldsymbol{P}(t)$, $\lambda(t)$, $\mu(t)$, and $\nu(t)$ are the sequences generated by* (2.32), (2.35), (2.36), (2.38), (2.40), *and* (2.41). *If the step size satisfies the diminishing rule (i.e., $\delta_t > 0$, $\sum_{t=1}^{\infty} \delta_t^2 < \infty$, $\sum_{t=1}^{\infty} \delta_t = \infty$), there actually exists a sufficiently large T_0 that $\forall t \geq T_0$, $x(t)$, $\upsilon(t)$, $\boldsymbol{P}(t)$, $\lambda(t)$, $\mu(t)$, and $\nu(t)$ converge to the globally optimal points.*

Proof. The proof is omitted, see [Bertsekas and Tsitsiklis, 2000], [Bertsekas and Tsitsiklis, 1989]. □

2.5 Simulation Results

This section represents examinations of the performance of the proposed method in comparison with the alternative framework [Guo et al., 2014]

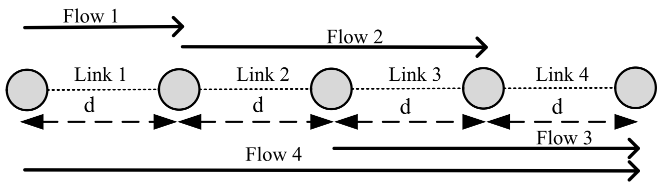


Figure 2.1: Physical and logical topology used for simulation.

2.5.1 Simulation Settings

We consider a MANET composed of five nodes, four links and four flows with the network topology as illustrated in Figure 1. Nodes are separated and placed equidistantly at d = 50 meters. The outage probability thresholds and SINR thresholds for links are set to (0.20, 0.20, 0.20, 0.20) and (1.0, 1.0, 1.0, 1.0), respectively. The fast-fading channel gain is assumed to be i.i.d. while the slow-fading channel gain is assumed to be $g_{lk} = g_0(d_{lk}/50)^{-AF}$, where $d_l k$ is the distance between transmitter of link k and receiver of link l, $AF = 4$ is the path loss attenuation factor, and g_0 is the reference channel gain at a distance 50 meters and meets a condition that the average receive SINR at 50 meters is 30 dB. Without loss of generality, weights ω_1 and ω_2 are assumed to be 1.

2.5.2 Performance of nRENUM and compared framework

We now compare the proposed method with the framework [Guo et al., 2014]. We use the Jains index, which is one of the most widely studied fairness measures and dened as $f(X) = \left(\sum_{n=1}^{N} x_n\right)^2 / \left(N \sum_{n=1}^{N} x_n^2\right)$ where $X = [x_1, x_2, ..., x_N]$ and $0 \leq f(X) \leq 1$. We keep the other parameters fixed and change the value, from 1 to 12, to examine its eects on the fairness. The fairness comparison is illustrated in Figure 2.2 where both fairness indices changes when varies. Specifically, nRENUMs fairness increases with the increment while that of RENUM decreases. In addition, nRENUM can achieve better fairness when $\alpha \geq 4$. Furthermore, we can observe that two fairness indices are almost the same when $\alpha = 4$; therefore, to compare the performance of RENUM and nRENUM, we use $\alpha = 4$.

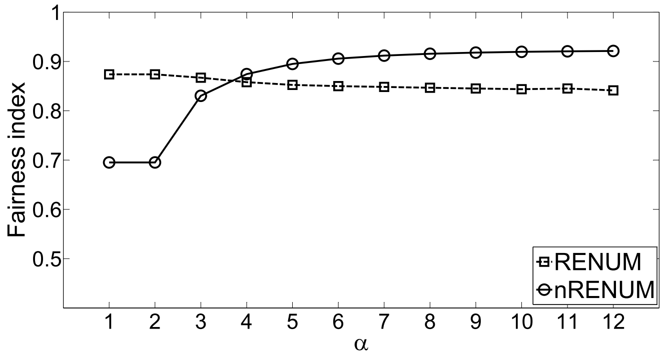
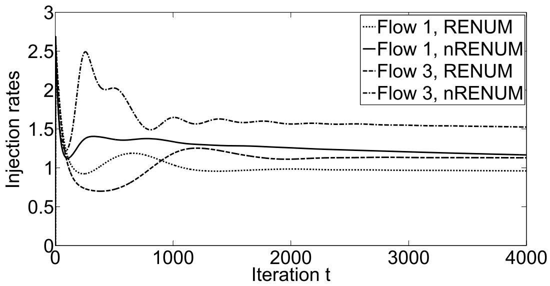


Figure 2.2: Fairness: nRENUM vs RENUM.

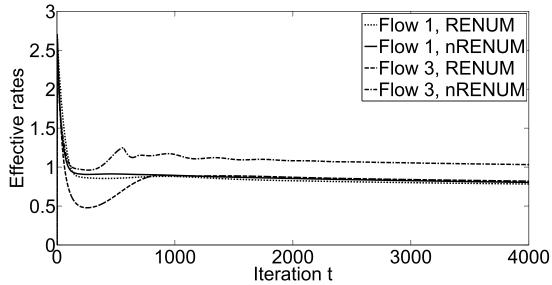
Figure 2.3a and 2.3b represent the comparison of the injection rates and effective rates. We can observe that the ow effective rates become smaller and smaller along its routes when compared to its injection rates. In addition, a ow traversing a larger number of hops suffers higher data rate loss compared to a ow traveling less hops (e.g., flow 1 traverses 1 link and flow 3 travels 2 links). It is due to the lossy nature of wireless links. The total transmission power is depicted in Figure 2.3c, where we can realize that RENUM consumes more powers than nRENUM. The flow-3 and flow-4 delays are demonstrated in Figure 2.3d, where packets in RENUM experience longer delays than that in nRENUM. To be more specific, we examine delays experienced by packets traveling through link 3 and link 4, which is represented in Figure 2.3e. For that, should a packet travel through both link 3 and link 4 in RENUM and nRENUM respectively, it will sustain much rate losses on link 3 and in RENUM.

The above results are since the constraint on rate outage probability in RENUM reduces the original solution region and uses the approximated form. In a meanwhile, the rate outage probability constraint in nRENUM is in the rightly closed-form; therefore, nRENUM can provide the globally optimal solutions. Moreover, the framework [Guo et al., 2014] does not take into consideration the SINR threshold γ_l^{th}, so RENUM can not vary the SINR thresholds for different links to get the appropriately optimal solutions.

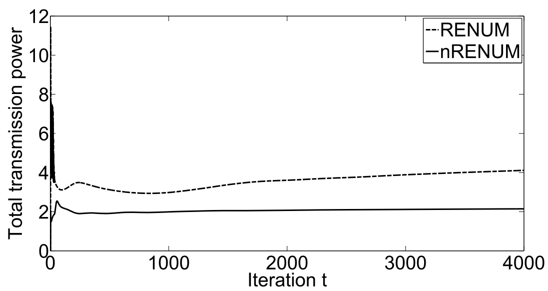
We continue exploring the impacts of the hop number on the performance of RENUM and nRENUM by varying flow-3 H_s, from 2 to 6. The ow-3 effective rates are described in Figure 2.4a, from which we can realize that the effective rates decrease as the number of hops increases and when H_s reaches the sufficiently large number (e.g., $H_s = 5$), the effective rates decrease to the minimum rates. This result can be explained by the lossy feature of the wireless links. Another observation shown in Figure 2.4b is that the ow delays increase with the increment of H_s. Obviously, the ow delays experienced in nRENUM are lower than that in RENUM and nRENUM provides better the effective rate in comparison to RENUM.



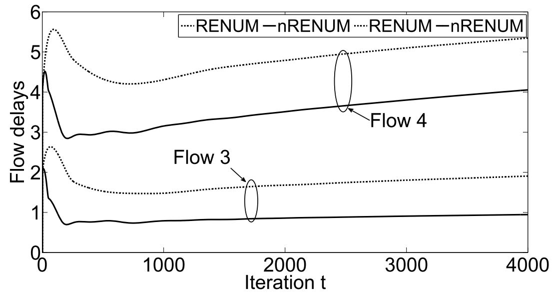
(a) Injection rates.



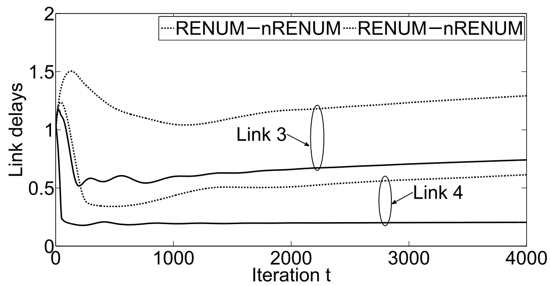
(b) Effective rates.



(c) Total transmission power.



(d) Flow delays.

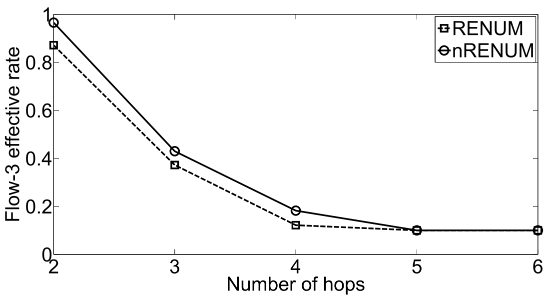


(e) Link delays.

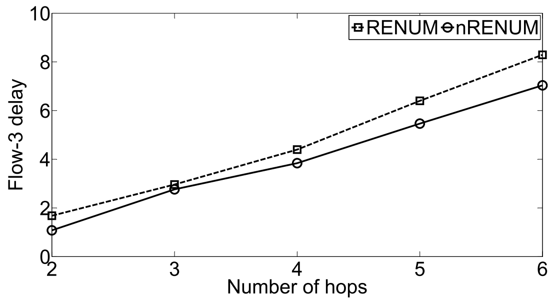
Figure 2.3: The performance comparison between nRENUM and RENUM.

2.6 Conclusion

This chapter studied the cross-layer problem of congestion control, link average delay and power allocation in fast-fading lossy delay-constrained multi-hop wireless networks. As opposed to RENUM framework, which cannot guarantee the globally optimal solutions, we proposed the



(a) Flow-3 effective rate.



(b) Flow-3 delay.

Figure 2.4: Effects of the hop number on the performance of nRENUM and RENUM.

nRENUM based on exactly closed form of the link outage probability, which can provide the optimal solutions to the problem. The non-convex original problem is converted into a convex one by logarithmic transformation and auxiliary variables. Then, the problem can be solved by the duality technique and implemented distributedly. Finally, the numerical results confirmed that the proposed method can achieve superior performance and significant improvements compared to the alternative design.

Chapter 3

A Multi-Timescale Cross-Layer Approach for Wireless Ad Hoc Networks

3.1 Introduction

Flow control in wired networks was firstly modeled as a network utility maximization (NUM) paradigm [Kelly et al., 1998; Low and Lapsley, 1999], as follows:

$$\max_{x \geq 0} \sum_{s} \mathrm{U_s}(x_s) \tag{3.1}$$

$$\text{s.t.} \quad Rx \leq c, \tag{3.2}$$

where x is the source rate vector, R is the routing matrix, and c is the fixed link capacity vector. In wireless networks, however, the link capacities c might be changed due to characteristics of wireless channel [Chen et al., 2009; Chiang, 2005]. Accordingly the problem (3.1) should be solved using Cross-Layer Designs (CLDs) by decomposing into subproblems, each of which corresponds to a layer. In fact, CLDs are essentially needed for wireless networks since CLDs can improve the network performance and reliability, for example, increasing throughput and reducing latency and bit error rate [Fu et al., 2014a].

There exist lots of works mainly focusing on CLDs for wireless networks [Chiang, 2005; Fu et al., 2014a; Gao et al., 2009; Guo et al., 2014; Papandriopoulos et al., 2008; Soldati and Johansson, 2009; Tran et al., 2013; Wang et al., 2013]. In [Guo et al., 2014], they not only considered a congestion control problem but also examined the effects of the lossy feature on the power control and link delay, namely Rate-Effective NUM (RENUM), with constraints on rate outage probability, data rate reduction, and delay-constrained traffics by taking them into consideration of the objective function. The rate outage probability is, however, based on the approximated form; therefore, RENUM may just produce suboptimal solutions to the problem [Tran et al., 2013]. In [Pham et al., 2015a], we studied a cross-layer design which can guarantee the globally optimal solutions to the RENUM. However, a common limitation of these works is that optimization variables at different layers are all updated simultaneously.

