Optimal Resource Allocation for Delay and Energy Constrained Wireless Networks

A thesis submitted in fulfilment of the requirements for the degree of

PhD in Electrical and Electronics Engineering

Koc University, November 2015

Yalçın Şadi

Graduate School of Sciences and Engineering

Koc University

Istanbul, Turkey

Koc University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a doctoral dissertation by

Yalçın Şadi

and have found it is complete and satisfactory in all respects, and that any and all revisions required by the final examining committee have been made.

Committee Members:

Asst. Prof. Dr. Sinem Çöleri Ergen (Advisor)

Prof. Dr. Özgür Barış Akan

Assoc. Prof. Dr. Öznur Özkasap

Asst. Prof. Dr. Ertuğrul Başar

Asst. Prof. Dr. Aktül Kavas

Bitaneme

Acknowledgements

I would like to express my sincere gratitude and profound respect to my Advisor Assist. Prof. Sinem Çöleri Ergen for her reliable guidance, inspiration, patience and valuable support. I am also grateful to members of my thesis committee Prof. Özgür Barış Akan, Assoc. Prof. Öznur Özkasap, Assist. Prof. Ertuğrul Başar and Assist. Prof. Aktül Kavas for their valuable comments and serving on my defense committee.

First, I thank to *bitanem* Pelin and my family for their endless faith, love and invaluable support in really hard times. Next, I would like to express my gratitudes to *Sonradan Gurmeler* Tolga, Seyhan, Kenan and Tuğtekin with whom I lived the best moments of a real friendship. I also thank to my office friends Bakhtiyar, Beakal, Saeed, Yahya, Fatima, Sanaz, Berkin and Naveed who provided me a funny and friendly research environment. It is not possible to forget all shared things with my *Özhakiki Canlar* Ümit, Bekir, Burak, Taylan, Caner, Umur, İskender, Nihan, Meltem, Remziye, Şebnem, Duygu, Şaziye, Merve E., Merve G. and İrem through the years.

I also acknowledge the support of Tubitak (The Scientific and Technological Research Council of Turkey) Grant #113E233 and Turk Telekom Collaborative Research Award #11315-10.

Abstract

In this thesis, the resource allocation in delay and energy constrained wireless networks is investigated. As the wireless applications and services in wireless networked control systems (WNCSs) and machine-to-machine communication (M2M) over cellular systems become widespread, energy and delay gain more importance in design of communication systems. Energy is one of the key concerns of communication system design since wireless sensors or devices mostly rely on batteries with limited capacity or energy harvesting techniques. Delay, on the other hand, is an important design metric since many wireless applications are time-critical in which timely delivery of information with high level of reliability is of paramount importance.

First, we provide a novel framework for the joint optimization of controller and communication systems encompassing efficient abstractions of both systems for WNCSs, which may lead to broader adoption and real-world deployment. For the control system model, stochastic maximum allowable transfer interval and maximum allowable delay constraints are used as efficient abstractions considering wireless network induced imperfections. Communication system is presented via power consumption model and schedulability constraint using the control system parameters. Using both analytical analysis on the optimality conditions of the optimization problems and fast heuristics, energy efficient resource allocation methods are proposed for any power-related objective which is a non-decreasing function of the power consumptions of the nodes, different modulation schemes including MQAM and MFSK and scheduling algorithms.

Second, we propose a new adaptivity metric for scheduling which exploits the periodic transmission nature of the sensor nodes in WNCSs. Using

this metric, we provide a novel optimization approach with the objective of providing maximum adaptivity in schedule accommodating packet losses and changes in network topology while considering the constraints of the individual sensor nodes in terms of periodic data generation, transmission delay, energy and reliability. In order to solve the joint optimization problem fast and efficiently, we propose a design framework for heuristic scheduling algorithms.

Third, we investigate the minimum length scheduling problem in wireless ad hoc networks considering packetized transmissions meaning that packet transmission of a node cannot be interrupted and divided into multiple chunks to be transmitted in different time slots due to impracticality and suboptimality reasons. Upon formulating the joint optimization problem and proving its NP-hardness, we show that the power control and rate adaptation problem can be separated from the scheduling problem. Then, we introduce a novel problem for the optimization of power control and rate adaptation where the time required for the concurrent transmission of a set of links each having an integer number of packets is minimized. Solving this problem allows formulation of the joint problem as a pure scheduling problem which is a large-scale Integer Programming (IP) problem. In order to solve this large-scale IP problem, we incorporate elegant Branch and Price Method and Column Generation method based heuristic algorithms.

Fourth, we propose optimal power control and rate allocation strategies and fast scheduling algorithms based on column generation method for delay minimization and delay constrained energy minimization problems in Ultra Wide Band (UWB) wireless networks.

Özetçe

Bu tezde, gecikme ve enerji kısıtlı kablosuz ağlarda kaynak özgüleme incelenmektedir. Kablosuz Ağlı Kontrol Sistemler (KAKS) ve hücresel ağlar üzerindeki Makineler Arası Haberleşme (MAH) alanlarındaki uygulamalar ve servisler yaygınlaştıkça, enerji ve gecikme gerekliliklerinin haberleşme sistemlerinin tasarlanmasındaki önemi artmaktadır. Kablosuz algılayıcılar ve cihazlar çoğunlukla kısıtlı kapasiteli bataryalarla ya da enerji toplama teknikleriyle çalıştığından, enerji haberleşme sistemlerinin tasarımındaki anahtar noktalardan biridir. Diğer taraftan gecikme, birçok kablosuz uygulama, bilginin zamanında ve yüksek güvenilirlikle gönderilmesinin çok önemli olduğu zaman-kısıtlı uygulamalar olduğundan, önemli bir tasarım ölçevidir.

İlk olarak, KAKS'lerin daha fazla yaygınlaşmasına öncülük edebilecek şekilde, kontrol ve haberleşme sistemlerinin her iki sistemin etkin soyutlamasını içeren ortak eniyilemesi için özgün bir çerçeve sunmaktayız. Kontrol sisteminin modellemesi için, olasılıksal maksimum izin verilebilir transfer aralığı ve maksimum izin verilebilir gecikme kısıtları kablosuz ağ kaynaklı kusurlar da dikkate alınarak etkin soyutlamalar olarak kullanılmıştır. Haberleşme sistemi ise kontrol sistemi parametrelerinin kullanıldığı güç tüketim modeli ve çizelgelenebilirlik kısıtıyla sunulmuştur. Hem eniyileme problemlerinin optimal koşullarının analitik çözümlemesi hem de hızlı buluşsal yöntemler kullanılarak, ağ düğümlerinin güç tüketiminin azalmayan fonksiyonu olan herhangi bir amaç işlevi, farklı çizelgeleme algoritmaları ve kipleme teknikleri için uygulanabilir enerji-verimli kaynak özgüleme yöntemleri önerilmiştir.

İkinci olarak, KAKS'lerde çizelgeleme için algılayıcıların periyodik veri gönderimi özelliğinden faydalanan yeni bir uyarlanabilirlik ölçevi önermekteyiz. Bu ölçevi kullanarak, amaç işlevi çizelgelerin veri paketi kayıplarına ve

ağ ilingesindeki değişikliklere maksimum seviyede uyarlanabilirlik göstermesini sağlamak olan ve aynı zamanda algılayıcıların periyodik veri gönderimi, veri gönderim süresi, enerji ve güvenilirlik kısıtlarını da dikkate alan özgün bir eniyileme yaklaşımı sunmaktayız. Bu eniyileme problemini hızlı ve etkin şekilde çözmek amacıyla, buluşsal çizelgeleme algoritmaları için bir tasarım çerçevesi önerilmiştir.

Üçüncü olarak, kablosuz tasarsız ağlarda minimum uzunlukta çizelgeleme problemi, bir veri paketinin gönderimi süresince durdurulamayacağı ve birden fazla zaman yuvasında gönderilmek üzere birden fazla veri yığınına ayrılamayacağı anlamına gelen paket-bazlı veri gönderimi ele alınarak incelenmiştir. Ortak eniyileme probleminin formüle edilmesi ve zorluğunun ispatlanmasının ardından, güç yönetimi ve veri transfer hızı adaptasyonu probleminin çizelgeleme probleminden ayrılabileceği gösterilmiştir. Daha sonra, güç yönetimi ve veri transfer hızı adaptasyonu problemi, her biri tamsayıda paketi aynı anda gönderecek linklerin eşzamanlı gönderimi için gereken sürenin enküçültmesi şeklinde özgün bir problem olarak verilmiştir. Bu problemin çözümü, ortak eniyileme probleminin büyük-ölçekli bir Tamsayı Programlama (TP) problemi olan saf bir çizelgeleme problemi olarak formüle edilmesini sağlamaktadır. Bu TP probleminin çözümü için eniyileme teorisinde kullanılan bazı algoritmaları temel alan buluşsal algoritmalar kullanılmıştır.

Dördüncü olarak, Ultra Geniş Bant kablosuz ağlarda gecikme enküçültme ve gecikme kısıtlı enerji enküçültme problemleri için, optimal güç kontolü ve veri gönderim hızı adaptasyonu stratejileri ve hızlı çizelgeleme algoritmaları önerilmiştir.

Contents

Chapter 1

Introduction

Wireless networks provide ubiquitous access to information. As the wireless applications and services in wireless networked control systems (WNCSs) and machine-to-machine communication (M2M) over cellular systems become widespread on this premise, energy and delay gain more importance in design of communication systems. Energy appeared as one of the key design concerns, mostly as wireless sensor networks proliferate with the advancements in sensor and radio module development, due to the fact that wireless sensors or devices are mostly operated via batteries with limited capacity or small amounts of harvested energy. Therefore, in order to achieve desired lifetime for the network, energy consumption of the nodes should be limited. Delay, on the other hand, is another important limitation since many wireless applications and services such as real-time surveillance and networked control systems requires timely delivery of critical messages with high level of reliability.

WNCSs are spatially distributed systems in which sensors, actuators, and controllers connect through a wireless network instead of traditional point-to-point links [1]. WNCSs have a tremendous potential to improve the efficiency of many large-scale distributed systems in industrial automation [2, 3], building automation [4], air transportation [5], and smart grid [6]. Transmitting sensor measurements and control commands over wireless links provide many benefits such as the ease of installation and maintenance, low complexity and cost, and large flexibility to accommodate the modification and upgrade of the components in many control applications. Several industrial organizations, such as International Society of Automation (ISA) [7], Highway Addressable Remote Transducer (HART) [8], and Wireless Industrial Networking Alliance (WINA) [9], have been actively pushing the application of wireless technologies in the control applications. However, building a WNCS is very challenging since control systems often have stringent requirements on timing and reliability, which are difficult to attain by wireless sensor networks due to the adverse properties of the wireless communication and limited battery resources of the nodes.

M2M communication is another promising area as the Internet of Things and Smart City paradigms emerge for the realization of full potential of wireless networking [10]. The next generation cellular systems are expected to support massive M2M applications such as remote and mobile healthcare, smart metering, elderly assistance, smart agriculture and traffic management providing solutions to the social and industrial demands [11]. Many of these applications are time-critical demanding the information to be transferred within a given time interval successfully. Moreover they require low power consumption since machine type communication (MTC) devices mostly have limited energy budgets considering the fact that they are mostly expected to function over years without any interruption. Deployment of MTC devices in scattered areas where battery replacements are not possible requires control of consumed energy.

In order to handle the energy and delay requirements of wireless networks, resource allocation schemes in which scheduling and allocation of transmission rates and powers of the nodes are determined are of paramount importance. This thesis investigates the resource allocation in delay and energy constrained wireless networks. It is organized as different chapters focusing on a variety of problems with different delay and energy related objectives and requirements.

1.1 Contributions

The main contributions of the thesis are listed as follows:

- A novel framework for the joint optimization of controller and communication systems encompassing efficient abstractions of both systems is provided for WNCSs, which may lead to broader adoption and real-world deployment. For the control system model, stochastic maximum allowable transfer interval and maximum allowable delay constraints are used as efficient abstractions considering wireless network induced imperfections. Communication system is presented via power consumption model and schedulability constraint using the control system parameters. Using both analytical analysis on the optimality conditions of the optimization problems and fast heuristics, energy efficient resource allocation methods are proposed for different power-related objectives including total power consumption and maximum power consumption, different modulation schemes including MQAM and MFSK and scheduling algorithms.

- For WNCSs, we introduce a new adaptivity metric for scheduling which exploits the periodic transmission nature of the sensor nodes. We provide a novel joint optimization framework with the objective of providing maximum adaptivity in schedule while considering the constraints of the individual sensor nodes in terms of periodic data generation, transmission delay, energy and reliability. In order to solve the joint optimization problem fast and efficiently, we propose a design framework for heuristic scheduling algorithms. Moreover, we formulate a novel power control and rate adaptation problem with the objective of minimizing the time required for the concurrent transmission of a set of sensor nodes while satisfying their transmission delay, reliability and energy consumption requirements for general wireless networks by adopting the Shannon's capacity formulation for the first time in the literature. We prove that the problem is non-convex and then propose a polynomial time algorithm that provides the optimal solution within a predetermined error bound following the derivation of the optimality conditions.

- For wireless ad hoc networks, minimum length scheduling problem is revisited considering packet traffic demands on the nodes where packet transmission of a node cannot be interrupted and divided into multiple chunks to be transmitted in different time slots. Using this packetized transmission constraint, we first formulate the joint optimization of the power control, rate adaptation and scheduling for minimizing the schedule length of a wireless ad hoc network and demonstrate the hardness of this problem. We then show that the power control and rate adaptation problem can be separated from the scheduling problem and introduce a novel problem for the optimization of power control and rate adaptation where the time required for the concurrent transmission of a set of links each having an integer number of packets is minimized. Solving this problem allows formulation of the joint problem as a pure scheduling problem in Integer Programming (IP) form. The resulting IP problem is large-scale with the number of variables exponential in the number of the nodes. We therefore incorporate elegant Branch and Price Method and Column Generation method based heuristic algorithms to solve this IP problem.

- For UWB wireless networks, fast scheduling algorithms are proposed along with the optimal power and rate allocation rules for delay minimization and delay constrained energy minimization problems.

1.2 Organization

The rest of the thesis is organized as follows.

- Chapter 2 presents the study on the joint optimization of controller and communication systems taking into account all the wireless network induced imperfections including packet error and delay; the parameters of the wireless communication system including the transmission power, rate and scheduling of the network nodes; and the parameters of the controller including the sampling period [12]. The objective of the optimization problem is to minimize the power consumption of the communication system whereas the constraints guarantee the stability of the control system and the schedulability in the communication system. The presented study proposes solutions for MQAM modulation and Earliest Deadline First (EDF) scheduling algorithm.

- Chapter 3 generalizes the optimization framework in Chapter 2 for any non-decreasing function of the power consumption of the nodes in the objective and many modulation schemes and scheduling algorithms satisfying certain properties [13, 14].

- Chapter 4 discusses a novel approach to provide maximum level of adaptivity accommodating packet losses and changes in network topology while exploiting periodic nature of the sensor node transmissions in WNCSs [15].

- Chapter 5 investigates minimum length scheduling problem in wireless ad hoc networks as determining the optimal power control, rate adaptation and scheduling with the objective of minimizing the schedule length given the packet traffic demands of the links and the constraints of packet transmissions [16].

- Chapter 6 studies the optimal power control, rate adaptation and scheduling for UWB wireless networks for two different problems which are delay minimization [17] and delay constrained energy minimization [18].

- Chapter 7 presents the concluding remarks and possible research directions.

Chapter 2

Minimum Energy Data Transmission for Wireless Networked Control Systems

In this study, we provide a framework for the joint optimization of controller and communication systems in WNCSs encompassing efficient abstractions of both systems. The objective of the optimization problem is to minimize the power consumption of the communication system due to the limited lifetime of the battery-operated wireless nodes. The constraints of the problem are the schedulability and maximum transmit power restrictions of the communication system, and the reliability and delay requirements of the control system to guarantee its stability. The formulation comprises communication system parameters including transmission power, rate and scheduling, and control system parameters including sampling period. The resulting problem is a Mixed-Integer Programming problem, which is NP-Hard. However, analyzing the optimality conditions on the variables of the problem allows us to reduce it to an Integer Programming problem for which we propose an efficient solution method based on its relaxation. Simulations demonstrate that the proposed method performs very close to optimal and much better than the traditional separate design of these systems.

2.1 Introduction

Wireless Networked Control Systems (WNCSs) are spatially distributed systems in which sensors, actuators, and controllers connect through a wireless network instead of traditional point-to-point links [1]. WNCSs have a tremendous potential to improve the efficiency of many large-scale distributed systems in industrial automation [2, 3], building automation [4], automated highway [19], air transportation [5], and smart grid [6]. Transmitting sensor measurements and control commands over wireless links provide many benefits such as the ease of installation and maintenance, low complexity and cost, and large flexibility to accommodate the modification and upgrade of the components in many control applications. Several industrial organizations, such as International Society of Automation (ISA) [7], Highway Addressable Remote Transducer (HART) [8], and Wireless Industrial Networking Alliance (WINA) [9], have been actively pushing the application of wireless technologies in the control applications.

Building a WNCS is very challenging since control systems often have stringent requirements on timing and reliability, which are difficult to attain by wireless sensor networks due to the adverse properties of the wireless communication and limited battery resources of the nodes. The wireless communication introduces non-zero packet error probability caused by the unreliability of the wireless transmissions, non-zero delay due to the packet transmission and shared wireless medium, and sampling and quantization errors since the signals are transmitted over the network via packets. Decreasing the packet error probability, delay and sampling period improves the performance of the control system at the cost of more energy consumed in the communication. There is an increasing need for methods and analysis tools that are able to quantify the joint performance of these systems in terms of the communication parameters, including the transmission power, rate and scheduling of the network nodes, the control parameters including the sampling period and wireless communication induced imperfections.

The communication system design for Networked Control Systems (NCS) has received little attention in the literature mainly due to the lack of efficient abstractions of both control and communication systems. The lack of abstractions led to either the simplification of the problem by excluding the key system parameters from the formulation or the solutions through numerical methods avoiding the widespread use of the formulations. Assuming no packet error occurs unless there is a collision in the network, the optimization of the scheduling given the sampling period and delay requirements of the sensors [20, 21] or the optimization of the sampling period and delay parameters of the sensors to minimize the overall performance loss while ensuring the system schedulability [22, 23] have been formulated. These formulations however cannot be applied to WNCS where the packet error probability is non-zero at all times. Some of the prior work on the communication system design for WNCS focus on ensuring low deterministic end-to-end delay and controlled jitter to real-time traffic across a very large mesh network distributed over a large area in a globally synchronized multi-channel Time Division Multiple Access (TDMA) [7, 8, 24, 25]. The optimization problem formulations for WNCS on the other hand aim to determine the best values of the sampling period and network scheduling parameters by using different objective and constraint functions. Some of them aim to minimize the energy consumption of the network subject to the packet loss probability and delay distribution that are derived from the desired control cost [26]. Others aim to maximize the performance of control systems subject to the wireless capacity constraints of the links and the delay requirement of the control system [27]. None of these works however consider the key parameters of the wireless communication system including the transmission power and rate of the links as a system variable to be optimized.

A more general framework where the optimal values of the parameters of the link layer, medium access control layer, and sampling period that minimize the control cost subject to the delay distribution and the packet error probability constraints is proposed in [28]. However, a suboptimal solution is obtained by an iterative numerical method due to the difficulty of dealing with the linear quadratic cost function representing the performance of the control system in the objective of the optimization problem. This numerical approach is hard to be generalized for different requirements of various control applications.

Joint optimization of the transmission power control, rate adaptation, and scheduling has been widely studied for delay constrained energy minimization in general purpose wireless networks [29]–[36]. Different algorithms have been proposed for the solution of the optimization problem depending on the delay constraint in terms of either a single deadline to all packets [29]–[32] or individual deadlines for each packet [33]–[35] and the usage of either a fixed capacity battery [29, 30, 33, 34, 35] or a finite storage capacity rechargeable battery [31, 32]. None of these algorithms however have been extended to formulate the joint optimization of power, rate and scheduling for energy minimization while meeting a certain controller performance.

The goal of this study is to study the joint optimization of controller and communication systems taking into account all the wireless network induced imperfections including packet error and delay; the parameters of the wireless communication system including the transmission power, rate and scheduling of the network nodes; and the parameters of the controller including the sampling period. The objective of the optimization problem is to minimize the power consumption of the communication system whereas the constraints guarantee the stability of the control system and the schedulability in the communication system. The original contributions of this study are listed as follows:

- We provide a framework for the joint optimization of controller and communication systems encompassing efficient abstractions of both systems, which may lead to broader adoption and real-world deployment.
- We formulate the joint optimization of controller and communication systems for MQAM modulation in a network containing multiple sensors communicating with their corresponding controllers as a Mixed Integer Programming problem.
- We propose an efficient solution method for the formulated optimization problem based on the analysis of the relations between the optimal values of the decision variables.

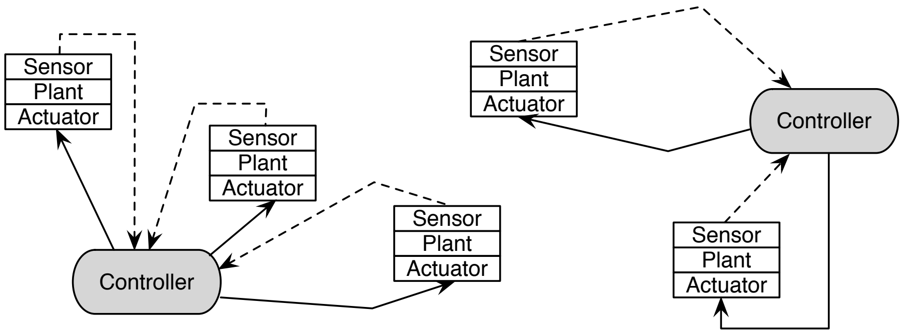

Figure 2.1: Overview of the WNCS setup.

- We prove the energy saving of the proposed joint optimization problem over the traditional separate design of controller and communication systems and the closeness to the optimality of the proposed solution method via extensive simulations.

The rest of this chapter is organized as follows. Section 2.2 describes the system model and the assumptions used throughout the study. Section 2.3 presents the control and communication system models. The joint optimization of controller and communication systems has been formulated in Section 2.4. Simulations are presented in Section 2.5. Finally concluding remarks are given in Section 2.6.

2.2 System Model and Assumptions

The system model and assumptions are detailed as follows.

1. Fig. 2.1 depicts the architecture of a WNCS where multiple plants are controlled over a wireless network. We assume that a sensor node is attached to each plant. Outputs of the plants are sampled at periodic intervals by the sensors and forwarded to the controller over a wireless network, which induces delays and packet errors. When the controller receives the measurements, a new control command is computed. The control command is forwarded to the actuator attached to the plant. We assume that the controller commands are successfully

received by the actuator. Many practical NCSs have several sensing channels, whereas the controllers are collocated with the actuators, as in heat, ventilation and air conditioning control systems, because the control command is very critical [37]. One of the controllers is assigned as the network manager.

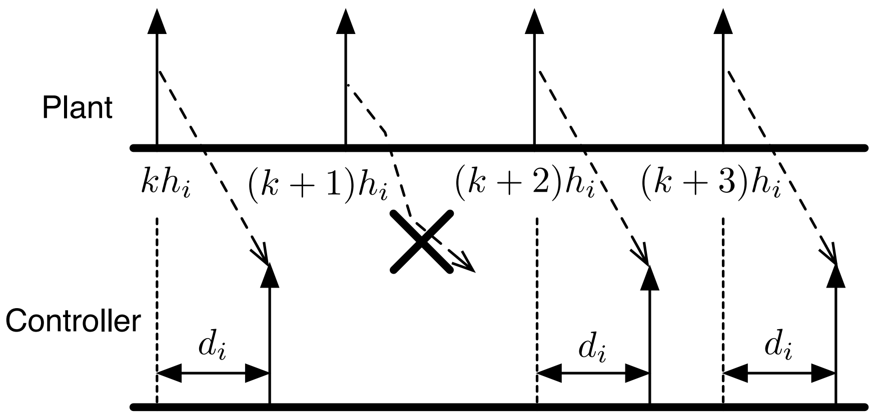

Figure 2.2: Timing diagram between a plant and a controller communicating over a wireless network.

2. Fig. 2.2 illustrates various possible situations of the information exchange between a plant and a controller. Let us denote the sampling period of node i by h_i, the transmission delay of the packet that includes the sample of the sensor node by d_i, and the packet error probability by p_i. We assume that $d_i \leq h_i$ to guarantee that the packets arrive to the controller in the correct order. The retransmission of outdated packets are generally not useful since the latest information of the plant state is the most critical information for control applications [1]. We also assume that the packet error is modeled as a Bernoulli random process with probability p_i for node i to simplify the problem.

3. We consider TDMA as MAC protocol since TDMA provides both delay guarantee and energy efficiency to the networks with predetermined topology and data generation patterns [38] thus is widely used in industrial control applications [7, 8]. The details of synchronization and

topology discovery mechanisms for energy efficient TDMA are out of scope of this study and can be found in [38].

4. The time is partitioned into frames. Each frame is further divided into a beacon and time slots. The beacon is used by the network manager to provide time synchronization within the WNCS and broadcast updates on the scheduling decisions. The scheduling decisions include the time slot allocation and the values of the optimal node parameters including the transmission power, rate and sampling period corresponding to each sensor node. We assume that no concurrent transmissions are scheduled. The network manager continually monitors the received power and the packet error rate over each link. If the channel conditions do not change, the beacon only provides synchronization information. Otherwise, the beacon also includes the updates on the scheduling decisions. Despite the updates on the schedule and optimal node parameters, the length of the frame is fixed as derived mathematically in Section 2.4.

5. The nodes are assumed to operate their radio in sleep mode when they are not scheduled to transmit or receive a packet, and transient mode when they switch from sleep mode to active mode to transmit or receive a packet and vice versa. We consider only the energy consumption in the transmission of the packets in the optimization problem because it is much larger than that in the sleep and transient modes [29], [36] and the energy consumption in the reception of the beacon packets is fixed.

2.3 Control and Communication System Models

This section provides control and communication system models based on the consideration of the stability of the control system, the restrictions related to the wireless communication and the power consumption of the communication system.

2.3.1 Control System Model

The performance and stability conditions for the control systems have been formulated in the form of Maximum Allowable Transfer Interval (MATI), defined as the maximum allowed time interval between subsequent state vector reports from the sensor nodes to the controller, and Maximum Allowable Delay (MAD), defined as the maximum allowed packet delay for the transmission from the sensor node to the controller, in [39, 40, 41, 42]. Such hard real-time guarantees can be satisfied by wireline networks but is an unreasonable expectation for wireless networks where the packet error probability is greater than zero at any point in time. Hence, many control applications such as wireless industrial automation [7], air transportation systems [5] and autonomous vehicular systems [43] set a stochastic MATI constraint in the form of keeping the time interval between subsequent state vector reports above the MATI value with a predefined probability to guarantee the stability of control systems. Stochastic MATI and MAD constraints are efficient abstractions of the performance of the control systems however have not been considered before in the joint design with the communication systems.

2.3.1.1 Stochastic MATI Constraint

Stochastic MATI constraint is formulated as

$$\Pr\left[\mu_i(h_i, d_i, p_i) \leq \Omega\right] \geq \delta \tag{2.1}$$

where μ_i is the time interval between subsequent state vector reports of node i as a function of h_i, p_i and d_i; Ω is the MATI; and δ is the minimum probability with which MATI should be achieved. The values of Ω and δ are determined by the control system.

In order that the time interval between subsequent state vector reports of node i is less than Ω, there should be at least one successful transmission within Ω. Given h_i and Ω, the number of reception opportunities of the state vector reports is $\left\lfloor \frac{\Omega}{h_i} \right\rfloor$. Based on the assumption that the packet error is modeled as a Bernoulli random process with probability p_i for node i, stochastic MATI constraint given in Eq. (2.1) can be rewritten as

$$1 - p_i^{\left\lfloor \frac{\Omega}{h_i} \right\rfloor} \geq \delta \,. \tag{2.2}$$

Many control applications and standards define the stochastic MATI requirements to guarantee the stability of the control systems. In industrial automation, closed-loop machine controls have specified $\Omega = 100$ ms and $\delta = 0.999$ [7, 44]. Moreover, to allow IEEE 802.15.4 devices [45] to support a wide range of industrial applications, IEEE 802.15.4e standard [46] specifies an amendment to the IEEE 802.15.4 standard to enhance its latency and reliability performances for industrial automation. They have specified $\Omega = 10$ ms and $\delta = 0.99$. In addition, the air transportation system requires $\Omega = 4.8$ s and $\delta = 0.95$ [5]. Furthermore, the requirements of cooperative vehicular safety applications are given by $\Omega = 100$ ms and $\delta = 0.95$ [43].

2.3.1.2 MAD Constraint

MAD constraint should be included in addition to the MATI constraint to guarantee the performance and stability of the control systems [42]. MAD constraint is formulated as

$$d_i \leq \Delta \tag{2.3}$$

where Δ is the MAD to stabilize the control system. The value of Δ is determined by the control system and typically on the order of a few tens of milliseconds for fast control applications [7, 44, 47].

2.3.2 Communication System Model

This section provides the formulations of the power consumption of the sensor nodes, the maximum transmit power and schedulability constraints of the communication system. The power consumption of the sensor node is formulated as a function of its sampling period, delay and packet error probability for both uncoded and trellis-coded MQAM modulation. The reason for choosing a specific modulation and coding strategy instead of the information theoretically optimal channel coding schemes is to include the trade-off between the packet error probability and power consumption in practical communication systems. The optimal channel coding schemes do not allow including such trade-off since they employ randomly generated codes with exponentially small probability of error for long block lengths [48]. Moreover, choosing a specific modulation strategy allows us to include the circuit power in addition to the transmission power consumption [36]. The proposed formulation for MQAM modulation can be extended for any modulation strategy.

2.3.2.1 Power Consumption

The key design concern of the communication protocol is to limit the energy consumed by the sensor devices. This will avoid battery replacement resulting in an affordable WNCS deployment.

The power consumption of the node i as a function of its sampling period, delay and packet error probability is formulated as

$$W_i\left(h_i, d_i(b_i), p_i\right) = \frac{\left(W_i^t\left(d_i(b_i), p_i\right) + W_i^c\right) d_i(b_i)}{h_i} \tag{2.4}$$

where b_i is the number of bits used per symbol and d_i is represented as a function of b_i for a given modulation scheme, W_i^t is the transmission power calculated as a function of the parameters d_i and p_i for a given modulation and channel coding, and W_i^c is the circuit power consumption in the active mode at the transmitter. The numerator provides the energy consumed for the transmission of duration $d_i(b_i)$ in one sampling period h_i. In the following, we present the formulation of $W_i^t(d_i(b_i), p_i)$ for MQAM modulation scheme based on the assumption of an additive white Gaussian noise (AWGN) channel by summarizing the derivation procedure in [36] for completeness.

Let us denote the number of bits sent by node i by L_i. $b_i = \log_2 M$ is the number of bits per symbol for MQAM and $b_i = 2$ is the minimum allowable value ensuring MQAM is well defined [36]. The number of MQAM symbols needed to send L_i bits is equal to both L_i/b_i and d_i/T^s, where T^s is the symbol period. If square pulses are used and $T^s \approx 1/B$, where B is the bandwidth, then

$$d_i(b_i) \approx \frac{L_i}{Bb_i}. \tag{2.5}$$

A bound on the probability of bit error for MQAM is given by [49]

$$P_i^b \leq \frac{4}{b_i}\left(1 - \frac{1}{\sqrt{2^{b_i}}}\right) e^{-\frac{3}{2^{b_i}-1}\frac{\gamma_i}{2}} \tag{2.6}$$

where P_i^b is the probability of bit error of node i, γ_i is the Signal-to-Noise Ratio (SNR) of node i defined as $\gamma_i = \frac{W_i^r}{2B\sigma^2 N^f}$, where W_i^r is the signal power received from node i, σ^2 is the power spectral density of the AWGN and N^f is the noise figure. By approximating this bound as an equality and using the relation between the transmission power and received power as $W_i^t = W_i^r G_i$ where G_i is the power gain factor, and the function mapping b_i to d_i in Eq. 2.5, the transmission power as a function of b_i and P_i^b is given by

$$W_i^t(b_i, p_i) \approx \frac{4}{3} N^f \sigma^2 B G_i \left(2^{b_i} - 1\right) \ln \frac{4\left(1 - 2^{-\frac{b_i}{2}}\right)}{b_i P_i^b}. \tag{2.7}$$

If no error control mechanism is used, a packet is considered to be in error in the presence of one or more bit errors. Assuming independent bit errors, the packet error probability p_i is given by

$$p_i = 1 - \left(1 - P_i^b\right)^{L_i} \approx P_i^b L_i \tag{2.8}$$

for small values of P_i^b where the approximation is obtained by finding the Taylor series expansion of $\left(1 - P_i^b\right)^{L_i}$ and ignoring higher order terms.