Each one of the five-layers in the TCP/IP network model takes its own networking functionalities and adheres to the distinct timescale. For an extreme example, the PHYsical layer (PHY) roles are to perform functions of power control and rate adaptation, while admission control, multi-flow control, and congestion control are performed at the transport layer. Practically, the TRANsport

(TRAN) layer is executed on the second timescale, while the data link control/MAC (DLC/MAC) layer and the PHY layer are executed on the scale of, respectively, milliseconds and microseconds. The precise timescales offer significant benefits of convergence speed and network performance [Soldati and Johansson, 2009]. Multi-timescale CLDs can be found in [Altman et al., 2012; Gamage et al., 2014; Kim et al., 2009; Nguyen et al., 2013; O'Neill et al., 2008; Pham and Hwang, 2014; Soldati and Johansson, 2009; Soldati et al., 2006; suk Kim et al., 2013; Zheng et al., 2009]. The authors in [Pham and Hwang, 2014; Soldati and Johansson, 2009] developed joint CLDs of congestion control and power allocation, which adheres to the natural timescale separation between rapid power control updates and slower end-to-end rate adjustments. In [Zheng et al., 2009], the authors reported that using the standard subgradient method in tackling the joint problem of MAC scheduling and congestion control might be not suitable for some circumstances [Kim et al., 2009; Soldati et al., 2006; suk Kim et al., 2013], for example, if the utility function is not concave at all primal variables such as time-share proportions of the allowed schedules in the case [Zheng et al., 2009], the primal variables may be oscillated, which can be avoided by proposing a two-timescale adaptive method. Nevertheless, no mentioned literature [Altman et al., 2012; Gamage et al., 2014; Kim et al., 2009; Nguyen et al., 2013; O'Neill et al., 2008; Pham and Hwang, 2014; Soldati and Johansson, 2009; Soldati et al., 2006; suk Kim et al., 2013; Zheng et al., 2009] has concurrently addressed the following issues (1) fast-fading, (2) lossy features of wireless networks, and (3) link delay requirement.

In this chapter, we study the problem of rate control, link delay, and power allocation for Wireless Ad Hoc NETworks (WANETs). Our objective is to find a cross-layer design that maximizes the overall utility and minimizes the total link delay and power consumption subject to constraints on outage probability, delay requirement and flow rate conservation, which not only guarantees globally optimal solutions to the underlying problem, but also adheres to the difference in layers' timescales. In a nutshell, the summary of features and contributions offered by our proposal are listed, as follows:

- In section 3.3, we present the network model and then formulate a joint optimization problem. Since the original optimization problem is non-convex, we first cast the underlying one into a convex one by auxiliary variables and log-transformations and then prove the convexity and strong duality properties of the transformed problem.

- We explore our proposed procedure, namely, Multi-TimeScale RENUM (MTSRENUM) in section 3.4 and make use of the primal vertical decomposition in order to derive three timescale-based subproblems: the Short-TimeScale (STS) subproblem (power control), the Mid-TimeScale (MTS) subproblem (link delay control), and the Long-TimeScale (LTS) subproblem (congestion control). Because the convexity and strong duality hold for the new optimization subproblems, each subproblem can be successively and optimally solved by the conventional duality technique and updated at its adhered timescale.

- The convergence of the proposed algorithm MTSRENUM is guaranteed, which is proven in subsection 3.4.4.

- Simulation results in section 3.5 confirm that our cross-layer design can provide large gains over the current frameworks. More specifically, by comparing the multi-timescale CLDs with their same-timescale counterparts, CLDs with multi-timescale controls illustrate the improvement in the convergence speed. In addition, it is observed that the MTSRENUM algorithm is better than the others in terms of suffered delay, consumed power, transmission rate as well as the trade-off between the aggregated effective rate and suffered delay and consumed power.

In the next section, we provide an overview of the state of the art related to CLDs and multi-timescale controls.

3.2 Related Work

Lots of literature whose principle focus on designing optimal CLD policies have been proposed [Bui et al., 2008; Chiang, 2005; Fu et al., 2014a; Gao et al., 2009; Guo et al., 2014; Kandukuri and Boyd, 2002; Kelly et al., 1998; Low and Lapsley, 1999; Palomar and Chiang, 2007; Papandriopoulos et al., 2008; Pham et al., 2015a; Soldati and Johansson, 2009; Tran et al., 2013; Vo et al., 2013; Wang et al., 2013], ranging from wired networks to wireless networks. The seminal researches on network resource allocation in wireline networks have been extensively investigated in [Kelly et al., 1998; Low and Lapsley, 1999]. A survey on and challenges of CLDs in wireless networks can be found in [Fu et al., 2014a; Srivastava and Motani, 2005b]. In [Fu et al., 2014a], the authors elaborated that CLDs can be categorized as non-manager and manager methods according to how to share information in one node and as centralized and distributed algorithms based on how to share cross-layer information among all nodes in a network. On another perspective, by the update timescale of the network, cross-layer designs can be classified into two groups: multi-timescale-based method and same-timescale-based method.

Most of popular works are the same-timescale methods. In [Chiang, 2005], the author presented a distributed power algorithm coupling with the existing congestion control protocols in order to increase end-to-end throughput and energy efficiency of the network. In this work, the high signal-to-interference-ratio (SIR) approximation, i.e., SIR is assumed to be much higher than than 1, is considered in order to cast the underlying non-convex problem into a convex one; however, this assumption just suits to the scenario where the wireless channels change very slowly while it may cause serious performance degradation in fast-fading wireless channels. Similar to [Chiang, 2005], Papandriopoulos et al. [Papandriopoulos et al., 2008] developed a distributed cross-layer design without any approximation that can attain the globally optimal solutions of link transmit powers and source rates for a mobile ad hoc network. A framework of congestion control and power allocation with an outage probability in fast-faded wireless channels has been studied in [Tran et al., 2013]. In [Gao et al., 2009], the authors firstly investigated the "leaky-pipe" flow model, called ENUM, where the transmission rate of a flow decreases along its route. Followed from [Gao et al., 2009], S. Guo et al. [Guo et al., 2014] proposed the RENUM framework, where the transmit power and link delay are also the optimization variables. In [Wang et al., 2013], the authors considered the problem of congestion control in interference-limited wireless networks with power control, namely ENUM with Power control (ENUMP). However, all of these works are assumed to execute at the same timescale, that of course leads to serious challenges.

Some recent works are devoted to the multi-timescale controls [Altman et al., 2012; Gamage et al., 2014; Kim et al., 2009; Nguyen et al., 2013; O'Neill et al., 2008; Pham and Hwang, 2014; Soldati and Johansson, 2009; Soldati et al., 2006; suk Kim et al., 2013; Zheng et al., 2009]. In [Soldati and Johansson, 2009], the authors considered a cross-layer design for congestion control and power control, respecting two timescales of the TRAN layer and PHY layer. In [O'Neill et al., 2008], the authors proposed a novel scheme, namely Multi-Period NUM/Adaptive Modulation (MPNUM/AP), which deals with optimal control policies at the PHY layer and upper layers at different timescales and can be solved by exploiting the Markov decision process. Literature [Nguyen et al., 2013] investigated a joint problem of rate adaption, power control and spectrum allocation in OFDMA-based multi-hop cognitive radio networks. The problem is decomposed into subproblems in which the spectrum allocation subproblem is updated centrally at the slower timescale compared to the others. A problem of congestion control and transmission scheduling in spatial-reuse time division multiple access wireless networks is studied in [Soldati et al., 2006] and that problem with the self-interference constraints, i.e., which is defined as the interference between different links of a same flow along a multihop path [Kim et al., 2009] and the constraints on rate requirements as minimal in wireless mesh networks is further developed in [suk Kim et al., 2013]. It is stated in [Kim et al., 2009; Soldati et al., 2006; suk Kim et al., 2013] that the conventional primal-dual approach cannot be used when considering the scheduling since (1) the primal-dual method assumes that all end-to-end links are active simultaneously while actually, only some links are active in an iteration due to the self-interference constraint (2) finding the optimum of the scheduling problem at every iteration requires a computational effort. To avoid these hurdles, they considered the cross-decomposition technique, which can be briefly described as the following. For a given end-to-end rate allocation, the scheduling and power updates are solved optimally through the traditional primal-dual technique at the slow-timescale. Then, with obtained solutions of the scheduling and power, end-to-end rates are updated based on the standard flow optimization problem [Kelly et al., 1998] at the fast-timescale. However, in [Kim et al., 2009; Soldati et al., 2006; suk Kim et al., 2013], the authors just elaborated the two-timescale cross-layer designs but did not examine them mathematically. In addition, as stated earlier, none of them has mutually addressed (1) fast-fading, (2) lossy features of wireless networks, and (3) link delay requirement.

3.3 System Model

3.3.1 Network Model

We consider a WANET with L logical links and S sources. Let $\Phi_s = \{1, 2, ..., S\}$ and $\Psi_l = \{1, 2, ..., L\}$ denote sets of sources and links, respectively. Let $L\,(s)$ be the set of links that flow s uses and $S(l)$ be the set of sources using link l. An example of a WANET is shown in Fig. 3.1, where there are five mobile nodes, four links $\Psi_l = \{1, 2, 3, 4\}$, four flows, each of which belongs to a source, then $\Phi_s = \{1, 2, 3, 4\}$. Accordingly, $L(3)$, the set of links that flows 3 uses, is $\{3, 4\}$, and similarly, $S(3)$, the set of sources using link 3, is $\{2, 3, 4\}$.

In this work, we assume that sources generate only elastic traffic, for example, file transfer, e-mail, Telnet, by which each flow associates with a strictly concave, nondecreasing, continuously differentiable utility function. The work involving an integration of heterogeneous traffic, i.e., elastic traffic and inelastic traffic, is not our focus in this chapter and is saved for extending works. We consider a well-known family of utility functions that has been discussed in [Mo and Walrand, 2000]

$$U_s(x_s) = \begin{cases} \frac{x_s^{1-\alpha}}{1-\alpha} & \text{if } \alpha \geq 0, \alpha \neq 1, \\ \ln(x_s) & \alpha = 1, \end{cases} \tag{3.3}$$

where parameter α corresponding to different properties of the utility function is the degree of fairness. The utility function can be interpreted as the satisfaction of source s at data rate x_s. In particular, the utility function is optimal fairness with $\alpha \to 0$, proportional fairness with $\alpha \to 1$, harmonic mean fairness with $\alpha \to 2$, and max-min fairness with $\alpha \to \infty$.

The instantaneous capacity on link l is modeled using the Shannon capacity $C_l(\mathbf{P}) = W \log\left(1 + \zeta\gamma_l(\mathbf{P})\right)$, where $\mathbf{P}$ is a vector of transmission powers, W is the baseband bandwidth, and ζ is the "SINR-gap" depending on particular modulation, coding scheme, and bit-error-rate. The instantaneous SINR is as

$$\gamma_l(\boldsymbol{P}) = \frac{p_l G_{ll} F_{ll}}{\sigma_l^2 + \sum_{k \neq l} p_k G_{lk} F_{lk}}, \tag{3.4}$$

where σ_l^2 is the thermal noise power at the receiver on link l and $G_{lk}F_{lk}$ is the instantaneous channel gain from the transmitter on link k to the receiver on link l. Each $G_{lk}F_{lk}$ consists of a slow-fading channel gain G_{lk} and a fast-fading channel gain F_{lk}.

Similar to [Tran et al., 2013], we consider the non-light-of-sight propagation where we can use the Rayleigh fast-fading model. For fast-fading wireless channels, the channel fades change very fast within milliseconds [Kandukuri and Boyd, 2002]; therefore, instead of using instantaneous SINR, we consider the average link SINR and capacity by utilizing the statistics of the SINR γ_l. Then

$$\overline{\gamma}_l(\mathbf{P}) = \frac{[p_l G_{ll} F_{ll}]}{\left[\sigma_l^2 + \sum_{k \neq l} p_k G_{lk} F_{lk}\right]} = \frac{p_l G_{ll}}{\sigma_l^2 + \sum_{k \neq l} p_k G_{lk}}, \tag{3.5}$$

where exponentially random variables F_{lk} are assumed to be independent and identically distributed (i.i.d) and $E\left[F_{lk}\right] = 1, \forall k$. Accordingly, $\overline{C}_l(\boldsymbol{P}) = W \log(1 + \zeta\overline{\gamma}_l(\mathbf{P}))$. Hence, if we use the network topology as in Fig. 3.1, the instantaneous SINR of, for example, link 3 and its corresponding average SINR are $\gamma_3(\boldsymbol{P}) = p_3 G_{33} F_{33} / \left(\sigma_3^2 + \sum_{k=\{1,2,4\}} p_k G_{3k} F_{3k}\right)$ and $\bar{\gamma}_3(\boldsymbol{P}) = p_3 G_{33} / \left(\sigma_3^2 + \sum_{k=\{1,2,4\}} p_k G_{3k}\right)$, respectively.

3.3.2 Average Delay

We consider the delay as a optimization variable. Generally, each link delay is the sum of four components: processing delay, queuing delay, transmission delay, and propagation delay. In this chapter, the processing delay and propagation delay are supposed to be zero.

Each link is modeled as a $M/M/1$ queuing system. Let $\tau(l)$ be the sum of the transmission delay and queuing delay on link l. The average packet delay[1] [Bertsekas et al., 1987; Guo et al., 2014] on link l is $E(\tau(l)) = K/\left(C_l(\mathbf{P}) - \sum_{s \in S(l)} x_s^l\right)$, where K (bits) is the mean of exponentially distributed rate of the input Poisson process and x_s^l is the transmission rate of flow s on link l. The link delay is required not to exceed the upper bound of the link delay threshold υ_l, i.e., we have $E(\tau(l)) \le \upsilon_l$. Therefore,

$$\sum_{s \in S(l)} x_s^l \le C_l(\mathbf{P}) - \frac{K}{\upsilon_l}. \tag{3.6}$$

3.3.3 Rate-Outage Probability and Effective Rate

Once channel states change, it is required to re-track the instant SINR and compelled to rerun the algorithm to seek the optimal solutions. For fast-fading networks where the channel might change very fast, it is not efficient and impractical. To overcome this isssue, we consider the outage probability [Goldsmith, 2005], which is $\Pr(\gamma_l < \gamma_l^{th})$ where γ_l^{th} is a target minimum SINR that below which performance becomes unacceptable.

[1]According to Little's result [Bertsekas et al., 1987], the average delay can be shown to be the average aggregate queue length divided by the average aggregate arrival rate.

The outage probability is rewritten as Pr$(\gamma_l \leq \gamma_l^{th}) = 1-\phi_l(\boldsymbol{P})$, where $\phi_l(\boldsymbol{P})$ can be regarded as the degradation of data rate over link l. With the outage probability, the optimal solution to the problem does not need to change when channel state wanders from one fading state to another one for a fraction of time. In the Rayleigh fading model, the closed-from expression [Kandukuri and Boyd, 2002; Papandriopoulos et al., 2008] of $\phi_l(\boldsymbol{P})$ is given as

$$\phi_l(\boldsymbol{P}) = \Pr\left(1 > \frac{\gamma_l^{th}}{p_l G_{ll}}\left(\sum_{k\neq l} p_k G_{lk} + \sigma_l^2\right)\right) = \exp\left(-\frac{\sigma_l^2\gamma_l^{th}}{p_l G_{ll}}\right)\prod_{k\neq l}\frac{1}{1+\gamma_l^{th}\frac{p_k G_{lk}}{p_l G_{ll}}}. \quad (3.7)$$

Let Pr$(\gamma_l \leq \gamma_l^{th}) \leq \epsilon_l$, which with (3.7) results in Eq. (3.8) as

$$\prod_{k\neq l}\left(1+\gamma_l^{th}\frac{p_k G_{lk}}{p_l G_{ll}}\right) \leq \Omega_l(p_l), \quad (3.8)$$

where $\Omega_l(p_l) = (1-\epsilon_l)^{-1}\exp\left(-\frac{\sigma_l^2\gamma_l^{th}}{p_l G_{ll}}\right)$.

We consider the leaky-pipe flow model [Gao et al., 2009], where the flow rate of each flow changes hop by hop and decrease along its route. For a flow s, the effective rate y_s at the destination is computed as the multiplication of the outage probability on links that flow s traverses and the injection rate x_s, i.e., $y_s = x_s \prod_{l\in L(s)}[1-\text{Pr}(l,\gamma_l)]$. Specifically, we assume that data generated by source s travels H_s hops and the data rate reduces along H_s hops before reaching the destination. The data rate at link i of flow s is x_s^i and given by

$$x_s^{i+1} = x_s^i\left(1-\Pr\left(l(s,i)\right)\right),\ i = 1, ..., H_s, \quad (3.9)$$

where Pr $(l(s,i))$ is the outage probability on link i, Eq. (3.9) is referred to the constraint on flow rate conservation.

3.3.4 Optimization Problem

From the constraints (3.6), (3.8), and (3.9), feasible region of power and flow rate, and the optimization objective intending to maximize the overall effective utility and minimize total consumed power and link delay, the optimization problem is considered, as follows:

$$\max_{x_s^l, p_l, \upsilon_l} \sum_{s\in\Phi_s} \mathrm{U}_s(x_s^{H_s+1}) - \omega_1 \sum_{l\in\Psi_l} p_l - \omega_2 \sum_{l\in\Psi_l} \upsilon_l \tag{3.10}$$

$$\text{s.t.} \quad \chi = \left\{ x_s^l | x_s^{l\,\min} \le x_s^l \le x_s^{l\,\max} \right\} \ \forall s, l, \tag{3.11}$$

$$\Gamma = \left\{ p_l | p_l^{\min} \le p_l \le p_l^{\max} \right\} \ \forall l, \tag{3.12}$$

$$\prod_{k\neq l} \left(1 + \gamma_l^{th} \frac{p_k G_{lk}}{p_l G_{ll}} \right) \le \Omega_l(p_l) \ \forall l, \tag{3.13}$$

$$\sum_{s\in S(l)} x_s^l \le C_l(\overline{\gamma}_l(\mathbf{P})) - \frac{K}{\upsilon_l} \ \forall l, \tag{3.14}$$

$$x_s^{i+1} \le x_s^i \left(1 - \Pr\left(l(s,i)\right)\right) \ i = 1, 2, ..., H_s, \forall s, \tag{3.15}$$

where ω_1 and ω_2 are prices per unit of consumed power and suffered delay, respectively.

The problem (3.13) is not an jointly convex problem in terms of $\left(x_s^l, p_l, \upsilon_l\right)$ due to the third and fifth constraints. Consequently, the locally optimal solution may be not the globally optimal solution and the KKT conditions is just necessary, not sufficient for solution optimality [Boyd and

Vandenberghe, 2009]. To ease (3.13), it is converted to a new one, as follows:

$$\max_{\hat{x}_s^l,\hat{p}_l,\hat{\upsilon}_l} \sum_{s\in\Phi_s} \mathrm{U}_s(e^{\hat{x}_s^{H_s+1}}) - \omega_1 \sum_{l\in\Psi_l} e^{\hat{p}_l} - \omega_2 \sum_{l\in\Psi_l} e^{\hat{\upsilon}_l} \tag{3.16}$$

$$\text{s.t.} \quad \sum_{k\neq l} \log\left(1+\gamma_l^{th}\frac{e^{\hat{p}_k}G_{lk}}{e^{\hat{p}_l}G_{ll}}\right) \le \log\Omega_l(e^{\hat{p}_l})\ \forall l, \tag{3.17}$$

$$\sum_{s\in S(l)} e^{\hat{x}_s^l} \le C_l(\overline{\gamma}_l(e^{\hat{\mathbf{P}}})) - \frac{K}{e^{\hat{\upsilon}_l}}\ \forall l, \tag{3.18}$$

$$\hat{x}_s^{i+1} \le \hat{x}_s^i - \psi_{l(s,i)}(\hat{\mathbf{P}})\ i=1,2,...,H_s, \forall s, \tag{3.19}$$

where auxiliary variables and transformation are:

$$x_s^l = e^{\hat{x}_s^l},\ \hat{\chi} = \left\{\hat{x}_s^l|\log\left(x_s^{l\,\min}\right) \le \hat{x}_s^l \le \log\left(x_s^{l\,\max}\right)\right\}, \tag{3.20}$$

$$p_l = e^{\hat{p}_l},\ \hat{\Gamma} = \left\{\hat{p}_l|\log\left(p_l^{\min}\right) \le \hat{p}_l \le \log\left(p_l^{\max}\right)\right\}, \tag{3.21}$$

$$\upsilon_l = e^{\hat{\upsilon}_l}, \tag{3.22}$$

$$x_s^{i+1} \le x_s^i\phi\left(l(s,i)\right),\ \psi_{l(s,i)}(\hat{\mathbf{P}}) = -\log\left(\phi\left(l(s,i)\right)\right). \tag{3.23}$$

The minus sign in (3.19) denotes the reduction of the effective rate compared with the injection rate on a link. For simplicity, $C_l(\overline{\gamma}_l(e^{\hat{\mathbf{P}}}))$ is adequately represented by $C_l(\hat{\overline{\gamma}}_l)$ or $C_l(e^{\hat{\mathbf{P}}})$ which means that we are dealing with the average SINR.

To prove $U_s(\exp(.))$ is a concave function, we make an assumption as $\frac{d^2U_s(x_s)}{dx_s^2}x_s + \frac{dU_s(x_s)}{dx_s} < 0$. This assumption is realistic and satisfied if we use α-fair utility function with the fair index $\alpha > 1$; in case of $\alpha = 1$, i.e., the log-utility function, the assumption now becomes $\frac{d^2U_s(x_s)}{dx_s^2}x_s + \frac{dU_s(x_s)}{dx_s} \le 0$ [Lee et al., 2007; Tran et al., 2013; Wang et al., 2013]. In addition, we need the

Lemma 1 and Lemma 2 to show convexity and strong duality of the problem (3.16).

Lemma 1. (3.16) *is a convex optimization problem.*

Proof. We have

$$\frac{d^2 u_s(e^{x_s})}{dx_s^2} = e^{\hat{x}_s}\left(\frac{d^2 U_s(x_s)}{dx_s^2}x_s + \frac{dU_s(x_s)}{dx_s}\right). \tag{3.24}$$

Since $\frac{d^2 U_s(x_s)}{dx_s^2}x_s + \frac{dU_s(x_s)}{dx_s} < 0$ is assumed to be negative, we have $\frac{d^2 f_x(x_s)}{dx_s^2} < 0$. As a result, $U_s(\exp(.))$ is a strictly concave function. Since the objective function is separate in terms of $\hat{p}_l$, $\hat{v}_l$, and $\hat{x}_s^{H_s+1}$, it is a concave function in $\hat{x}_s^{H_s+1}$, $\hat{p}_l$ and $\hat{v}_l$. Constraints (3.20) and (3.21) are line-segments, so they are convex. For the third constraint, its Hessian [Tran et al., 2013] meets the inequality that $\mathrm{v}^T\nabla^2 f_l(\hat{\mathbf{P}})\mathrm{v} \geq 0$. Consequently, the constraint (3.17) is convex in $\hat{\mathbf{P}}$.

Let $f_l(\hat{\mathbf{P}}, \hat{v}_l) = C_l(e^{\hat{\mathbf{P}}}) - \frac{K}{e^{\hat{v}_l}}$. The Hessian of the right-hand side of constraint (3.18) in $\hat{v}_l$ is $\nabla^2 f_l^{(\hat{v}_l)}(\hat{\mathbf{P}}, \hat{v}_l) = -K/e^{\hat{v}_l}$ which is nonpositive. We now define $f(x) = \log(\overline{\gamma}_l(\hat{\mathbf{P}}))$ and $f(x)$ can be reformed as follows:

$$f(x) = \log\left(\exp(\hat{p}_l)G_{ll}\right) - \log\left(\sigma_l^2 + \sum_{k\neq l}\exp(\hat{p}_k)G_{lk}\right). \tag{3.25}$$

Clearly, it is a concave function in $\hat{\mathbf{P}}$ due to the sum of linear and log-sum-exp terms. According to [Boyd and Vandenberghe, 2009], composition with an affine mapping, $f(1+\zeta x)$, is also concave. We therefore conclude that the constraint (3.18) is concave. The remaining work is to prove the right-hand side of (3.19) is concave in $\hat{\mathbf{P}}$.

We have $\psi_{l(s,i)}(\hat{\mathbf{P}}) = -\log\left(\phi\left(l(s,i)\right)\right)$ according to equation (3.23) which can be rewritten as

$$\psi_{l(s,i)}(\hat{\mathbf{P}}) = -\frac{\sigma_l^2 \gamma_l^{th}}{e^{\hat{p}_l} G_{ll}} + \sum_{k \neq l} \log \left(1 + \gamma_l^{th} \frac{e^{\hat{p}_k} G_{lk}}{e^{\hat{p}_l} G_{ll}} \right). \tag{3.26}$$

In the above equation, it is clear that the first term is convex in $\hat{\mathbf{P}}$. The Hessian of the second term is just the Hessian of the left-hand side of constraint (3.17). Thus, the right-hand side of constraint (3.19) is concave. Consequently, we finally state that (3.16) is a convex optimization problem. □

Lemma 2. *Strong duality holds for the problem* (3.16).