If forward error correction codes are used, the required value of SNR to meet a target bit error probability decreases by a factor called coding gain denoted by G^c at the cost of resulting bandwidth expansion in order to communicate the extra redundant bits. Fortunately, bandwidth expansion can be circumvented when the channel coding and modulation processes are jointly designed, for instance, in trellis-coded MQAM [49]. The decrease in the required SNR results in the decrease in the required transmission power W_i^t for the same bit error probability.

By substituting Eqs. 2.7 and 2.8 into Eq. 2.4, the power consumption for trellis-coded MQAM is formulated as

$$\begin{aligned} W_i(h_i, b_i, p_i) = & \frac{4N^f \sigma^2 G_i L_i}{3G^c h_i b_i}(2^{b_i} - 1) \ln \frac{4L_i(1 - 2^{-\frac{b_i}{2}})}{p_i b_i} \\ & + \frac{W_i^c L_i}{B h_i b_i}. \end{aligned}$$

2.3.2.2 Maximum Transmit Power Constraint

We assume that there exists a maximum power level, denoted by $W^{t,\max}$, that a node can use for transmission. This is enforced by the limited weight and size of the sensor nodes. The maximum transmit power constraint is formulated as

$$W_i^t(b_i, p_i) \leq W^{t,\max} \tag{2.9}$$

where

$$W_i^t(b_i, p_i) = \frac{4N^f \sigma^2 G_i B}{3G^c}(2^{b_i} - 1) \ln \frac{4L_i(1 - 2^{-\frac{b_i}{2}})}{p_i b_i}. \tag{2.10}$$

2.3.2.3 Schedulability Constraint

The schedulability constraint represents the allocation of the transmission times of multiple sensor nodes in the network in the absence of concurrent transmissions using pre-emptive Earliest Deadline First (EDF) scheduling algorithm. EDF is a dynamic scheduling algorithm where the task closest to its deadline is scheduled whenever a scheduling event occurs. EDF has been proven to be an optimal uniprocessor scheduling algorithm for periodic tasks with certain deadlines using pre-emption, in the following sense: If a real-time task set cannot be scheduled by EDF, then this task set cannot be scheduled by any algorithm [50]. For the case where there exists a node such that its deadline is not equal to its period, i.e. $\exists i \in [1, N], \Delta \neq h_i$, there exists an exact schedulability analysis based on the simulation of the EDF algorithm for a time duration formulated as a function of d_i, h_i, Δ and task arrival times a_i for all $i \in [1, N]$: The EDF schedule is generated for the specified time duration and given values of d_i, h_i, Δ and a_i for all $i \in [1, N]$ are declared feasible if and only if no deadlines are missed in this schedule [51, 52]. This simulation based schedulability analysis however cannot be used in an optimization framework where the explicit formulations of the necessary and sufficient conditions for the schedulability are required. Since such explicit formulations do not exist, we propose the use of the schedulability constraint given by

$$\sum_{i=1}^{N} \frac{d_i}{h_i} \leq \beta \tag{2.11}$$

where β is the utilization bound satisfying $0 < \beta \leq 1$ in our optimization framework. The use of this new schedulability constraint is demonstrated to provide near-optimal solutions via simulations in Section 2.5 when the value of β is adapted to the network topology and requirements as described in the following.

Lemma 1. *For $\beta = \beta_{nec} = 1$, the schedulability constraint given by Eq. (2.11) is a necessary condition for a feasible schedule.*

Proof: Every term $\frac{d_i}{h_i}$ in Eq. (2.11) represents the ratio of the total time duration allocated for the transmission of node i to the schedule length. Since there are no concurrent transmissions, the sum of these terms gives the ratio of the total time duration allocated to all the nodes $i \in [1, N]$ to the schedule length. The value of this ratio cannot exceed 1 since the sum of the time durations allocated to all the nodes in a schedule cannot exceed the schedule length. $\square$

Lemma 2. *For $\beta = \beta_{suf} = \min\{1, \min_{i\in[1,N]} \frac{\Delta}{h_i}\}$, the schedulability constraint given by Eq. (2.11) is a sufficient condition for a feasible schedule.*

Proof: For the specified value of $\beta = \beta_{suf}$, Eq. (2.11) can be reformulated as

$$\sum_{i=1}^{N} \frac{d_i}{h_i \min\{1, \min_{i\in[1,N]} \frac{\Delta}{h_i}\}} \leq 1 \tag{2.12}$$

Since

$$\sum_{i=1}^{N} \frac{d_i}{\min\{\Delta, h_i\}} \leq 1 \tag{2.13}$$

is a sufficient condition for a feasible schedule [53], and

$$\sum_{i=1}^{N} \frac{d_i}{h_i \min\{1, \min_{i\in[1,N]} \frac{\Delta}{h_i}\}} \geq \sum_{i=1}^{N} \frac{d_i}{\min\{\Delta, h_i\}} \tag{2.14}$$

Eq. (2.11) is also a sufficient condition for $\beta = \beta_{suf}$. □

β_{nec} and β_{suf} in Lemmas 1 and 2 provide the upper and lower values of the utilization bound respectively. Using a β value larger than β_{nec} expands the feasible region of the corresponding optimization problem by infeasible schedules whereas using a β value smaller than β_{suf} shrinks the feasible region by removing the schedules satisfying $\beta < \sum_{i=1}^{N} \frac{d_i}{h_i} \leq \beta_{suf}$ from the feasible region of the corresponding optimization problem. Moreover, increasing the value of β expands the feasible region of the optimization problem. Therefore, the maximum value of $\beta \in [\beta_{suf}, \beta_{nec}]$ that yields a feasible schedule for an optimization problem needs to be determined as explained in detail in Section 2.4.2.

2.4 Joint Optimization of Control and Communication Systems

This section formulates the problem of the joint optimization of control and communication systems with the objective of minimizing the power consumption of the network subject to the stochastic MATI and MAD constraints guaranteeing the stability of the control systems and maximum transmit power and schedulability constraints of the wireless communication system. The key in formulating this optimization problem is the trade-off between the control performance and the power consumption of the wireless communication network. Decreasing the packet error probability, delay and sampling period improves the performance of the control system at the cost of more energy consumed in the communication. The sampling period, packet error probability, and delay must therefore be flexible design parameters that need to be adequate for the control requirements.

The joint optimization of control and communication systems is formulated as

$$
\begin{aligned}
\min_{h_i, b_i, p_i, i \in [1,N]} \quad & \sum_{i=1}^{N} W_i(h_i, b_i, p_i) && \text{(2.15a)} \\
\text{s.t.} \quad & \sum_{i=1}^{N} \frac{d_i(b_i)}{h_i} \leq \beta \,, && \text{(2.15b)} \\
& 1 - p_i^{\left\lfloor \frac{\Omega}{h_i} \right\rfloor} \geq \delta, \quad \forall i \in [1, N] \,, && \text{(2.15c)} \\
& 0 < d_i(b_i) \leq \min \{\Delta, h_i\}, && \text{(2.15d)} \\
& 0 < h_i \leq \Omega, && \text{(2.15e)} \\
& 0 < p_i < 1, && \text{(2.15f)} \\
& W_i^t(b_i, p_i) \leq W^{t,\max}, && \text{(2.15g)}
\end{aligned}
$$

where N is the number of nodes in the network.

The goal of the optimization problem is to minimize the total power consumption in the network. Eq. (2.15b) represents the schedulability constraint. Eqs. (2.15c) and (2.15d) represent the stochastic MATI and MAD constraints respectively. Eq. (2.15e) states that the sampling period of the nodes must be less than or equal to the MATI. Eq. (2.15f) states the lower and upper bounds for the packet error probability. Finally, Eq. (2.15g) represents the maximum transmit power constraint. The variables of the problem are $h_i, i \in [1, N]$, the sampling period of the nodes; $b_i, i \in [1, N]$, the number of bits used per symbol for each node; and $p_i, i \in [1, N]$, the packet error probability of the nodes.

This optimization problem is non-convex Mixed-Integer Programming problem thus difficult to solve for the global optimum [54]. We now analyze the optimality conditions for this problem and propose efficient solution methods for the network containing one sensor node, i.e. N=1, and multiple sensor nodes in Sections 2.4.1 and 2.4.2 respectively.

2.4.1 One Sensor Case

The joint optimization of control and communication systems for the network containing one sensor node is formulated as

$$
\begin{aligned}
\min_{h_i, b_i, q_i} \quad & C_{i1} \frac{2^{b_i} - 1}{h_i b_i} \left(\ln \frac{4L_i \left(1 - 2^{\frac{-b_i}{2}}\right)}{b_i} - q_i \right) + \frac{W_i^c C_{i2}}{h_i b_i} && \text{(2.16a)} \\
\text{s.t.} \quad & \left\lfloor \frac{\Omega}{h_i} \right\rfloor q_i - \ln(1-\delta) \leq 0, && \text{(2.16b)} \\
& 0 < d_i(b_i) \leq \min\{\Delta, h_i\}, && \text{(2.16c)} \\
& 0 < h_i \leq \Omega, && \text{(2.16d)} \\
& q_i \leq 0, && \text{(2.16e)} \\
& \frac{C_{i1}}{C_{i2}} \left(2^{b_i} - 1\right) \left(\ln \frac{4L_i \left(1 - 2^{-\frac{b_i}{2}}\right)}{b_i} - q_i \right) \leq W^{t,\max}. && \text{(2.16f)}
\end{aligned}
$$

where $C_{i1} = \frac{4N^f \sigma^2 G_i L_i}{3G^c}$ and $C_{i2} = \frac{L_i}{B}$. The new variable q_i is equal to $\ln p_i$. The objective function in Equation (2.16a) includes the power consumption formulation defined in Equation (2.15a) for $N = 1$ and MQAM. The constraint given in Eq. (2.15b) is eliminated since the constraint in Eq. (2.15d) is already sufficient for schedulability when $N = 1$. The constraints in Eqs. (2.16b), (2.16c), (2.16d), (2.16e) and (2.16f) correspond to the constraints given in Eqs. (2.15c), (2.15d), (2.15e), (2.15f) and (2.15g) respectively for $N = 1$ and MQAM.

The optimization problem is again a non-convex Mixed-Integer Programming problem thus difficult to solve for the global optimum [54]. Next, we will investigate the relations between the optimal values of h_i and q_i, denoted by h_i^* and q_i^* respectively, to that of b_i, denoted by b_i^*, so that we can rewrite this problem as an Integer Programming (IP) problem, including only b_i as a variable and eliminating h_i and q_i, for which there are efficient approximation algorithms [54].

Lemma 3. *The optimal value of the sampling period is given by*

$$\frac{\Omega}{h_i^*} = k_i \tag{2.17}$$

where k_i is a positive integer.

Proof: We prove the Lemma by contradiction. Suppose that $\frac{\Omega}{h_i^*}$ is not a positive integer then $\left\lfloor \frac{\Omega}{h_i^*} \right\rfloor < \frac{\Omega}{h_i^*}$. If h_i^* increases such that the equality $\left\lfloor \frac{\Omega}{h_i^*} \right\rfloor = \frac{\Omega}{h_i^*}$ holds for the first time while satisfying the upper bound given in Eq. (2.16d), the stochastic MATI constraint given in Eq. (2.16b) still holds since the value of $\left\lfloor \frac{\Omega}{h_i^*} \right\rfloor$ does not change. The remaining constraint including h_i given in Eq. (2.16c) also still holds with this change. However, the power consumption given in Eq. (2.16a) decreases since it is a monotonically decreasing function of h_i. □

Lemma 4. *In the optimal solution, the stochastic MATI is satisfied with equality such that*

$$\frac{\Omega}{h_i^*} = \frac{\ln(1-\delta)}{q_i^*} = k_i \tag{2.18}$$

where k_i is a positive integer.

Proof: We prove the Lemma by contradiction. Suppose that

$$\frac{\Omega}{h_{i^*}} > \frac{\ln(1-\delta)}{q_i^*} \tag{2.19}$$

If q_i^* increases such that the stochastic MATI constraint is satisfied with equality, the constraints given in Eqs. (2.16e) and (2.16f) still hold. However, the power consumption given in Eq. (2.16a) decreases since it is a monotonically decreasing function of q_i. Then the result follows when combined with Lemma 3. □

Next, we eliminate the variables h_i and q_i from the optimization problem (2.16) by using the expressions derived in Lemma 4 for their optimal values as a function of the single variable k_i. Note that k_i is the number of transmissions within Ω. The joint optimization of control and communication systems is then reformulated as

$$\begin{aligned}
\min_{b_i,k_i} \quad & C_{i1}\frac{(2^{b_i}-1)k_i}{\Omega b_i}\left(\ln\frac{4L_i(1-2^{\frac{-b_i}{2}})}{b_i}-\frac{\ln(1-\delta)}{k_i}\right) \\
& +\frac{W^c C_{i2} k_i}{\Omega b_i} && \text{(2.20a)} \\
\text{s.t.} \quad & 0 < d_i(b_i) \le \min\left\{\Delta, \frac{\Omega}{k_i}\right\}, && \text{(2.20b)} \\
& \frac{C_{i1}}{C_{i2}}\left(2^{b_i}-1\right)\left(\ln\frac{4L_i\left(1-2^{-\frac{b_i}{2}}\right)}{b_i}-\frac{\ln(1-\delta)}{k_i}\right) \\
& \le W^{t,\max}, && \text{(2.20c)}
\end{aligned}$$

where the constraints given in Eqs. (2.20b) and (2.20c) correspond to those in Eqs. (2.16c) and (2.16f) respectively and the remaining constraints in the optimization problem (2.16) are removed due to the additional constraint of k_i being a positive integer. The following lemma expresses the optimal value of k_i in terms of b_i so that the above optimization problem can be formulated with the variable b_i only.

Lemma 5. *Based on the assumption that the optimization problem (2.20) is feasible, the optimal value of k_i denoted by k_i^* is expressed as a function of b_i as*

$$k_i^* = k_i(b_i) = \max\left\{1, \left\lceil \frac{\ln(1-\delta)}{\ln\frac{4L_i(1-2^{-\frac{b_i}{2}})}{b_i} - \frac{W^{t,\max}C_{i2}}{G^c C_{i1}(2^{b_i}-1)}} \right\rceil\right\} \tag{2.21}$$

Proof: Since the power consumption given by Eq. (2.20a) is a monotonically increasing function of k_i, k_i^* is the minimum positive integer satisfying Eqs. (2.20b) and (2.20c). The second term of the maximum in Eq. (2.21) is the minimum k_i value satisfying Eq. (2.20c) since Eq. (2.20b) is satisfied by this minimum k_i value if there exists a feasible solution. □

Note that k_i is a non-decreasing function of b_i. Let us call $b_i^{\min}$ and $b_i^{\max}$ the minimum and maximum value of b_i satisfying the constraints given in Eqs. (2.20b) and (2.20c). In the following, Lemma 6 and Theorem 1 derive the relation between the optimal value of k_i and these boundary values $b_i^{\min}$ and $b_i^{\max}$.

Lemma 6. *For two feasible consecutive constellation size values b_i and b_i+1, if $k_i(b_i) \geq 2$, then $k_i(b_i+1) \geq \frac{3}{2}k_i(b_i)$.*

Proof: If $k_i(b_i) \geq 2$ then the second term in the maximum operator in Eq. (2.21) is effective such that

$$k_i(b_i) = \left\lceil \frac{\ln(1-\delta)}{\ln \frac{4L_i(1-2^{-\frac{b_i}{2}})}{b_i} - \frac{W^{t,\max}C_{i2}}{G^c C_{i1}(2^{b_i}-1)}} \right\rceil \tag{2.22}$$

Since $k_i(b_i) \geq 2$, the denominator of the above expression is negative. If b_i increases by 1, the first term in the denominator decreases by a factor of at most 2 whereas the second term in the denominator decreases by a factor of at least 2 resulting in a factor of at least 2 decrease in the denominator such that

$$\frac{\ln \frac{4L_i(1-2^{-\frac{b_i}{2}})}{b_i} - \frac{W^{t,\max}C_{i2}}{G^c C_{i1}(2^{b_i}-1)}}{\ln \frac{4L_i(1-2^{-\frac{b_i+1}{2}})}{b_i+1} - \frac{W^{t,\max}C_{i2}}{G^c C_{i1}(2^{b_i+1}-1)}} \geq 2 \tag{2.23}$$

Due to the ceiling operation in the expression of $k_i(b_i)$, a factor of at least 2 decrease in the denominator results in a factor of at least $\frac{3}{2}$ increase in $k_i(b_i)$ for $k_i(b_i) \geq 2$.

□

Theorem 1. *The optimal value of $k_i(b_i)$ is given by*

$$k_i^*(b_i) = k_i(b_i^{\min}) \tag{2.24}$$

Proof: We prove the Theorem by contradiction. Suppose that $k_i^*(b_i)$ is not equal to $k_i(b_i^{\min})$ such that $k_i^*(b_i) \geq 2$. Decreasing b_i by one decreases the power consumed in the actual data transmission; i.e., first term of Eq. (2.20a), since it is an increasing function of both b_i and k_i. Decreasing b_i by one also decreases the circuit power consumption; i.e., second term of Eq. (2.20a), since it is an increasing function of b_i for $b_i \geq 2$ based on Lemma 6 that states that $k_i(b_i-1) \leq \frac{2k_i^*(b_i)}{3}$ if $k_i(b_i-1) \geq 2$ and the fact that $k_i(b_i-1) \leq \frac{2k_i^*(b_i)}{3}$ if $k_i(b_i-1) = 1$ and $k_i^*(b_i) \geq 2$. The power consumption can therefore further decrease if we decrease b_i by one. This is a contradiction. □

The joint optimization of control and communication systems is then reformulated as

$$\begin{aligned}
\min_{b_i} \quad & C_{i1}\frac{(2^{b_i}-1)k_i(b_i^{\min})}{\Omega b_i}\left(\ln\frac{4L_i(1-2^{\frac{-b_i}{2}})}{b_i} - \frac{\ln(1-\delta)}{k_i(b_i^{\min})}\right) \\
& + \frac{W_i^c C_{i2} k_i(b_i^{\min})}{\Omega b_i} && \text{(2.25a)} \\
\text{s.t.} \quad & b_i^{\min} \leq b_i \leq b_i^{\max}, && \text{(2.25b)}
\end{aligned}$$

where $b_i^{\min}$ is determined by using Eq. (2.20b) and $b_i^{\max}$ is the maximum value of b_i such that the corresponding $k_i(b_i)$ value determined by using Eq. (2.21) is equal to $k_i(b_i^{\min})$.

This optimization problem is an IP problem for which there is no known polynomial-time algorithm [55]. We therefore propose the following heuristic algorithm. We first use the upper bound

$$\ln\frac{4L_i(1-2^{\frac{-b_i}{2}})}{b_i} \leq \ln\frac{4L_i(1-2^{\frac{-b_i^{\min}}{2}})}{b_i^{\min}}. \tag{2.26}$$

in the objective function (2.25a). The relaxation of the problem (2.25) in which the integrality constraint on the constellation size b_i is removed is then a convex optimization problem and can be solved optimally by using interior point method [54]. However, the solution of the relaxed problem is generally not integer. Since the constellation size must be integer, the optimal solution is one of the two neighboring integer values of the solution of the relaxed problem. The best integer solution can be determined by evaluating the power consumption corresponding to these two integer values and choosing the one with the minimum power.

In the following, we summarize the entire procedure for solving the joint optimization of control and communication systems for the network containing one sensor node:

1. Determine $b_i^{\min}$ and $k_i^*(b_i) = k_i(b_i^{\min})$: We determine the minimum b_i value based on Eq. (2.20b), which requires satisfying both $0 < d_i(b_i) \leq \Delta$ and $0 < d_i(b_i) \leq \frac{\Omega}{k_i(b_i)}$ constraints. We first determine the minimum b_i value satisfying $0 < d_i(b_i) \leq \Delta$. Then, $k_i(b_i)$ value is determined based on Eq. (2.21). If $0 < d_i(b_i) \leq \frac{\Omega}{k_i(b_i)}$ is also satisfied using these values of b_i and $k_i(b_i)$ then minimum feasible constellation size $b_i^{\min}$ is equal to that particular b_i. Otherwise, b_i is incremented by 1 until $0 < d_i(b_i) \leq \frac{\Omega}{k_i(b_i)}$ constraint is satisfied. Once $b_i^{\min}$ is determined, $k_i^*(b_i) = k_i(b_i^{\min})$.

2. Determine $b_i^{\max}$: Given the optimal value $k_i^*(b_i)$, $b_i^{\max}$ is the maximum value of b_i such that $k_i(b_i) = k_i(b_i^{\min})$ considering Eq. (2.21).

3. Determine h_i^* and q_i^*: Given the optimal value $k_i^*(b_i)$, the optimal h_i^* and q_i^* values are determined based on Eq. (2.18).

4. Determine b_i^*: Solve the relaxation of the optimization problem (2.25) using the bound given by Eq. (2.26) and determine the best integral solution as b_i^*.

The only suboptimality of this procedure comes from the use of the bound given by Eq. (2.26). The use of this bound results in an error less than 5% in the total power consumption when b_i is within the range $[2, 20]$ (which is a reasonable range for practical MQAM systems) and the other parameters are set as given in Table 2.1. However, the effect of this error on the optimal total power consumption is much less than 5% in most scenarios as illustrated through simulations in Section 2.5.

2.4.2 Multiple Sensor Case

The joint optimization of control and communication systems for the network containing multiple sensor nodes only brings the additional schedulability constraint given by Eq. (2.15b). The schedulability constraint is the only constraint requiring a joint decision for multiple sensor nodes since the remaining constraints represent the individual constraints of the nodes already considered in Section 2.4.1.

The findings derived in Lemmas 3-6 for the network containing one sensor node still hold in the existence of the schedulability constraint of multiple sensor nodes since they are not related to the schedulability constraint. The schedulability constraint of multiple sensor nodes can therefore be rewritten as

$$\sum_{i=1}^{N} \frac{C_{i2} k_i(b_i)}{\Omega b_i} \leq \beta \tag{2.27}$$

This schedulability constraint can be simplified based on the following Lemma.

Lemma 7. *The optimal value of $k_i(b_i)$ in the schedulability constraint given by Eq. (2.27) is equal to $k_i(b_i^{min})$ for each node $i \in [1, N]$.*

Proof: Using minimum $k_i(b_i)$ value for every $i \in [1, N]$ minimizes each term $\frac{C_{i2} k_i(b_i)}{\Omega b_i}$ in the schedulability constraint and each power consumption term in the objective function of the joint optimization problem (2.15) based on Theorem 1. Therefore, the sum of these terms are also minimized with $k_i(b_i^{min})$ for every $i \in [1, N]$.$\square$

The schedulability constraint can therefore be reformulated as

$$\sum_{i=1}^{N} \frac{C_{i2} k_i(b_i^{min})}{\Omega b_i} \leq \beta \tag{2.28}$$

The optimal solution procedure derived for one sensor case can be extended to the multiple sensor case by determining first $b_i^{\min}$, $k_i^*(b_i) = k_i(b_i^{\min})$, $b_i^{\max}$, h_i^* and q_i^* separately for each sensor $i \in [1, N]$ then b_i^* for $i \in [1, N]$ via solving the joint optimization problem formulated as

$$\min_{b_i, i \in [1,N]} \quad \sum_{i=1}^{N} \frac{C_{i1}(2^{b_i} - 1) k_i^{\min}}{\Omega b_i} \left(\ln \frac{4L_i(1 - 2^{\frac{-b_i}{2}})}{b_i} - \frac{\ln(1-\delta)}{k_i^{\min}} \right) + \frac{W_i^c C_{i2} k_i^{\min}}{\Omega b_i} \tag{2.29a}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} \frac{C_{i2} k_i(b_i^{\min})}{\Omega b_i} \leq \beta \,, \tag{2.29b}$$

$$b_i^{\min} \leq b_i \leq b_i^{\max} \,, \quad \forall i \in [1, N]. \tag{2.29c}$$

This optimization problem is an IP problem for which there is no known polynomial-time algorithm [55]. Similar to the one sensor case, the following heuristic algorithm is proposed to solve this problem. We first use the bound given in Eq. (2.26). The relaxation of this problem in which the integrality constraint on the constellation sizes is removed is again a convex optimization problem thus can be solved optimally by using interior point method [54]. The resulting optimal values of the constellation sizes are possibly non-integer therefore ceiled to obtain an integral solution while avoiding the violation of the schedulability constraint.

Algorithm 1 Energy Minimizing Schedule Generation Algorithm (EMSA)

$\beta_{suf} = \min\{1, \min_{i\in[1,N]} \frac{\Delta}{h_i^*}\}$; $\beta_{nec} = 1$;
$\beta_{low} = \beta_{suf}$; $\beta_{up} = \beta_{nec}$;
$\beta = \beta_{up}$;
$\{d_i^*, i \in [1,N]\} =$ **solveOptim**(β);
if isSchedulable($\{h_i^*, d_i^*, i \in [1,N]\}$) **then**
    $\beta_{opt} = \beta$;
else
    while $\beta_{up} - \beta_{low} >= \epsilon$ **do**
        $\beta = (\beta_{up} + \beta_{low})/2$;
        $\{d_i^*, i \in [1,N]\} =$ **solveOptim**(β);
        if isSchedulable($\{h_i^*, d_i^*, i \in [1,N]\}$) **then**
            $\beta_{low} = \beta$;
        else
            $\beta_{up} = \beta$;
        end if
    end while
    $\beta_{opt} = \beta_{low}$;
end if
$\{d_i^*, i \in [1,N]\} =$ **solveOptim**(β_{opt});
schedule $=$ **EDFSchedule**($\{h_i^*, d_i^*, i \in [1,N]\}$);

Based on the above observations, the length of the scheduling frame defined in Section 2.2 can be chosen to be equal to $n_{MATI}\Omega$ where n_{MATI} is a positive integer fixed over time. Since the optimal period h_i^* of each node $i \in [1,N]$ is given by $\frac{\Omega}{k_i^*}$ for some positive integer k_i^* due to Lemmas 3 and 5, the scheduling of the nodes repeats every time duration of length Ω if the channel conditions do not change. An integer multiple of the time duration Ω is therefore suitable to update the schedule. Since a beacon is transmitted at the beginning of each frame, the value of n_{MATI} is chosen depending on the synchronization accuracy required by the system and the speed at which the channel conditions change.

As explained in Section 2.3.2.3, we need to determine the maximum value of the utilization bound $\beta \in [\beta_{suf}, \beta_{nec}]$ for which the solution of the joint optimization problem (2.29) yields a feasible EDF schedule, which is denoted by β_{opt}. The overall procedure of generating the feasible EDF schedule corresponding to β_{opt} given in Algorithm 1 is described next. The upper and lower bounds of β_{opt}, denoted by β_{up} and β_{low} respectively, are initialized to $\beta_{nec} = 1$ and $\beta_{suf} = \min\{1, \min_{i\in[1,N]} \frac{\Delta}{h_i^*}\}$ based on Lemmas 1 and 2 respectively (Lines 1-2). The joint optimization problem (2.29) with the value of β initialized to β_{up} (Line 3) is solved for the optimal delay values d_i^* for all $i \in [1, N]$ (Line 4, **solveOptim**(β) function returns the optimal delay values d_i^* corresponding to the optimal constellation size values b_i^* for all $i \in [1, N]$ obtained via solving the optimization problem (2.29) with a particular value of β). The existence of a feasible schedule using these optimal delay values and the optimal period values obtained by the procedure described in Section 2.4.1 is then checked by performing the exact schedulability analysis of the EDF scheduling algorithm (Line 5, **isSchedulable**($\{h_i, d_i, i \in [1, N]\}$) function returns true if there exists a feasible schedule corresponding to the given period and delay values and false otherwise. The exact schedulability analysis is based on the simulation of the EDF algorithm for the time duration τ formulated as

$$\tau = \frac{c}{1-c} \max_{i\in[1,N]} \{h_i - \Delta\} \tag{2.30}$$

where $c = \sum_{i=1}^{N} \frac{d_i}{h_i}$ [51]. However, since the scheduling of the nodes repeats every time duration of length Ω given that the optimal period $h_i^* = \frac{\Omega}{k_i^*}$ for some positive integer k_i^* for all $i \in [1, N]$ due to Lemmas 3 and 5, it is enough to perform the schedulability analysis over the time duration of $\min\{\tau, \Omega\}$. **isSchedulable**($\{h_i, d_i, i \in [1, N]\}$) simply generates the EDF schedule for the time duration equal to $\min\{\tau, \Omega\}$ and declares that there exists a feasible schedule if and only if no deadlines are missed in this schedule. The schedulability analysis has a pseudo-polynomial complexity of $O(N\sum_{i=1}^{N}\frac{\min\{\tau,\Omega\}}{h_i})$ [53]). If such a feasible schedule exists, $\beta_{opt} = \beta$ (Line 6). Otherwise, β is set to $\frac{\beta_{up}+\beta_{low}}{2}$ (Line 9). In each iteration of the algorithm, the optimization problem (2.29) with the current β value is solved to determine the optimal delay values for all $i \in [1, N]$ (Line 10). Then, the feasibility of constructing a schedule with the optimal delay and period values is checked (Line 11). If there exists a feasible schedule, β_{low} is set to β (Line 12); otherwise, β_{up} is set to β (Line 14). β is then updated with the value $\frac{\beta_{up}+\beta_{low}}{2}$ (Line 9). The algorithm stops when β_{low} and β_{up} get sufficiently close to each other such that $\beta_{up} - \beta_{low} < \epsilon$ (Line 8) where ϵ is a predetermined arbitrarily small constant. β_{opt} is then set to β_{low} (Line 17). Finally, the joint optimization problem (2.29) is solved for $\beta = \beta_{opt}$ (Line 19). The corresponding feasible schedule is constructed by the EDF scheduling algorithm over one scheduling frame (Line 20, **EDFSchedule(**$\{h_i, d_i, i \in [1, N]\}$**)** function constructs the EDF schedule corresponding to the given period and delay values over one scheduling frame. The function first generates the EDF schedule over the time duration of length Ω and then repeats the schedule n_{MATI} times since the length of the scheduling frame is equal to $n_{MATI}\Omega$).

EMSA algorithm requires at most $K = \lceil \log_2[(\beta_{nec} - \beta_{suf})/\epsilon] \rceil$ exact schedulability analyses performed by **isSchedulable(**$\{h_i, d_i, i \in [1, N]\}$**)** function since the value of $(\beta_{nec} - \beta_{suf})$ is decreased by half at each iteration until it becomes less than ϵ and one EDF schedule construction performed by **EDFSchedule(**$\{h_i, d_i, i \in [1, N]\}$**)**. For example, for an ϵ value of 0.001, the maximum number of exact schedulability analyses in EMSA is $K = 10$ since the maximum value of $(\beta_{nec} - \beta_{suf})$ is equal to 1.