Proof. There exist feasible solutions $\hat{x}_s^l \in \hat{\chi}$, $\hat{p}_l \in \hat{\Gamma}$, and $\hat{\upsilon}_l$ that the constraints (3.17)-(3.19) hold with strict inequalities. Accordingly, Slater's condition holds [Boyd and Vandenberghe, 2009]. Moreover, the problem is convex (Lemma 1), so strong duality holds. □

In [Pham et al., 2015a], we presented the nRENUM algorithm which can guarantee the globally optimal solutions to the problem (3.16). However, updates in nRENUM are all executed on the same timescale. Concretely, this work assumes that parameters are computed online and updated simultaneously, requires synchronization, and coordination among layers. In practice, there is a difference in timescales among layers to ease implementations of NUM. The link delay is normally updated at the millisecond timescale, that is the slow timescale update. In the meanwhile, update timescales of congestion control and power control are, respectively, the second scale and the microsecond scale. Moreover, the same timescale operation might create more signaling overhead and stems the system convergence [Soldati and Johansson, 2009]. That is why our focus is in studying the NUM framework considering the difference in layers' timescales in WANETs. In the next section, our design is described in detail.

3.4 MTSRENUM Distributed Algorithm

This section presents an attractive aspect to solve the optimization problem (3.16). We make use of the primal decomposition and difference in layers' timescales to derive three subproblems. For each subproblem, we define the dual problem, which are solved by exploiting the gradient method. Here, subproblems are the resource control subproblem executed on the short-timescale t_1, the delay control subproblem derived at the mid-timescale t_2 and the congestion control subproblem executed on the long-timescale t_3.

Instead of directly solving the problem (3.16), we first consider the STS subproblem that aims to minimize the transmit power for a given rate vector $\{\hat{\boldsymbol{x}} = \hat{x}_s^l | \hat{x}_s^l \in \hat{\chi}, s \in \Phi_s, l \in \Psi_l\}$ and link delay vector $\{\hat{\boldsymbol{v}} = \hat{v}_l | l \in \Psi_l\}$.

$$\begin{aligned}
\text{STS:} \quad & \max_{\hat{p}_l} \; -\omega_1 \sum_{l \in \Psi_l} e^{\hat{p}_l} \qquad (3.27)\\
\text{s.t.} \quad & \hat{p}_l \in \hat{\Gamma}_l \; \forall l,\\
& \sum_{k \neq l} \log\left(1 + \gamma_l^{th} \frac{e^{\hat{p}_k} G_{lk}}{e^{\hat{p}_l} G_{ll}}\right) \leq \log \Omega_l(e^{\hat{p}_l}) \; \forall l,\\
& \sum_{s \in S(l)} e^{\hat{x}_s^l} \leq C_l(e^{\hat{\mathbf{P}}}) - \frac{K}{e^{\hat{v}_l}} \; \forall l,\\
& \hat{x}_s^{i+1} \leq \hat{x}_s^i - \psi_{l(s,i)}(\hat{\mathbf{P}}) \; i = 1, 2, ..., H_s, \forall s.
\end{aligned}$$

Solving the subproblem (3.27) at the short-timescale t_1, we can obtain the minimum cost $\varphi(t_1(\mathbf{P}))$ in term of transmit power and the optimal power vector $\mathbf{P}^*(t_1)$. Then, the MTS subproblem which intends to minimize the total link average delay for a given source rate vector $\{\hat{\boldsymbol{x}} = \hat{x}_s^l | \hat{x}_s^l \in \hat{\chi}, s \in \Phi_s, l \in \Psi_l\}$, the optimum cost $\varphi(t_1(\mathbf{P}))$, and the optimal power $\mathbf{P}^*(t_1)$ is considered at the timescale t_2, as follows:

$$\text{MTS:} \quad \max_{\hat{\upsilon}_l} \; -\omega_2 \sum_{l \in \Psi_l} e^{\hat{\upsilon}_l} - \varphi(t_1(\mathbf{P})) \tag{3.28}$$

$$\text{s.t.} \quad \sum_{s \in S(l)} e^{\hat{x}_s^l} \leq C_l(e^{\hat{\mathbf{P}}^*}) - \frac{K}{e^{\hat{\upsilon}_l}} \; \forall l. \tag{3.29}$$

Similarly, by solving the subproblem (3.28) we obtain the minimum cost $\varphi(t_2(\boldsymbol{\upsilon}, t_1(\mathbf{P})))$ in term of both transmission power and link average delay and the optimal link delay vector $\boldsymbol{\upsilon}^*(t_2) = [\upsilon_1^*(t_2), ..., \upsilon_L^*(t_2)]^T$. Finally, the TRAN layer subproblem which aims at maximizing the total effective utility for the given minimum cost $\varphi(t_2(\boldsymbol{\upsilon}, t_1(\mathbf{P})))$, the optimal vectors $\boldsymbol{\upsilon}^*(t_2)$ and $\mathbf{P}^*(t_1)$ is considered at the timescale t_3 as

$$\text{LTS:} \quad \max_{\hat{x}_s^l} \sum_{s \in \Phi_s} \mathrm{U_s}(e^{\hat{x}_s^{H_s+1}}) - \varphi(t_2(\upsilon, t_1(\mathbf{P}))) \tag{3.30}$$

$$\text{s.t.} \quad \hat{x}_s^l \in \hat{\chi} \; \forall s, l, \tag{3.31}$$

$$\sum_{s \in S(l)} e^{\hat{x}_s^l} \leq C_l(e^{\hat{\mathbf{P}}^*}) - \frac{K}{e^{\hat{\upsilon}_l^*}} \; \forall l. \tag{3.32}$$

$$\hat{x}_s^{i+1} \leq \hat{x}_s^i - \psi_{l(s,i)}(\hat{\mathbf{P}}^*) \; i = 1, 2, ..., H_s, \forall s. \tag{3.33}$$

The update timescale of the LTS subproblem is much slower than that of the MTS subproblem and both of them are much slower than that of the STS subproblem. Accordingly, for given vectors $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{\upsilon}}$, optimizing the STS subproblem can be well treated to be instantaneous, and is also true for the MTS subproblem for a given $\hat{\boldsymbol{x}}$.

Lemma 3. *Convexity and strong duality properties hold for the new optimization subproblems.*

Proof. As convexity and strong duality of the optimization problem (3.16) are proved in Lemma 1 and Lemma 2, three optimization subproblems (3.27), (3.28), (3.30) and their constraints are derived without any modifications. As a result, convexity and strong duality property hold for three subproblems. □

3.4.1 Short-timescale Iterative Subalgorithm

We first define the Lagrangian, as follows:

$$\begin{aligned} L_p(\hat{\mathbf{P}}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}) &= \sum_s \sum_{i=1}^{H_s} \nu_s^i \left(\hat{x}_s^i - \hat{x}_s^{i+1} - \psi_{l(s,i)}(\hat{\mathbf{P}}) \right) \\ &+ \sum_l \lambda_l \left(\log \Omega_l \left(e^{\hat{p}_l} \right) - \sum_{k \neq l} \log \left(1 + \gamma_l^{th} \frac{e^{\hat{p}_k} G_{lk}}{e^{\hat{p}_l} G_{ll}} \right) \right) \\ &+ \sum_l \mu_l \left(C_l(e^{\hat{\mathbf{P}}}) - \sum_{s \in S(l)} e^{\hat{x}_s^l} - \frac{K}{e^{\hat{v}_l}} \right) - \omega_1 \sum_l e^{\hat{p}_l}, \end{aligned} \tag{3.34}$$

where $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$, and $\boldsymbol{\nu}$ are the Lagrange multipliers. Then, we define the Lagrange dual function $g_P(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu})$ as

$$g_P(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}) = \max_{\hat{\mathbf{P}}} L_p(\hat{\mathbf{P}}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}), \tag{3.35}$$

which can be specified as the maximum of the Lagrangian (3.34) over the primal variable $\hat{\mathbf{P}}$. Accordingly, that leads to

$$\min_{\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}} g_P(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}). \tag{3.36}$$

The dual problem (3.36) is a convex problem. We can use the primal-dual method to develop the STS iterative subalgorithm, as follows.

Power allocation:[1] the link transmit power is updated by

$$p_l(t_1+1) = \left[\frac{M_l(t_1) + \left(\lambda_l(t_1) + \nu_s^l(t_1)\right)\sigma_l^2 m_l^{th}(t_1)}{\sum\limits_{k\neq l} G_{kl}\left(A_{kl}(t_1) + m_k(t_1)\right) + \omega_1}\right]_\Gamma , \tag{3.37}$$

where $M_n(t_1) = W\mu_n(t_1)\frac{\zeta\overline{\gamma}_n(t_1)}{1+\zeta\overline{\gamma}_n(t_1)}$, $m_n(t_1) = \frac{\overline{\gamma}_n(t_1)}{G_{nn}p_n(t_1)}M_n(t_1)$, $m_n^{th}(t_1) = \frac{\overline{\gamma}_n^{th}(t_1)}{G_{nn}p_n(t_1)}$, and $A_{kl}(t_1) = \left(\lambda_k(t_1) + \nu_s^k(t_1)\right)\frac{m_k^{th}(t_1)}{1+G_{kl}m_k^{th}(t_1)p_l(t_1)}$.

Lagrange multipliers:[2]

Outage prices: each link l updates its dual variable λ_l as

$$\lambda_l(t_1+1) = \left[\lambda_l(t_1) - \delta_p(t_1)\nabla g_P^{(\lambda_l)}(t_1)\right]^+ , \tag{3.38}$$

where $\nabla g_P^{(\lambda_l)}(t_1) = \sum\limits_{k\neq l}\log\left(1+\gamma_l^{th}\frac{e^{\hat{p}_k(t_1)}G_{lk}}{e^{\hat{p}_l(t_1)}G_{ll}}\right) - \log\Omega_l(e^{\hat{p}_l(t_1)})$.

Delay prices: the dual variable $\mu_l(t_1+1)$ can be updated as

$$\mu_l(t_1+1) = \left[\mu_l(t_1) - \delta_p(t_1)\nabla g_P^{(\mu_l)}(t_1)\right]^+ , \tag{3.39}$$

where $\nabla g_P^{(\mu_l)}(t_1) = \sum_{s\in S(l)} e^{\hat{x}_s^l(t_3)} - C_l(e^{\hat{\mathbf{P}}(t_1)}) + \frac{K}{e^{\hat{v}_l(t_2)}}$.

[1]Notation $[x]_\Gamma$ is the projection of x on the set Γ.
[2]$[z]^+ = \max\{z, 0\}$.

Lossy prices: at each link i of flow s, ν_s^i can be updated by

$$\nu_s^i(t_1+1) = \left[\nu_s^i(t_1) - \delta_p(t_1)\nabla g_P^{(\nu_s^i(t_1))}(t_1)\right]^+, \tag{3.40}$$

where $\nabla g_P^{(\nu_s^i(t_1))}(t_1) = \hat{x}_s^{i-1}(t_3) - \hat{x}_s^i(t_3) + \psi_{l(s,i)}(\hat{\mathbf{P}}(t_1))$.

Remark 2.

- *Here, transmission power and its dual variables of the STS subproblem are updated for given link delay $\hat{\upsilon}_l(t_2)$ and data rate $\left\{\hat{x}_s^i(t_3)|i=1,...,H_s\right\}$.*
- *The STS iterative subalgorithm can be implemented distributedly through message passing (e.g., $m_k^{th}(t_1)$, $m_k(t_1)$, $\lambda_k(t_1)$ and $\nu_s^k(t_1)$).*

3.4.2 Mid-timescale Iterative Subalgorithm

The Lagrangian of subproblem (3.28) is defined as follows:

$$L_{\hat{\boldsymbol{\upsilon}}}(\hat{\boldsymbol{\upsilon}}, \boldsymbol{\mu}) = -\omega_2 \sum_l e^{\hat{\upsilon}_l} - \varphi(t_1(\mathbf{P})) + \sum_l \mu_l \left(C_l(e^{\hat{\mathbf{P}}^*}) - \sum_{s\in S(l)} e^{\hat{x}_s^l} - \frac{K}{e^{\hat{\upsilon}_l}} \right), \tag{3.41}$$

where $\boldsymbol{\mu}$ is the delay prices for the constraint (3.29). The dual function and the dual problem are, respectively, as follows.

$$g_{\boldsymbol{\upsilon}}(\boldsymbol{\mu}) = \max_{\hat{\boldsymbol{\upsilon}}} L_{\hat{\boldsymbol{\upsilon}}}(\hat{\boldsymbol{\upsilon}}, \boldsymbol{\mu}). \tag{3.42}$$

$$\min_{\boldsymbol{\mu}}\ g_{\boldsymbol{\upsilon}}(\boldsymbol{\mu}). \tag{3.43}$$

Making use of the gradient method, we consider the following MTS iterative subalgorithm.