2.5 Performance Evaluation

The goal of this section is to evaluate the energy saving of the proposed joint optimization problem over the traditional separate design of controller and communication systems. In the traditional separate design of these systems, which is denoted by "TS", the constellation size and sampling period of the sensor nodes are predetermined. For instance, WirelessHart [8] and ISA100.11a [7], which are the two competing wireless standards for industrial control applications, employ O-QPSK (Offset Quadrature Phase Shift Keying) without optimizing the constellation size nor the sampling period. In "TS", the values of the fixed constellation size and sampling period of the sensor nodes are determined such that the existence of a solution is guaranteed for the worst case scenario with no adjustment to different network and channel conditions. For instance, when we analyze the variability of the energy consumption as a function of MAD, we choose one of the constellation size values that is feasible for all MAD values. The proposed heuristic algorithm for the joint optimization of these systems, which is denoted by "HS", follows the procedure described in Section 2.4.1 to determine the optimal parameters including $b_i^{\min}$, $k_i^*(b_i) = k_i(b_i^{\min})$, $b_i^{\max}$, h_i^* and q_i^* for each $i \in [1, N]$ separately and the optimal constellation size b_i^* for all $i \in [1, N]$ by solving the relaxation of the IP problem (2.29) with the optimal utilization bound β_{opt} and ceiling the non-integral solution values to obtain a feasible integral solution. To understand the maximum deviation of the HS algorithm from the optimal formulation, we have also included the optimal solution, denoted by "OPT". This optimal solution is obtained by exhaustive search where every feasible solution of the IP problem (2.29) with $\beta = 1$ that yields a feasible EDF schedule is enumerated and the one with the minimum power consumption is determined. This exhaustive search has exponential computational complexity in the network size.

σ^2	-174 dBm/Hz	B	10 KHz
$W^{t,\max}$	250 mW	W^c	50mW
$L_i, i \in [1, N]$	100 bits	δ	0.95
N^f	10 dB	G^c	1 (uncoded) [36]

Table 2.1: Simulation Parameters

Simulation results are obtained based on 1000 independent random network topologies where the sensor nodes are uniformly distributed within a circular area of radius r transmitting to a controller located in the center of the area. The parameters used in the simulations are given in Table-2.1.

The attenuation of the links are determined considering both large scale statistics that arise primarily from the free space loss and the environment affecting the degree of refraction, diffraction, reflection and absorption, and small scale statistics that occur due to multipath propagation and variations in the environment. The dependence of the path loss on distance summarizing large scale statistics is modeled as

$$PL(d) = PL(d_0) + 10\alpha \log(d/d_0) + Z \tag{2.31}$$

where d is the distance between the transmitter and receiver, $PL(d)$ is the path loss at distance d in decibels, $PL(d_0) = 70$ dB is the path loss at reference distance $d_0 = 1$ m, $\alpha = 3.5$ is the path loss exponent [36] and Z is a Gaussian random variable with zero mean and standard deviation equal to 4 dB [56]. The small-scale fading on the other hand has been modeled by using Nakagami fading with scale parameter Ω set to the mean power level determined by using Eq. (2.31) and shape parameter m chosen from the set $\{1, 3, 5\}$ [56, 57]. Note that Nakagami fading with shape parameter equal to 1 corresponds to Rayleigh fading. In the first part of the simulations including Figs. 2.3- 2.7, the channel is assumed fixed at the mean value determined based on the large scale statistics. In the second part of the simulations including Figs. 2.8- 2.10, the robustness of the proposed algorithm to the time-varying channel conditions is analyzed by considering small-scale statistics.

Fig. 2.3 shows the average power consumption in a network of 20 nodes at different average distances from the controller. The average distance is calculated by taking the average of the distances of the nodes randomly distributed within a circle of different radii. The MAD and MATI values are chosen as $\Delta = 5$ ms and $\Omega = 100$ ms. The effect of the average distance on the average power consumption is twofold. First, as the distance increases, the transmit power required to compensate for the increasing attenuation increases. Since the power consumption is proportional to the transmit power, increasing the transmit power also increases the power consumption. Second, as the distance increases, the maximum feasible constellation size imposed by the maximum transmit power constraint decreases shrinking the feasible

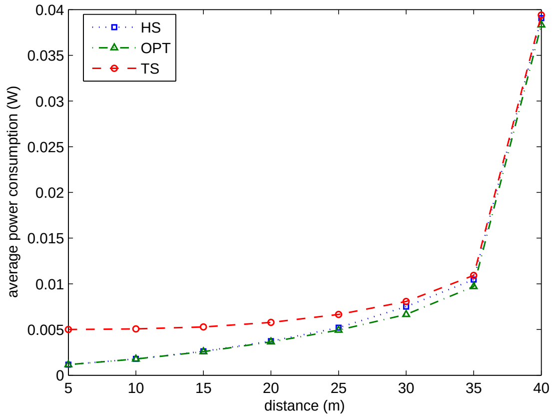

Figure 2.3: Average power consumption in a network of 20 nodes at different average distances from the controller where $\Delta = 5$ ms and $\Omega = 100$ ms.

region over which the power consumption is minimized. This accelerates the increase in the power consumption by increasing distance. The constellation size for the TS algorithm is determined such that it is feasible for the maximum distance value since feasibility for maximum distance guarantees feasibility for lower distance values. Hence, the HS algorithm outperforms the TS algorithm significantly for relatively small average distance values. Moreover, the HS algorithm performs very close to the OPT independent of the distance value, by an approximation ratio of around 1.01 where approximation ratio is defined as the ratio of the solution of the HS algorithm to the optimal solution.

Fig. 2.4 shows the average power consumption in a network of 20 nodes for different MAD values. The nodes are uniformly distributed within a circular area of radius 10 m and the MATI value is chosen as $\Omega = 200$ ms. The MAD constraint determines the minimum constellation size together with $\Omega/k_i(b_i^{\min})$ as explained in detail in Section 2.4.1. As the MAD increases up to a certain value, around 2 ms, the average power consumption decreases since for smaller MAD values, the nodes in the network are forced to choose greater constellation size values, which increases the power consump-

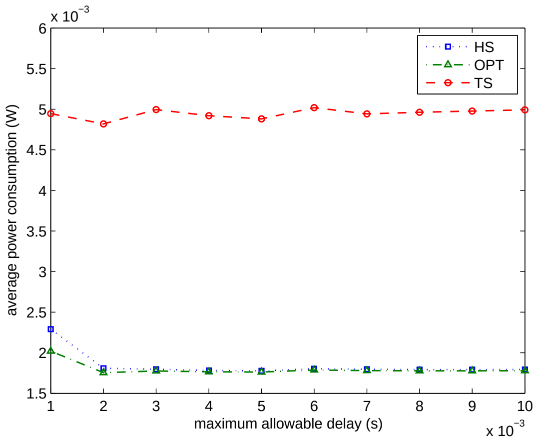

Figure 2.4: Average power consumption in a network of 20 nodes for different MAD values where nodes are uniformly distributed within a circular area of radius 10 m and $\Omega = 200$ ms.

tion dramatically. On the other hand, as the MAD increases further, the average power consumption stays constant due to the fact that the optimal constellation size remains constant although the feasible region expands. The constellation size for the TS algorithm is determined such that it is feasible for the minimum MAD value since then feasibility of all MAD values is ensured. Since the power consumption function does not depend on the MAD value, the resulting power consumption of the TS algorithm remains constant for different MAD values and is dramatically worse than the HS algorithm which performs very close to the OPT, by an average approximation ratio of around 1.02.

Fig. 2.5 shows the average power consumption in a network of 20 nodes for different MATI values. The nodes are uniformly distributed within a circular area of radius 10 m and the MAD value is chosen as $\Delta = 10$ ms. Since the power consumption is a decreasing function of MATI, the average power consumption decreases as the MATI increases. However, the effect of the MATI on the average power consumption in a network is not limited to this functional dependency for the HS algorithm. The schedulability constraint given by Eq. (2.29b) suggests that for small MATI values, the objective of power

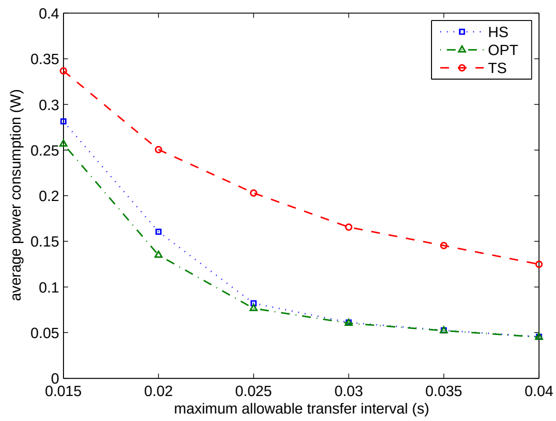

Figure 2.5: Average power consumption in a network of 20 nodes for different MATI values where nodes are uniformly distributed within a circular area of radius 10 m and $\Delta = 10$ ms.

minimization requires a joint decision among multiple nodes in the network since minimizing the power consumption independently for each node in the network may result in the violation of the schedulability constraint. Hence, as the MATI decreases, the amount of increase in the power consumption is much more than the functional dependency of the power consumption on the MATI. For example, when we increase the MATI from 15 ms to 20 ms, the power consumption is expected to decrease by 33% but actually decreases by 50%. On the other hand, for the TS algorithm, the effect of the MATI on the power consumption is limited to the functional dependency of the power consumption on the MATI. Again, the HS algorithm performs very close to the OPT, by an approximation ratio of around 1.05, and outperforms the TS algorithm which has an approximation ratio of more than 2 for large MATI values. The performance of the TS algorithm is relatively better for smaller MATI values since the predetermined constellation size value is adjusted considering the feasibility for the minimum MATI value but still much worse than the performance of the HS algorithm.

Fig. 2.6 shows the average power consumption for different number of nodes. The nodes are uniformly distributed within a circular area of radius

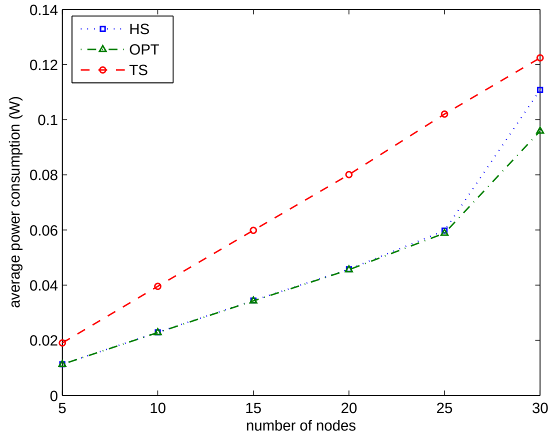

Figure 2.6: Average power consumption for different number of nodes where nodes are uniformly distributed within a circular area of radius 5 m, $\Delta = 25$ ms and $\Omega = 25$ ms.

5 m. The MAD and MATI values are chosen as $\Delta = 25$ ms and $\Omega = 25$ ms. For the HS algorithm, as the number of nodes increases up to a specific value, i.e. 25 in this case, depending on the MATI, the power consumption increases linearly since the objective of the power minimization does not require a joint decision among multiple nodes in the network meaning that minimizing the power consumption independently for each node does not result in the violation of the schedulability constraint. However, as the number of nodes increases further, a joint decision is necessary to satisfy the schedulability constraint. The nodes in the network are forced to choose higher constellation size than they would choose when they minimize their power consumption independently. This causes a much faster increase in the power consumption. The average approximation ratio of the HS algorithm is 1.03. The resulting power consumption of the TS algorithm on the other hand increases linearly as expected since the predetermined constellation size is constant. The performance of the TS algorithm is much worse than that of the HS algorithm with an average approximation ratio of 1.65.

Fig. 2.7 shows the average power consumption in a network of 20 nodes for different path loss exponent values. The nodes are uniformly distributed within a circular area of radius 5 m. The MAD and MATI values are chosen as $\Delta = 10$ ms and $\Omega = 100$ ms. The constellation size of the TS algorithm is determined such that it is feasible for the highest path loss exponent and highest distance from the controller since then feasibility for all path loss exponent values and all the nodes in the network is guaranteed. For fixed values of the constellation size and remaining communication parameters used in the TS algorithm, the power consumption increases exponentially with the path loss exponent. The reason for this exponential increase is that the transmit power required to compensate for the increasing attenuation in an environment of larger path loss exponent increases exponentially based on the path loss model given in Eq. (2.31) and the power consumption is proportional to the transmit power. The average power consumption of the HS algorithm on the other hand increases linearly with the path loss exponent since the constellation size is optimized at every node in the network for all the path loss exponent values. The average approximation ratio of the HS algorithm is 1.01. The difference between TS and HS algorithms increases as the path loss exponent increases since higher path loss exponent creates more variation of the optimal constellation size of the nodes in the network increasing the value of optimizing the communication parameters of the sensor nodes in the network.

Figs. 2.8, 2.9 and 2.10 show the robustness of the proposed algorithm over time by considering the time-varying channel condition modeled using Rayleigh fading, Nakagami fading with shape parameter $m = 3$ and $m = 5$ respectively. We assume that the fading level is available to the transmitter through a causal Channel State Information (CSI) feedback at the beginning of each transmission interval. As the distance of the node from the controller and fading level increases, the maximum feasible constellation size imposed by the maximum transmit power constraint decreases shrinking the feasible region over which the power consumption is minimized. Therefore, the constellation size of the TS algorithm is chosen as the minimum value in the feasibility region. Similar to the analysis of the effect of the path loss exponent on the power consumption, the power consumption of the TS algorithm employing fixed constellation over all fading levels increases signi-

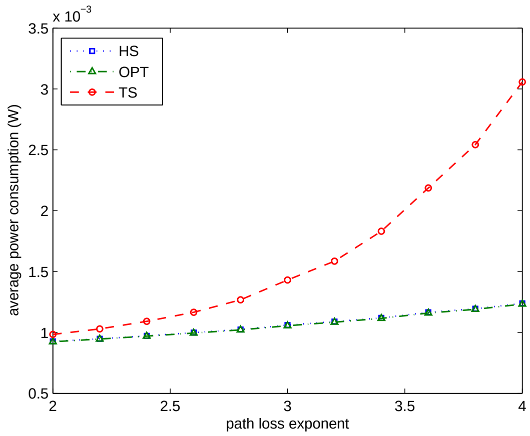

Figure 2.7: Average power consumption for different path loss exponent values in a network of 20 nodes where nodes are uniformly distributed within a circular area of radius 5 m, $\Delta = 10$ ms and $\Omega = 100$ ms.

ficantly when the fading level is high. The reason for this increase is that the transmit power required to compensate for the increasing attenuation increases and the power consumption is proportional to the transmit power. This also causes large fluctuations in the power consumption of the TS algorithm over time. The amount of these fluctuations on the other hand depends on the variance of the fading distribution. The fluctuations increases when the fading distribution changes from Nakagami fading with $m = 5$ to Nakagami fading with $m = 3$ and from Nakagami fading with $m = 3$ to Rayleigh fading. Furthermore, the average power consumption of the HS algorithm is stable over time and very close to the optimal with average approximation ratio less than 1.01 for all three fading distributions since the values of the parameters are optimized according to the channel information at the beginning of each interval.

2.6 Conclusion and Discussion

A joint design of communication and control application layers is studied by considering a constrained optimization problem, for which the objective

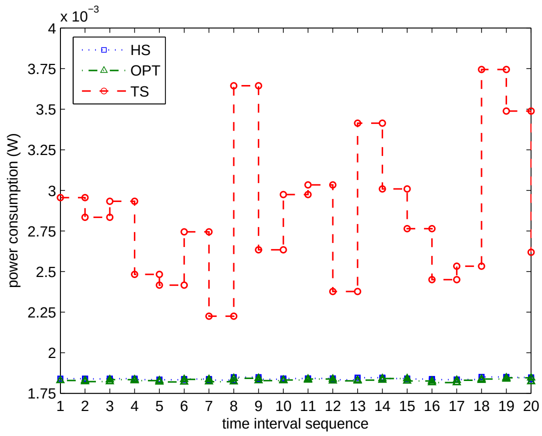

Figure 2.8: Power consumption in a network of 20 nodes under Rayleigh fading where nodes are uniformly distributed within a circular area of radius 5 m, $\Delta = 10$ ms and $\Omega = 100$ ms.

function is the power consumption of the network and the constraints are the reliability and delay requirements for the stability of the control systems and the schedulability and maximum transmit power constraints of the communication systems. The decision variables of this optimization problem are the sampling period of the control layer, the constellation size of the modulation and the probability of error of the communication layer. The optimization problem is first formulated as a Mixed-Integer Programming problem which is NP-Hard. Upon deriving the relation of the optimal sampling period and packet error probability to the optimal constellation size, the problem is reduced to an IP problem. We propose an efficient solution method based on the relaxation of this IP problem. Extensive simulation results show that the performance of the proposed solution procedure is very close to the optimal solution and much better than the traditional separate design of controller and communication systems for varying network and control system parameters and under different fading channel models.

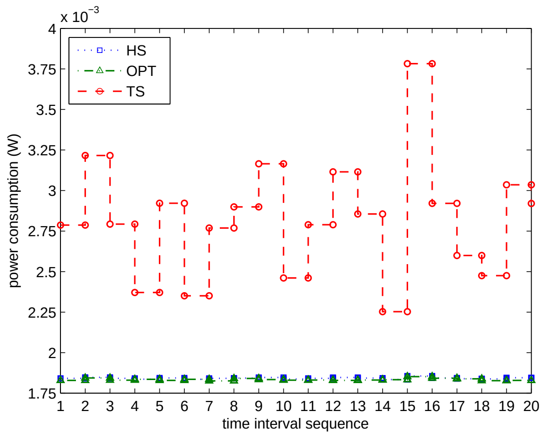

Figure 2.9: Power consumption in a network of 20 nodes under Nakagami fading with $m = 3$ where nodes are uniformly distributed within a circular area of radius 5 m, $\Delta = 10$ ms and $\Omega = 100$ ms.

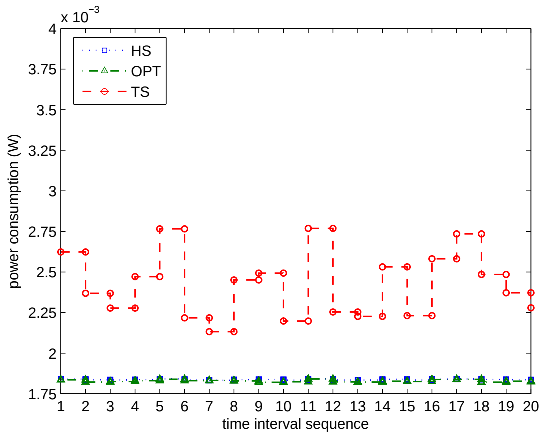

Figure 2.10: Power consumption in a network of 20 nodes under Nakagami fading with $m = 5$ where nodes are uniformly distributed within a circular area of radius 5 m, $\Delta = 10$ ms and $\Omega = 100$ ms.

Chapter 3

Joint Optimization of Wireless Network Energy Consumption and Control System Performance in WNCSs

In the previous chapter, a novel framework for the joint optimization of control and communication systems encompassing efficient abstractions of both systems is proposed with the objective of minimizing the total power consumption in the network while adopting M-ary Quadrature Amplitude Modulation (MQAM) modulation and Earliest Deadline First (EDF) scheduling. In this study, we generalize this optimization framework with any non-decreasing function of the power consumption of the nodes in the objective, any modulation with monotonically decreasing power consumption in the transmit power and any scheduling algorithm. Upon the formulation of the joint optimization as a non-convex Mixed-Integer Programming problem, we first introduce an exact solution method based on the analysis of the optimality conditions and smart enumeration techniques. Then, we propose two polynomial-time heuristic algorithms employing a utilization based search space reduction technique. Extensive simulations demonstrate that the proposed algorithms outperform the existing solution methods for various network sizes, modulation schemes, objective functions, and control system parameters.

3.1 Introduction

The joint optimization of controller and communication systems considering all the wireless network induced imperfections including packet error and delay, and all the parameters of both the wireless communication system and the control system has been recently studied in [12] with the objective of minimizing the total power consumption of the communication system while guaranteeing the performance and stability of the control system and the schedulability of the communication system for MQAM modulation and EDF scheduling. The goal of this study is to extend this study by generalizing the objective function, the modulation scheme and the scheduling algorithm adopted by the WNCS. The original contributions are listed as follows:

- We provide a generalized framework for the joint optimization of controller and communication systems with any non-decreasing function of the power consumption of the nodes in the objective, any modulation with monotonically decreasing power consumption in the transmit power and any scheduling algorithm. This framework is expected to lead to broader adoption in many real-world control applications.

- We propose an optimal solution method for the formulated generalized optimization problem based on the optimality analysis on the decision variables and smart enumeration techniques.

- We propose two polynomial-time heuristic algorithms based on a utilization based search space reduction technique which decreases the complexity of the algorithms compared to the optimal solution significantly while keeping the performance very close to optimal.

- We illustrate the superiority of the proposed generalized optimization framework to previously proposed solution methods and the closeness to the optimality via extensive simulations.

The rest of the chapter is organized as follows. Section 3.2 describes the system model and assumptions used throughout the chapter. The generalized joint optimization of controller and communication systems has been formulated and reduced to Integer Programming (IP) problem in Section 3.3. Section 3.4 presents an optimal smart enumeration based algorithm. Section 3.5 provides polynomial-time heuristic solution methods employing a utilization based search space reduction algorithm. Simulation results are presented in Section 3.6. Finally, concluding remarks are given in Section 3.7.

3.2 System Model and Assumptions

The system model and assumptions are detailed as follows.

1. The WNCS consists of multiple plants; multiple controllers; multiple sensors attached to the plants communicating to the controllers over a wireless channel; and multiple actuators that are mostly collocated with the controllers. Each sensor samples and transmits its sensed data to the corresponding controller periodically. The sampling period, transmission delay and packet error probability of node i are denoted by h_i, d_i, and p_i, respectively, as depicted in Fig. 2.2. We assume that $d_i \leq h_i$ since the packets are outdated and replaced with the new sampled data beyond d_i delay. Moreover, the packet error model is assumed to be a Bernoulli random process with probability p_i for node i. Upon reception of sensor data, the controller computes a control command and forwards it to the actuator. We assume that actuators successfully receive commands since they are collocated with the controllers, as in heat, ventilation and air conditioning control systems, due to their high criticality [37]. One controller is assigned as the coordinator.

2. Time Division Multiple Access (TDMA) is considered as medium access control (MAC) protocol. Explicit scheduling of the node transmissions in TDMA allows meeting the strict delay and reliability requirement of control systems while minimizing the energy consumption of the

sensor nodes by putting their radio in sleep mode when they are not scheduled to transmit or receive any packet. Moreover, mostly predetermined topology and data generation patterns of the sensor nodes in a WNCS decreases the synchronization and topology learning overhead associated with TDMA. TDMA is commonly used in industrial control applications [7, 8].

3. The time is divided into scheduling frames of fixed length, each of which is further partitioned into a beacon and variable number of variable-length time slots. Coordinator controller transmits the beacon periodically to maintain synchronization among the elements of the WNCS. Besides, in case of any change in the scheduling decisions upon variations of the channel conditions or network topology, the beacon additionally includes the updated schedule, and the transmission power, rate and sampling period of sensor nodes. The coordinator controller continually monitors the network topology and channel conditions by keeping track of the received power and packet error rate in the wireless links.

4. We assume a multi-modal operation for the sensor nodes such that they operate in sleep mode if they are not scheduled to transmit or receive a packet, in active mode if they are scheduled to transmit or receive a packet, and in transient mode while switching from active mode to sleep mode and vice versa. However, we consider only the power consumption in the transmission of the packets in the optimization problem since it is much larger than those in the sleep and transient modes, and the energy consumed in the reception of beacon packets is fixed [29], [36].

5. The performance and stability conditions for the WNCS have been formulated in the form of Stochastic Maximum Allowable Transfer Interval (MATI), defined as keeping the time interval between subsequent state vector reports from the sensor nodes to the controller above MATI value with a predefined probability, and Maximum Allowable Delay (MAD), defined as the maximum allowed packet delay for the trans-

mission from the sensor node to the controller. Stochastic MATI and MAD constraints are efficient abstractions of the performance of the control systems however have been considered only recently in the joint design with the communication systems [12].

(a) Stochastic MATI constraint is formulated as

$$\Pr\left[\mu_i(h_i, d_i, p_i) \leq \Omega\right] \geq \delta \tag{3.1}$$

where μ_i is the time interval between subsequent state vector reports of node i as a function of h_i, p_i and d_i; Ω is the MATI; and δ is the minimum probability with which MATI should be achieved. The number of reception opportunities of the state vector reports is equal to $\left\lfloor \frac{\Omega}{h_i} \right\rfloor$ within each Ω. Based on the assumption above on the modeling of packet error as a Bernoulli random process with probability p_i for node i, Eq. (3.1) can be rewritten as

$$1 - p_i^{\left\lfloor \frac{\Omega}{h_i} \right\rfloor} \geq \delta\,. \tag{3.2}$$

The values of Ω and δ are determined by the control system. For instance, in industrial automation, closed-loop machine controls have a stochastic MATI requirement that can be as low as 100 ms with user availability expectations at 99.9% [7, 44]. The air transportation system requires $\Omega = 4.8$ s and $\delta = 0.95$ [5]. In addition, the cooperative vehicular safety applications requires $\Omega = 100$ ms and $\delta = 0.95$ [43].

(b) MAD constraint is formulated as

$$d_i \leq \Delta \tag{3.3}$$

where Δ is the MAD value to stabilize the control system. Typical Δ values are on the order of of a few tens of milliseconds for fast control applications [7, 44, 47].

6. The average power consumption of sensor node i is formulated as a function of its sampling period, transmission delay and packet error probability as

$$W_i\left(h_i, d_i(b_i), p_i\right) = \frac{\left(W_i^t\left(b_i, p_i\right) + W_i^c\right) d_i(b_i)}{h_i} \tag{3.4}$$

where b_i is the number of bits used per symbol or the constellation size and d_i is represented as a function of b_i for a predetermined modulation scheme, W_i^t is the transmission power calculated as a function of the parameters b_i and p_i for a given modulation and channel coding, and W_i^c is the circuit power consumption in the active mode at the transmitter. In the following, we will use the notation of $W_i\left(h_i, b_i, p_i\right)$ instead of $W_i\left(h_i, d_i(b_i), p_i\right)$ for convenience. We assume that the average power consumption $W_i\left(h_i, b_i, p_i\right)$ and transmit power $W_i^t(b_i, p_i)$ satisfy the following properties:

   (a) *$W_i\left(h_i, b_i, p_i\right)$ is a monotonically decreasing function of h_i.*

   (b) *$W_i^t(b_i, p_i)$ is a monotonically decreasing function of p_i.*

   Property (a) follows from Eq. (3.4) and holds for any modulation scheme. Property (b) implies that a lower power consumption can be achieved at the expense of a higher packet error probability, keeping the remaining parameters fixed. It can be verified that these properties are satisfied in many modulation schemes including QAM and FSK (Frequency Shift Keying) [36].

7. We assume that the transmit power of a sensor node cannot exceed a maximum allowed power level $W^{t,\max}$ due to the limited weight and size of the sensor nodes. The maximum transmit power constraint is formulated as

$$W_i^t(b_i, p_i) \leq W^{t,\max} \tag{3.5}$$

8. The schedulability constraint represents the feasibility of the allocation of the time slots corresponding to given constellation size and sampling period of the nodes in the network, while concurrent transmissions

of the sensor nodes are not allowed, for a pre-determined scheduling algorithm. The schedulability constraint is formulated as

$$\{(d_1(b_1), h_1), ..., (d_N(b_N), h_N)\} \in S^{feasible} \tag{3.6}$$

where $S^{feasible}$ denotes the set of $\{(d_1(b_1), h_1), ..., (d_N(b_N), h_N)\}$ values with which a feasible schedule can be constructed. Any scheduling algorithm including EDF, Least Laxity First, Rate Monotonic and Deadline Monotonic scheduling algorithms [58] can be adopted in this framework. For instance, the schedulability constraint has been formulated as $\sum_{i=1}^{N} \frac{d_i}{h_i} \leq \beta$, where β is the utilization bound in the range $(0, 1]$, for pre-emptive EDF scheduling algorithm in [12].

9. We assume that the objective function $f(W_1(h_1, b_1, p_1), ..., W_N(h_N, b_N, p_N))$ is a non-decreasing function of $W_i(h_i, b_i, p_i)$ for all $i \in [1, N]$, where N is the number of nodes in the network. This assumption holds for many commonly used objective functions. Some examples can be listed as follows:

$$f(W_1(h_1, b_1, p_1), ..., W_N(h_N, b_N, p_N)) = \sum_{i\in[1,N]} W_i(h_i, b_i, p_i) \tag{3.7a}$$

$$f(W_1(h_1, b_1, p_1), ..., W_N(h_N, b_N, p_N)) = \max_{i\in[1,N]} W_i(h_i, b_i, p_i) \tag{3.7b}$$

$$f(W_1(h_1, b_1, p_1), ..., W_N(h_N, b_N, p_N)) = \log \prod_{i\in[1,N]} W_i(h_i, b_i, p_i) \tag{3.7c}$$

3.3 Joint Optimization of Control and Communication Systems

The joint optimization of control and communication systems aims to minimize a generalized non-decreasing function of the power consumption of the nodes while satisfying the stochastic MATI and MAD constraints guaranteeing the stability of the control systems and maximum transmit power and schedulability constraints of the wireless communication system. The optimization problem is formulated as

$$\begin{aligned}
\min_{h_i, b_i, p_i, i\in[1,N]} \quad & f(W_1(h_1, b_1, p_1), ..., W_N(h_N, b_N, p_N)) && \text{(3.8a)}\\
\text{s.t.} \quad & \left\lfloor \frac{\Omega}{h_i} \right\rfloor \ln p_i - \ln(1-\delta) \le 0, \quad i \in [1, N], && \text{(3.8b)}\\
& 0 < d_i(b_i) \le \min\{\Delta, h_i\}, \quad i \in [1, N], && \text{(3.8c)}\\
& 0 < h_i \le \Omega, \quad i \in [1, N], && \text{(3.8d)}\\
& 0 < p_i < 1, \quad i \in [1, N], && \text{(3.8e)}\\
& W_i^t(b_i, p_i) \le W^{t,\max}, \quad i \in [1, N], && \text{(3.8f)}\\
& \{(d_1(b_1), h_1), ..., (d_N(b_N), h_N)\} \in S^{feasible}. && \text{(3.8g)}
\end{aligned}$$

Eq. (3.8a) represents the generalized objective function of the power consumption of the nodes. Eqs. (3.8b) and (3.8c) represent the stochastic MATI and MAD constraints, respectively. Eq. (3.8d) states that the sampling period of the nodes must be less than or equal to the MATI. Eq. (3.8e) states the lower and upper bounds of the packet error probability. Eq. (3.8f) provides the maximum transmit power constraint. Finally, Eq. (3.8g) represents the schedulability constraint. The variables of the problem are $h_i, i \in [1, N]$, the sampling period of the nodes; $b_i, i \in [1, N]$, the constellation size of the nodes; and $p_i, i \in [1, N]$, the packet error probability of the nodes.

This optimization problem is a non-convex Mixed-Integer Programming problem thus difficult to solve for the global optimum [54]. We now analyze the optimality conditions, which allow us to reduce the problem to an IP problem.

Lemma 8. *The optimal sampling period and packet error probability, denoted by h_i^* and p_i^* respectively, satisfy*

$$\frac{\Omega}{h_i^*} = \frac{\ln(1-\delta)}{\ln p_i^*} = k_i, \tag{3.9}$$

where k_i is a positive integer, for all $i \in [1, N]$

Proof: This result has been proven for the objective of minimizing total power consumption in the network, MQAM modulation and EDF scheduling in [12]. Similarly, we first prove $\frac{\Omega}{h_i^*} = k_i$ by contradiction. Suppose that $\frac{\Omega}{h_i^*}$ is not a positive integer then $\left\lfloor \frac{\Omega}{h_i^*} \right\rfloor < \frac{\Omega}{h_i^*}$. If h_i^* increases such that the equality $\left\lfloor \frac{\Omega}{h_i^*} \right\rfloor = \frac{\Omega}{h_i^*}$ holds for the first time while satisfying the upper bound given in Eq. (3.8d), the stochastic MATI constraint given in Eq. (3.8b) still holds since the value of $\left\lfloor \frac{\Omega}{h_i^*} \right\rfloor$ does not change. The remaining constraints including h_i given in Eqs. (3.8c) and (3.8g) also still hold with this change. However, the objective cost function given in Eq. (3.8a) does not increase since it is a non-increasing function of h_i for each node $i \in [1, N]$.