Link average delays: at each link l, υ_l can be updated as

$$\upsilon_l(t_2+1)=\sqrt{\mu_l(t_2)\frac{K}{\omega_2}}. \tag{3.44}$$

Delay prices: at each link l, the dual price can be updated as

$$\mu_l(t_2+1)=\left[\mu_l(t_2)-\delta_\upsilon(t_2)\nabla g_{\boldsymbol{\upsilon}}^{(\mu_l)}(t_2)\right]^+, \tag{3.45}$$

where δ_υ is a positive step size and

$$\nabla g_{\boldsymbol{\upsilon}}^{(\mu_l)}(t_2)=C_l(e^{\hat{\mathbf{P}}^*(\hat{\boldsymbol{\upsilon}}(t_2))})-\sum_{s\in S(l)}e^{\hat{x}_s^l(t_3)}-\frac{K}{e^{\hat{\upsilon}_l(t_2)}}, \tag{3.46}$$

where $\hat{\mathbf{P}}^*(\hat{\boldsymbol{\upsilon}}(t_2))$ is the optimal transmission power vector obtained from the STS iterative subalgorithm.

Remark 3.

- *The link delay update* (3.44) *depends on only local information on delay prices. Consequently, the MTS subalgorithm can be implemented distributedly*
- *The link delay and delay price of the MTS subproblem are updated for given the optimal solutions obtained from the STS iterative subalgorithm and given data rate* $\left\{\hat{x}_s^i(t_3)|i=1,...,H_s\right\}$.

3.4.3 Long-timescale Iterative Subalgorithm

Similar to subsections 3.4.1 and 3.4.2, we first construct the Lagrangian for the LTS subproblem, as follows:

$$\begin{aligned} L_{\hat{\boldsymbol{x}}}\left(\hat{\boldsymbol{x}}, \boldsymbol{\mu}, \boldsymbol{\nu}\right) = & \sum_{s \in \Phi_s} \mathrm{U}_s(e^{\hat{x}_s^{H_s+1}}) - \varphi(t_2(\upsilon, t_1(\mathbf{P}))) \\ & + \sum_l \mu_l \left(C_l(e^{\hat{\mathbf{P}}^*}) - \frac{K}{e^{\hat{\upsilon}_l^*}} - \sum_{s \in S(l)} e^{\hat{x}_s^l} \right) + \sum_s \sum_{i=1}^{H_s} \nu_s^i \left(\hat{x}_s^i - \psi_{l(s,i)}(\hat{\mathbf{P}}^*) - \hat{x}_s^{i+1} \right). \end{aligned} \tag{3.47}$$

Then the dual function is given as

$$g_{\boldsymbol{x}}\left(\boldsymbol{\mu}, \boldsymbol{\nu}\right) = \max_{\hat{\boldsymbol{x}}} L_{\hat{\boldsymbol{x}}}\left(\hat{\boldsymbol{x}}, \boldsymbol{\mu}, \boldsymbol{\nu}\right). \tag{3.48}$$

Accordingly, the dual problem is, as follows:

$$\min_{\boldsymbol{\mu}, \boldsymbol{\nu}} g_{\boldsymbol{x}}\left(\boldsymbol{\mu}, \boldsymbol{\nu}\right). \tag{3.49}$$

Similar to the previous subsections, we exploit the gradient method to solve the dual problem (3.49) and develop the following LTS iterative subalgorithm.

Data rate:

The data rate at link i of flow s, can be updated, as follows:

$$\hat{x}_s^i(t_3+1) = \left[\hat{x}_s^i(t_3) + \delta_x(t_3) \nabla g_{\boldsymbol{x}}^{(\hat{x}_s^i)}(t_3) \right]_{\hat{\mathcal{X}}}, \tag{3.50}$$

where δ_x is a positive scalar step size and $\nabla g_{\boldsymbol{x}}^{(\hat{x}_s^i)}$ is the gradient of $L_{\hat{\boldsymbol{x}}}$ with respect to $\hat{x}_s^i$ and specified by $\nabla g_{\boldsymbol{x}}^{(\hat{x}_s^i)}(t_3) = \nu_s^i(t_3) - \nu_s^{i-1}(t_3) - \mu_{l(s,i)} e^{\hat{x}_s^i(t_3)}$ for $i = 1, ..., H_s$. For $i = H_s + 1$, $\nabla g_{\boldsymbol{x}}(\hat{x}_s^i)(t_3) = U_s'\left(e^{\hat{x}_s^{H_s+1}(t_3)}\right) e^{\hat{x}_s^{H_s+1}(t_3)} - \nu_s^{H_s}(t_3)$.

Delay prices: each link l update its delay price $\mu_l(t_3)$ as

$$\mu_l(t_3 + 1) = \left[\mu_l(t_3) - \delta_x(t_3)\nabla g_x^{(\mu_l)}(t_3)\right]^+, \tag{3.51}$$

where $\nabla g_{\boldsymbol{x}}^{(\mu_l)}(t_3) = C_l(e^{\hat{\mathbf{P}}^*(\hat{\boldsymbol{x}}(t_3))}) - \frac{K}{e^{\hat{v}_l^*(\hat{\boldsymbol{x}}(t_3))}} - \sum_{s \in S(l)} e^{\hat{x}_s^l(t_3)}$.

Lossy prices: at each link i of source s, the lossy price $\nu_s^i(t_3)$ is updated by

$$\nu_s^i(t_3 + 1) = \left[\nu_s^i(t_3) - \delta_x(t_3)\nabla g_{\boldsymbol{x}}^{(\nu_s^i)}(t_3)\right]^+, \tag{3.52}$$

where $\nabla g_{\boldsymbol{x}}^{(\nu_s^i)}(t_3) = \hat{x}_s^i(t_3) - \hat{x}_s^{i+1}(t_3) - \psi_{l(s,i)}(\hat{\mathbf{P}}^*(\hat{\boldsymbol{x}}(t_3)))$.

Remark 4.

- *The link rate, effective rate and dual variables can be updated based on the optimal solutions obtained from the STS subalgorithm and the MTS subalgorithm.*
- *The LTS subalgorithm can be implemented distributively through message passing (i.e., dual variables) among links and sources.*

To summarize, the key aspect of the MTSRENUM iterative algorithm is illustrated in Algorithm 2.

Algorithm 2 MTSRENUM algorithm

Initialization: Set t_1, t_2, t_3, $\boldsymbol{x}(0)$, $\boldsymbol{\upsilon}(0)$, $\boldsymbol{P}(0)$, $\boldsymbol{\lambda}(0)$, $\boldsymbol{\mu}(0)$, and $\boldsymbol{\nu}(0)$, which are required to be nonnegative.

Iteration

1) For a given $\boldsymbol{x}(t_3)$ and $\boldsymbol{\upsilon}(t_2)$, solve the STS subproblem iteratively until convergence.

2) For a given $\boldsymbol{x}(t_3)$ and optimal solutions $(\mathbf{P}^*(\boldsymbol{\upsilon}(t_2)), \boldsymbol{\lambda}^*(\boldsymbol{\upsilon}(t_2)), \boldsymbol{\mu}^*(\boldsymbol{\upsilon}(t_2)), \boldsymbol{\nu}^*(\boldsymbol{\upsilon}(t_2)))$ obtained from the STS subalgorithm, repeat until convergence.

3) For a given $\boldsymbol{x}(t_3)$, we can obtain the optimal solutions $\boldsymbol{\upsilon}^*(\boldsymbol{x}(t_3))$ from the MTS iterative subalgorithm and $(\mathbf{P}^*(\boldsymbol{\upsilon}(t_2), \boldsymbol{x}(t_3)), \boldsymbol{\lambda}^*(\boldsymbol{\upsilon}(t_2), \boldsymbol{x}(t_3)), \boldsymbol{\mu}^*(\boldsymbol{\upsilon}(t_2), \boldsymbol{x}(t_3)), \boldsymbol{\nu}^*(\boldsymbol{\upsilon}(t_2), \boldsymbol{x}(t_3)))$ from the STS subalgorithm, update $\boldsymbol{x}(t_3+1)$ via the LTS subalgorithm.

4) Set $t_3 = t_3 + 1$ and go to step 1 until satisfying the termination condition.

3.4.4 Convergence Analysis

The MTSRENUM algorithm can be regarded as a tight loop in which the STS, MTS and LTS subalgorithms respond to the inner loop, the mid loop and the outer loop, respectively. The STS subalgorithm operates at the short-timescale t_1 for given data rate and delay vectors. The mid loop is updated at the mid-timescale t_2 for a given data rate vector, optimal solutions obtained from the STS subalgorithm. The last one is executed on the timescale t_3 for stable link delay vector, stable power vector and the corresponding prices.

We now analyze the convergence of the STS subalgorithm. Denote $\mathbf{P}^*$ as the optimal solutions to the STS subproblem and $\|.\|$ as the Euclidean distance. We have the following theorem.

Theorem 2. ***Convergence of the STS subalgorithm****: Let $\boldsymbol{P}(0) \in \Gamma$, $\boldsymbol{\lambda}(t_1 = 0)$, $\boldsymbol{\mu}(t_1 = 0)$, and $\boldsymbol{\nu}(t_1 = 0)$ are initial values of power and dual variables. Here, $\boldsymbol{P}(t_1)$, $\boldsymbol{\lambda}(t_1)$, $\boldsymbol{\mu}(t_1)$, and $\boldsymbol{\nu}(t_1)$ are the sequences generated by* (3.37)*,* (3.38)*,* (3.39)*,* (3.40)*. With the chosen step size satisfying the diminishing rule (i.e., $\delta_t \geq 0$, $\sum_{t=0}^{\infty} \delta_t = \infty$, $\sum_{t=0}^{\infty} (\delta_t)^2 < \infty$), there actually exists a sufficiently large T_1 that $\forall t_1 \geq T_1$, $\boldsymbol{P}(t_1), \boldsymbol{\lambda}(t_1), \boldsymbol{\mu}(t_1)$, and $\boldsymbol{\nu}(t_1)$ converge to the globally optimal points such that $\lim_{t_1 \to T | T > T_1} \|\boldsymbol{P}(t_1) - \boldsymbol{P}^*(t_1)\| < \epsilon_{t_1}$ for any $\epsilon_{t_1} > 0$.*

Proof. The proof is omitted, see [Bertsekas and Tsitsiklis, 1989]. □

Next, we carry out the convergence of the MTS subalgorithm as provided in the Theorem 3.

Theorem 3. ***Convergence of the MTS subalgorithm:*** *For each link delay vector and given data rate vector, the STS subalgorithm is assumed to converge instantaneously. Here $\boldsymbol{\upsilon}(0)$ and $\boldsymbol{\mu}(0)$ are initial values of link delays and its delay prices. Let $\boldsymbol{\upsilon}(t_2)$ and $\boldsymbol{\mu}(t_2)$ are the sequences generated by the MTS subalgorithm. There exists a set $\Omega^*_{t_2}$ and a sufficiently large T_2 such that $\boldsymbol{\upsilon}(t_2)$ approaches $\Omega^*_{t_2}$ and the smallest distance from $\boldsymbol{\upsilon}(t_2)$ to any point in $\Omega^*_{t_2}$, i.e., dist $\left(\boldsymbol{\upsilon}(t_2), \Omega^*_{t_2}\right) = \inf_{\boldsymbol{\upsilon}^*(t_2)\in\Omega^*_{t_2}} \|\boldsymbol{\upsilon}(t_2) - \boldsymbol{\upsilon}^*(t_2)\| < \epsilon_{t_2}$ as $t > T_2$ for any $\epsilon_{t_2} > 0$.*

Proof. Recall that $L(\hat{\mathbf{P}}, \hat{\boldsymbol{\upsilon}}, \hat{\boldsymbol{x}}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu})$ is the Lagrangian of the underlying problem. For a given data rate vector, we make use of the duality techniques to define $g(\hat{\boldsymbol{\upsilon}})$, as follows:

$$g(\hat{\boldsymbol{\upsilon}}) = \min_{\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu}} D(\hat{\boldsymbol{\upsilon}}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}) = \min_{\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu}} \max_{\hat{\mathbf{P}}} L(\hat{\mathbf{P}}, \hat{\boldsymbol{\upsilon}}, \hat{\boldsymbol{x}}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}). \tag{3.53}$$

The Danskin's theorem [Başar and Bernhard, 2008; Bertsekas, 1999] gives $\frac{dg(\hat{\boldsymbol{\upsilon}})}{d\hat{\boldsymbol{\upsilon}}} = \dot{D}(\hat{\boldsymbol{\upsilon}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\nu}})$, where $\overline{\boldsymbol{\lambda}}, \overline{\boldsymbol{\mu}}, \overline{\boldsymbol{\nu}}$ achieve the minimum of $g(\hat{\boldsymbol{\upsilon}})$ and

$$\dot{D}(\hat{\boldsymbol{\upsilon}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\nu}}) = -\omega_2 \sum_l e^{\dot{\hat{\upsilon}}_l} - \min_{\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu}} \sum_l \mu_l \frac{K}{e^{\dot{\hat{\upsilon}}_l}} = -\omega_2 \sum_l e^{\dot{\hat{\upsilon}}_l} - \sum_l \overline{\mu}_l \frac{K}{e^{\dot{\hat{\upsilon}}_l}}. \tag{3.54}$$

We recall the delay update in the MTS iterative subalgorithm, as follows:

$$\upsilon_l(t_2 + 1) = \left[\mu_l(t_2)\frac{K}{\omega_2}\right]^{1/2}. \tag{3.55}$$

It is fairly straightforward to see that $\dot{\hat{\upsilon}}_l \mu_l \geq 0$ which with (3.54) leads to $\frac{dg(\hat{\boldsymbol{\upsilon}})}{d\hat{\boldsymbol{\upsilon}}} \leq -\sum_l \overline{\mu}_l \frac{K}{e^{\dot{\hat{\upsilon}}_l}}$

or $\frac{dg(\hat{\boldsymbol{v}})}{d\hat{\boldsymbol{v}}} \leq 0$, i.e., $\dot{g}(\hat{\boldsymbol{v}}) \geq 0$. Invoking Lasalle's invariance principle [Khalil and Grizzle, 2002], the sequence $\upsilon_l(t_2)$ generated by the MTS iterative subalgorithm converges to the largest invariant set inside or an equilibrium $\boldsymbol{v}^*(t_2)$ such that $\Omega^*(t_2)$ $\{\boldsymbol{v}^*(t_2) : \dot{g}(\boldsymbol{v}^*(t_2)) = 0\}$ and the following property holds: $\forall \epsilon_{t_2} > 0, \exists T_2, \forall t > T_2$, dist $\left(\boldsymbol{v}(t_2), \Omega^*_{t_2}\right) = \inf_{\boldsymbol{v}^*(t_2) \in \Omega^*_{t_2}} \|\boldsymbol{v}(t_2) - \boldsymbol{v}^*(t_2)\| < \epsilon_{t_2}$. Therefore, the convergence of the MTS iterative subalgorithm is guaranteed. we now prove that the largest invariant set $\Omega^*(t_2)$ also solves the optimization problem (3.16).

Since the STS and MTS iterative subalgorithms are assumed to be converged instantaneously for a given data rate vector, in this case we only set the presence of power and link delay. For each $\boldsymbol{v}^*(t_2)$, let $\mathbf{P}(\boldsymbol{v}^*(t_2))$, $\boldsymbol{\lambda}(\boldsymbol{v}^*(t_2))$, $\boldsymbol{\mu}(\boldsymbol{v}^*(t_2))$, and $\boldsymbol{\nu}(\boldsymbol{v}^*(t_2))$ are the optimal solutions to the STS iterative subalgorithm. Making use of the duality and KKT optimality conditions [Boyd and Vandenberghe, 2009] for the STS subproblem we have

$$\hat{\mathbf{P}}(\boldsymbol{v}^*(t_2)) = \underset{\hat{\mathbf{P}}}{\operatorname{argmax}} \Bigg\{ -\omega_1 \sum_l e^{\hat{p}_l} + \sum_l \lambda_l(\boldsymbol{v}^*(t_2)) \log \Omega_l \left(e^{\hat{p}_l}\right)$$
$$- \sum_l \lambda_l(\boldsymbol{v}^*(t_2)) \log \left(1 + \gamma_l^{th} \frac{e^{\hat{p}_k} G_{lk}}{e^{\hat{p}_l} G_{ll}}\right) + \sum_l \mu_l(\boldsymbol{v}^*(t_2)) C_l(e^{\hat{\mathbf{P}}}) - \sum_s \sum_{i=1}^{H_s} \nu_s^i(\upsilon^*(t_2)) \psi_{l(s,i)}(\hat{\mathbf{P}}) \Bigg\}, \tag{3.56}$$

$$\sum_l \lambda_l(\boldsymbol{v}^*(t_2)) \left(\sum_{k \neq l} \log \left(1 + \gamma_l^{th} \frac{e^{\hat{p}_k(\boldsymbol{v}^*(t_2))} G_{lk}}{e^{\hat{p}_l(\boldsymbol{v}^*(t_2))} G_{ll}}\right) - \log \Omega_l \left(e^{\hat{p}_l(\boldsymbol{v}^*(t_2))}\right) \right) = 0, \tag{3.57}$$

$$\sum_l \mu_l(\boldsymbol{v}^*(t_2)) \left(\sum_{s \in S(l)} e^{\hat{x}_s^l} - C_l(e^{\hat{\mathbf{P}}(\boldsymbol{v}^*(t_2))}) + \frac{K}{e^{\hat{\boldsymbol{v}}_l^*(t_2)}} \right) = 0, \tag{3.58}$$

$$\sum_{s}\sum_{i=1}^{H_s}\nu_s^i(\boldsymbol{\upsilon}^*(t_2))\left(\hat{x}_s^{i+1}-\hat{x}_s^i+\psi_{l(s,i)}\left(\hat{\mathbf{P}}(\boldsymbol{\upsilon}^*(t_2))\right)\right)=0. \tag{3.59}$$

It can be seen that (3.56), (3.57), (3.58), (3.59) are the optimality conditions for the optimization problem (3.16). As a consequence, we finally state that the MTS iterative subalgorithm also solves the underlying problem. □

Finally, we study the convergence of the LTS subalgorithm.

Theorem 4. ***Convergence of the LTS subalgorithm**: For each data rate vector, as assumed that the STS and MTS subalgorithms instantly converge. Here, $\hat{\boldsymbol{x}}(0) \in \hat{\chi}$, $\boldsymbol{\mu}(t_3 = 0)$, and $\boldsymbol{\nu}(t_3 = 0)$ are initial values. Let $\hat{\boldsymbol{x}}(t_3)$, $\boldsymbol{\mu}(t_3)$, and $\boldsymbol{\nu}(t_3)$ are the sequences generated by the STS subalgorithm. There exists a set $\Omega_{t_3}^*$ and a sufficiently large T_3 such that $\hat{\boldsymbol{x}}(t_3)$ approaches $\Omega_{t_3}^*$ and the smallest distance from $\hat{\boldsymbol{x}}(t_3)$ to any point in $\Omega_{t_3}^*$, i.e., dist $\left(\hat{\boldsymbol{x}}(t_3), \Omega_{t_3}^*\right) = \inf_{\hat{\boldsymbol{x}}^*(t_3)\in\Omega_{t_3}^*}\|\hat{\boldsymbol{x}}(t_3)-\hat{\boldsymbol{x}}^*(t_3)\| < \epsilon_{t_3}$ as $t > T_3$ for any $\epsilon_{t_3} > 0$.*

Proof. Similar to the one of Theorem 3, therefore omitted. □

3.5 Simulation Results

In this section, we evaluate the proposed method with numerical examples, where we highlight comparative results in terms of convergence speed, power consumption, flow delay, injection rate and effective rate. We then compare our schemes with the alternative frameworks [Gao et al., 2009; Guo et al., 2014; Pham et al., 2015a; Soldati and Johansson, 2009; Wang et al., 2013].

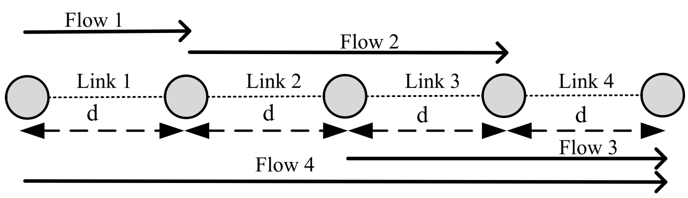


Figure 3.1: Physical and logical topology used for simulation.

3.5.1 Simulation Settings

We consider a network with topology illustrated in Fig. 3.1. Equidistantly, nodes are positioned at d = 50 meters. The outage probability and SINR thresholds are set to (0.20, 0.20, 0.20, 0.20) and (1.0, 1.0, 1.0, 1.0), respectively. The fast-fading channel gain is assumed to be i.i.d. while the slow-fading channel gain is $g_{lk} = g_0(d_{lk}/50)^{-\mathrm{AF}}$, where d_{lk} is the distance between transmitter of link k to receiver of link l, $\mathrm{AF} = 4$ is the path loss attenuation factor, and g_0 is the reference channel gain at a distance 50 meters and meets a condition that the average receive SINR at 50 meters is 30 dB. The power spectral density of white noise is further assumed to be -50 dBm/Hz. Without loss of generality, weights ω_1 and ω_2 are assumed to be 1. Unless other specified, we consider α-fair utility function $U_s(x_s) = (1-\alpha)^{-1}x^{1-\alpha}$ with $\alpha = 5$.

3.5.2 Performance of MTSRENUM

We first define the certainty-equivalent margin (CEM) of link l as the ratio of the average SINR $\bar{\gamma}_l$ to the threshold SINR γ_l^{th}, i.e., $\mathrm{CEM} = \bar{\gamma}_l/\gamma_l^{th}$. According to the relationship between CEM and outage probability [Kandukuri and Boyd, 2002; Tran et al., 2013], the upper and the lower bounds of the outage probability is given, as follows:

$$\frac{1}{1+\text{CEM}} \leq \Pr(\gamma_l \leq \gamma_l^{th}) \leq 1 - e^{-1/\text{CEM}}. \tag{3.60}$$

For simplicity, we consider the lower bound of (3.60) which with the condition of the rate outage probabiity ϵ_l results in

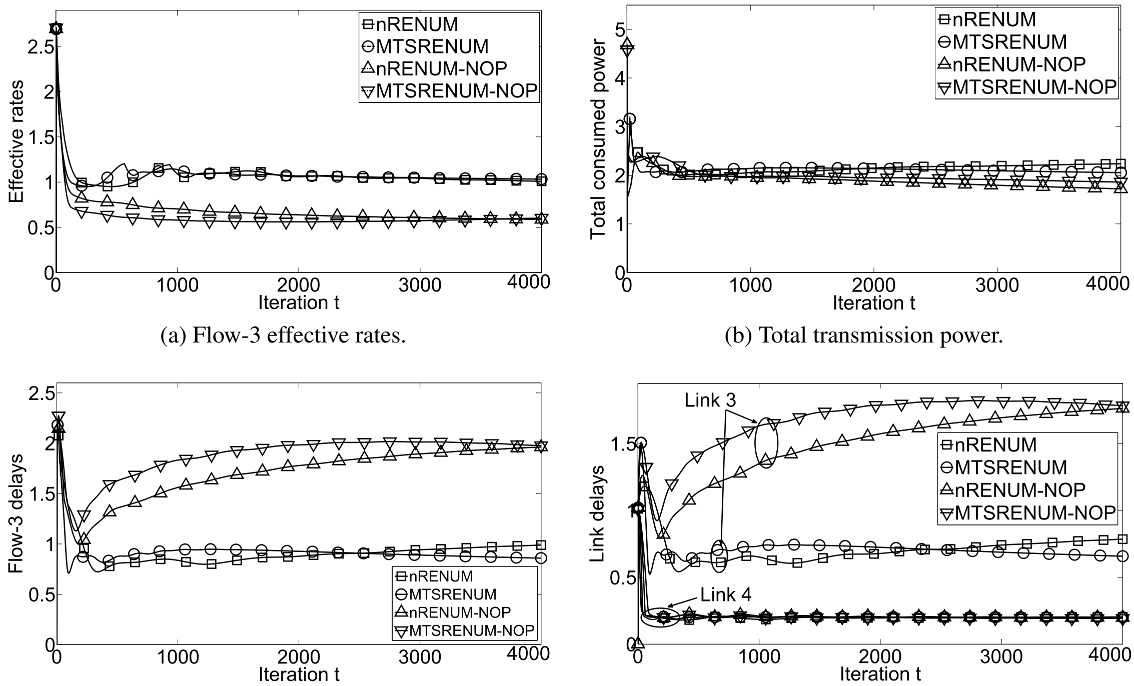
$$\bar{\gamma}_l^{th} \leq \bar{\gamma}_l, \tag{3.61}$$