We also prove $\frac{\Omega}{h_i^*} = \frac{\ln(1-\delta)}{\ln p_i^*}$ by contradiction. Suppose that $\frac{\Omega}{h_{i^*}} > \frac{\ln(1-\delta)}{\ln p_i^*}$. If p_i^* increases such that the stochastic MATI constraint is satisfied with equality, the constraint given in Eq. (3.8f) still holds since the power consumption of node i is assumed to be a monotonically decreasing function of p_i. However, the objective cost function given in Eq. (3.8a) does not increase since it is a non-increasing function of p_i for each node $i \in [1, N]$. □

Since h_i and p_i can be represented as functions of a single variable k_i in the optimal solution based on Lemma 8, we eliminate them from the optimization problem (3.8), which is then reformulated as

$$\min_{b_i, k_i, i \in [1,N]} \quad f(W_1(b_1, k_1), ..., W_N(b_N, k_N)) \tag{3.10a}$$

$$\text{s.t.} \quad 0 < d_i(b_i) \le \min\left\{\Delta, \frac{\Omega}{k_i}\right\}, \quad i \in [1, N], \tag{3.10b}$$

$$W_i^t(b_i, k_i) \le W^{t,\max}, \quad i \in [1, N], \tag{3.10c}$$

$$\left\{(d_1(b_1), \frac{\Omega}{k_1}), ..., (d_N(b_N), \frac{\Omega}{k_N})\right\} \in S^{feasible}, \tag{3.10d}$$

where $W_i(b_i, k_i)$ and $W_i^t(b_i, k_i)$ are obtained by replacing h_i and p_i by their expression of their optimal values as a function of k_i in $W_i(h_i, b_i, p_i)$ and $W_i^t(b_i, p_i)$, respectively, based on Lemma 8. The constraints given in Eqs. (3.10b), (3.10c) and (3.10d) correspond to those in Eqs. (3.8c), (3.8f) and (3.8g), respectively, and the remaining constraints in the optimization problem (3.8) are removed by exploiting the additional constraint of k_i being a positive integer. The following lemma expresses the relation between the optimal values of k_i and b_i so that the above optimization problem can be formulated with the variable b_i only.

Lemma 9. *The optimal value of k_i is the minimum positive integer satisfying Eq. (3.10c) and can be expressed as a function of b_i.*

Proof: Since the objective function is a non-increasing function of h_i and p_i, it is a non-decreasing function of k_i due to Lemma 8. Therefore, minimizing the objective function requires finding the minimum positive integer satisfying the constraints of the optimization problem given in Eqs. (3.10b), (3.10c) and (3.10d). Decreasing k_i does not shrink the feasibility regions for b_i determined by the constraints (3.10b) and (3.10d). Hence k_i^* is the minimum positive integer satisfying Eq. (3.10c) and can therefore be represented as a function of b_i given the transmit power function $W_i^t(b_i, k_i)$. $\square$

Next, we can determine the minimum and maximum values of b_i for each sensor node i, denoted by $b_i^{\min}$ and $b_i^{\max}$, respectively, by evaluating the constraints given in Eqs. (3.10b) and (3.10c). Then, based on Lemma 9, the optimization problem can be further simplified as

$$\min_{b_i, i\in[1,N]} \quad f(W_1(b_1, k_1^*(b_1)), ..., W_N(b_N, k_N^*(b_N))) \tag{3.11a}$$

$$\text{s.t.} \quad b_i^{\min} \le b_i \le b_i^{\max}, \quad i \in [1, N], \tag{3.11b}$$

$$\left\{(d_1(b_1), \frac{\Omega}{k_1^*(b_1)}), ..., (d_N(b_N), \frac{\Omega}{k_N^*(b_N)})\right\} \in S^{feasible}. \tag{3.11c}$$

Since the constellation size b_i is integer for all $i \in [1, N]$, the optimization problem is an IP. The relaxation of the problem is non-convex in general, hence Branch and Bound techniques that are efficient to solve IPs are not applicable. Next, we propose an optimal algorithm to solve the problem in reasonable runtime using smart enumeration techniques and heuristic algorithms to achieve close-to-optimal solutions in polynomial runtime.

3.4 Optimal Algorithm

The IP problem formulated in the previous section can be solved by an exhaustive search algorithm that simply calculates the objective value for each solution in the feasible region and determines the one minimizing the objective function. Such a search algorithm however is intractable for even medium network sizes. For example, for the number of nodes and the number of possible constellation size values for each sensor node i given by $N = 10$ and $K_i = b_i^{\max} - b_i^{\min} + 1 = 10$, respectively, 10^{10} possible constellation size vectors need to be checked for feasibility based on Eq. (3.11c) and value of objective function given in Eq. (3.11a). Next, we present the Optimal Fast Enumeration (OFE) Algorithm which employs smart enumeration techniques based on the main characteristics listed as

1. ordering the set of constellation sizes in increasing power consumption,
2. starting the evaluation of the feasibility and objective function with the constellation size vector corresponding to the minimum power consumption of each node,
3. pruning the feasible constellation size vectors,
4. regenerating the constellation size vectors for evaluation without repetitions covering all vectors in the case of the existence of no feasible vectors by associating each vector with a number denoting the number of vectors it is branched into.

OFE Algorithm given by Algorithm 2 is described in detail as follows. The input of the algorithm are K_i possible constellation size values for each sensor node $i \in [1, N]$ resulting from Eq. (3.11b) in the simplified IP problem given in previous section. Suppose that for each sensor node $i \in [1, N]$, K_i constellation size values are sorted in increasing power consumption $W_i(b_i)$ and b_{ij} denotes the constellation size corresponding to the j^{th} minimum power consumption for node i. Let $\mathbf{b}$ be the constellation size vector including the constellation size of node j as the j-th element. Let $deg(\mathbf{b})$ denote the degree of $\mathbf{b}$, which is defined as the number of vectors that vector $\mathbf{b}$ is branched into.

Algorithm 2 **Optimum Fast Enumeration (OFE) Algorithm**

Input: $b_{ij},\ i \in [1, N],\ j \in [1, K_i]$;
Output: $\mathbf{b}^*, f^*$;

```
f* = ∞;
b = (b_11, b_21, ..., b_N1);
deg(b) = N;
B = {b};
while B ≠ ∅ do
    B+ = ∅;
    for each b ∈ B do
        if f(b) ≤ f* then
            if isSchedulable(b) then
                b* = b;
                f* = f(b);
            else
                for j = 1 : deg(b) do
                    b+ = b;
                    set constellation size b+(N − j + 1) to next value;
                    deg(b+) = j;
                    B+ = B+ + {b+};
                end for
            end if
        end if
    end for
    B = B+;
end while
```

The assignment of the degree to each constellation size vector assures that the algorithm generates a particular vector $\mathbf{b}$ only once, and all possible vectors if there does not exist a schedulable vector, as proven in Lemma 10. Algorithm starts with the constellation size of minimum power consumption for each node, resulting in the vector $\mathbf{b} = (b_{11}, b_{21}, ..., b_{N1})$ (Line 2). Degree of root vector $(b_{11}, b_{21}, ..., b_{N1})$ is set to N (Line 3). Vector set $\mathbf{B}$ is defined as the set of constellation size vectors to be evaluated in the next iteration of the algorithm and initially contains the vector $(b_{11}, b_{21}, ..., b_{N1})$ (Line 4). f^* denotes the minimum value of objective function corresponding to a feasible constellation size vector and initialized to ∞ (Line 1). For each vector $\mathbf{b}$ in $\mathbf{B}$, the algorithm first determines whether it can improve the best solution so far with a smaller value of objective function (Lines $7-8$). If so, the algorithm checks the schedulability of vector $\mathbf{b}$. If $\mathbf{b}$ is also schedulable, the best constellation size vector $\mathbf{b}^*$ and best solution f^* are updated with $\mathbf{b}$ and the value of objective function corresponding to $\mathbf{b}$, respectively (Lines $9-11$). Note that schedulable vectors are not branched into new vectors since the objective value of the new vectors branched from a vector do not decrease based on the fact that the objective function is a non-decreasing function of the power consumption of the nodes. If $\mathbf{b}$ is not schedulable, the algorithm branches the vector $\mathbf{b}$ into $deg(\mathbf{b})$ vectors (Lines $12-19$). For every j value in $[1, deg(\mathbf{b})]$ interval, a vector $\mathbf{b}^+$ is generated by setting the constellation size of the $N-j+1$-th value in vector $\mathbf{b}$ to the next constellation size and the degree to j. Each newly generated vector $\mathbf{b}^+$ is included in the set $\mathbf{B}^+$, which will be equalized to the set $\mathbf{B}$ at the end for the evaluation in the next iteration of the algorithm (Lines 17 and 22). Algorithm terminates when all the vectors in $\mathbf{B}$ are schedulable or have a higher objective value than the best solution so far, generating $\mathbf{B} \neq \emptyset$ in the following iteration (Line 5). OFE algorithm is illustrated in Fig. 3.1.

Lemma 10. *OFE algorithm generates all possible constellation size vectors unless there is a schedulable vector and each vector is generated only once.*

Proof: OFE algorithm generates a descendant vector $\mathbf{b}^+$ of a particular vector $\mathbf{b}$ with $deg(\mathbf{b})$ by changing the constellation size $\mathbf{b}(N-j+1)$ to the

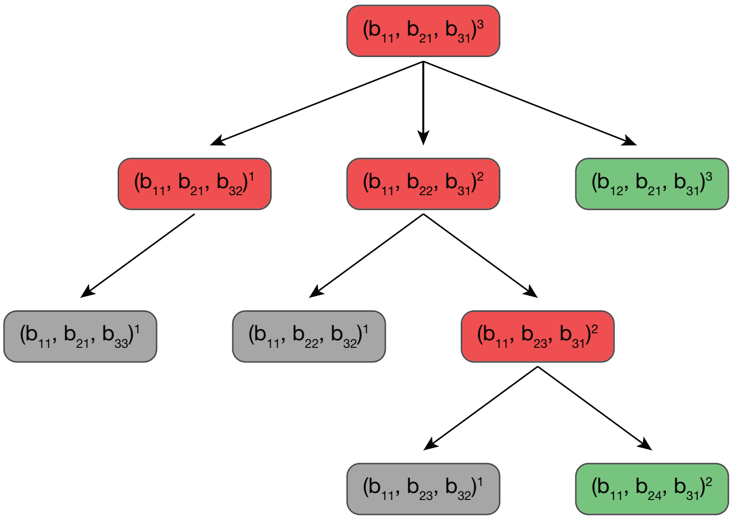

Figure 3.1: OFE algorithm illustration for the case where $N = 3$ and $C_i = 5$, for all $i \in [1, N]$. The constellation size vectors are evaluated in the following order: (b_{11}, b_{21}, b_{31}), (b_{11}, b_{21}, b_{32}), (b_{11}, b_{22}, b_{31}), (b_{12}, b_{21}, b_{31}), (b_{11}, b_{21}, b_{33}), (b_{11}, b_{22}, b_{32}), (b_{11}, b_{23}, b_{31}), (b_{11}, b_{23}, b_{32}), (b_{11}, b_{24}, b_{31}). The superscripts represent the degree of the constellation size vectors. Green-colored constellation size vectors are not branched into new vectors since they are evaluated as schedulable. Grey-colored constellation size vectors are the vectors that are not branched into new vectors since their objective is higher than the best solution so far. Red-colored constellation size vectors are branched into new vectors since they are not schedulable and their objective is less than or equal to the best solution obtained so far.

next value for a particular $j \in [1, deg(\mathbf{b})]$ (Lines 13-15). The degree of a descendant node $\mathbf{b}^+$ generated by changing $\mathbf{b}(N - j + 1)$ is set as j (Line 14). Therefore, on a particular path starting with root constellation size vector and continueing with subsequent descendant nodes, the degrees either decrease or stay constant when a new vector is generated. Hence, degree of a particular vector $\mathbf{b}$ except the root constellation size vector is equal to minimum j such that $\mathbf{b}(N - j + 1)$ is not equal to $b_{(N-j+1)1}$ and the unique ascendant of that vector from which it is generated can be determined by setting the constellation size $\mathbf{b}(N - j + 1)$ to the lower value reversing the generation rule of descendant nodes.

Now, consider any possible constellation size vector $\overline{\mathbf{b}}$ except the root constellation size vector and let $k \in [1, N]$ be the degree of vector $\overline{\mathbf{b}}$. Therefore, the elements of vector $\overline{\mathbf{b}}$ with indices $i \in [N-k+2, N]$ are equal to $b_{(i)1}$ and $\overline{\mathbf{b}}(N-k+1)$ is not equal to $b_{(N-k+1)1}$ from the previous argument. The ascendant of $\overline{\mathbf{b}}$ can be found by setting $\overline{\mathbf{b}}(N-k+1)$ to the lower constellation size value and if the lowered value is equal to $b_{(N-k+1)1}$, which will eventually happen on the path, the degree of the ascendant will be larger than $k+1$. Subsequent ascendants of the ascendant vector can be determined similarly. As long as any element $i \in [1, N-k+1]$ of a particular vector $\mathbf{b}$ on the path is not equal to $b_{(i)1}$, the ascendant of $\mathbf{b}$ can be determined by lowering $\mathbf{b}(N - deg(\mathbf{b}) + 1)$ which eventually leads the path to the root constellation size vector whose all elements $i \in [1, N]$ are equal to $b_{(i)1}$. Hence, since each vector except the root constellation size vector has a unique ascendant, a unique path from any possible vector $\overline{\mathbf{b}}$ to the root constellation size vector exists. This completes the proof.$\square$

The complexity of OFE algorithm is $O(\prod_{i\in[1,N]} K_i \times F)$ where F is the complexity of schedulability analysis for the specific scheduling algorithm used since in the worst case the algorithm evaluates all possible constellation size vectors and checks the schedulability of each vector. For EDF scheduling algorithm, the complexity required by the exact schedulability analysis is given by

$$F = N \sum_{i=1}^{N} \frac{\min\{\frac{c}{1-c} \max_{i\in[1,N]}\{h_i - \Delta\}, \Omega\}}{h_i} \tag{3.12}$$

where $c = \sum_{i=1}^{N} \frac{d_i(b_i)}{h_i}$. The details of the schedulability and complexity analysis of EDF scheduling algorithm can be found in [12]. A categoric analysis of the schedulability analysis of other scheduling algorithms can be found in [58].

3.5 Polynomial-time Heuristic Algorithms

In the previous section, we have proposed an optimal algorithm that efficiently spans the search space of the constellation size vectors till optimality of a constellation size vector is guaranteed. Although the OFE algorithm reaches the optimality without searching whole search space using the relations among the constellation size vectors, the worst case complexity of the algorithm is still exponential. For large network sizes and large number of feasible constellation sizes for the nodes in the network, faster solutions may be desired depending on the application at the expense of achieving sub-optimal results. This requires design of efficient heuristic algorithms. In the following, we first present a technique that reduce the problem search space significantly by first defining and exploiting utilization and energy consumption based dominance relations of the constellation size of each sensor node separately in Section 3.5.1. We then propose two polynomial time heuristic algorithms based on moving either in the direction of maximum improvement in the power consumption related objective function while keeping feasibility starting with the constellation size vector corresponding to the maximum power consumption or in the direction of minimum degradation in the objective function searching for the feasibility starting with the constellation size vector associated with the minimum power consumption in Sections 3.5.2 and 3.5.3, respectively.

3.5.1 Utilization based Search Space Reduction (USR)

Definition 1. *The* ***utilization*** $u_i(b_i)$ *of a sensor node* i *for a constellation size value* b_i *is defined to be the ratio of its transmission delay* $d_i(b_i)$ *to its sampling period* $h_i(b_i) = \frac{\Omega}{k_i^*(b_i)}$*; i.e.,* $u_i(b_i, h_i) = \frac{d_i(b_i)k_i^*(b_i)}{\Omega}$.

Utilization is a very important metric used commonly in determining the schedulability of real-time periodic tasks [59, 60]. In the real-time scheduling of periodic task sets, the utilization u_i of task i with execution time c_i and task period t_i is analytically expressed as $u_i = \frac{c_i}{t_i}$. The utilization of a task in a real-time system is the fraction of time used by that task in the scheduling frame. Note that task i refers to the transmission of sensor node i, and c_i and t_i correspond to transmission time d_i and sampling period h_i in our model. The total utilization U of a periodic task set is then defined as the sum of the utilizations of tasks; i.e., $U = \sum_i u_i$. This can again be interpreted as the fraction of time used by the entire periodic task set. Clearly, for $U > 1$, there exists no feasible schedule for the periodic task set with any scheduling algorithm. On the other hand, if $U \leq 1$, there may exist a feasible schedule depending on the parameters of the task set and the scheduling algorithm used in the system.

The total utilization of a periodic task set is used to determine the schedulability conditions for commonly used scheduling algorithms. For Rate Monotonic (RM) scheduling algorithm, which is the main fixed priority algorithm used in the real time scheduling of periodic task sets, where the tasks are assigned fixed priorities such that the task with smaller period receives higher priority, a sufficient schedulability condition for a set of n tasks is defined as

$$\sum_i u_i \leq n(2^{1/n} - 1) \tag{3.13}$$

under the assumption that relative deadlines of the tasks are equal to periods. For EDF scheduling algorithm, which is the main dynamic priority algorithm in the real time scheduling of periodic task sets, where priorities are assigned dynamically and are inversely proportional to the absolute deadlines of the tasks, a set of n tasks is schedulable if and only if

$$\sum_i u_i \leq 1 \tag{3.14}$$

under the assumption that relative deadlines of the tasks are equal to periods. Similar schedulability conditions employing only the utilizations of the tasks for periodic task systems can be found in [59]. The foregoing discussion shows the significance of utilization in schedulability analysis of the real time periodic systems. Moreover, all utilization based schedulability conditions indicate that a smaller utilization for a task is better in terms of the schedulability of a system. In fact, for EDF algorithm, decreasing the utilization of a task keeps the schedulability of the system if the original system is schedulable.

While utilization based schedulability conditions are mainly based on the assumption that the task deadlines are equal to task periods, exact schedulability analysis for periodic task systems without this assumption employ different task parameters in addition to utilizations of the tasks. Response Time Analysis proposed for RM algorithm and Processor Demand Criterion (PDC) technique proposed for EDF algorithm [59, 60] describe exact schedulability conditions employing execution time, deadline and periods of the tasks. However, following results on the sustainability of these scheduling analysis present a strong basis to employ utilization as a heuristic to evaluate the impact of a single task on the schedulability of the periodic task set, or for our system, the impact of the utilization of a sensor node on the schedulability of the entire network of nodes. A schedulability test for a scheduling algorithm is defined to be sustainable if any system deemed schedulable by that schedulability test remains schedulable when the parameters of one or more individual tasks are changed in any of the directions such as decreased execution requirements, larger periods, larger deadlines. Both RTA proposed for RM schedulability and PDC proposed for EDF schedulability are sustainable with respect to execution requirements, task periods and deadlines as presented in Theorems 1 and 10 of [60], respectively. The sustainability of these exact schedulability tests indicates that decreasing the transmission delay or increasing the sampling period of a sensor node keeps the schedulability of a system if the original system is schedulable. Considering the definition of utilization given by Definition 1 which is the ratio of transmission delay to sampling period, both decreasing transmission delay and increasing sampling period map to a decreasing utilization for a sensor node. Hence, although not necessarily, a decreasing utilization for a sensor node is expected to preserve the schedulability of a system along with the fact that a smaller utilization corresponds to a smaller fraction of time allocated by that sensor node in the entire schedule. Considering this and the utilization based schedulability conditions discussed previously leads us use the utilization to define a dominance relation among the constellation sizes of a sensor node and propose a search space reduction algorithm that will reduce the number of constellation sizes to be considered for a sensor node.

Definition 2. *A constellation size value $b_i^{'}$ for a sensor node i is said to be* ***dominating*** *another constellation size value $b_i^{''}$ if*

1. *the power consumption corresponding to* $b_i^{'}$ *is less than or equal to the power consumption corresponding to* $b_i^{''}$*; i.e.,* $W_i(b_i^{'}, k_i^*(b_i^{'})) \leq W_i(b_i^{''}, k_i^*(b_i^{''}))$*.*

2. *the utilization corresponding to* $b_i^{'}$ *is less than or equal to the utilization corresponding to* $b_i^{''}$*; i.e.,* $u_i(b_i^{'}) \leq u_i(b_i^{''})$*.*

Algorithm 3 Utilization based Search Space Reduction (USR) Algorithm

---

Input: $b_i^{\min}$,$b_i^{\max}$, $K_i = b_i^{\max} - b_i^{\min} + 1,\ i \in [1, N]$
Output: $b_{ij},\ i \in [1, N],\ j \in [1, C_i]$

1: **for** $i = 1 : N$ **do**
2:    order $b_i \in [b_i^{\min}, b_i^{\max}]$ in increasing $W_i(b_i)$: $S_i^p = [b_{i1}^p, ..., b_{iK_i}^p]$;
3:    order $b_i \in [b_i^{\min}, b_i^{\max}]$ in increasing $u_i(b_i)$: $S_i^u = [b_{i1}^u, ..., b_{iK_i}^u]$;
4:    $C_i = 1$;
5:    **while** $S_i^p \neq \emptyset$ **do**
6:       $b_{iC_i} = S_i^u(C_i)$;
7:       determine k s.t. $b_{ik}^p = b_{iC_i}$;
8:       remove $\{b_{ij}^p | j \geq k\}$ from S_i^p;
9:       remove $\{b_{ij}^p | j > k\}$ from S_i^u;
10:       $C_i = C_i + 1$;
11:    **end while**
12: **end for**

Utilization based Search Space Reduction (USR) Algorithm is based on reducing the number of constellation size values corresponding to each sensor node separately by exploiting the dominance relations among the constellation size values, as described in detail next (Algorithm 3). For each sensor node $i \in [1, N]$, the algorithm starts by sorting the corresponding constellation size values in the feasible interval $[b_i^{\min}, b_i^{\max}]$ in increasing power consumption $W_i(b_i)$ and utilization $u_i(b_i)$, and storing them in the sets S_i^p and S_i^u, respectively (Lines $1-2$). Here, b_{ij}^p and b_{ij}^d denote the constellation size with the j-th smallest power consumption and delay, respectively. The algorithm initializes C_i to 1 and increases C_i by 1 at each iteration of the algorithm such that the constellation size corresponding to the C_i-th smallest utilization in S_i^u is considered at the C_i-th iteration (Line $4, 6, 10$). At the

C_i-th iteration, the index of b_{iC_i} in the set S_i^p is determined and denoted by k (Line 7). Based on Definition 2, b_{iC_i} optimally dominates all the constellation sizes $\{b_{ij}^p | j > k\}$ since it has both lower utilization and energy consumption. $\{b_{ij}^p | j > k\}$ therefore need not be considered in the rest of the algorithm and are removed from S_i^u (Line 9). The total number of iterations corresponding to each node i given by the value of C_i at the output of the algorithm then gives the reduced number of constellation sizes for consideration. The algorithm stops when all the constellation sizes are either included in the output of the algorithm or removed by considering dominance relations (Lines $5, 8$).

Table 3.1: Example search space reduction by USR.

Iteration #	S_i^u	S_i^p
0	$\{4, 10, 9, 3, 7, 8, 6, 1, 2, 5\}$	$\{3, 5, 9, 8, 2, 4, 1, 10, 6, 7\}$
1	$\{4, 9, 3, 8, 2, 5\}$	$\{3, 5, 9, 8, 2\}$
2	$\{4, 9, 3, 5\}$	$\{3, 5\}$
4	$\{4, 9, 3\}$	$\emptyset$

Consider the example execution of the algorithm in Table 3.1. Let $b_i^{\min} = 1$ and $b_i^{\max} = 10$ for sensor node i and the sorted lists of the constellation sizes in increasing orders of utilization and power consumption be given by $\{4, 10, 9, 3, 7, 8, 6, 1, 2, 5\}$ and $\{3, 5, 9, 8, 2, 4, 1, 10, 6, 7\}$, respectively. At the first iteration of the algorithm, $b_{iC_i} = 4$ gives minimum utilization. The constellation size values $\{1, 10, 6, 7\}$ yield both greater utilization and power consumption than 4. Hence, these values are removed from S_i^u, resulting in $\{4, 9, 3, 8, 2, 5\}$, and together with $b_{iC_i} = 4$ value itself from S_i^p, resulting in $\{3, 5, 9, 8, 2\}$. At the second iteration of the algorithm, the constellation size value with the second minimum utilization, given by $b_{iC_i} = 9$, is considered. The constellation size values $\{8, 2\}$ yield both greater utilization and power consumption. S_i^u is updated to exclude these values, resulting in $\{4, 9, 3, 5\}$, and S_i^p is updated by excluding these values and the constellation size 9, resulting in $\{3, 5\}$. At the last iteration of the algorithm, when the constellation size value with the third minimum utilization, given by $b_{iC_i} = 3$, is considered, both the optimally dominating set given by $\{5\}$ and the constellation size itself equal to 3 are excluded from S_i^p. S_i^p then becomes empty set, which is the end of the algorithm for sensor node i. When we order the constellation size values in the resulting set $S_i^u = \{4, 9, 3\}$ in increasing power consumption, the resulting set is given by $\{3, 9, 4\}$. These sets are reversed versions of each other. This result is given as a Lemma next and exploited in the heuristic algorithms in Sections 3.5.2 and 3.5.3.

Lemma 11. *In the set S_i^u at the output of USR algorithm, in the direction of increasing power consumption, the utilization decreases.*

Proof: S_i^u initially consists of all possible constellation size values $j \in [1, K_i]$ in increasing order of utilization. At each iteration of the algorithm, the algorithm eliminates the constellation size values from the set S_i^u with greater indices and higher power consumptions than the constellation size considered at that iteration. Hence, for any constellation size in the set S_i^u at the output of USR algorithm, the constellation sizes with greater indices have lower power consumptions. Therefore, in the direction of increasing utilization, the power consumption decreases and vice versa. $\square$

In the following, we propose two heuristic algorithms such that the input of the algorithms is the constellation size values b_{ij} in increasing order of utilization for each node $i \in [1, N]$ and $j \in [1, C_i]$ obtained as the result of the USR Algorithm. As briefly discussed previously, since the constellation sizes that are dominated by others are eliminated, a list in the increasing order of power consumptions will be in decreasing order of the utilizations. The following heuristics exploit this special characteristic in two different directions. The first algorithm starts with the solution which yields the worst result and whose feasibility is most probable if the problem is feasible and attempts to find solutions yielding better results while keeping the feasibility. The second algorithm on the other hand starts with the best solution which is most probably infeasible since utilizations are at maximum and attemps to find a feasible solution while degrading the solution.

3.5.2 Keep Feasible Improve Maximum (KFIM) Algorithm

Algorithm 4 **Keep Feasible Improve Maximum (KFIM) Algorithm**

Input: b_{ij}, $\forall i \in [1, N]$, $\forall j \in [1, C_i]$;

1: $\mathbf{b} = (b_{11}, b_{21}, ..., b_{N1})$;
2: **if isSchedulable(b) then**
3:     **while** $\mathbf{b} \neq (b_{1C_1}, b_{2C_2}, ..., b_{NC_N})$ **do**
4:         $j = \text{argmin}_i \, f(\mathbf{b}_i^{++})$;
5:         **if isSchedulable(**$\mathbf{b}_j^{++}$**) then**
6:             $\mathbf{b} = \mathbf{b}_j^{++}$;
7:         **else**
8:             break;
9:         **end if**
10:     **end while**
11:     **return** **b**;
12: **else**
13:     **return** no feasible solution exists;
14: **end if**

Keep Feasible Improve Maximum (KFIM) algorithm, given by Algorithm 4, is described as follows. Let the i-th element of the constellation size vector $\mathbf{b}$ given by $\mathbf{b}_i$ correspond to the constellation size of node i. Let $\mathbf{b}_i^{++}$ denote the constellation size vector obtained by keeping all the elements of the vector $\mathbf{b}$ the same except the i-th element, which is set to the next constellation size in the increasing utilization direction if it is not equal to b_{iC_i} and also kept the same otherwise. The algorithm starts with the constellation size vector $\mathbf{b}$ corresponding to the minimum utilization and maximum power consumption for each node in the network; i.e., $\mathbf{b}_i = b_{i1}$ for all $i \in [1, N]$ (Line 1). If this vector is not schedulable, then the algorithm terminates stating no feasible solution exists since the tutilization of each sensor node is at minimum (Lines $12 - 13$). If it is schedulable, then the algorithm determines the sensor node j that minimizes the objective value f when the constellation size of that sensor node is set to the next value in the direction of increasing utilization or, in other words, decreasing power consumption (Line $2, 4$). If the resulting constellation size vector $\mathbf{b}_j^{++}$ is schedulable, then it is set to the current best solution $\mathbf{b}$ (Lines $5-6$). The algorithm terminates if $\mathbf{b}_j^{++}$ is not schedulable or $\mathbf{b} = (b_{1C_1}, b_{2C_2}, ..., b_{NC_N})$ (Lines 3 and $8-9$). An example for the KFIM algorithm is given in Table 3.2.

Table 3.2: KFIM Algorithm example for the case where $N = 3$ and $C_i = 5$ for all $i \in [1, N]$. The highlighted vector is the solution of the algorithm.

Iteration #	Constellation Size Vector	Schedulability
0	(b_{11}, b_{21}, b_{31})	+
1	(b_{11}, b_{22}, b_{31})	+
2	$(\mathbf{b_{11}}, \mathbf{b_{23}}, \mathbf{b_{31}})$	+
3	(b_{11}, b_{24}, b_{31})	-

3.5.3 Seek Feasible Degrade Minimum (SFDM) Algorithm

Algorithm 5 **Seek Feasible Degrade Minimum (SFDM) Algorithm**

Input: b_{ij}, $\forall i \in [1, N]$, $\forall j \in [1, C_i]$;

$\mathbf{b} = (b_{1C_1}, b_{2C_2}, ..., b_{NC_N})$;
while $\mathbf{b} \neq (b_{11}, b_{21}, ..., b_{N1})$ **do**
    if isSchedulable(b) then
        break;
    else
        $j = \text{argmin}_i \, f(\mathbf{b}_i^{--})$;
        $\mathbf{b} = \mathbf{b}_j^{--}$;
    end if
end while
if isSchedulable(b) then
    return $\mathbf{b}$;
else
    return no feasible solution exists;
end if

Seek Feasible Degrade Minimum (SFDM) algorithm, given by Algorithm 5, is described as follows. Let $\mathbf{b}_i^{--}$ denote the constellation size vector obtained by keeping all the elements of the vector $\mathbf{b}$ the same except the i-th element, which is set to the previous constellation size in the decreasing utilization direction if it is not equal to b_{i1} and also kept the same otherwise. The algorithm starts with the constellation size vector $\mathbf{b}$ corresponding to the minimum power consumption and maximum utilization for each node in the network; i.e., $\mathbf{b}_i = b_{iC_i}$ for all $i \in [1, N]$. If this constellation size vector is schedulable, the algorithm terminates and solution is returned (Lines $3-5$ and $9-10$). Otherwise, the algorithm starts to seek a feasible solution in the direction of minimum increase in the value of the objective function (Lines $5-7$). At each iteration, the algorithm determines the sensor node j that minimizes the objective value f when the constellation size of that sensor node is set to the next value in the direction of increasing power consumption or, in other words, decreasing utilization. The resulting constellation size vector $\mathbf{b}_j^{--}$ is set to the current solution $\mathbf{b}$. The algorithm terminates if the resulting $\mathbf{b}$ is schedulable or $\mathbf{b} = (b_{11}, b_{21}, ..., b_{N1})$ (Lines $2-4$). An example for the SFDM algorithm is given in Table 3.3.

Table 3.3: SFDM Algorithm example for the case where $N = 3$ and $C_i = 5$ for all $i \in [1, N]$. The highlighted vector is the solution of the algorithm.

Iteration #	Constellation Size Vector	Schedulability
0	(b_{15}, b_{25}, b_{35})	-
1	(b_{15}, b_{25}, b_{34})	-
2	(b_{14}, b_{25}, b_{34})	-
3	(b_{13}, b_{25}, b_{34})	-
4	(b_{13}, b_{25}, b_{33})	-
5	(b_{12}, b_{25}, b_{33})	-
6	(b_{12}, b_{24}, b_{33})	-
7	(b_{11}, b_{24}, b_{33})	-
8	(b_{11}, b_{24}, b_{32})	-
9	(b_{11}, b_{24}, b_{31})	-
10	$(\mathbf{b_{11}}, \mathbf{b_{23}}, \mathbf{b_{31}})$	+

The complexity of both heuristic algorithms is $O(\sum_{i\in[1,N]} C_i \times F)$ since the algorithm checks the schedulability of the constellation size vector with complexity $O(F)$ and varies the constellation size of one node at each iteration, resulting in $\sum_{i\in[1,N]} C_i$ iterations at maximum.

3.6 Performance Evaluation

The goal of this section is to evaluate the performance of the heuristic algorithms compared to the traditional separate design of controller and communication systems, previously proposed heuristic algorithms and optimal solution for different modulation schemes and energy consumption related objective functions. In the separate design of controller and communication systems, denoted by “TS”, the constellation size and sampling period of the sensor nodes are predetermined such that a feasible solution is guaranteed for the worst case scenario without adjustments to different simulation scenarios such as varying network and channel conditions. For example, WirelessHart [8] and ISA100.11a [7], which are the two competing wireless standards for industrial control applications, employ O-QPSK (Offset Quadrature Phase Shift Keying) without including any mechanism for the optimization of constellation size and sampling period. The heuristic algorithm previously proposed for the joint optimization of controller and communication systems with the objective of minimizing the total power consumption of the communication system while guaranteeing the performance and stability of the control system and the schedulability of the communication system for MQAM modulation and EDF scheduling, is denoted by “HS” [12]. This algorithm is based on analyzing the optimality conditions on the variables of the problem to reduce it to an IP problem and proposing an efficient solution method based on its relaxation. The optimal solution is obtained by OFE algorithm and denoted by “OFE”. The proposed heuristic algorithms that combine utilization based search space reduction algorithm, denoted by USR, with efficient search algorithms, denoted by KFIM and SFDM, are called “USR-KFIM” and “USR-SFDM”, respectively. For schedule construction and schedulability analysis, we use EDF scheduling algorithm.

Simulation results are obtained by averaging the performance over 1000 independent random network topologies where the sensor nodes are uniformly distributed within a circular area and transmit to a controller located in the center of the area. The channel attenuation is calculated by using Rayleigh fading with scale parameter set to the mean power level determined by using the large scale statistics modeled as $PL(d) = PL(d_0) + 10\alpha \log(d/d_0) + Z$, where d is the distance between the node and the controller, $PL(d)$ is the path loss at distance d, $PL(d_0)$ is the path loss at reference distance d_0, Z is the Gaussian random variable with zero mean and standard deviation σ_z [56, 57]. Each sensor node i has a packet length of L_i bits to be transmitted periodically. Table-3.4 lists the parameters used in the simulations.

$W^{t,\max}$	250 mW	W^c	50mW
$L_i, i \in [1, N]$	100 bits	δ	0.95
$PL(d_0)$	70 dB	d_0	1
α	3.5 [36]	σ_z^2	4 dB [56]

Table 3.4: Simulation Parameters

Figs. 3.2-3.3 show the total power consumption and log sum power consumption for different number of nodes where the modulation scheme is MQAM. The nodes are uniformly distributed within a circular area of radius 5 m. The MAD and MATI values are chosen as $\Delta = 25$ ms and $\Omega = 25$ ms. The resulting power consumption values of the TS algorithm increase linearly as expected as the number of nodes increases since the predetermined constellation size is constant. For the other algorithms on the other hand, as the number of nodes increases up to a specific value which is around 25 in this case, the power consumption increases linearly since the objective of the power minimization does not require a joint decision among multiple nodes in the network meaning that minimizing the power consumption independently for each node does not result in the violation of the schedulability constraint. However, as the number of nodes in the WNCS increases above this value, a joint decision becomes necessary to satisfy the schedulability constraint. Note that, this threshold value may change depending on the specific scheduling algorithm used. The nodes in the WNCS tend to select constellation size values corresponding to smaller transmission delay values than they would select when they minimize their power consumption cost independently of the other nodes in the WNCS. This causes an acceleration affect in the increase in the power consumption cost. However, this acceleration affect is greater for HS algorithm which moves the resulting power consumption value away from the optimal solution dramatically. On the other hand, the heuristic algorithms USR-KFIM and USR-SFDM perform very close to optimal and resemble the optimal solution while the joint decision among the nodes in the WNCS is strictly required. The average approximation ratio values of the USR-KFIM and USR-SFDM algorithms are below 1.01 where the approximation ratio is obtained as the ratio of the solution of a particular algorithm to the optimal solution.

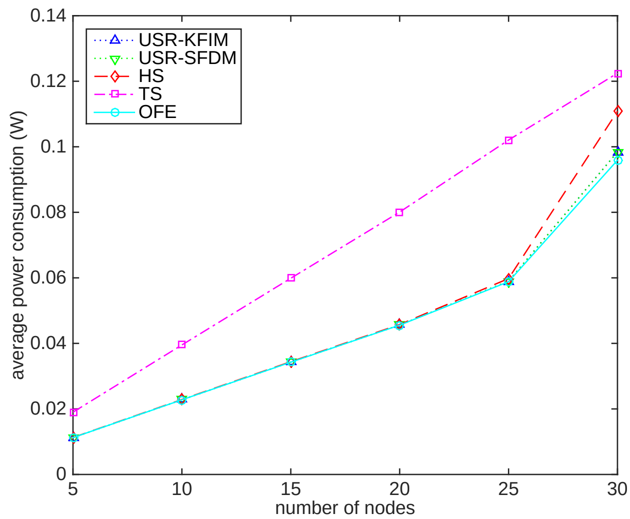

Figure 3.2: Total power consumption in networks with different number of nodes for MQAM modulation scheme.

Figs. 3.4-3.5 show the total power consumption and log sum power consumption for different number of nodes where the modulation scheme is MFSK. The nodes are uniformly distributed within a circular area of radius 5 m. The MAD and MATI values are chosen as $\Delta = 25$ ms and $\Omega = 25$ ms. The performance of the algorithms are very similar to the results depicted in Figs. 3.2-3.3 for MQAM modulation scheme. The proposed heuristic algorithms perform very close to optimal yielding average approximation ratio values below 1.01. This illustrates the robustness of the proposed algorithms to different modulation schemes.

Fig. 3.6 illustrates the total power consumption in a network of 20 nodes

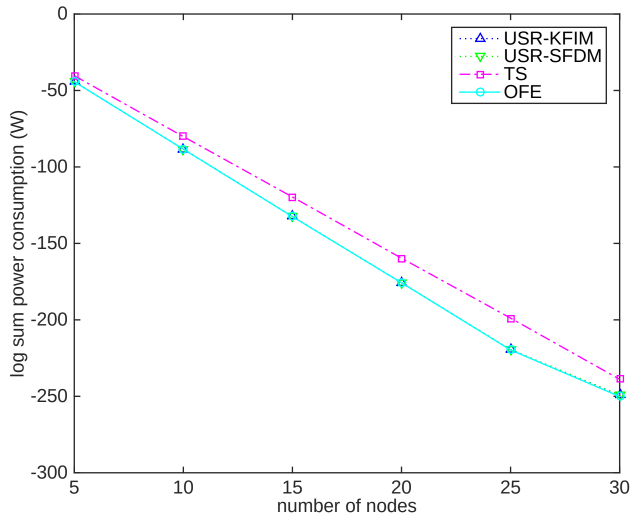

Figure 3.3: Log sum power consumption in networks with different number of nodes for MQAM modulation scheme.

for different MAD values where the modulation scheme is MQAM. The nodes are uniformly distributed within a circular area of radius 10 m and the MATI value is chosen as $\Omega = 200$ ms. While the MAD constraint constrains either the minimum or maximum constellation size depending on the specific modulation scheme used, it specifies the minimum constellation size for MQAM since the transmission delay is inversely proportional to MAD value for MQAM modulation scheme. As the MAD increases up to a certain value, around 2 ms, the total power consumption decreases since for smaller MAD values, the nodes in the network are forced to select greater constellation size values, which increases the power consumption cost. While the optimality gap for the HS algorithm widens for the MAD values below this threshold value, the performance proposed heuristic algorithms follow the optimal solution very closely. On the other hand, beyond the threshold value, increasing MAD value has no effect on the average power consumption due to the fact that the optimal constellation size remains constant although the feasible region expands. Since increasing MAD value expands the feasible region, the constellation size value for TS is determined such that it yields a

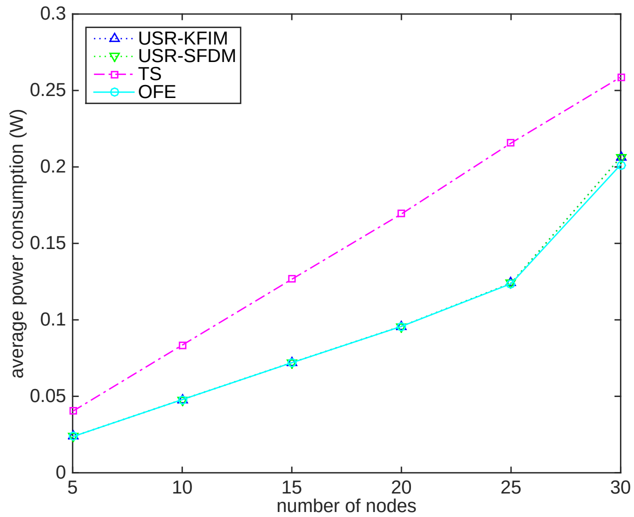

Figure 3.4: Total power consumption in networks with different number of nodes for MFSK modulation scheme.

feasible schedule for the minimum MAD value since then schedulability for all MAD values is ensured. Since the power cost function does not depend on the MAD value, the resulting total power consumption obtained by TS algorithm remains constant for different MAD values and dramatically worse than the other algorithms. The average approximation ratio of the proposed heuristic algorithms are less than 1.01 and outperforms the HS algorithm for all MAD values.

Fig. 3.7 illustrates the total power consumption in a network of 20 nodes for different MATI values where the modulation scheme is MQAM. The nodes in the WNCS are uniformly distributed within a circular area of radius 10 m and the MAD value is chosen as $\Delta = 10$ ms. The total power consumption of the WNCS decreases as the MATI value increases since the total power consumption is a monotonically decreasing function of the MATI value. Moreover, as the MATI value decreases, the feasible region determined by the schedulability region shrinks yielding a faster increase in the total power consumption than the expected increase from the functional dependency of the total power consumption on MATI. This effect depending on the

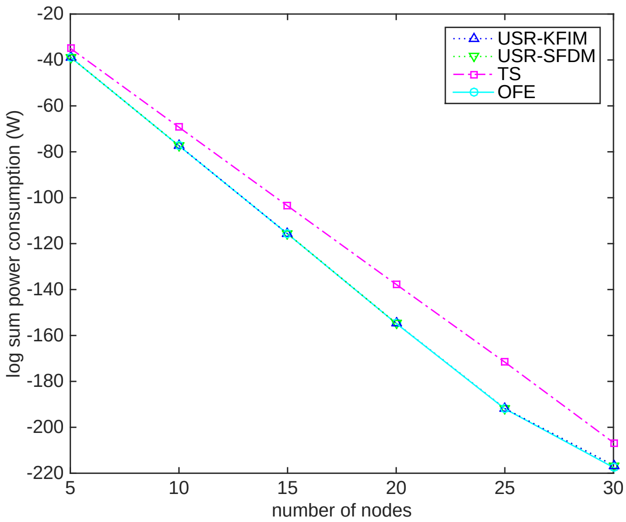

Figure 3.5: Log sum power consumption in networks with different number of nodes for MFSK modulation scheme.

joint optimization of the constellation size values of the sensor nodes in thw WNCS is not visible for TS algorithm since the constellation size values of the nodes are predetermined as constant to ensure a feasible schedule for all MATI values. Again, when the joint decision among the nodes is required for smaller MATI values, the performance of the HS algorithm differentiates from the optimal solution significantly. On the other hand, the performances of the proposed heuristic algorithms are robust to varying MATI values and very close to optimal with an average approximation ratio of around 1.01 outperfoming the HS algorithm yielding an approximation ratio of around 1.05.

Fig. 3.8 shows the average runtime of the algorithms for different number of nodes and MQAM modulation scheme where the objective is total power consumption in the network. The nodes are uniformly distributed within a circular area of radius 5 m. The MAD and MATI values are chosen as $\Delta = 25$ ms and $\Omega = 25$ ms. As expected, USR-KFIM and USR-SFDM heuristic algorithms have the least average runtime values due to the fact that USR technique reduces the search space for constellation size vectors

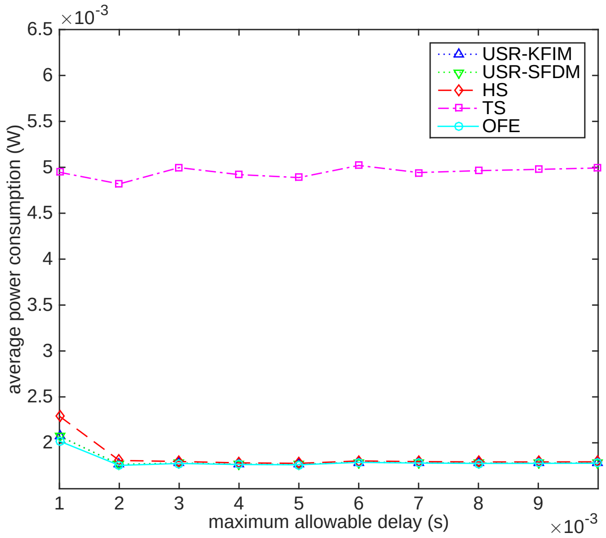

Figure 3.6: Total power consumption in a network of 20 nodes for different MAD values and MQAM modulation scheme.

to a greater extent. The average runtime of USR-KFIM and USR-SFDM increase almost linearly as the number of nodes increases and outperform the HS algorithm. On the other hand, the average runtime of OFE algorithm increases exponentially as the number of nodes increases and is much more than the average runtime of the heuristic algorithms since it does not employ the search space reduction technique.

3.7 Conclusion and Discussion

In this study, we have extended our previous work on the joint design of communication and control systems in WNCS in which the joint optimization problem is formulated with the objective of minimizing the average power consumption of the network for MQAM modulation scheme. We have generalized the optimization problem for wide range of objectives including total power consumption of the network, maximum power consumption among the nodes in the network and log sum of the power consumptions of the nodes in the network and for any modulation scheme satisfying certain properties

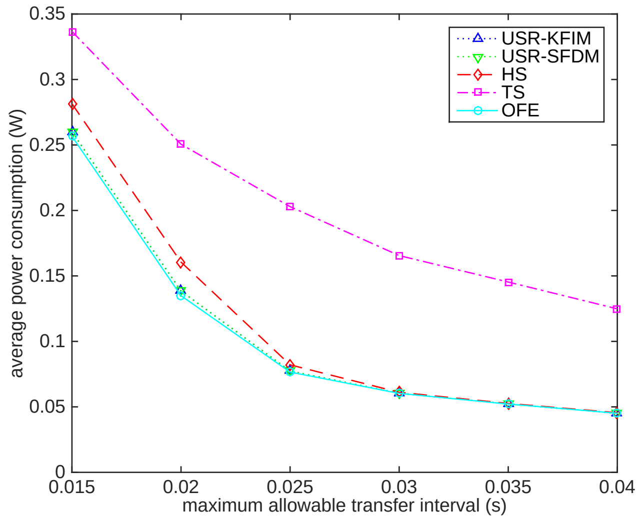

Figure 3.7: Total power consumption in a network of 20 nodes for different MATI values and MQAM modulation scheme.

including mostly-used MQAM and MFSK modulation schemes. The generalization for different power related objectives and different modulation schemes leads to broader adoption of the proposed solution mechanisms to different NCS applications. We propose an exact algorithm for the generalized problem based on the optimality relations among the decision variables of the problem and smart enumeration techniques. Moreover, we propose two polynomial time heuristic algorithms which employ a utilization based search space reduction technique decreasing the complexity of the algorithms compared to the exact algorithm significantly. Extensive simulations illustrate that the proposed solution methods outperform the existing solution methods significantly for different network scenarios and varying control system parameters including MAD and MATI values.

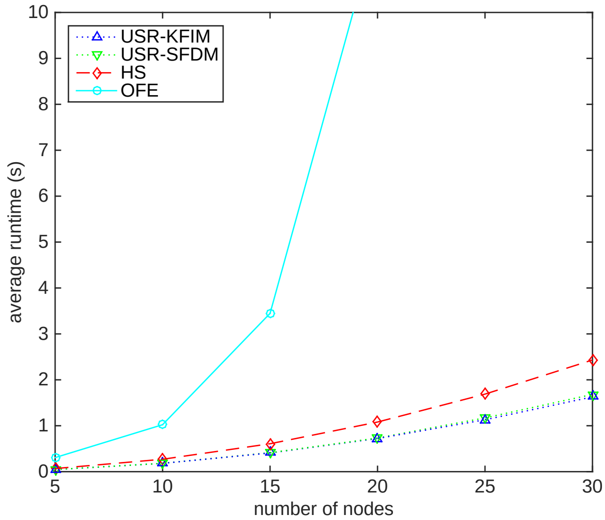

Figure 3.8: Average runtime of the algorithms for different number of nodes and MQAM modulation scheme where the objective is total power consumption in the network.

Chapter 4

Maximum Adaptive Schedule for Wireless Networked Control Systems

Communication system design for Wireless Networked Control Systems (WNCSs) is very challenging since the strict timing and reliability requirements of control systems should be met by the wireless communication systems that introduce non-zero packet error probability and non-zero delay at all times. Particularly, the scheduling algorithms for WNCSs should be designed to provide maximum level of adaptivity accommodating packet losses and changes in network topology while exploiting periodic nature of the sensor node transmissions. Creating such a schedule has been previously studied for an Ultra-Wideband (UWB) based WNCS. In this study, we extend the joint optimization problem of power control, rate adaptation and scheduling with the objective of providing maximum adaptivity for general WNCSs employing continuous rate transmission model in which Shannon's channel capacity formulation is used for the achievable transmission rate. Upon proving the NP-hardness of the problem, we provide a framework for the design of a heuristic algorithm for scheduling and propose an optimal polynomial time algorithm for the power control and rate adaptation problem following the derivation of the optimality conditions. We demonstrate via extensive simulations that the proposed algorithms outperform the existing algorithms with performance close to optimal solution and average runtime admissible for practical WNCSs.

4.1 Introduction

WNCSs are spatially distributed feedback control systems in which sensor nodes transmit information to controllers via a wireless network [1]. The usage of wireless communication in information exchange brings various advantages such as the ease of installation and maintenance, low complexity and cost, and large flexibility to accommodate the modification and upgrade of sensor nodes. WNCSs have been studied for a wide range of applications in industrial automation [61], building automation [4], automated highway [19], smart grid [62] and vehicle control systems [63], and backed up by several leading industrial organizations such as International Society of Automation (ISA) [7], Highway Addressable Remote Transducer (HART) [8], and Wireless Industrial Networking Alliance (WINA) [9].

The scheduling algorithms of WNCSs have been studied for event-triggered controllers acting in response to spatially distributed time-triggered sensor nodes. The main challenge in the scheduling algorithm design is the exploitation of the periodic transmission nature of the sensor nodes in meeting the contradicting requirements and capabilities of the control and wireless communication systems: The strict timing and reliability requirements of control systems should be met by the wireless communication systems that introduce non-zero packet error probability and non-zero delay at all times. Previous work however either did not exploit the periodic nature of the sensor node transmissions or adopt the algorithms designed for the scheduling of the periodic controller tasks running on a processor to WNCSs without considering the wireless communication imperfections. Recent communication standards for WNCSs such as WirelessHART [64], ISA-100.11a [65] and IEEE 802.15.4e [46] adopt a globally synchronized multi-channel Time Division Multiple Access (TDMA) with a multi-hop multi-path routing protocol. Scheduling algorithms based on these standards however are designed mostly to provide low deterministic end-to-end delay and controlled jitter for real-time traffic across a very large mesh network distributed over a large area without considering the periodic nature of the transmissions [24, 25]. On the other hand, the algorithms designed for the scheduling of periodic controller tasks running on a processor, such as Earliest Deadline First (EDF), Least Laxity First (LLF) and Deadline Monotonic (DM) scheduling algorithms [66], have been adopted for the scheduling of the direct transmission of the periodic data packets of the sensor nodes to their corresponding controllers for the case where no concurrent transmissions are allowed [26, 27, 67]. However, since they assign time slots to the tasks as soon as they are available, none of these algorithms provide any adaptivity to the packet losses.

In addition to accommodating packet losses, the WNCS schedule should provide adaptivity to the changes in the requirements of the sensors and network topology. The adaptivity of the WNCS schedule can be attained by distributing the packet transmissions of the sensors as uniformly as possible over time as illustrated in Fig. 4.1. Fig. 4.1 (a) shows the schedule generated by EDF, which has been shown to be optimal for deadline constrained scheduling under various modeling assumptions [66], whereas Fig. 4.1 (b) shows an alternative schedule based on distributing the time slots of the sensors uniformly over time. The example network consists of 4 sensor nodes. Sensor nodes 1 and 2 have packet generation periods of 1 ms; i.e., they have to transmit one packet in each 1 ms time interval, and their packet transmissions require time slot allocations of lengths $t_1 = 0.15$ ms and $t_2 = 0.25$ ms respectively. Sensor nodes 3 and 4 have packet generation periods of 2 ms and their packet transmissions require time slot allocations of lengths $t_3 = 0.25$ ms and $t_4 = 0.30$ ms respectively. The schedule in Fig. 4.1 (b) is more adaptive to packet losses, topology and sensor requirement changes than EDF schedule in Fig. 4.1 (a) as illustrated below:

- Suppose that the transmission of the data packet of sensor 2 in the first 1 ms failed. The schedule in Figure 4.1 (b) includes enough space to allocate the retransmission of sensor 2 whereas the schedule in Figure 4.1 (a) does not.

- Suppose that the transmission rate of sensor 3 needs to decrease such that the time slot length is doubled as $t_3 = 0.5$ ms. The allocation in Figure 4.1 (b) is able to accommodate the new change whereas the one in Figure 4.1 (a) cannot.

- Suppose that in addition to the periodic data packet generation of the scheduled sensor nodes, an additional packet of 0.3 ms time slot length is generated by an event-triggered sensor node at the beginning of the scheduling frame. Then the time slot of the event-triggered packet can be allocated with a delay of 0.65 ms in the schedule of Figure 4.1 (b) compared to 1.4 ms in the schedule of Figure 4.1 (a).

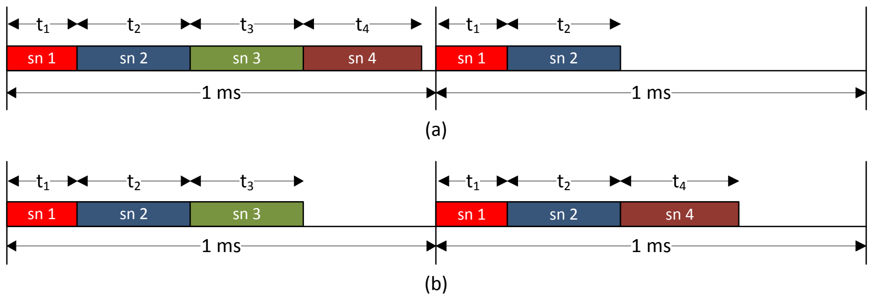

Figure 4.1: Example schedule used for the illustration of the adaptivity requirement. (a) EDF schedule. (b) Alternative schedule based on distributing the time slots of the sensors uniformly over time. The network consists of 4 sensor nodes denoted by sn i for $i \in [1, 4]$. The blocks with the same color represent the packet transmissions of the same sensor node.

Besides providing maximum adaptivity, the schedule designed for WNCSs should satisfy the requirements of the individual sensors in terms of *periodic packet generation*, *transmission delay*, *reliability* and *energy consumption*. Satisfying the packet generation period, transmission delay and reliability requirement of the sensors given the reliability of the underlying wireless channel is necessary to maintain a certain control system performance [68, 69]. Transmission energy requirement on the other hand ensures achieving the desired lifetime of the sensor nodes, which are mostly battery operated or rely on energy harvesting techniques.

The transmission delay and energy consumption of a sensor node are determined by the transmission power and rate of that node, which depend on the transmission power and rate of the sensor nodes that are scheduled to transmit concurrently with that particular sensor node. Therefore, the schedule should be determined together with the transmission power and rate allocation of the sensor nodes in a joint framework. Joint optimization of the transmission power control, rate adaptation, and scheduling has been widely studied for delay constrained energy minimization in general purpose wireless networks. Different algorithms have been proposed for the solution of the optimization problem depending on the delay constraint in terms of either a single deadline to all packets [30, 32] or individual deadlines for each packet [33, 35]. However, none of the previously proposed algorithms consider the joint optimization of scheduling, power and rate allocation to satisfy the periodic data generation, transmission delay, reliability and energy consumption requirements of the sensor nodes and the adaptivity requirement of the network.

The goal of this chapter is to study the joint optimization of power control, rate adaptation and scheduling for WNCS with the objective of providing maximum adaptivity while considering the constraints of the individual sensors in terms of periodic packet generation, transmission delay, reliability and energy consumption. This problem has been studied earlier for Ultra Wide Band (UWB) based Intra-Vehicular Wireless Sensor Network (IVWSN), which is a WNCS deployed inside vehicles where the sensor nodes transmit their data to their corresponding electronic control unit to be used in the real-time control of the mechanical parts in several compartments of the vehicle [63]. This study extends it by generalizing the underlying communication technology and the WNCS application. The original contributions of this study are listed as follows:

- We generalize the framework of the joint optimization of power control, rate adaptation and scheduling proposed for IVWSNs [63] to WNCSs. We prove the NP-hardness of the optimization problem. Since solving an NP-hard problem optimally requires exponential runtime in the network size and is intractable even for moderate network sizes, in order to find fast and efficient solutions, we provide a framework for the design of a heuristic algorithm for scheduling for the first time in the literature. The proposed scheduling algorithm consists of a node assignment algorithm that assigns the nodes to the subframes in a certain order considering their transmission times and connection to the controllers, and a concurrency allocation algorithm that determines the subsets of the nodes that can transmit concurrently from the node sets with the same packet generation period allocated to the same subframe by the node assignment algorithm by using power control and rate adaptation algorithm.

- We formulate a novel power control and rate adaptation problem with the objective of minimizing the time required for the concurrent transmission of a set of sensor nodes while satisfying their transmission delay, reliability and energy consumption requirements for general wireless networks by adopting the Shannon's capacity formulation for the first time in the literature. We show that this is a non-convex optimization problem and then propose a polynomial time algorithm that provides the optimal solution within a predetermined error bound following the derivation of the optimality conditions.

- The superior adaptivity and close to optimality of the proposed framework compared to previously proposed algorithms have been demonstrated for different node assignment and concurrency allocation algorithms of different complexities under various network scenarios via extensive simulations.

The rest of the chapter is organized as follows. Section 4.2 describes the system model and the assumptions used throughout the chapter. In Section 4.3, the objective and the constraints of the joint optimization of power control, rate adaptation and scheduling problem are formulated and the NP-hardness of this problem is proved. Section 4.4 provides the framework for the design of a heuristic algorithm for scheduling. Section 4.5 presents the optimization problem for power control and rate adaptation and describes the optimal polynomial time algorithm upon deriving optimality conditions. Simulations and performance evaluation are presented in Section 4.6. Finally, concluding remarks and future work are given in Section 4.7.

4.2 System Model and Assumptions

The system model and assumptions are detailed as follows:

1. The architecture of a WNCS is illustrated in Fig. 4.2. Sensor nodes attached to the plants sample and transmit their sensed data either periodically or upon a triggering event, depending on whether the sensor node is time-triggered or event-triggered respectively, to the corresponding controller via a wireless channel. Controllers do not have multi-reception capability; i.e., they cannot receive packets from more than one sensor node at a time. Upon reception of the information from a sensor node, controller computes a new control command, which is then forwarded to the actuator. We assume that the actuators receive the controller command successfully. Many practical NCSs have several sensing channels, whereas the controllers are collocated with the actuators because the control command is very critical [37].

2. We use the terms *node* and *link* interchangeably, where link l refers to the connection between sensor node l and its controller.

3. We assume that each sensor node directly communicates with its corresponding controller. The proposed framework can be extended to include the routing in the optimization problem for multi-hop transmission of the packets from the sensor nodes to the controller. To avoid complexity in the first step of the study and focus better on the scheduling, power control and rate adaptation problems, the extension for multi-hop networks is out of scope of this study and subject to future work.

4. One of the controllers in the WNCS is selected as the central coordinator hence responsible for the scheduling and resource allocation de-

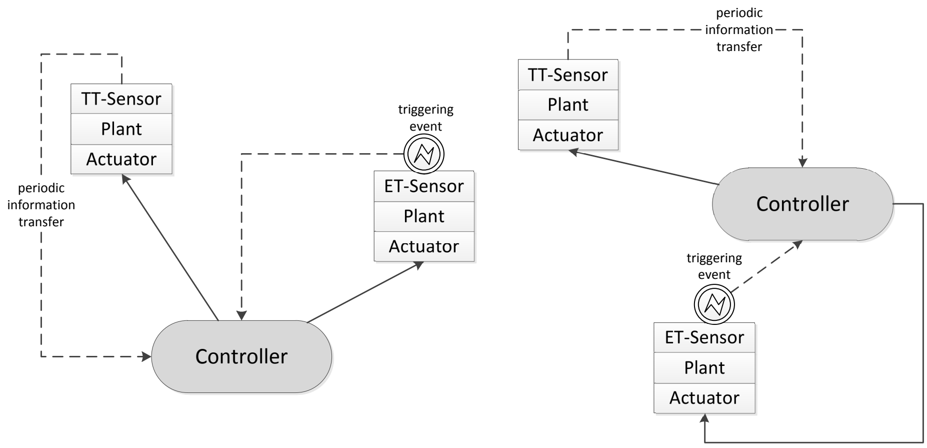

Figure 4.2: Overview of the WNCS setup. TT-Sensor and ET-Sensor represent time-triggered and event-triggered sensor respectively.

cisions and network synchronization. Such a centralized architecture is frequently envisioned in industrial control applications [7, 8]. The central coordinator is assumed to have complete topology information. However, the mechanisms that can be employed for topology discovery and synchronization are out of scope of this study and can be found in [38] in detail. The central coordinator continually monitors the received power and the packet error rate over each link to adjust the scheduling and resource allocation decisions of the sensor nodes.

5. As Medium Access Control (MAC) protocol, we consider TDMA. TDMA is commonly used in industrial control applications [7, 8] since it yields delay guarantee and less energy consumption for the networks with predetermined topology and data generation characteristics [38].

6. The time is divided into fixed-length scheduling frames (super frame or beacon interval are also commonly used as an alternative to scheduling frame in wireless standards such as IEEE 802.15.4e [46]) each of which is further divided into a beacon and a number of variable-length time slots. The number of time slots in each scheduling frame and the length of the time slots depend on the scheduling and resource allocation decisions. A guard time exists between consecutive time slots to

compensate for synchronization errors. The beacon is used to manage time synchronization among the elements of the WNCS and broadcast the scheduling and resource allocation decisions to the sensor nodes. In case all sensor nodes are in direct communication range of the central coordinator, the beacon is only broadcast by the central coordinator. If there are sensor nodes outside the direct communication range of the central coordinator, it communicates the related beacon information to their corresponding controllers, which transmit the beacon to the sensor nodes in their domains. In case of packet failures, an optional beacon is immediately transmitted to indicate the required packet retransmissions. If the packet error rate is below an acceptable level then the scheduling and resource allocation decisions of the sensor nodes need to be updated and broadcast by the central coordinator.

7. The WNCS contains time-triggered and event-triggered sensors. Time-triggered sensors transmit their data to their associated controllers periodically. For time-triggered sensors, the packet generation period, transmission delay and reliability providing a certain control performance requirement, denoted by T_l, d_l and r_l respectively for link l, are assumed to be given. Event-triggered sensors, on the other hand, are activated upon a predetermined event acting as a trigger and transit to active mode from inactive mode to transmit their data to their controller about the triggering event. For event-triggered sensors, the packets are either transmitted in the unallocated parts of the schedule as they arrive or assigned periodic time slots with T_l, d_l and r_l values chosen such that $kT_l + d_l$ is less than the maximum response time requirement guaranteeing the arrival of packets within maximum response time with confidence $1 - (1 - r_l)^k$ for high criticality applications.