where $\bar{\gamma}_l^{th} = \gamma_l^{th}(1/\epsilon_l - 1)$. From this point, we construct the new optimization problem with the outage probability replaced by (3.61). Similar to [Pham et al., 2015a], the updates of new optimization problem are exactly the same forms, excepting for power and outage price which are updated, as follows:

$$p_l(t+1) = \left[\frac{\lambda_l(t) + M_l(t) + \frac{\nu_s^l(t)\gamma_l^{th}}{\bar{\gamma}_l(t)+\gamma_l^{th}}}{\omega_1 + \sum\limits_{k\neq l} G_{kl} m_k(t)}\right]_{\Gamma}, \tag{3.62}$$

where $M_n(t) = W\mu_n(t)\frac{\zeta\overline{\gamma}_n(t)}{1+\zeta\overline{\gamma}_n(t)}$, $m_n(t) = \frac{\overline{\gamma}_n(t)}{G_{nn}p_n(t)}\left(\lambda_n(t) + \mu_n(t)M_n(t) + \frac{\nu_s^n(t)\gamma_n^{th}}{\overline{\gamma}_n(t)+\gamma_n^{th}}\right)$, $\lambda_l(t+1) = \left[\lambda_l(t) - \delta(t)\left(\log(\overline{\gamma}_l^{th}) - \log(\overline{\gamma}_l(t))\right)\right]^+$, and $\psi_{l(s,i)}(\hat{\mathbf{P}}) = -\log(\phi_{l(s,i)}(\hat{\mathbf{P}}))$, where $\phi_{l(s,i)}(\hat{\mathbf{P}}) = \frac{\hat{\gamma}_l(\hat{\mathbf{P}})}{\hat{\gamma}_l(\hat{\mathbf{P}})+\gamma_l^{th}}$. We call the new resulting schemes, near-optimal nRENUM (nRENUM-NOP) and MTSRENUM-NOP algorithms.

Secondly, we show the results of nRENUM [Pham et al., 2015a], MTSRENUM, nRENUM-NOP, and MTSRENUM-NOP in order to capture the significant effects of the near-optimal schemes and timescales on the performance. Fig. 3.2a, 3.2b, 3.2c, and 3.2d represent performance comparisons between these above algorithms in terms of flow-3 effective rate, total power consumption,



(a) Flow-3 effective rates.

(b) Total transmission power.

(c) Flow-3 delays.

(d) Link delays.

Figure 3.2: The performance comparison between nRENUM, MTSRENUM, nRENUM-NOP, and MTSRENUM-NOP.

flow-3 delay, and link delay, respectively. These figures indicate that MTSRENUM and nRENUM, nRENUM-NOP and MTSRENUM-NOP in turn have almost the same performance. However, since MTSRENUM and MTSRENUM-NOP consider the difference in layers' timescales, they can improve the convergence speed in comparison with the corresponding same-timescale-based algorithms (nRENUM and nRENUM-NOP, respectively). Fig. 3.2d reveals the performance comparison of link-3 and link-4 delays. We can see from Fig. 3.2d that link-3 delays converge slower than these of corresponding link-4 delays. This can be explained that there are three flows through link 3 and it is, therefore, highly affected by outage probability variations due to channel fadings. Moreover, nRENUM and MTSRENUM outperform near-optimal schemes in effective rates, flow delays, and link delays and they all have almost the same level of power consumption. This is suitable since nRENUM and MTSRENUM require much more message-passing (see [Pham et al., 2015a]) to update power control than near-optimal schemes which require only $m_n(t)$. Actually, there exists a trade-off between performance and signaling overhead, this is, however, not a concern of this chapter.

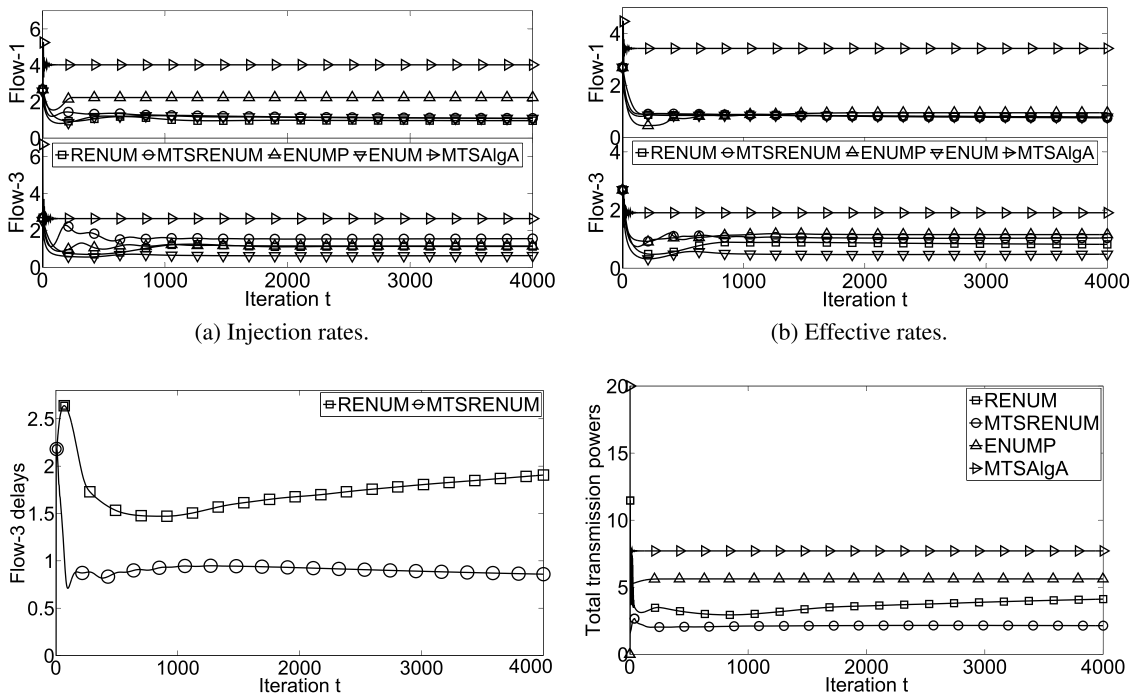
Table 3.1: Comparison between nRENUM and MTSRENUM in term of convergence speed.

Stopping criterion ε	MTSRENUM	nRENUM
1e-3	61.0	207.0
5e-4	106.0	406.0
1e-4	117.0	441.0
5e-5	120.0	2306.0
1e-5	716.0	6831.0
5e-6	721.0	8493.0
1e-6	1189.0	12048.0

To further confirm the improvement in the convergence speed of MTSRENUM compared to that of nRENUM, we make use of a stopping criterion, as the following:

$$\max_{l \in \Psi_l} \left| \frac{p_l(t) - p_l(t-1)}{p_l(t-1)} \right| < \varepsilon, \tag{3.63}$$

where ε is an arbitrarily small number used to stop the algorithms. Table 3.1 shows the number of iterations needed for convergence of MTSRENUM and nRENUM. We can see that when the stopping criterion becomes more strict, the number of iterations increase almost linearly and the increase rate of nRENUM is much higher than that of MTSRENUM. It is in addition observed that in all cases, MTSRENUM converges faster than nRENUM since in MTSRENUM, the difference in layers' timescales is considered; which was already validated in [Soldati and Johansson, 2009] for the case where the joint cross-layer design of congestion control and power allocation in slow-fading wireless multihop networks is investigated.



(a) Injection rates. (b) Effective rates. (c) Flow-3 delays. (d) Total transmission power.

Figure 3.3: The performance comparison between MTSRENUM, RENUM, ENUMP, and ENUM.

Finally, we compare the performance of MTSRENUM, RENUM, ENUMP, ENUM, and the algorithm proposed in [Soldati and Johansson, 2009], which is called MTSAlgA, to further show MTSRENUM's improvements and reliabilities. To evaluate [Gao et al., 2009; Wang et al., 2013], we set $\gamma_l^{th} = 1.0$ and $\bar{\gamma}_l = 2.25 * \gamma_l^{th}$. Fig. 3.3a and Fig. 3.3b illustrate the comparison of the injection rates and effective rates, respectively. It is observed that the effective rates reduce and become smaller and smaller along the routes when compared to the injection rates. In addition, a flow traversing a larger number of hops suffers higher data rate loss than that of a flow traveling less hops (e.g. flow 1 traverses 1 link and flow 3 travels 2 links). It is due to the lossy nature of wireless links. Fig. 3.3c and Fig. 3.3d depict that MTSRENUM significantly outperforms RENUM in terms of power consumption and flow delays. This is since the constraint on rate outage probability in RENUM reduces the original solution region and uses the approximated form. In a meanwhile, the rate outage probability constraint in MTSRENUM is in the rightly closed-form; therefore, MTSRENUM can guarantee the globally optimal solutions. Moreover, the framework [Guo et al., 2014] does not take into consideration the SINR threshold γ_l^{th}, so RENUM cannot vary the SINR theshold for different links to get the appropriately optimal solutions. Obviously, ENUMP efficiently and effectively performs in terms of the injection rates and effective rates over RENUM and MTSRENUM; however, a serious drawback of ENUMP is that it consumes much more powers than these of RENUM and MTSRENUM. The reasons are that (1) the transmission power and link average delay are taken into consideration of both RENUM and MTSRENUM; therefore, they try to balance the overall profit between the gained network benefit and total consumed power and suffered delay. These trade-offs can be parameterized by weights ω_1 and ω_2, which are assumed to be 1 for simplicity (2) ENUMP does not consider the transmission power in the objective function and fix rigidly links' SINR and (3) since ENUMP problem is to maximize the overall utility and the transmission power is just varied appropriately to maximize its utility, ENUMP maybe yields higher rates, i.e., injection rate and effective rate, compared to RENUM and MTSRENUM with the great sacrifice to the consumed power. Another observation shown in Fig. 3.3a and 3.3b is that ENUM yields less injection rates and effective rates than the others do. This is because ENUM does not consider γ_l^{th} and each link does not adjust its transmission power even through the channel quality is slow and still generate low data rates. Moreover, the transmit power and link delay are not the optimization variables and ENUM does not analyze the effects of the lossy feature on powers and delays. In contrast to the others, MTSAlgA yields the highest values in term of injection rate and effective rate; however, the total consumed power by MTSAlgA is largest. We can explain these observations as (1) the network model in [Soldati and Johansson, 2009] is slowly varying-fading while in the others, the link rate outage is taken into consideration to capture effects of fast-fading channels; therefore, the transmit power in MTSRENUM, RENUM, ENUMP can be used more appropriately and effectively to compensate for the link outage and (2) because of the participation of the new optimization variable, average link delay, in MTSRENUM and RENUM and of the rate outage constraints in these frameworks, links' effective capacity reduce and then the flow rates (injection rate and effective rate) generated by these procedures are lower than flow rates of MTSAlgA. We in addition see that the convergence speed of MTSAlgA is fastest. Besides considering the difference in layers' timescale, it is also since the constraints on flow rate conservation and/or the constraints on link rate outage in MTSRENUM, RENUM, ENUMP, and ENUM make the optimization problems more complicated than [Soldati and Johansson, 2009].

3.6 Conclusion

This chapter studied a cross-layer problem in wireless ad hoc networks. The non-convex formulated optimization problem is converted to a convex one by log-transformed and auxiliary variables. We called the resulting design MTSRENUM where the timescale difference is taken into consideration. We made use of the primal decomposition to derive three subproblems. Each subproblem can be solved by the duality techniques and updated in the distinct timescale. We proved the convergence of each scheme. Through simulation, we demonstrated that our design can achieve significant improvements.

Chapter 4

Discussion and Future Work

The final chapter summarizes the important contributions of this thesis and highlights the future research directions.

4.1 Contributions

In this thesis, we propose two cross-layer designs for congestion control at the transport layer, link delay at the link layer, and power control at the physical layer, using either the primal decomposition technique or the dual decomposition technique.

We first study the cross-layer problem in fast-fading lossy delay-constrained MANETs. In this work, we investigate the joint problem of congestion control, link delay, and power allocation with the exactly closed form of rate-outage probability. To avoid the non-convexity of the original problem, we cast it into a convex one by using the log-transformation. Then we decouple the primal optimization problem into joint rate control, link delay, and power allocation subproblems.