8. We assume that the packet generation period of each sensor is either a multiple or aliquot of the other packet generation periods; e.g. $\{1, 5, 10, 50...\}$. This can be given as a constraint to the control applications. Sensors generating data periodically at various frequencies exist in many WNCS applications such as industrial automation [61], smart grid [62] and

vehicle control [63].

9. To achieve fixed determinism, which is often preferred in control systems over bounded determinism [70], the duration of the time slot allocated to a sensor node is fixed over all subframes and the time difference between consecutive time slot allocations is fixed at its packet generation period.

10. The sensor nodes are assumed to operate in active mode when they are scheduled to transmit or receive a packet, sleep mode when they are not scheduled, and transient mode when they switch from sleep mode to active mode and vice versa. The energy consumption in active mode is much larger than that in the sleep and transient modes [36]. Moreover, the energy consumption in the reception of the beacon packets is assumed to be constant hence not subject to optimization. We therefore consider only the energy consumption during the packet transmissions in our optimization framework. The energy requirement e_l is defined as the maximum allowed energy consumption to transmit one packet for sensor l.

11. We use *continuous power* model in which the transmit power can take any value below a maximum level p_{max}. The continuous power model is commonly used in most of the previous work such as [71, 72] due to the simplification of the optimization problems and high approximation accuracy with the large number of discrete power levels that can be supported by the existing radios.

12. Let g_{kl} denote the channel gain from the transmitter of link k to the receiver of link l. We assume that the channel gain between each pair of transmitter and receiver is constant during the scheduling frame. Extension of this work considering fast fading wireless channels is beyond the scope of this study and subject to future work.

13. We use *continuous rate* transmission model [72, 73] in which Shannon's channel capacity formulation for an Additive White Gaussian Noise

(AWGN) wireless channel is used in the calculation of the maximum achievable rate as a function of Signal-to-Interference-plus-Noise Ratio (SINR) as

$$x_l \leq W \times \log\left(1 + \frac{1}{W} \times \frac{p_l g_{ll}}{N_0 + \sum_{k \neq l} p_k g_{kl}}\right) \tag{4.1}$$

where x_l is the transmission rate of link l, p_l is the transmission power of link l, W is the channel bandwidth, N_0 is the background noise power.

4.3 Optimization Model

4.3.1 Objective of the Problem

The objective of the joint power control, rate adaptation and scheduling problem is *to provide maximum level of adaptivity*, which has been formulated in [63]. In the following, we summarize the quantification of the adaptivity metric for the sake of completeness.

Providing maximum level of adaptivity requires distribution of data transmissions as uniformly as possible over the scheduling frame as illustrated in Fig. 4.1. This first necessitates the choice of a time unit over which the transmissions are distributed. We call this time unit *subframe*. Then, we define *total active length of a subframe l*, denoted by a_l, as the sum of the length of the time slots allocated to subframe l. The subframe length is determined as the minimum packet generation period of all sensor nodes in [63] and denoted by S here. Given that the schedule is repeated with the maximum packet generation period, the frame length is equal to the maximum packet generation period, denoted by F, containing $M = F/S$ subframes. The objective can then be quantified as minimizing the maximum total active length of all the M subframes in a frame.

Fig. 4.3 depicts the scheduling frame structure for a small network scenario. Packet generation period of sensor node 1 is 1 ms whereas the packet generation period of sensor nodes 2, 3 and 4 is 2 ms. The scheduling frame and subframe lengths are equal to maximum packet generation period 2 ms and minimum packet generation period 1 ms respectively and hence the scheduling frame contains 2 subframes. Total active length of subframe 1, a_1, is equal to the sum of the length of the time slots allocated to sensor nodes 1 and 2; i.e., $a_1 = t_1 + t_2$, and total active length of subframe 2, a_2, is equal to the sum of the length of the time slots allocated to sensor node 1 and concurrent transmission of sensor nodes 3 and 4; i.e., $a_2 = t_1 + t_{3,4}$.

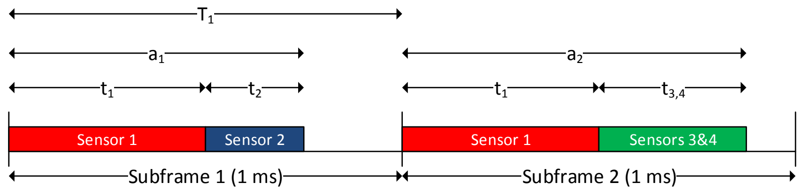

Figure 4.3: Scheduling frame structure for a network of 4 sensor nodes. The blocks with the same color represent the packet transmissions of the same sensor node.

4.3.2 Constraints of the Problem

The constraints of the optimization problem correspond to the individual requirements of the sensor nodes in terms of periodic data generation, transmission delay and transmission energy, and concurrent transmission requirement as follows:

1. Periodic data generation and fixed determinism requirements together necessitate that the duration of the time slot allocated to sensor node j be fixed over all subframes and the time difference between consecutive time slot allocations be equal to its packet generation period T_j as stated in Section 4.2. Fig. 4.3 illustrates this requirement for sensor node 1. The duration of the time slot allocated to sensor node 1 is fixed at t_1 and the time difference between its consecutive time slots allocated to subframes 1 and 2 is equal to its packet generation period T_1. This requirement can also be restated considering the subframe definition as follows: Let s_j be the ratio of the packet generation period T_j to the subframe length S. The allocation of a fixed length time slot

with period T_j can be achieved by the allocation of the fixed length time slot every s_j subframes then arranging the time slots within each subframe in increasing order of packet generation periods. To keep the time difference between consecutive time slot allocations the same when concurrent transmissions are considered, only the nodes with the same packet generation period are chosen for concurrent transmissions [63].

2. Transmission delay requirement necessitates that the length of the time slot allocated to sensor node j, which is kept constant over the schedule to maintain fixed determinism, be less than the required transmission delay d_j.

3. Transmission energy requirement necessitates that the energy consumed in the transmission of the packet of sensor node j be less than its energy requirement e_j.

4. Concurrent transmission requirement necessitates that the transmission power and rate of the sensor nodes assigned for concurrent transmission satisfy Shannon's channel capacity constraint given in Eqn. (4.1) and maximum transmit power condition.

4.3.3 NP-hardness of the Problem

Theorem 2. *The joint power control, rate adaptation and scheduling problem with the objective given in Section 4.3.1 and constraints given in Section 4.3.2 is NP-hard.*

Proof: Consider a problem instance where there exists only one controller in the WNCS. Then, only one sensor node can be allocated to a time slot. The length of the corresponding time slot is equal to the transmission time of that particular sensor node, which can be calculated independently of the other nodes since there are no concurrent allocations. Given the transmission time of the sensor nodes, the problem reduces to the scheduling problem where the time slots corresponding to the sensor nodes are allocated to the subframes such that their maximum total active length is minimized while considering the periodicity requirement of these sensor nodes. This specific problem instance is the same as the scheduling problem in Theorem 1 of [63] in which the one electronic control unit case in IVWSNs is investigated, which is proven to be NP-hard by the reduction of the NP-hard Minimum Makespan Scheduling Problem (MSP) on identical machines to the scheduling problem. □

4.4 Scheduling

The joint power control, rate adaptation and scheduling problem as described in detail in Section 4.3 is NP-hard. In order to solve this joint problem fast and efficiently, we propose a framework for the design of a heuristic scheduling algorithm. Scheduling algorithm assigns the nodes to the subframes through the node assignment (NA) algorithm and determines the subsets of nodes that concurrently transmit through the concurrency allocation (CA) algorithm.

The NA algorithm may consider the transmission time of each sensor node while satisfying its delay and energy requirement in its assignment to the subframes whereas the CA algorithm uses whether the concurrent transmission of a subset of sensor nodes while satisfying their transmission delay and energy requirements is feasible and their transmission time if feasible. Given the subset of nodes that are assigned to a subframe for concurrent transmission, Power Control and Rate Adaptation (PCRA) Algorithm determines their transmission power and transmission rate with the goal of minimizing their concurrent transmission time to minimize the total active length of the subframes. Since the problem of minimizing the transmission time of a node subset is independent of the other subsets, PCRA algorithm can be used as a black box in both NA and CA algorithms. The detailed formulation and solution of the PCRA algorithm is given in Section 4.5.

The general structure of the scheduling algorithm is depicted in Fig. 4.4. Let L sensor nodes in the WNCS have Z distinct packet generation periods such that $T_1 < T_2 < ... < T_Z$. Let L_j denote the set of sensors with packet generation period T_j. For each packet generation period T_j in the increasing order, the NA algorithm assigns each node $i \in L_j$ to the subframe with minimum total active length. Let L_{jk} denote the subset of the nodes in L_j allocated to the same subframe $k \in [1, s_j]$. Following the assignment of the nodes by NA algorithm, for each subframe $k \in [1, s_j]$, the CA algorithm determines the node subsets in L_{jk} that concurrently transmit. The reason for assigning the nodes in the increasing packet generation period and considering only the nodes with the same packet generation period for concurrent transmission is to keep the time difference between consecutive time slot allocations fixed at their packet generation period. The NA algorithm uses PCRA algorithm to determine the transmission time of a sensor node whereas the CA algorithm uses PCRA algorithm to determine the feasibility and if so concurrent transmission time of a subset of nodes. These subsets are then periodically allocated every s_j subframes to meet the periodicity requirement. Next, we propose two algorithms of different complexities for NA and CA.

4.4.1 Node Assignment (NA) Algorithms

In this section, we propose two NA algorithms that assign the nodes to the subframes considering different priority ordering.

4.4.1.1 Greedy Node Assignment (GNA) Algorithm

GNA algorithm greedily assigns each sensor node $i \in L_j$ to the subframe $k \in [1, s_j]$ with the minimum total active length in random order. The complexity of the GNA algorithm is $O(|L_j|f(1)+|L_j|M)$ where $f(N)$ denotes the complexity of the PCRA algorithm as a function of the number of nodes N on which it is performed: each sensor node $i \in L_j$ whose time slot length is determined using PCRA algorithm is assigned to the subframe with the minimum total active length among $s_j \leq M$ subframes. The complexity of the PCRA algorithm proposed in Section 4.5 is stated in Lemma 13.

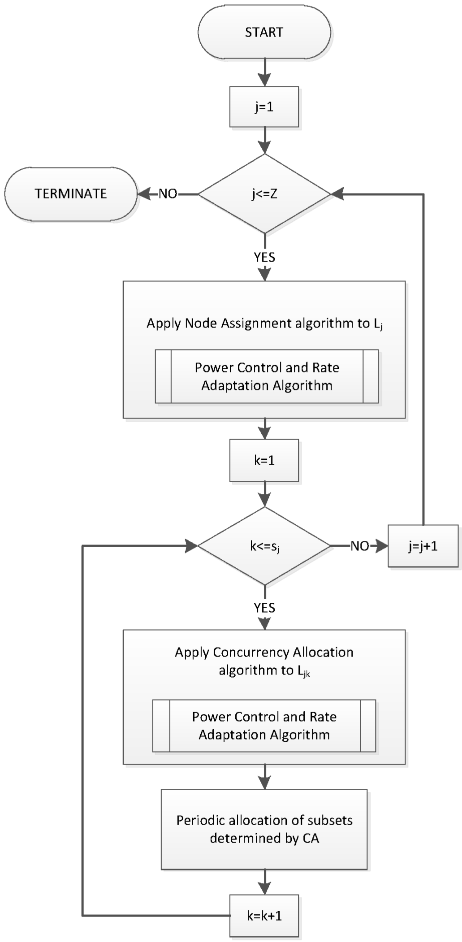

Figure 4.4: General structure of the scheduling algorithm.

4.4.1.2 Sorted Node Assignment (SNA) Algorithm

SNA algorithm assigns each node to the subframe $k \in [1, s_j]$ with the minimum total active length giving priority to the nodes with higher transmission times and considering the nodes connected to different controllers separately. For each controller c in the WNCS, each sensor node $i \in L_j$ transmitting to controller c with maximum transmission time is assigned to the subframe with minimum total active length. The ordering of transmission times used in SNA has $O(|L_j|log(|L_j|))$ complexity to be added to the complexity required by GNA which uses random ordering. Hence, the complexity of the

SNA algorithm is $O(|L_j|log(|L_j|) + |L_j|f(1) + |L_j|M)$.

4.4.2 Concurrency Allocation (CA) Algorithms

In this section, we propose two CA algorithms of different complexities to determine the subset of nodes for concurrent transmission.

4.4.2.1 Minimum Length Allocation (MLA) Algorithm

MLA algorithm, illustrated as Algorithm 6, consists of two phases. In Phase-I, Fast Tree Enumeration (FTE), all feasible subsets of the set L_{jk} are enumerated (Lines $1-16$). In Phase-II, Minimum Price Covering (MPC), a heuristic algorithm is proposed with the goal of selecting some of these feasible subsets of L_{jk} such that every node in L_{jk} is covered at least once and the total time allocated to L_{jk} is minimized (Lines $17-24$).

FTE is described in detail as follows. Each node of the tree corresponds to a subset of nodes and is characterized by a vector $\mathbf{p}$ whose j-th element is equal to 1 if sensor node j is included in the set and 0 otherwise. $\mathbf{F}$ and $\mathbf{TF}$ denote the set of feasible subsets of nodes and their transmission times respectively and are initialized to empty set (Lines $2-3$). $\mathbf{P}^+$ denotes the set of nodes that will be evaluated in the next level of the tree and initially contains only the root node $(0, 0, ..., 0)$ (Line 4). At each level i of the tree, each node $\mathbf{p} \in \mathbf{P}^+$ is branched into two nodes by setting $\mathbf{p}(i)$ to 1 and 0 (Lines $5-9$). However, since the latter is actually equivalent to $\mathbf{p}$, it is not evaluated in the current level. Then, the feasibility of the subset of nodes generated by setting $\mathbf{p}(i)$ to 1 is checked by using PCRA algorithm (Line 10). If it is feasible, it is included in the set $\mathbf{P}^+$ to be branched in the next level of the tree and in the set $\mathbf{F}$ with its transmission time included in the set $\mathbf{TF}$ (Lines $11-13$). The complexity of the FTE phase of the MLA algorithm is $O(2^{|L_{jk}|}f(|L_{jk}|))$: In the worst case all possible subsets of the nodes in L_{jk} are generated and PCRA algorithm is run on each subset.

MPC is described as follows. $\mathbf{S}$ and $\mathbf{A}$ denote the set of selected subsets from the set $\mathbf{F}$ and the set of included sensor nodes in the subsets in $\mathbf{S}$ respectively, and are initialized to empty sets (Lines $18-19$). Let $\alpha_{\mathbf{f}}$ be the

Algorithm 6 Minimum Length Allocation (MLA) Algorithm

PHASE-I: Fast Tree Enumeration (FTE)
$\mathbf{F} = \emptyset$;
$\mathbf{TF} = \emptyset$;
$\mathbf{P}^+ = \{(0, 0, ..., 0)\}$;
for $i = 1 : |L_{jk}|$ **do**
  $l_{\mathbf{P}^+} = |\mathbf{P}^+|$;
  for $j = 1 : l_{\mathbf{P}^+}$ **do**
    $\mathbf{p} = \mathbf{P}^+(j)$;
    set $\mathbf{p}(i) = 1$;
    if isFeasiblePCRA(p) then
      $\mathbf{P}^+ = \mathbf{P}^+ + \{\mathbf{p}\}$;
      $\mathbf{F} = \mathbf{F} + \{\mathbf{p}\}$;
      $\mathbf{TF} = \mathbf{TF} + \{\mathbf{PCRA(p)}\}$;
    end if
  end for
end for

PHASE-II: Minimum Price Covering (MPC)
$\mathbf{S} = \emptyset$;
$\mathbf{A} = \emptyset$;
while $\mathbf{A} \neq L_{jk}$ **do**
  $\mathbf{s} = \min_{\mathbf{f} \in \mathbf{F}} \frac{\mathbf{TF(f)}}{|\mathbf{f} \setminus \mathbf{A}|}$;
  $\mathbf{A} = \mathbf{A} \cup \mathbf{s}$;
  $\mathbf{S} = \mathbf{S} + \{\mathbf{s}\}$;
end while

price of a subset $\mathbf{f}$ equal to the ratio of its time slot length $\mathbf{TF(f)}$ to the number of nodes that are included in that subset but not included in the set $\mathbf{A}$, denoted by $|\mathbf{f} \setminus \mathbf{A}|$; i.e., $\alpha_{\mathbf{f}} = \frac{\mathbf{TF(f)}}{|\mathbf{f} \setminus \mathbf{A}|}$. In each iteration, the subset $\mathbf{f}$ in $\mathbf{F}$ with minimum $\alpha_{\mathbf{f}}$ is included in $\mathbf{S}$ with the corresponding additional sensor nodes included in $\mathbf{A}$ (Lines $21 - 23$). This continues until $\mathbf{A}$ includes all sensor nodes in L_{jk} (Line 20). MPC is equivalent to the Greedy Set Cover Algorithm (GSCA) proposed for the Weighted Set Cover Problem (WSCP) [74]. Given a collection P of subsets whose union is a set U; i.e., $U \doteq \cup_{p \in P} p$, where each subset $p \in P$ has a specified non-negative weight w_p, GSCA determines a set cover $Q \subseteq P$ whose union is U by repeatedly choosing a set p that minimizes the weight w_p divided by the number of elements in p not yet covered by chosen sets. The complexity of GSCA is $O(|U|^2|P|)$ since the algorithm picks at least one element of set U in each iteration and determining the subset p minimizing the weight w_p divided by the number of elements in p not yet covered by chosen sets among $|P|$ number of subsets has $O(|U||P|)$ complexity. Therefore, the complexity of the MPC phase of the MLA algorithm is $O(|L_{jk}|^2 2^{|L_{jk}|})$ since there are $2^{|L_{jk}|}$ feasible subsets of the nodes in L_{jk} in the worst case. Moreover, GSCA is an α-approximation algorithm for WSCP, where $\alpha = \sum_{i=1}^{n} 1/n$ and n is the size of the largest subset in P [74]. Hence, MPC is also an α-approximation algorithm where $\alpha = \sum_{i=1}^{n} 1/n$ and n is the size of the largest subset in $\mathbf{F}$.

The overall complexity of the MLA algorithm is then $O(2^{|L_{jk}|} f(|L_{jk}|) + 2^{|L_{jk}|}|L_{jk}|^2)$.

4.4.2.2 Maximum Utility Allocation (MUA) Algorithm

MUA algorithm, illustrated as Algorithm 7, is described as follows. $\mathbf{D}$ and $\mathbf{S}$ denote the set of sensor nodes considered for concurrent transmissions but not allocated yet and the set of subsets of sensor nodes allocated for concurrent transmissions respectively. $\mathbf{D}$ and $\mathbf{S}$ are initialized to L_{jk} and $\emptyset$ respectively (Lines $1-2$). In each iteration of the algorithm, one subset $\mathbf{G}$ of the sensor nodes that have not been allocated is determined and included in the set $\mathbf{S}$ (Lines $3-18$). First, an arbitrary node that has not been allocated

$i \in \mathbf{D}$ is included in $\mathbf{G}$ (Lines $4-7$). Then, the nodes that maximize utility value are included in $\mathbf{G}$ one by one unless there exists no sensor node to be allocated in $\mathbf{D}$ or no sensor node improves the utility value by its addition to $\mathbf{G}$ (Lines $8-16$). The utility of a set $\mathbf{G}$ is defined as the decrease in the transmission time of a set of sensor nodes when they transmit concurrently and formulated as

$$u(\mathbf{G}) = \sum_{n \in \mathbf{G}} t(n) - t(\mathbf{G}) \tag{4.2}$$

where $t(\mathbf{G})$ is the minimum time slot length required for the concurrent transmission of the nodes in the set $\mathbf{G}$ and determined by the PCRA algorithm. Algorithm terminates when all sensor nodes are allocated to a subset $\mathbf{G}$ (Line 3).

The complexity of the MUA algorithm is $O(|L_{jk}|^2 f(|L_{jk}|))$.

Algorithm 7 Maximum Utility Allocation (MUA) Algorithm

1: $\mathbf{D} = L_{jk}$;
2: $\mathbf{S} = \emptyset$;
3: **while** $\mathbf{D} \neq \emptyset$ **do**
4:     $\mathbf{G} = \emptyset$;
5:     pick an arbitrary node $i \in \mathbf{D}$;
6:     $\mathbf{G} = \mathbf{G} + \{i\}$;
7:     $\mathbf{D} = \mathbf{D} - \{i\}$;
8:     **while** $\mathbf{D} \neq \emptyset$ **do**
9:         $j = \text{argmax}_{i \in \mathbf{D}} u(\mathbf{G} + \{i\})$;
10:         **if** $u(\mathbf{G} + \{j\}) > u(\mathbf{G})$ **then**
11:             $\mathbf{G} = \mathbf{G} + \{j\}$;
12:             $\mathbf{D} = \mathbf{D} - \{j\}$;
13:         **else**
14:             **break**;
15:         **end if**
16:     **end while**
17:     $\mathbf{S} = \mathbf{S} + \{\mathbf{G}\}$;
18: **end while**

4.4.3 Heuristic Scheduling Algorithms

Combining each NA algorithm with each CA algorithm, we create four heuristic scheduling algorithms. The GNA-MLA scheduling algorithm is created using GNA algorithm for the NA and MLA algorithm for the CA. Similarly, GNA-MUA, SNA-MLA and SNA-MUA scheduling algorithms are created. Based on the complexities of the corresponding NA and CA algorithms, the worst case complexity of GNA-MLA and SNA-MLA algorithms is $O(ZM2^L(L^2 + f(L)))$ whereas the complexity of SNA-MUA and GNA-MUA is $O(ZML^2 f(L))$.

GNA-MUA and SNA-MUA algorithms resemble the Maximum Utility based Concurrency Allowance (MUCA) scheduling algorithm proposed in [63] with the main improvements listed as follows. Instead of exploiting concurrent transmissions of the sensor nodes after all the sensor nodes transmitting at different packet generation periods are allocated to the subframes, both GNA-MUA and SNA-MUA algorithms exploit concurrent transmissions among the nodes with the same packet generation period just after their allocation to the subframes. Although the NA algorithm of GNA-MUA is the same as that of the MUCA algorithm, the NA algorithm of SNA-MUA considers the nodes connected to different controllers separately and gives priority to the nodes with higher transmission times. The sensor nodes transmitting to the same controller are assigned to different subframes to increase the possibility of finding more node subsets for concurrency allocation. In addition, the nodes with higher transmission times are given higher priority to obtain smaller total active lengths in the subframes to which the sensor nodes are assigned.

4.5 Power Control and Rate Adaptation Problem

4.5.1 Mathematical Formulation

The power control and rate adaptation problem aims to minimize the time required for the concurrent transmission of a set of sensor nodes while satisfying their requirements in terms of transmission delay and energy consumption.

Given the power allocation of a subset of the links, the maximum achievable transmission rate of a link given by the upper bound in Eqn. (4.1) formulated as a function of the transmission powers of the links in the set is optimal for the following reasons: First, the transmission rate of a link is independent of that of the other links given their power allocation. Moreover, the higher the transmission rate of a link, the smaller its transmission delay, its energy consumption and the time slot length.

The optimal power control and rate adaptation problem is therefore formulated as

minimize

$$t \tag{4.3a}$$

subject to

$$t_l \leq t, \quad l \in S \tag{4.3b}$$

$$t_l \leq d_l, \quad l \in S \tag{4.3c}$$

$$t_l p_l \leq e_l, \quad l \in S \tag{4.3d}$$

$$p_l \leq p_{max}, \quad l \in S \tag{4.3e}$$

$$t_l = \frac{R_l}{W \times \log\left(1 + \frac{1}{W} \times \frac{p_l g_{ll}}{N_0 + \sum_{k \in S, k \neq l} p_k g_{kl}}\right)}, \quad l \in S \tag{4.3f}$$

variables

$$t \geq 0,\ p_l \geq 0,\ l \in S \tag{4.3g}$$

where t is the time slot length required for the concurrent transmission of the links in the set S, t_l is the transmission time of link l and R_l is the length of the packet of node l.

The objective of the optimization problem is minimizing the length of the time slot. Eqn. (4.3b) states that the time slot length is equal to the maximum of the transmission time of the links in the set S. Eqns. (4.3c) and (4.3d) represent the delay and energy consumption requirement of the transmission of the links respectively. Eqn. (4.3e) states that the transmit power of link l cannot exceed maximum allowed transmit power. Eqn. (4.3f) provides the relation between the transmission time of a link and the transmission power of the concurrently transmitting links.

This optimization problem is a non-convex problem due to the non-convexity of the energy consumption function, given in Eqn. (4.3d), in the transmission power of the nodes hence cannot be solved for global optimum by the algorithms proposed for the convex optimization problems [54]. Next, we first present the optimality conditions on the variables of the problem and then propose an optimal algorithm exploiting these optimality conditions.

4.5.2 Optimality Conditions

Theorem 3. *Let us enumerate the links in the set S from 1 to $|S|$. Let $\boldsymbol{t} = [t_1, t_2, ..., t_{|S|}]$ be a vector containing the transmission time of the links in the set S. Suppose that $\boldsymbol{p}^{\min} = [p_1, p_2, ..., p_{|S|}]$ is the component-wise minimum power vector corresponding to the transmission time vector $\boldsymbol{t}$. If the energy and maximum power constraints given by Eqns. (4.3d) and (4.3e) respectively cannot be satisfied by $\boldsymbol{p}^{\min}$, no power vector yielding the transmission time vector $\boldsymbol{t}$ satisfies these constraints.*

Proof: Consider an arbitrary power vector $\mathbf{p}^a = [p_1^a, p_2^a, ..., p_{|S|}^a]$ yielding the transmission time vector $\mathbf{t}$. Suppose that $\mathbf{p}^{\min}$ violates the energy constraint given by Eqn. (4.3d). Then, there exists at least one link l for which the energy consumption is greater than the energy requirement; i.e., $t_l p_l > e_l$. Since $\mathbf{p}^{\min}$ is the minimum power vector, $p_l^a \geq p_l$, therefore, $t_l p_l^a \geq t_l p_l$, which is greater than the energy requirement e_l. Hence, $\mathbf{p}^a$ cannot satisfy the energy constraint.

Similarly, suppose that $\mathbf{p}^{\min}$ violates the maximum power constraint given by Eqn. (4.3e). Then, there exists at least one link l for which the maximum power constraint is violated; i.e., $p_l > p_{max}$. Since $\mathbf{p}^{\min}$ is the minimum power vector, $p_l^a \geq p_l$, which is greater than p_{max}. Hence, $\mathbf{p}^a$ cannot satisfy the maximum power constraint. □

Lemma 12. *Let us assume the transmission time vector* $\boldsymbol{t}$ *satisfies the delay requirement given by Eqn. (4.3c). If and only if there exists a transmit power vector corresponding to* $\boldsymbol{t}$*, which is determined by testing Perron-Frobenius conditions [75], and the resulting component-wise minimum power vector* $\boldsymbol{p}^{\min}$ *satisfies the energy and maximum power constraints given by Eqns. (4.3d) and (4.3e) respectively,* $\boldsymbol{t}$ *is feasible.*

Proof: The feasibility of the transmission time vector $\mathbf{t}$ first requires the existence of a transmit power vector corresponding to $\mathbf{t}$. Since the transmission time of each link in the set S is predetermined by $\mathbf{t}$, this requires that the SINR constraint of each link $l \in S$ as given by

$$\frac{p_l g_{ll}}{N_0 + \sum_{k \neq l} p_k g_{kl}} \geq \beta_l \tag{4.4}$$

where

$$\beta_l = W\left(2^{R_l/(Wt_l)} - 1\right) \tag{4.5}$$

is satisfied. The existence of such transmit power vector is determined by testing Perron-Frobenius conditions [75] described as follows: Let $\mathbf{B}_S$ be $|S| \times |S|$ relative channel gain matrix such that the element in the l-th row and k-th column of B_S takes value g_{kl}/g_{ll} for $l \neq k$ and 0 for $l = k$. Let $\mathbf{D}_S$ be $|S| \times |S|$ diagonal matrix with the l-th diagonal entry equal to β_l. Let $\mathbf{V}_S$ be $|S| \times 1$ normalized noise power vector with the l-th element equal to $\beta_l N_0/g_{ll}$. Then, the SNIR requirements given by Eqn. (4.4) can be written in matrix form as

$$(\mathbf{I} - \mathbf{D}_S\mathbf{B}_S)\mathbf{p} \succeq \mathbf{V}_S \tag{4.6}$$

Perron-Frobenius conditions state that there exists a power vector $\mathbf{p}$ yielding the transmission time vector $\mathbf{t}$ if and only if the largest real eigenvalue of $\mathbf{D}_S\mathbf{B}_S$ is less than 1. Then the component-wise minimum power vector $\mathbf{p}^{\min}$ is given as

$$\mathbf{p}^{\min} = (\mathbf{I} - \mathbf{D}_S\mathbf{B}_S)^{-1}\mathbf{V}_S \tag{4.7}$$

Furthermore, if $\mathbf{p}^{\min}$ satisfies the energy and maximum power constraints given by Eqns. (4.3d) and (4.3e) respectively; i.e., $p_l^{\min} \leq p_{max}$, and $t_l p_l^{\min} \leq e_l$ for all $l \in S$, $\mathbf{t}$ is feasible. Otherwise, it is infeasible based on Theorem 3. $\square$

Theorem 4. *Let* $\boldsymbol{t}^* = [t_1^*, t_2^*, ..., t_{|S|}^*]$ *be the optimal transmission time vector for the optimization problem (4.3) with the time slot length* $t^* = \max_{l \in S} t_l^*$*. Then, for the links with delay requirement less than the time slot length, i.e.* $d_j \leq t^*$*, their optimal transmission time is equal to their delay requirement; i.e.,* $t_j^* = d_j$*. For the remaining links with delay requirement greater than the time slot length, i.e.* $d_k > t^*$*, their optimal transmission time is equal to the time slot length; i.e.* $t_k^* = t^*$*.*

Proof: We prove this theorem by contradiction. First, suppose that there exists a link j with delay requirement less than the time slot length; i.e., $d_j \leq t^*$, and optimal transmission time less than its delay requirement; i.e., $t_j^* < d_j$. If we decrease the transmit power of link j while keeping the transmit powers of the remaining links in the set S same such that its transmission time t_j^* becomes equal to d_j, the energy consumption constraint of link j is still satisfied since the energy consumption decreases as the transmit power decreases. Moreover, the transmission time of each link $l \neq j$ decreases due to the decreasing interference created by link l, which decreases the time slot length $t^* = \max_{i \in S} t_i^*$. This is a contradiction.

Second, suppose that there exists a link k with delay requirement greater than the time slot length; i.e., $d_k > t^*$, and optimal transmission time less than the time slot length; i.e. $t_k^* < t^*$. If we decrease the transmit power of link k while keeping the transmit powers of the remaining links in the set S same, the transmission time of link k increases whereas the transmission time of each link $l \neq k$ decreases due to the decreasing interference. Moreover, the energy consumption of all the links in the set S decreases while preserving the feasibility of the solution. Hence, we can decrease the time slot length by decreasing the transmit power of link k. This is a contradiction. $\square$

Theorem 5. *Let* $\boldsymbol{t}^1 = [t_1^1, t_2^1, ..., t_{|S|}^1]$ *and* $\boldsymbol{t}^2 = [t_1^2, t_2^2, ..., t_{|S|}^2]$ *be two arbitrary transmission time vectors satisfying the delay requirements of the links in the set* S *such that* $t_l^1 \geq t_l^2$ *for all* $l \in S$*. If* $\boldsymbol{t}^1$ *is infeasible, then* $\boldsymbol{t}^2$ *is infeasible.*

Proof: Suppose that $\mathbf{p}^1 = [p_1^1, p_2^1, ..., p_{|S|}^1]$ and $\mathbf{p}^2 = [p_1^2, p_2^2, ..., p_{|S|}^2]$ are the component-wise minimum power vectors corresponding to $\mathbf{t}^1$ and $\mathbf{t}^2$ respectively. First, we will prove that $p_l^1 \leq p_l^2$ for all $l \in S$. Let β_l^1 and β_l^2 for a link $l \in S$ be the SNIR levels corresponding to t_l^1 and t_l^2 obtained by using Eqn. (4.5) respectively. Since $t_l^1 \geq t_l^2$ for all $l \in S$, $\beta_l^1 \leq \beta_l^2$ for all $l \in S$. Then, any feasible power vector satisfying SNIR requirements represented by Eqn. (4.6) for SNIR levels $\{\beta_l^2, l \in S\}$ also satisfies them for SNIR levels $\{\beta_l^1, l \in S\}$. Hence, $\mathbf{p}^2$ yielding $\mathbf{t}^2$ satisfies the SNIR requirements for SNIR levels $\{\beta_l^1, l \in S\}$. Suppose that for an arbitrary link k, $p_k^1 > p_k^2$. Then, $\mathbf{p}^1$ cannot be the component-wise minimum power vector satisfying the SNIR requirements for SNIR levels $\{\beta_l^1, l \in S\}$. This is a contradiction.

If $\mathbf{p}^1$ is infeasible due to the maximum power constraint, then there exists at least one link k such that $p_k^1 > p_{max}$. Since $p_k^2 \geq p_k^1$, $p_k^2 > p_{max}$. Hence, $\mathbf{p}^2$ is infeasible due to the maximum power constraint.

If $\mathbf{p}^1$ is infeasible due to the energy constraint, then there exists at least one link j for which the energy consumption is greater than the energy requirement; i.e., $t_j^1 p_j^1 > e_j$. Let α denote the ratio of p_j^2 to p_j^1. Then the ratio of the resulting transmission times t_j^1 to t_j^2 is less than α since increasing the transmission power by a factor of α increases the transmission rate by a factor less than α due to the logarithm in the transmission rate formulation given by Eqn. (4.1) and increasing interference by the increasing power of the remaining links. Hence, $t_j^2 p_j^2 > \frac{t_j^1}{\alpha} \alpha p_j^1 = t_j^1 p_j^1 > e_j$. $\mathbf{p}^2$ is infeasible due to the energy constraint. $\square$

4.5.3 Optimal Polynomial-time Algorithm

Based on the findings stated by Theorems (3)-(5) and Lemma (12), we propose the optimal Power Control and Rate Adaptation (PCRA) Algorithm as follows. Transmission time vector $\mathbf{t}^*$ is initialized such that $\mathbf{t}_l^* = d_1$ for each link $l \in S$ (Line 1). At each iteration of the algorithm, the feasibility of the transmission time vector $\mathbf{t}^*$ is determined by first testing the existence of a transmit power vector corresponding to $\mathbf{t}^*$ using Perron-Frobenius conditions [75], and then checking whether the resulting component-wise minimum power vector $\mathbf{p}^{\min}$ satisfies the energy and maximum power constraints given by Eqns. (4.3d) and (4.3e) respectively based on Lemma 12. If $\mathbf{t}^*$ is not feasible, then the transmission time of each link $l \in [i+1, |S|]$ is set to

Algorithm 8 Power Control and Rate Adaptation (PCRA) Algorithm

Input: Links sorted in increasing delay requirements $d_1 \leq d_2 \leq ... \leq d_{|S|}$;
Output: Optimal transmission time vector $\mathbf{t}^* = [t_1^*, t_2^*, ..., t_{|S|}^*]$ if S is feasible

$\mathbf{t}^* = [d_1, d_1, ..., d_1]$;
$i = 1$;
while ($\mathbf{t}^*$ is not feasible) and ($i < |S|$) **do**
  $\mathbf{t}^*_{\{i+1:|S|\}} = d_{i+1}$;
  $i = i + 1$;
end while
if ($\mathbf{t}^*$ is not feasible) and ($i = |S|$) **then**
  no feasible solution exists;
else
  $t^*_{up} = d_i$;
  $t^*_{lw} = d_{i-1}$;
  while $\frac{t^*_{up} - t^*_{lw}}{t^*_{lw}} > \epsilon$ **do**
    $t^* = \frac{t^*_{up} + t^*_{lw}}{2}$;
    $\mathbf{t}^*_{\{i:|S|\}} = t^*$;
    if $\mathbf{t}^*$ is not feasible **then**
      $t^*_{lw} = t^*$;
    else
      $t^*_{up} = t^*$;
    end if
  end while
  $\mathbf{t}^*_{\{i:|S|\}} = t^*_{up}$;
end if

d_{i+1} to test the feasibility of the time slot length d_{i+1} while satisfying the delay constraints of all the links in the following iteration (Lines $4-5$). In each iteration of the algorithm, the delay constraint of the links $l \leq i$ are not updated since their transmission time are equal to their delay constraint based on Theorem 4. Also, once $\mathbf{t}^*$ is determined to be infeasible, only the transmission time vectors larger than $\mathbf{t}^*$ are checked based on Theorem 5. This iterative mechanism continues until $\mathbf{t}^*$ becomes feasible or all the delay constraints have been evaluated for the time slot length (Line 3). If the transmission time vector $\mathbf{t}^*$ in which $\mathbf{t}_l^* = d_l$ for all links $l \in S$ is infeasible, then no feasible solution exists and the algorithm terminates (Lines $7-9$). Otherwise, since the time slot length d_i is feasible, minimum feasible $t^* \in (d_{i-1}, d_i)$ is determined by using a binary search mechanism (Lines $10-21$). The search interval is initialized by $(t_{lw}^*, t_{up}^*) = (d_{i-1}, d_i)$, where t_{lw}^* and t_{up}^* correspond to the currently determined maximum infeasible and minimum feasible time slot lengths respectively. In each iteration of the search, the feasibility of the middle point of the interval is checked updating the value of t_{lw}^* or t_{up}^* by this middle point thus decreasing the search interval by half. When the search interval is small enough, which is determined by the value ϵ, $\mathbf{t}^*$ is updated with the minimum feasible value t_{up}^*.

Lemma 13. *For a fixed predetermined relative error bound ϵ, the complexity of PCRA is $O(|S|^4 + K|S|^3)$ where $K = \max_{i=1:|S|} \lceil \log \frac{d_i - d_{i-1}}{\epsilon d_{i-1}} \rceil$.*

Proof: PCRA algorithm checks the feasibility of the transmission time vector at most $|S|$ times in Lines $(3-9)$. In addition, in order to decrease the ratio $\frac{t_{up}-t_{lw}}{t_{lw}}$ below ϵ (Line 12), PCRA requires $K = \max_{i=1:|S|} \lceil \log \frac{d_i - d_{i-1}}{\epsilon d_{i-1}} \rceil$ iterations since the ratio is reduced at least by a factor of 2 in each iteration by setting either t_{lw} or t_{up} to the middle point of the interval $[t_{lw}, t_{up}]$ (Lines 13-19). In each iteration, PCRA checks the feasibility of the transmission time vector by evaluating the Perron-Frobenius conditions of complexity $O(|S|^3)$. Hence, the complexity of PCRA is $O(|S|^4 + K|S|^3)$. $\square$

Theorem 6. *PCRA determines optimal time slot length within a relative error bound ϵ.*

Proof: Considering the optimality conditions stated in Theorem 4, in Lines (3-6) of Algorithm 8, PCRA determines the nodes $j \in S' \subseteq S$ for which the transmission time t_j is equal to transmission delay requirement d_j; i.e., $t_j = d_j$. The transmission times of the rest of nodes $j \in S \setminus S'$ are equal to each other and less than or equal to d_i; i.e., $t_j \leq d_i$. In Lines (10-11), the range for the optimal time slot length is determined such that the lower and upper bounds are set to the currently determined maximum infeasible and minimum feasible value of the time slot lengths respectively. If a particular t^* value is infeasible, then all $t < t^*$ are infeasible and if a particular t^* value is feasible, then all $t > t^*$ are feasible due to Theorem 5. Based on this finding, algorithm iteratively shrinks the range for the optimal t^* in Lines (12-20) until the selected t^* falls into the desired neighborhood of the optimal value, which is specified by the relative error bound value ϵ. $\square$

4.6 Performance Evaluation

The goal of this section is to evaluate the performance of the proposed power control and scheduling algorithms compared to the optimal and previously proposed scheduling algorithms in terms of maximum total active length; maximum aperiodic packet delay, which is defined as the worst-case delay that an aperiodic packet experiences from the packet generation until the transmission in the unallocated part of the schedule; and average runtime.

Simulation results are obtained based on 100 independent random network topologies, where the sensor nodes and the controllers are uniformly distributed within a square area such that the side of the square is adjusted for network density 5 nodes/m^2 unless otherwise stated, in MATLAB. We assume that the sensor nodes communicate to the nearest controller. The packet generation periods and packet lengths of the sensor nodes are uniformly chosen from the set $\{1, 2, 4, 8\}$ ms and the interval $[50, 100]$ bits, respectively.

The attenuations of the links are determined considering both large and small scale statistics. The attenuation of the links considering the large scale

Table 4.1: Simulation Parameters

$PL(d_0)$	70 dB	p_{max}	250 mW
α	3.5	σ_z	4 dB
N_0	10^{-8} W/Hz	W	100 MHz

statistics is modeled as

$$PL(d) = PL(d_0) + 10\alpha \log(d/d_0) + Z \tag{4.8}$$

where d is the distance between the sensor node and the controller, $PL(d)$ is the path loss at distance d in decibels, $PL(d_0)$ is the path loss at reference distance $d_0 = 1$ m, α is the path loss exponent [36] and Z is a Gaussian random variable with zero mean and standard deviation σ_z [56]. The small-scale fading has been modeled by using Rayleigh fading with scale parameter Ω set to the mean power level determined by using Eqn. (4.8) [56, 57]. The simulation parameters are listed in Table 4.1.

4.6.1 Maximum Total Active Length Performance

Fig. 4.5 shows the maximum total active length of the subframes of different scheduling algorithms including EDF, LLF, MUCA [63], OPT denoting the optimal solution, the proposed scheduling algorithms GNA-MLA, GNA-MUA, SNA-MLA and SNA-MUA for different number of nodes transmitting to one controller. OPT is obtained by enumerating all feasible solutions of the optimization problem given in Section 4.3 and determining the one with the minimum value of its objective function defined as maximum total active length of the subframes in a frame. This one controller scenario is analyzed separately since EDF and LLF scheduling algorithms do not support concurrent transmissions so can only be included for comparison in the simulation of the transmission to one controller. Since the concurrency allocation mechanisms of the proposed algorithms are not used due to the existence of only one controller with no multi-reception capability, the proposed algorithms are different only in the node assignment mechanism. The performances of the proposed algorithms and the MUCA algorithm are very similar to each

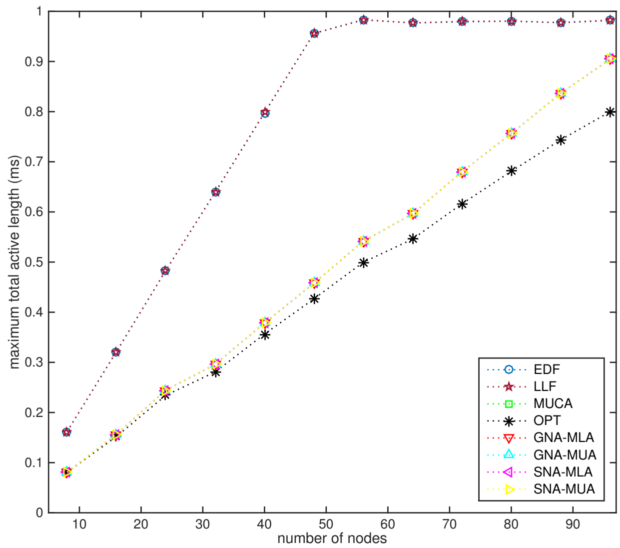

Figure 4.5: Maximum total active length of the subframes of the scheduling algorithms for different number of nodes transmitting to one controller.

other for one controller case and all outperform the EDF and LLF scheduling algorithms that assign the nodes to the nearest available time slots based on their priorities without considering any adaptivity metric. The proposed algorithms also perform very close to the optimal solution with an approximation ratio below 1.13 for the networks containing up to 100 nodes, where the approximation ratio is defined as the ratio of the maximum total active length of the scheduling algorithm to that of the optimal solution.

Fig. 4.6 shows the approximation ratio of the proposed scheduling algorithms and the MUCA algorithm for different number of nodes and number of controllers. Note that the number used next to the algorithm name in the legend denotes the number of controllers. As opposed to the one controller case in which the concurrency allocation algorithm cannot be used, the proposed algorithms have superior performance compared to the MUCA algorithm. The proposed scheduling algorithms employing the MLA concurrency allocation outperform the ones using MUA algorithm. Considering the node assignment algorithms, the proposed algorithms using SNA perform

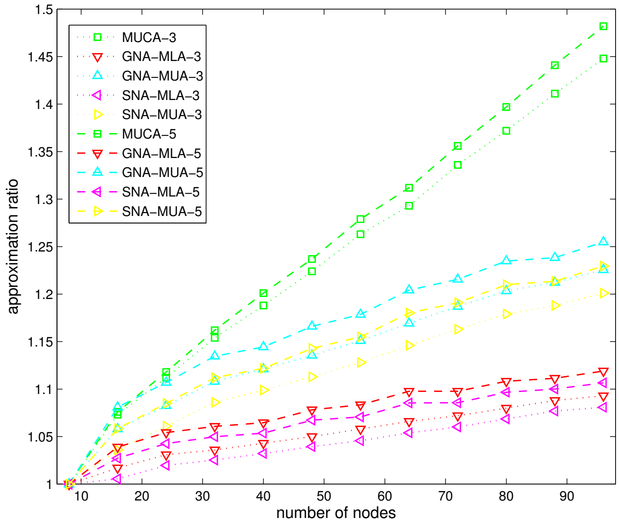

Figure 4.6: Approximation ratio of the scheduling algorithms for different number of nodes and different number of controllers.

slightly better than those adopting GNA algorithm. Moreover, changing the concurrency allocation algorithm from MUA to MLA creates a higher performance increase than changing the node assignment algorithm from GNA to SNA. For example, GNA-MLA outperforms SNA-MUA which performs better than GNA-MUA. For all the proposed algorithms, as the number of controllers increases, the approximation ratio performance degrades only slightly. As the number of controllers increases, the maximum number of sensors that can be allocated in a particular time slot increases since sensor nodes transmitting to the same controller cannot be allocated concurrently. However, the harmful mutual interference among the nodes limits the number of sensors that should actually concurrently transmit.

Fig. 4.7 shows the approximation ratio of the scheduling algorithms for different path-loss exponent values and different number of controllers in a network of 100 nodes. Similar to Fig. 4.6, the proposed scheduling algorithms outperform the MUCA algorithm and the performance of the proposed scheduling algorithms employing MLA and SNA are better than those using MUA and GNA respectively. Increasing the number of controllers in

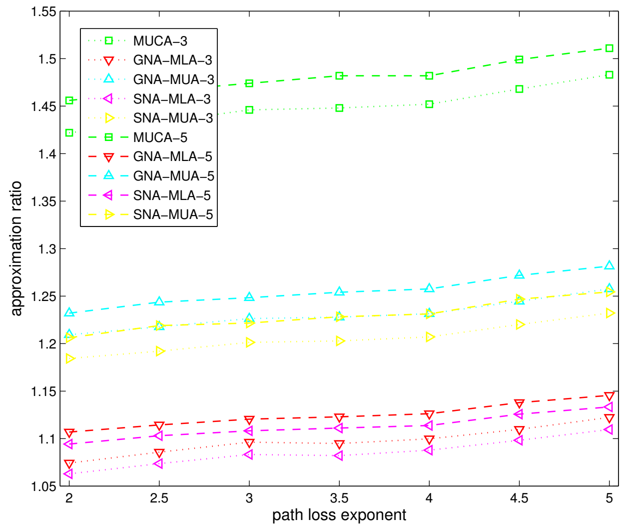

Figure 4.7: Approximation ratio of the scheduling algorithms for different path-loss exponents and different number of controllers in a network of 100 nodes.

the network does not affect the approximation ratio of the algorithms considerably. Moreover, the approximation ratios of the algorithms are robust to the increasing path loss exponent value with slight increase in the approximation ratio as the path loss exponent increases due to the increase in the variance of link attenuations.

Fig. 4.8 shows the approximation ratio of the scheduling algorithms for different network densities and different number of controllers in a network of 100 nodes. The figure validates the superiority of the proposed algorithms over the MUCA algorithm and the performance order among the proposed algorithms illustrated by the foregoing results. Furthermore, the figure reveals that the scheduling algorithms presented perform worst at a critical network density value around 3 nodes/m^2 and the performance improves as the network density increases and decreases from this point. At low and high network densities, the performance of all the algorithms get closer to each other and the optimal solution due to the diminishing effect of the concurrency allocation mechanisms in the performance of the algorithms: At low network density, all the nodes in the network tend to transmit concur-

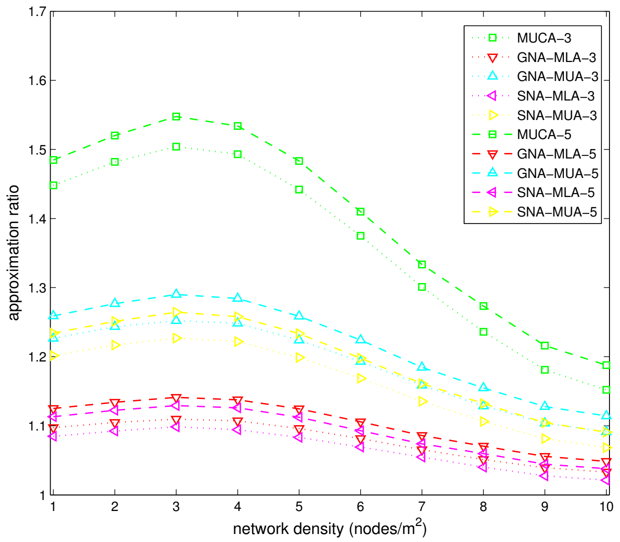

Figure 4.8: Approximation ratio of the scheduling algorithms for different network densities and different number of controllers in a network of 100 nodes.

rently since the interference among the links is negligible. At high density networks on the other hand, since the nodes are located very close to each other, most of them generate excessive interference to each other. Therefore, the number of sets of nodes that are distributed sufficiently far from each other so decrease their total transmission time by concurrent transmission is limited. The decrease in the number of such sets reduces the performance gap between the optimal solution and the proposed algorithms. At the extreme high density case, all the nodes are located very close to each other without any gain from concurrent transmission eliminating the advantage of any concurrency allocation algorithm.

4.6.2 Maximum Aperiodic Packet Delay Performance

Fig. 4.9 shows the adaptivity performance of the scheduling algorithms for different number of nodes transmitting to one controller in terms of the maximum aperiodic packet delay. Since the simulation is performed for one controller case, the concurrent transmission of the sensor nodes in the network

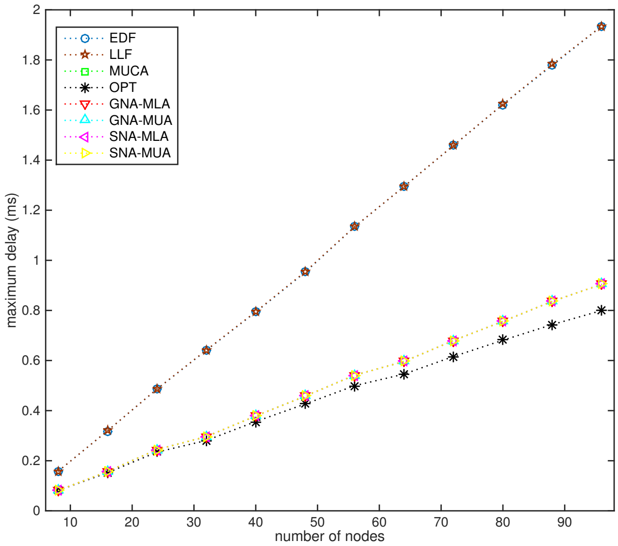

Figure 4.9: Maximum aperiodic packet delay for different scheduling algorithms and different number of nodes transmitting to one controller.

is not possible. Hence, similar to Fig. 4.5, the performance of the proposed scheduling algorithms and MUCA algorithm are very close to each other. They also all outperform EDF and LLF scheduling algorithms. EDF and LLF scheduling algorithms allocate the data packets of the sensor nodes as soon as they become available. Since the total active length of some of the subframes may be very large, no space exists for the allocation of additional packets. As a result, as the number of nodes increases, the allocation of an aperiodic packet can be postponed for multiple subframes until a large enough unallocated part of the schedule becomes available. On the other hand, the proposed algorithms and MUCA algorithm employ node assignment mechanisms in which the sensor nodes are allocated over the subframes in the scheduling frame as uniformly as possible. This mechanism leaves enough unallocated part in each subframe in the schedule for the allocation of aperiodic packets or packet retransmissions. Therefore, the proposed algorithms can allocate aperiodic packets mostly within the subframe in which they are generated.

4.6.3 Average Runtime Performance

Fig. 4.10 shows the average runtime of the scheduling algorithms for different number of nodes and different number of controllers. The average runtime of the MUCA algorithm and the proposed algorithms employing MUA concurrency allocation algorithm, namely SNA-MUA and GNA-MUA, increases almost linearly as the number of nodes increases without considerable runtime increase as the number of controllers increases. These algorithms require much less runtime than the proposed algorithms using MLA concurrency allocation algorithm, namely SNA-MLA and GNA-MLA, due to the exponential-time enumeration mechanism used in the MLA algorithm to determine the concurrent transmissions. However, the average runtime required by the SNA-MLA and GNA-MLA algorithms is still much lower than that of the optimal solution, especially for large number of controllers and large number of nodes in the network. Besides, as the number of controllers increases, the exponent of the runtime increases in both MLA based scheduling algorithms and optimal solution.

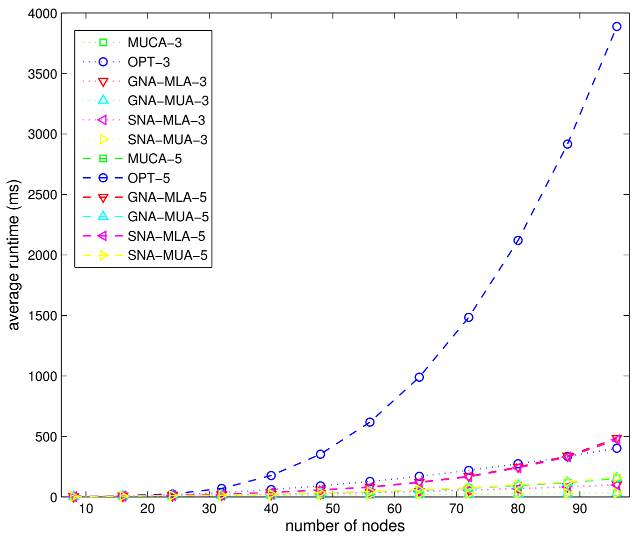

Figure 4.10: Average runtime of the scheduling algorithms for different number of nodes and different number of controllers.

4.7 Conclusion and Discussion

In this chapter, we study the optimal power control, rate adaptation and scheduling with the goal of providing maximum level of adaptivity while meeting the packet generation period, transmission delay, reliability and energy requirements of the sensor nodes for WNCSs. We have extended our work on the optimal power control, rate adaptation and scheduling for UWB-based IVWSNs [63] to general WNCSs for continuous rate transmission model in which Shannon's capacity formulation is used for the achievable transmission rate. We prove that this optimization problem is NP-hard and provide a framework for the design of a heuristic algorithm for scheduling. This framework consists of a node assignment algorithm that assigns the nodes to the subframes in a certain order, and a concurrency allocation algorithm that determines the subsets of the nodes that can transmit concurrently from the node sets with the same packet generation period allocated to the same subframe by the node assignment algorithm. Both node assignment and concurrency allocation algorithms use the optimal polynomial time algorithm proposed as a solution to a novel power control and rate adaptation problem, which is formulated with the objective of minimizing the time required for the concurrent transmission of a set of sensor nodes while satisfying their individual requirements. Extensive simulations for different network scenarios evaluating different performance metrics illustrate that the proposed scheduling algorithms outperform the scheduling algorithm we have proposed for IVWSNs with performance close to optimal solution and average runtime admissible for practical WNCSs. Moreover, the use of the algorithms of different complexities for node assignment and concurrency allocation allows the trade-off between performance and runtime.

In the future, we are planning to extend the proposed framework for multi-hop networks and fast fading channels to provide realistic solutions for a wider range of WNCS applications.

Chapter 5

Minimum Length Scheduling in Wireless Ad Hoc Networks

Traditional approach to the minimum length scheduling problem ignores the packet level details of transmission protocols, meaning that a packet transmission can be divided into several data chunks each of which is transmitted at different rate due to the difference in the set of concurrently transmitting nodes. This solution requires including packet headers for each data chunk resulting in both an increase in the system overhead and underutilization of the time slots. In this study, we extend the previous works on minimum length scheduling by considering the transmission of the packets of arbitrary sizes in the time slots of arbitrary lengths. Given the packet traffic demands on the links, we formulate the joint optimization of the power control, rate adaptation and scheduling for minimizing the schedule length of a wireless ad hoc network and demonstrate the hardness of this problem. Upon solving the power control and rate adaptation problem separately, we formulate the scheduling problem as an integer programming (IP) problem where the number of variables is exponential in the number of the links. In order to solve this large-scale IP problem fast and efficiently, we propose Branch and Price Method and Column Generation Method based heuristic algorithms.

5.1 Introduction

Scheduling the wireless links and controlling their transmission power and rate in Spatial-reuse Time Division Multiple Access (STDMA) wireless networks has been investigated to minimize the detrimental effects of the interference for various objectives including maximizing total throughput [76, 77], maximizing minimum throughput [78, 79], minimizing total transmit power [80, 81] and minimizing schedule length [82, 83, 84, 85, 86, 87, 88, 89, 90, 71, 91, 92, 93, 94, 95, 72]. Among these objectives, minimizing schedule length given the traffic demands of the links is increasingly used with the proliferation of the time critical applications of wireless networks such as real-time surveillance, networked control [96]. Moreover, minimum length schedule determines the traffic-carrying capability of the wireless network since minimization of the total time required for a set of links to transmit their data packets is effectively maximizing the total data transmission rate while providing fairness among the links, which are allocated enough time to transmit their packets.

Earlier work on minimum length scheduling adopts Protocol Interference Model, which describes the interference constraints among the active links according to a conflict graph, where nodes within a certain distance can communicate as long as the receiver is separated by at least a fixed distance from any other active transmitter. Since this interference model does not allow controlling the transmission power and rate of the links due to the predetermined conflict graph, the time is partitioned into slots of fixed length allowing the transmission of a fixed length packet. The algorithms developed for minimum length scheduling using Protocol Interference Model aim to satisfy either the uniform traffic demand of the links activating each link for at least one time slot during the schedule [82, 83] or non-uniform traffic demand considering many-to-one transmission activating each link for multiple time slots during the schedule [84, 85]. Although the scheduling algorithms developed for Protocol Interference model take into account the packetized transmission allocating one packet transmission to each time slot, this model is inadequate in modeling today's radios where multiple power and rate levels are supported. Besides, this interference model does not take into account the cumulative effects of the interference due to simultaneous transmissions thus requiring the overcompensation of the interference by using a large distance of separation between the active transmitters and receivers.

Recently, *Physical Interference Model* that considers the cumulative effect of the interference in the form of Signal-to-Interference-plus-Noise Ratio (SINR) and allows including the transmission power and rate of the links as a variable in minimizing the length of the schedule has gained wider acceptance. Despite the flexibility introduced by this interference model to include variable transmission power and rate, some of the algorithms developed for this optimization problem restricts the partition of the time into slots of fixed length. The length of the time slot corresponds to the transmission time of either a fixed length packet at fixed transmission rate [87, 88] or an integer number of packets the value of which is determined as a function of the rate chosen from a finite set of possible transmission rates [89, 90]. These algorithms generate sub-optimal solutions since the fixed length time slots may be underutilized by the links. The algorithms that partition the time into slots of variable length on the other hand ignore the packet level details of the transmission protocols by allocating any amount of data to the time slots with the only goal of satisfying the total traffic demand of each node [71, 91, 92, 93, 94, 95, 72]. This means that in the optimal solution, a packet transmission can be divided into several data chunks each of which is transmitted at different rate due to the difference in the set of concurrently transmitting nodes. Figure 5.1 shows the solution to the problem of minimizing the length of the schedule formulated in [72] for a scenario where each link has one packet to transmit. Since the traffic demand in this problem is specified as the total number of bits ignoring packet level details, in the optimal solution, all the packet transmissions are transmitted in several data chunks requiring interrupting the packet transmission. For instance, packet transmission of link 1 starts at $T0$ but is interrupted at $T1$ and restarts at $T3$. Moreover, the data rate of the links is different for each data chunk due to the different set of nodes concurrently transmitting in each time interval. For instance, the data rate of link 1 is different in $T0 - T1$ and $T3 - T4$

intervals since link 2 is transmitting concurrently with link 1 in $T0 - T1$ interval whereas link 4 is transmitting concurrently in $T3 - T4$ interval. The data chunks of the same packet must therefore be transmitted as individual packets by including a packet header at the beginning in the implementation. This additional packet header increases the system overhead. Moreover, since the data rates of the links allocated to the same time slot are different, the transmission time of these headers are different for each link resulting in the underutilization of the time slots.

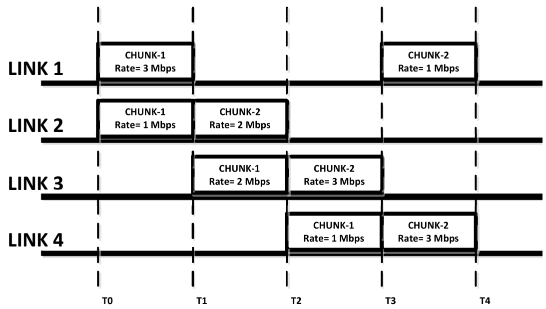

Figure 5.1: Solution to the traditional formulation of minimizing the length of the schedule for the scenario where each of 4 links has one packet of 100 bits to transmit; the links are of 1 m length and uniformly distributed over an area of 15 m$\times$ 15 m; the links can only transmit at discrete rates $1, 2, 3$ Mbps if the SINR value at the receiver of the link exceeds the SINR level $10, 20, 30$ dB respectively [72]. Time axis values are $T0 = 0\mu$s, $T1 = 23.0769\mu$s, $T2 = 61.5384\mu$s, $T3 = 69.2307\mu$s and $T4 = 100\mu$s.

The goal of this study is to determine the optimal power control, rate adaptation and scheduling with the objective of minimizing the schedule length given the *packet traffic demands* of the links and the *constraints of packet transmissions* in a wireless ad hoc network. We first formulate the joint optimization of the power control, rate adaptation and scheduling for minimizing the schedule length of a wireless ad hoc network and demonstrate the hardness of this problem. We then show that the power control and rate adaptation problem can be separated from the scheduling problem and introduce a novel problem for the optimization of power control and rate adaptation where the time required for the concurrent transmission of a set of links each having an integer number of packets is minimized. This problem determines the length of the time slot corresponding to this link set and the corresponding number of packets to be transmitted during the slot. Solving this power control and rate adaptation problem for all possible subsets of the links in the network transmitting any possible number of packets then allows formulating the problem of minimizing the schedule length as an IP problem. The resulting IP problem is large-scale with the number of variables exponential in the number of the links. We therefore incorporate elegant Branch and Price Method and Column Generation method based heuristic algorithms to solve this IP problem. Column generation method has been previously used for the minimum length scheduling problem in different contexts: fixed transmission rate and variable power [71] and variable transmission rate and variable power [93, 94, 95, 72] but without considering packetized transmission; variable transmission rate and variable power but considering only a limited set of packet transmission scenarios for fixed length time slots [89, 90]. In this study, we extend these previous works on minimum length scheduling problem by considering the transmission of the packets of arbitrary sizes in the time slots of arbitrary lengths in the optimization of power control, rate adaptation and scheduling.

The rest of this chapter is organized as follows. Section 5.2 describes the system model and the assumptions used throughout the study. In Section 5.3, the joint optimization of power control, rate adaptation and scheduling with the objective of minimizing the schedule length is formulated considering the constraints of packet transmissions for constant, continuous and discrete rate transmission models and the solution strategy based on the decomposition of the power control and rate adaptation, and scheduling problems is described. Section 5.4 presents the optimal power control and rate adaptation problem with the goal of determining the minimum time-slot length for the concurrent packet transmissions of a subset of the links in the network and propose both optimal and heuristic algorithms for each transmission model. Section 5.5 formulates the optimal scheduling problem as an IP problem where the number of variables is exponential in the number of the links whereas Sections 5.6 and 5.7 introduce Branch and Price and Column Generation Method based heuristic algorithms respectively to solve this IP problem fast and efficiently. Simulations and performance evaluation are presented in Section 5.8. Finally, concluding remarks and future work are given in Section 5.9.