The globally optimal solutions can be obtained with the helps of message passing.

The second cross-layer design we study is the same problem as the previous chapter; however, the design in this chapter not only guarantees the globally optimal solutions to the original problem, but also adheres the the timescale difference among layers. In this work, we use the primal decomposition technique in order to derive three subproblems, which operate and update at the corresponding timescales. Through simulation results, we show that the convergence of the proposed algorithm is guaranteed and proved.

4.2 Future Work

As a future research direction for the work in chapter 2, we will consider a cross-layer optimization of congestion control, scheduling, and power control in OFDMA-based fast-fading wireless multihop networks. For implementation, we will try to validate the network-wide performance of the proposed cross-layer schemes. In addition, the analyses of frameworks, which can support the coexistence of elastic traffic and inelastic traffic or consider multipath communication networks and complex communication networks will be our future research work. Finally, since the publication point of the Kelly's paper, a large number of researches have concentrated on the problem of congestion control; however, there is a lack of a system-reviewed paper. Consequently, conducting a survey on congestion control cross-layer design will be really important.

References

Eitan Altman, Konstantin Avrachenkov, and Sreenath Ramanath. Multiscale fairness and its application to resource allocation in wireless networks. *Computer Communications*, 35(7):820 – 828, 2012. ISSN 0140-3664. doi: http://dx.doi.org/10.1016/j.comcom.2012.01.013. 29, 32

Tamer Başar and Pierre Bernhard. *H-infinity optimal control and related minimax design problems: a dynamic game approach*. Springer, 2008. 51

Dimitri P Bertsekas. Nonlinear programming. 1999. 51

Dimitri P Bertsekas and John N Tsitsiklis. *Parallel and distributed computation: numerical methods*. Prentice-Hall, Inc., 1989. 21, 51

Dimitri P Bertsekas and John N Tsitsiklis. Gradient convergence in gradient methods with errors. *SIAM Journal on Optimization*, 10(3):627–642, 2000. 21

Dimitri P Bertsekas, Robert G Gallager, and Pierre Humblet. *Data networks*, volume 2. Prentice-hall Englewood Cliffs, NJ, 1987. 12, 36

Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2009. 15, 38, 40, 41, 52

Dang Quang Bui, Myeonggil Choi, and Won Joo Hwang. Joint scheduling and rate optimization in multi-channel multi-radio wireless networks with contention-based mac. *Journal of Korea Multimedia Society*, 11(12):1716–1721, 2008. 10, 31

G. Carneiro, J. Ruela, and M. Ricardo. Cross-layer design in 4g wireless terminals. *Wireless Communications, IEEE*, 11(2):7–13, Apr 2004. ISSN 1536-1284. doi: 10.1109/MWC.2004.1295732. 2

Ling-Jyh Chen, Tony Sun, Guang Yang, M.Y. Sanadidi, and Mario Gerla. Adhoc probe: end-to-end capacity probing in wireless ad hoc networks. *Wireless Networks*, 15(1):111–126, 2009. ISSN 1022-0038. doi: 10.1007/s11276-007-0047-4. 28

Mung Chiang. Balancing transport and physical layers in wireless multihop networks: jointly optimal congestion control and power control. *Selected Areas in Communications, IEEE Journal on*, 23(1):104–116, Jan 2005. ISSN 0733-8716. doi: 10.1109/JSAC.2004.837347(410)23. 10, 28, 31, 32

F. Foukalas, V. Gazis, and N. Alonistioti. Cross-layer design proposals for wireless mobile networks: A survey and taxonomy. *Communications Surveys Tutorials, IEEE*, 10(1):70–85, First 2008. ISSN 1553-877X. doi: 10.1109/COMST.2008.4483671. 2, 3

Bo Fu, Yang Xiao, Hongmei Deng, and Hui Zeng. A survey of cross-layer designs in wireless networks. *Communications Surveys Tutorials, IEEE*, 16(1):110–126, First 2014a. ISSN 1553-877X. doi: 10.1109/SURV.2013.081313.00231. 10, 28, 31

Bo Fu, Yang Xiao, Hongmei Deng, and Hui Zeng. A survey of cross-layer designs in wireless networks. *Communications Surveys & Tutorials, IEEE*, 16(1):110–126, 2014b. 1, 2, 3, 4, 5

A.T. Gamage, Hao Liang, and Xuemin Shen. Two time-scale cross-layer scheduling for cellular/wlan interworking. *Communications, IEEE Transactions on*, 62(8):2773–2789, Aug 2014. ISSN 0090-6778. doi: 10.1109/TCOMM.2014.2337890. 29, 32

Qinghai Gao, Junshan Zhang, and S.V. Hanly. Cross-layer rate control in wireless networks with lossy links: leaky-pipe flow, effective network utility maximization and hop-by-hop algorithms. *Wireless Communications, IEEE Transactions on*, 8(6):3068–3076, June 2009. ISSN 1536-1276. doi: 10.1109/TWC.2009.080604. 7, 9, 10, 11, 12, 14, 28, 31, 32, 37, 53, 58

Andrea Goldsmith. *Wireless communications*. Cambridge university press, 2005. 13, 36

S. Guo, C. Dang, and Y. Yang. Joint optimal data rate and power allocation in lossy mobile ad hoc networks with delay-constrained traffics. *Computers, IEEE Transactions on*, PP(99):1–1, 2014. ISSN 0018-9340. doi: 10.1109/TC.2013.2296043. 6, 7, 9, 10, 11, 13, 22, 23, 24, 28, 31, 32, 36, 53, 59

S. Kandukuri and S. Boyd. Optimal power control in interference-limited fading wireless channels with outage-probability specifications. *Wireless Communications, IEEE Transactions on*, 1(1): 46–55, Jan 2002. ISSN 1536-1276. doi: 10.1109/7693.975444. 13, 31, 35, 37, 54

Frank P Kelly, Aman K Maulloo, and David KH Tan. Rate control for communication networks:

shadow prices, proportional fairness and stability. *Journal of the Operational Research society*, pages 237–252, 1998. 8, 10, 27, 31, 33

Hassan K Khalil and JW Grizzle. *Nonlinear systems*, volume 3. Prentice hall Upper Saddle River, 2002. 52

Tae-Suk Kim, Yong Yang, J.C. Hou, and S.V. Krishnamurthy. Joint resource allocation and admission control in wireless mesh networks. In *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, 2009. WiOPT 2009. 7th International Symposium on*, pages 1–10, June 2009. doi: 10.1109/WIOPT.2009.5291636. 29, 32, 33

J Kuros and Keith Ross. Computer networking: a top-down approach, 2007. 1, 4

Jang-Won Lee, Mung Chiang, and A.R. Calderbank. Utility-optimal random-access control. *Wireless Communications, IEEE Transactions on*, 6(7):2741–2751, July 2007. ISSN 1536-1276. doi: 10.1109/TWC.2007.05991. 39

S.H. Low and D.E. Lapsley. Optimization flow control. i. basic algorithm and convergence. *Networking, IEEE/ACM Transactions on*, 7(6):861–874, Dec 1999. ISSN 1063-6692. doi: 10.1109/90.811451. 27, 31

Jeonghoon Mo and J. Walrand. Fair end-to-end window-based congestion control. *Networking, IEEE/ACM Transactions on*, 8(5):556–567, Oct 2000. ISSN 1063-6692. doi: 10.1109/90.879343. 11, 34

Van Mui Nguyen, LEE Sungwon, and Choong Seon Hong. Joint rate adaption, power control, and spectrum allocation in ofdma-based multi-hop crns. *IEICE Transactions on Communications*, 96(1):242–253, 2013. 29, 32, 33

D. O'Neill, A.J. Goldsmith, and S. Boyd. Optimizing adaptive modulation in wireless networks via utility maximization. In *Communications, 2008. ICC '08. IEEE International Conference on*, pages 3372–3377, May 2008. doi: 10.1109/ICC.2008.634. 29, 32

D.P. Palomar and Mung Chiang. Alternative distributed algorithms for network utility maximization: Framework and applications. *Automatic Control, IEEE Transactions on*, 52(12):2254–2269, Dec 2007. ISSN 0018-9286. doi: 10.1109/TAC.2007.910665. 31

J. Papandriopoulos, S. Dey, and J. Evans. Optimal and distributed protocols for cross-layer design of physical and transport layers in manets. *Networking, IEEE/ACM Transactions on*, 16(6): 1392–1405, Dec 2008. ISSN 1063-6692. doi: 10.1109/TNET.2008.918099. 10, 13, 28, 31, 32, 37

Quoc Viet Pham and Won Joo Hwang. A novel cross-layer design for fast-fading multihop wireless networks. In *Proceeding of the Fall Conference of the Korea Multimedia Society*, Dongguk University, Korea, 2014. KMMS. 10, 29, 32

Quoc-Viet Pham, Hoon Kim, and Won-Joo Hwang. Globally optimal solutions for cross-layer design in fast-fading lossy delay-constrained manets. *Journal of Korea Multimedia Society*, 18 (2):168–177, Feb 2015a. 7, 9, 28, 31, 41, 53, 55, 57

Quoc-Viet Pham, Hoang-Linh To, and Joo-Hwang Won. A multi-timescale cross-layer approach for wireless ad hoc networks. *Computer Networks*, under revision, 2015b. 7

P. Soldati and M. Johansson. Reducing signaling and respecting time-scales in cross-layer protocols design for wireless networks. In *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pages 1–8, Nov 2009. doi: 10.1109/GLOCOM.2009.5425714. 28, 29, 31, 32, 41, 53, 58, 60

Pablo Soldati, Björn Johansson, and Mikael Johansson. Proportionally fair allocation of end-to-end bandwidth in stdma wireless networks. In *Proceedings of the 7th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, MobiHoc '06, pages 286–297, New York, NY, USA, 2006. ACM. ISBN 1-59593-368-9. doi: 10.1145/1132905.1132937. 29, 32, 33

V. Srivastava and M. Motani. Cross-layer design: a survey and the road ahead. *Communications Magazine, IEEE*, 43(12):112–119, Dec 2005a. ISSN 0163-6804. doi: 10.1109/MCOM.2005.1561928. 2, 3, 4

V. Srivastava and M. Motani. Cross-layer design: a survey and the road ahead. *Communications Magazine, IEEE*, 43(12):112–119, Dec 2005b. ISSN 0163-6804. doi: 10.1109/MCOM.2005.1561928. 31

Tae suk Kim, Yong Yang, J.C. Hou, and S.V. Krishnamurthy. Resource allocation for qos support in wireless mesh networks. *Wireless Communications, IEEE Transactions on*, 12(5):2046–2054, May 2013. ISSN 1536-1276. doi: 10.1109/TWC.2013.021213.120285. 29, 32, 33

N.H. Tran, Choong Seon Hong, and Sungwon Lee. Cross-layer design of congestion control and power control in fast-fading wireless networks. *Parallel and Distributed Systems, IEEE*

Transactions on, 24(2):260–274, Feb 2013. ISSN 1045-9219. doi: 10.1109/TPDS.2012.118. 10, 12, 28, 31, 32, 35, 39, 40, 54

Phuong Luu Vo, Nguyen Hoang Tran, Choong Seon Hong, and Sungwon Lee. Network utility maximisation framework with multiclass traffic. *Networks, IET*, 2(3):152–161, Sept 2013. ISSN 2047-4954. doi: 10.1049/iet-net.2012.0120. 31

Fei Wang, Xiaofeng Liao, Songtao Guo, and Hongyu Huang. Jointly optimal congestion and power control for rayleigh-faded channels with outage constraints. *Wireless Personal Communications*, 77(1):101–125, 2013. ISSN 0929-6212. doi: 10.1007/s11277-013-1497-x. 7, 9, 10, 28, 31, 32, 39, 53, 58

D. Xenakis, N. Passas, L. Merakos, and C. Verikoukis. Mobility management for femtocells in lte-advanced: Key aspects and survey of handover decision algorithms. *Communications Surveys Tutorials, IEEE*, 16(1):64–91, First 2014. ISSN 1553-877X. doi: 10.1109/SURV.2013.060313.00152. 4

Xiaoying Zheng, Feng Chen, Ye Xia, and Yuguang Fang. A class of cross-layer optimization algorithms for performance and complexity trade-offs in wireless networks. *Parallel and Distributed Systems, IEEE Transactions on*, 20(10):1393–1407, Oct 2009. ISSN 1045-9219. doi: 10.1109/TPDS.2008.245. 29, 32