5.2 System Model and Assumptions

The system model and assumptions are detailed as follows:

1. The wireless ad hoc network consists of a set of L directed links with positive traffic demand.

2. A central controller executes the algorithm for the joint optimization of power control, rate adaptation and scheduling based on the available information on the network topology, traffic demands and channel characteristics of the links. This centralized framework is appropriate for the communication of the routers in wireless mesh networks (WMNs) that have emerged recently to partially replace the costly wired network infrastructures [97]. The changes in WMN topology is quite infrequent since the locations of the WMN routers are fixed. Moreover, the bandwidth requests of the WMN links are typically slowly varying over time and can be guaranteed with an almost static resource allocation. Besides, the centralized optimal solution provides an upper bound on the performance of any distributed algorithm since distributed algorithms rely on the local topology information in contrast to the centralized algorithms using the knowledge of the complete topology. Such an upper bound allows evaluating the performance of the distributed algorithms. Introducing distributed algorithms within the proposed framework is out of scope of this study and subject to future work.

3. Link l has traffic demand of G_l packets of length R_l bits. Although the mathematical formulations derived in the study allow having packets of different lengths for the same link, the formulations are easier to express with this assumption. The link traffic demands can be either given for single-hop networks or calculated by using end-to-end traffic demands in a multi-hop network with predetermined routing. The

proposed framework can be extended to include routing in the optimization problem by either a two phase approach in which scheduling and routing are solved iteratively [98, 99] or a joint approach in which routing is included via flow-balance equations that map the end-to-end flow demands to the link traffic demands [100, 89, 101]. To avoid complexity in the first step of the study and better focus on the packetized scheduling, the inclusion of the routing in the optimization problem is out of scope of this study and subject to future work.

4. We consider Time Division Multiple Access (TDMA) as MAC protocol since TDMA provides delay guarantee to the networks with predetermined topology and data generation patterns [38]. The time is partitioned into frames, which are further partitioned into slots of possibly variable lengths. Two possible interpretations can be considered for studying the problem of minimum length scheduling given traffic demands as detailed in [72]: In the first scenario, the traffic demand, i.e. G_l packets for link l, is defined as the number of packets to be transmitted in every frame. On the other hand, in the second scenario, the traffic demand is defined as the long-run average data transfer rate of G_l^s packets per second. This can then be accomplished by dividing the time axis into frames of arbitrary positive constant duration T requiring the satisfaction of $G_l = G_l^s T$ packet transmissions for every link l during each frame. Minimizing the frame length given these traffic demands is then useful for both scenarios because it permits a larger number of frames per unit time. Extending this work for the dynamic adaptation of the schedule, power and rate of the links to the arrival of packets at any instant in time is beyond the scope of this study and subject to future work.

5. A node cannot receive and transmit simultaneously. Also, a node cannot receive from or transmit to more than one node simultaneously.

6. The transmit power can take any value below a maximum level p_{max}. Although the radios are restricted to support a finite number of trans-

mit power levels in practice, the assumption of continuous power is commonly used in most of the previous work such as [71, 87, 94, 95, 72] due to the simplification of the resulting problem and high approximation accuracy with the large number of discrete power levels that can be supported by the existing radios.

7. Let g_{ll} and g_{kl} be the channel gain of link l and the channel gain from the transmitter of link k to the receiver of link l respectively. We assume that the fading is slow such that the channel gain between every transmitter and receiver is fixed during the schedule. This is a common assumption used in the prior formulations of the minimum length scheduling problem in wireless ad hoc networks [71, 87, 91, 94, 95, 72]. Extending this work for fast fading wireless networks by either minimizing the mean value of the schedule length as in [93] or providing an optimal policy by employing the principles of dynamic programming while incorporating a lot of overhead for collecting the channel information at the granularity of the time slot duration as in [92] is beyond the scope of this study and subject to future work.

8. First transmission model used in the formulations is *constant rate* transmission model in which each active link l assigned to a time slot n transmits at constant rate r if the SINR of link l is above a fixed threshold as
$$\frac{p_l^{(n)} g_{ll}}{N_0 + \sum_{k \neq l} p_k^{(n)} g_{kl}} \geq \beta_l \tag{5.1}$$
where $p_l^{(n)}$ is the transmit power of link l in time slot n, N_0 is the background noise and β_l is the SINR threshold to be kept by link l [71, 87].

9. Second transmission model used in the formulations is *continuous rate* transmission model in which Shannon's channel capacity formulation for an Additive White Gaussian Noise (AWGN) wireless channel is used in the calculation of the maximum achievable rate as a function

of SINR as

$$x_l^{(n)} \leq W \times \log\left(1 + \frac{1}{W} \times \frac{p_l^{(n)} g_{ll}}{N_0 + \sum_{k \neq l} p_k^{(n)} g_{kl}}\right) \tag{5.2}$$

where $x_l^{(n)}$ is the transmission rate of link l in time slot n, W is the channel bandwidth [93, 94, 95, 72, 102, 73].

10. Third transmission model used in the formulations is *discrete rate* transmission model in which a finite set of transmit rates $r = (r^1, r^2, ..., r^M)$ and a finite set of SINR levels $\boldsymbol{\gamma} = (\gamma^1, \gamma^2, ..., \gamma^M)$ are determined such that link l can transmit at rate r^c in a time slot n if the SINR achieved at the link

$$\gamma_l^{(n)} = \frac{p_l^{(n)} g_{ll}}{N_0 + \sum_{k \neq l} p_k^{(n)} g_{kl}} \tag{5.3}$$

is greater than or equal to γ^c from the set $\boldsymbol{\gamma}$ [91, 92]. Here $r^1 = 0$ and $\gamma^1 = -\infty$ dB so that if $\gamma_l^{(n)}$ is below the threshold γ^2, the transmission rate is 0. This is in general more realistic model than the *continuous rate* model since a wireless transmitter can only work with a limited number of rate levels as suggested by the practical realization of multiple data rates in [103].

5.3 Packetized Minimum Length Schedule Problem

5.3.1 Problem Formulation

The joint optimization of power control, rate adaptation and scheduling with the objective of minimizing the schedule length given the packet traffic demands and packetized transmission constraints is mathematically formulated for constant rate, continuous rate and discrete rate transmission models.

5.3.1.1 Constant Rate Transmission Model

The formulation of the optimization problem for constant rate transmission model is as follows:

minimize

$$\sum_{n=1}^{N} t^{(n)} \tag{5.4}$$

subject to

$$x_l^{(n)} t^{(n)} \geq \alpha_l^{(n)} R_l, \quad l \in [1, L], n \in [1, N] \tag{5.5}$$

$$\sum_{n=1}^{N} \alpha_l^{(n)} \geq G_l, \quad l \in [1, L] \tag{5.6}$$

$$p_l^{(n)} \leq p_{max}, \quad l \in [1, L], n \in [1, N] \tag{5.7}$$

$$\frac{p_l^{(n)} g_{ll}}{N_0 + \sum_{k \neq l} p_k^{(n)} g_{kl}} \geq \beta_l 1_{\{p_l^{(n)} > 0\}}, \quad l \in [1, L], n \in [1, N] \tag{5.8}$$

$$x_l^{(n)} = 1_{\{p_l^{(n)} > 0\}} r, \quad l \in [1, L], n \in [1, N] \tag{5.9}$$

$$a_{lk} + 1_{\{p_l^{(n)} > 0\}} + 1_{\{p_k^{(n)} > 0\}} \leq 2, \quad l, k \in [1, L], n \in [1, N] \tag{5.10}$$

variables

$$\begin{aligned} t^{(n)} \geq 0,\ p_l^{(n)} \geq 0,\ x_l^{(n)} \geq 0,\ \alpha_l^{(n)} \in [0, G_l], \\ l \in [1, L],\ n \in [1, N] \end{aligned} \tag{5.11}$$

where N is the number of time slots, a_{lk} is a constant that takes the value 1 if links l and k share a common node and 0 otherwise. The variables of the optimization problem are $t^{(n)}$, the length of the n-th time slot; $p_l^{(n)}$, the transmit power of link l in time slot n; $x_l^{(n)}$, the transmission rate of link l in time slot n and $\alpha_l^{(n)}$, the number of packets sent by link l in time slot n.

The goal of the optimization problem is to minimize the length of the schedule. Equation (5.5) represents the packetized transmission requirement where each node transmits an integer number of packets within each time slot. Equation (5.6) gives the constraint on the total number of packets generated at each node. Note that in the case where the packetized transmission within each time slot is not taken into account, Equations (5.5) and (5.6) are replaced by one equation only representing the total number of bits that needs to be transmitted by each node as $\sum_{n=1}^{N} x_l^{(n)} t^{(n)} \geq G_l R_l$. Equation (5.7) states the upper bound for the transmit power of the links. Equation (5.8) represents the condition that the SINR of link l is above a fixed threshold if it is active. Equation (5.9) represents the transmission at fixed rate r for the active links. Equation (5.10) states that any node in the network cannot transmit to and receive from more than one node simultaneously.

This optimization problem is a Mixed Integer Non-Linear Programming problem thus difficult to solve for the global optimum [54]. We prove the hardness of the problem next.

Theorem 7. *The packetized minimum length schedule problem for constant rate transmission model is NP-hard.*

Proof: The packetized minimum length schedule problem as formulated by Eqs. (5.4)-(5.11) for constant rate transmission model is equivalent to the integrated link scheduling and power control problem in [87], which is shown to be NP-hard, for the instance in which the packet length of all the links in the network is the same. □

5.3.1.2 Continuous Rate Transmission Model

The formulation of the optimization problem for continuous rate transmission model is very similar to that formulated for constant rate transmission model except that Equations (5.8) and (5.9) are replaced by the following constraint:

$$x_l^{(n)} \leq W \times \log\left(1 + \frac{1}{W} \times \frac{p_l^{(n)} g_{ll}}{N_0 + \sum_{k \neq l} p_k^{(n)} g_{kl}}\right), \qquad l \in [1, L], n \in [1, N] \tag{5.12}$$

This problem is a non-convex optimization problem for which there is no known polynomial time algorithm [54].

5.3.1.3 Discrete Rate Transmission Model

The formulation of the optimization problem for discrete rate transmission model is also very similar to that formulated for constant rate transmission model except that Equations (5.8) and (5.9) are replaced by the following constraints:

$$p_l^{(n)} g_{ll} - z_{li}^{(n)} \gamma^i (N_0 + \sum_{k \neq l} p_k^{(n)} g_{kl}) \geq 0, \quad l \in [1, L], \qquad n \in [1, N], i \in [1, M] \tag{5.13}$$

$$\sum_{i=1}^{M} z_{li}^{(n)} = 1, \quad l \in [1, L], n \in [1, N] \tag{5.14}$$

$$x_l^{(n)} = \sum_{i=1}^{M} z_{li}^{(n)} r^i, \quad l \in [1, L], n \in [1, N] \tag{5.15}$$

and the inclusion of the variable $z_{li}^{(n)} \in \{0, 1\}$, where $l \in [1, L], n \in [1, N], i \in [1, M]$, that takes value 1 if link l transmits at rate r^i in time slot n and 0 otherwise. Equation (5.13) states that the assigned rate r^i to link l satisfies the corresponding SINR requirement as $\gamma_l^{(n)} \geq \gamma^i$. Equations (5.14) and (5.15) state that each link l is assigned to exactly one rate.

Corollary 1. *The packetized minimum length schedule problem for discrete rate transmission model is NP-hard.*

Proof: Discrete rate transmission model is equivalent to constant rate transmission model for the instance in which $M = 2$. Since the packetized minimum length schedule problem for constant rate transmission model is NP-hard based on Theorem 7, the packetized minimum length schedule problem for discrete rate transmission model is NP-hard. □

5.3.2 Solution Strategy

The exponential complexity of the packetized minimum length schedule problem for different transmission models necessitates the use of heuristic algorithms. The solution strategy proposed in this study is based on the decomposition of the power control and rate adaptation, and scheduling problems as described below:

- The constraints of the joint power control, rate adaptation and scheduling problem provided in Equations (5.5)-(5.15) include the variables for the transmit power and rate of the links and the slot length corresponding to only one time slot at a time except the constraint on the total number of packets generated at each node given in Equation (5.6). Therefore, given the number of packets to be transmitted by each link, the length of the time slot should be minimized independent of the other slots in order to attain minimum total schedule length. A novel power control and rate adaptation problem formulated as the minimization of the time slot length required for the concurrent transmission of a set of simultaneously active links each transmitting an integer number of packets is formulated and solved for constant rate, continuous rate and discrete rate transmission models.

- The optimal scheduling contains the time slots each of which corresponds to a subset of the links with each link in the subset sending a certain number of packets. Therefore, given the minimum time slot lengths corresponding to all the subset of the links concurrently transmitting any possible number of packets, calculated by using the formulation of the optimal power control and rate adaptation problem described previously, the scheduling problem is formulated as an Integer Programming (IP) problem where the number of variables is exponential in the number of the links.

- To solve the large-scale IP formulation for the scheduling problem, Branch and Price Method and Column Generation Method based heuristics are proposed.

5.4 Power Control and Rate Adaptation Problem

The power control and rate adaptation problem aims to minimize the length of the time slot given the number of packets to be transmitted by each active link in that time slot. Given the power allocation of the links in a time slot, the feasible and achievable rate region for a link given in Equations (5.9), (5.12) and (5.15) for constant, continuous and discrete rate transmission models respectively is independent of the rate of the concurrently transmitting links. Since higher rate results in smaller transmission delay, the optimal transmission rate is the maximum achievable rate for the minimization of the time slot length. Power control and rate adaptation problem can therefore be reduced to a pure power control problem where the transmission rates of the links are formulated as a function of the transmit powers of the links.

Let us define the set of active links in a time slot S. We assume that none of the links in the set S have common end point to satisfy the requirement given in Equation (5.10). Let t be the length of the time slot to be allocated to the transmission of α_l packets by each link l within the set S. The objective of the optimal power control problem under all transmission models is to minimize t. We remove the superscript of the variables, which represent the time slot in Section 5.3, to simplify the expressions in this section.

5.4.1 Constant Rate Transmission Model

The optimal power control problem for constant rate transmission model is formulated as

minimize

$$t \tag{5.16}$$

subject to

$$\frac{\alpha_l R_l}{x_l} \leq t, \quad l \in S \tag{5.17}$$

$$p_l \leq p_{max}, \quad l \in S \tag{5.18}$$

$$\frac{p_l g_{ll}}{N_0 + \sum_{k \neq l} p_k g_{kl}} \geq \beta_l, \quad l \in S \tag{5.19}$$

$$x_l = r, \quad l \in S \tag{5.20}$$

variables

$$t \geq 0,\ p_l \geq 0,\ l \in S \tag{5.21}$$

The constraints given in Equations (5.17), (5.18), (5.19) and (5.20) correspond to the constraints given in Equations (5.5), (5.7), (5.8) and (5.9) respectively.

The power control problem for constant rate transmission model first requires testing the feasibility of the concurrent transmission of the links in S. For the feasibility of the concurrent transmission of the links in S, there should exist a set of transmit power allocations to the links in the set S such that the SINR constraint of each link $l \in S$ as given in Equation (5.19) is satisfied while meeting the maximum transmit power constraint in Equation (5.18). The existence of such a set of transmit power allocations is determined by testing Perron-Frobenius conditions [75] described as follows: Let B_S be $|S| \times |S|$ relative channel gain matrix such that the element in the l-th row and k-th column of B_S takes value g_{kl}/g_{ll} for $l \neq k$ and 0 for $l = k$. Let D_S be $|S| \times |S|$ diagonal matrix with the l-th diagonal entry equal to β_l. Let V_S be $|S| \times 1$ normalized noise power vector with the l-th element equal to $\beta_l N_0/g_{ll}$. Perron-Frobenius conditions state that there exists a feasible power vector if and only if the largest real eigenvalue of $D_S B_S$ is less than 1 and every element of the component-wise minimum power vector $(I - D_S B_S)^{-1} V_S$ is less than or equal to p_{max}. If the concurrent transmission of the links within the set S is feasible, then the optimal value of t is equal to $\max_{l \in S} \frac{\alpha_l R_l}{r}$.

Since the complexity of evaluating Perron-Frobenius conditions for $|S|$ links is $O(|S|^3)$ [75], the complexity of solving the power control problem for constant rate transmission model is $O(|S|^3)$.

5.4.2 Continuous Rate Transmission Model

The optimal power control problem for continuous rate transmission model is very similar to that formulated for constant rate transmission model except Equations (5.19) and (5.20) are replaced by the constraint given in Equation (5.12) for the links in the set S. This power control problem is a non-convex optimization problem for which there is no known polynomial time algorithm

[54]. In the following, we first state the theorem on the equality of the optimal time duration required for the transmission of all the links assigned to the same time slot then provide an iterative algorithm that gets close to the optimal time slot length within a certain accuracy exponentially fast.

Theorem 8. *Let t^* denote the optimal length of the time slot allocated to the transmission of α_l packets by each link l within the set S such that $t^* = \max_{l \in S} t_l^*$ where t_l^* is the transmission time required for link l. Then, the transmission times of all the links within S are equal; i.e., $t^* = t_l^* \; \forall l \in S$.*

Proof: We will prove the theorem by contradiction. Suppose that $t^* \neq t_l^*$ for an arbitrary $l \in S$. Let $t_k^* = \max_{j \in S} t_j^*$. Then $t_k^* > t_l^*$. If we decrease the transmission power of link l by an arbitrarily small amount such that t_l^* is still less than t_k^*, the transmission time of all the links except link l decreases due to the decreasing amount of interference created by node l. Note that the transmission rate of a link j is an increasing function of the transmission power of link j and decreasing function of the transmission power of link $i \neq j$. As a result, time slot length required for the transmission of all the links, $t^* = \max_{j \in S} t_j^*$ decreases. Hence, as long as there exists a node j such that $t_j < \max_{l \in S} t_l$, we can improve the solution by decreasing the transmission power of link j. Hence, at the optimal solution, the transmission time of every link in the set S must be equal.

□

If we fix the value of the length of the allocated time slot, the optimal power control problem reduces to determining whether there exists a set of transmit power assignment to the links in the set S such that the SINR constraint of each link $l \in S$ as given in

$$\frac{p_l g_{ll}}{N_0 + \sum_{k \neq l, k \in S} p_k g_{kl}} \geq \beta_l \tag{5.22}$$

where

$$\beta_l = W \left(2^{\alpha_l R_l / (W t)} - 1 \right) \tag{5.23}$$

is satisfied while meeting the maximum transmit power constraint in Equation (5.18). The existence of such a set of transmit power vector is determined by testing Perron-Frobenius conditions as explained in Section 5.4.1.

Fast-Iterative Power Control Algorithm (FIPCA) is based on performing an intelligent search for the minimum feasible value of t.

FIPCA algorithm determines an interval for the optimal value of the length of the time slot corresponding to the transmission of α_l packets by each link l within the feasible set S and then reduces the length of this interval by half in each iteration. The lower and upper end points of this interval determined as the lowest infeasible value and highest feasible value for the time slot length t are denoted by t^{lw} and t^{up} respectively. The algorithm starts by setting the lower bound to the maximum value of the time slot lengths required when only one link transmits at maximum transmit power p_{max} and upper bound to the length of the time slot when every link $l \in S$ transmits at maximum transmit power p_{max} (Lines $1-2$). In each iteration, the algorithm checks the feasibility of the middle point of the interval (t^{lw}, t^{up}) (Lines $4-5$). If the value $t = (t^{lw}+t^{up})/2$ is feasible then any value greater than t is a worse feasible solution than t thus requiring the search for the optimal time slot over interval $(t^{lw}, (t^{lw} + t^{up})/2)$ by updating the upper bound (Lines $6-7$). On the other hand, if the value $t = (t^{lw} + t^{up})/2$ is infeasible then any value less than t is also infeasible requiring the search for the optimal time slot over interval $((t^{lw} + t^{up})/2, t^{up})$ by updating the lower bound (Lines $8-9$). The termination criteria of the algorithm is achieving a predetermined relative error bound level represented by ϵ (Line 3). The value of the time slot length is then set to t^{up} (Line 12). Since the optimal time slot length t^{opt} is between t^{lw} and t^{up}, the relative error bound of the optimal time slot length given by $\frac{t^{up}-t^{opt}}{t^{opt}}$ is less than $\frac{t^{up}-t^{lw}}{t^{lw}}$ so less than ϵ meaning that the value picked by the algorithm t^{up} lies in the interval $[t^{opt}, (1+\epsilon)t^{opt}]$.

Lemma 14. *For a fixed predetermined relative error bound level ϵ, the complexity of the FIPCA algorithm is $O(K|S|^3)$, where $K = \lceil \log(\frac{t^{up}_{init}-t^{lw}_{init}}{\epsilon\, t^{lw}_{init}}) \rceil$, t^{up}_{init} and t^{lw}_{init} are the initial values of the upper and lower bounds for the optimal time slot length set by the algorithm (Lines $1-2$).*

Proof: The FIPCA algorithm reduces the initial ratio $\frac{t^{up}_{init}-t^{lw}_{init}}{t^{lw}_{init}}$ by a factor of at least two in each iteration since $t^{up} - t^{lw}$ is reduced by half by setting either t^{up} or t^{lw} to the middle point of the interval $[t^{up}, t^{lw}]$ (Lines $4-9$).

Algorithm 9 Fast Iterative Power Control Algorithm (FIPCA)

1: $t^{lw} = \max_{l \in S} \frac{\alpha_l R_l}{W \times \log\left(1 + \frac{1}{W} \times \frac{pmax\, g_{ll}}{N_0}\right)}$;

2: $t^{up} = \max_{l \in S} \frac{\alpha_l R_l}{W \times \log\left(1 + \frac{1}{W} \times \frac{pmax\, g_{ll}}{N_0 + \sum_{k \neq l, k \in S} pmax\, g_{kl}}\right)}$

3: **while** $\frac{t^{up} - t^{lw}}{t^{lw}} > \epsilon$ **do**

4: $\quad t = (t^{up} + t^{lw})/2$;

5: $\quad$ check feasibility for time-slot length t;

6: $\quad$ **if** t is feasible **then**

7: $\quad\quad t^{up} = t$;

8: $\quad$ **else**

9: $\quad\quad t^{lw} = t$;

10: $\quad$ **end if**

11: **end while**

12: $t = t^{up}$;

Hence, the number of iterations required to decrease the ratio $\frac{t^{up} - t^{lw}}{t^{lw}}$ to ϵ is $K = \lceil \log(\frac{t^{up}_{init} - t^{lw}_{init}}{\epsilon\, t^{lw}_{init}}) \rceil$. Since the algorithm checks the feasibility of the existence of a transmit power vector satisfying SINR and maximum power constraints in each iteration evaluating the Perron-Frobenius conditions (Line 5) of $O(|S|^3)$ complexity, the complexity of the algorithm is $O(K|S|^3)$.$\square$

5.4.3 Discrete Rate Transmission Model

The optimal power control problem for discrete rate transmission model is very similar to that formulated for constant rate transmission model except Equations (5.19) and (5.20) are replaced by the constraints given in Equations (5.13), (5.14) and (5.15) for the links in the set S. This power control problem is a Mixed Integer Non-linear Programming thus difficult to solve for the global optimum [54]. Note that unlike the continuous rate transmission model, the optimal time duration required for the transmission of the links assigned to the same time slot are not necessarily equal in this case due to the discrete nature of the rates.

A straightforward search algorithm enumerates all possible rate allocations of the links in the set S and check their feasibility by testing the existence of a power vector satisfying the SINR requirements corresponding to these rate allocations and maximum power constraints by evaluating Perron-Frobenius conditions. The optimal solution is then the minimum of the time slot lengths of the feasible rate allocations. The complexity of this search is $\mathcal{O}(M^{|S|}|S|^3)$.

We will now propose a fast fathoming-based smart enumeration algorithm based on the following two lemmas. In this algorithm, we use a tree structure for the rate index vector $x_r = (i_1, i_2, ..., i_{|S|})$ where the links are enumerated from 1 to $|S|$ and the j-th link is assigned to the i_j-th rate level where $i_j \in [1, M]$. The root of the tree is $(M, M, ..., M)$. The children of the rate index vector $x_r = (i_1, i_2, ..., i_{|S|})$ are obtained by updating the j-th element of x_r as $i_j - 1$ and keeping the remaining elements the same for every $j \in [1, |S|]$ such that $i_j > 2$. A rate index vector $(i_1, i_2, ..., i_{|S|})$ is a descendant of another rate index vector $(m_1, m_2, ..., m_{|S|})$ if $i_j \leq m_j$ for every $j \in [1, |S|]$.

Lemma 15. *If a rate index vector $x_r = (i_1, i_2, ..., i_{|S|})$ is feasible with the corresponding time slot length t_r then the time slot length corresponding to any descendant of x_r is greater than or equal to t_r.*

Proof: Let us denote the descendant of x_r by $y_r = (p_1, p_2, ..., p_{|S|})$. Since $p_j \leq i_j$ for $j \in [1, |S|]$, the transmission duration of link j at rate r^{p_j} is greater than or equal to that at rate r^{i_j} for every $j \in [1, |S|]$. Since the time slot length is the maximum of the transmission duration of all the links in the set S, the time slot length corresponding to any descendant of x_r is greater than or equal to t_r. □

Lemma 16. *If a rate index vector $x_r = (i_1, i_2, ..., i_{|S|})$ is infeasible and $\max_{j \in [1,|S|]} \frac{\alpha_j R_j}{r^{i_j}}$ is greater than or equal to the time slot length corresponding to a feasible rate index vector denoted by t_f, then the time slot length corresponding to any feasible descendant of x_r is greater than or equal to t_f.*

Proof: Let us denote the descendant of x_r by $y_r = (p_1, p_2, ..., p_{|S|})$. Since $p_j \leq i_j$ for $j \in [1, |S|]$, the time slot length corresponding to y_r that is given by $\max_{j \in [1,|S|]} \frac{\alpha_j R_j}{r^{p_j}}$ is greater than or equal to $\max_{j \in [1,|S|]} \frac{\alpha_j R_j}{r^{i_j}}$, which is greater than or equal to t_f. □

Based on the foregoing lemmas, we propose the Fast Tree-Based Fathoming Algorithm (FTFA) as described next. x_r^{opt} and t^{opt} correspond to the

Algorithm 10 Fast Tree-Based Fathoming Algorithm (FTFA)

```
A = {(M, M, ..., M)};
F = ∅;
C = ∅;
t^{opt} = ∞;
x_r^{opt} = ∅;
while A ≠ ∅ do
    x_r = (i_1, i_2, ..., i_{|S|}) = first rate index vector in A;
    discard x_r from set A;
    if x_r is not descendant of a node in F and not inside C then
        t = max_{j∈[1,|S|]} (α_j R_j) / (r^{i_j});
        if x_r is feasible then
            if t < t^{opt} then
                t^{opt} = t;
                x_r^{opt} = x_r;
            end if
            add x_r to set F; {due Lemma 15}
        else
            if x_r = (2, 2, ..., 2) then
                break; {no feasible solution exists}
            end if
            if t >= t^{opt} then
                add x_r to set F; {due Lemma 16}
            else
                include children of x_r at the beginning of set A;
                include x_r in set C;
            end if
        end if
    end if
end while
```

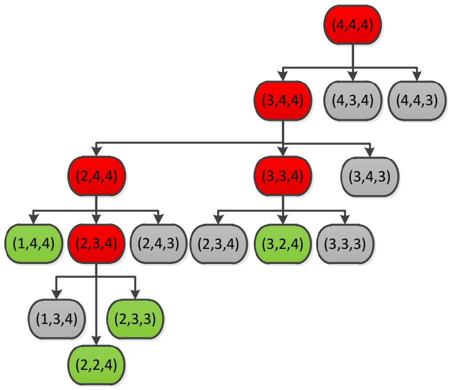

Figure 5.2: FTFA algorithm illustration on the tree structure for the case where $|S| = 3$ and $M = 4$. The rate vectors are evaluated in the following order: $(4,4,4)$, $(3,4,4)$, $(2,4,4)$, $(1,4,4)$, $(2,3,4)$, $(1,3,4)$, $(2,2,4)$, $(2,3,3)$, $(2,4,3)$, $(3,3,4)$, $(2,3,4)$, $(3,2,4)$, $(3,3,3)$, $(3,4,3)$, $(4,3,4)$ and $(4,4,3)$. Green-colored rate index vectors are the fathomed vectors that are evaluated as feasible and the descendants of which are pruned by the algorithm based on Lemma 15 (Line 16). Grey-colored rate index vectors are the vectors that are not evaluated since they are either a descendant of a node in F or inside C (Line 9) or fathomed since they are infeasible rate index vectors for which t is greater than or equal to the current optimal time slot length t^{opt} based on Lemma 16 (Lines 21−22). Red-colored rate index vectors are the vectors that are evaluated as infeasible for which t is less than the current optimal time slot length t^{opt} and whose child nodes are generated as a consequence (Lines 24−25). For example, vector $(2,4,3)$ is fathomed since the corresponding t value is higher than the current optimal time slot length associated with the best feasible rate vector $(2,2,4)$, vector $(1,3,4)$ is not evaluated since it is a descendant of a previously fathomed vector $(1,4,4)$ and vector $(2,3,4)$ which is generated as a child node of rate vector $(3,3,4)$ is not evaluated since it is inside C.

optimal rate index vector and optimal time slot length in each iteration respectively. x_r^{opt} and t^{opt} are initialized to $\emptyset$ and ∞ respectively (Lines $4-5$). Three sets are updated in each iteration of the algorithm. A is the set of rate index vectors to be evaluated in order. A is initialized to $\{(M, M, ..., M)\}$ (Line 1) and extended to include the children of the infeasible rate index vector for which $t = \max_{j\in[1,|S|]} \frac{\alpha_j R_j}{r^{i_j}} < t^{opt}$ since these children are potential candidates for optimal rate index vectors (Line 24). F is initialized to $\emptyset$ (Line 2) and extended to include feasible rate index vectors since any descendant of these vectors is also feasible and cannot give a time slot length smaller than the current time slot length t due to Lemma 15 (Line 16) and infeasible rate index vectors for which t is greater than or equal to the current optimal time slot length t^{opt} since any descendant of these vectors cannot give a time slot length smaller than t^{opt} due to Lemma 16 (Lines $21-22$). C is the set of rate index vectors that are evaluated but not fathomed. C is initialized to $\emptyset$ (Line 3). The children of the rate index vector placed into C in each iteration are included in A to be checked in the following iterations (Line 25). In each iteration of the algorithm, the first rate index vector from the set A is chosen and evaluated for feasibility by checking Perron-Frobenius conditions if it is not descendant of a node in F and not inside C (Lines $7-9$). If the rate index vector evaluated in the current iteration is feasible, its corresponding time slot length is checked for optimality (Lines $11-12$). If the time slot length is less than the current optimal value t^{opt}, x_r^{opt} and t^{opt} are updated (Lines $13-14$). If there are no rate index vectors to be evaluated in A (Line 6) or the smallest rate index vector where all the nodes in the set S are active, i.e. $x_r = (2, 2, ..., 2)$, is infeasible (Lines 18-20), the algorithm stops.

The complexity of FTFA algorithm is $O(M^{|S|}|S|^3)$ since the algorithm checks the feasibility of all of the $M^{|S|}$ possible rate index vectors by evaluating Perron-Frobenius conditions with complexity $O(|S|^3)$ in the worst case. However, the average runtime of FTFA is much smaller due to the fathoming mechanism as will be demonstrated via extensive simulations in Section 5.8.

5.5 Optimal Scheduling Problem Formulation

Once the minimum time slot length of all possible simultaneous packet transmission scenarios, i.e. the transmission of α_l packets by link l where $\alpha_l \in [0, G_l]$ for $l \in [1, L]$, is determined by solving the power control and rate adaptation problem described in Section 5.4, the minimum length scheduling problem can be formulated by assigning a variable to each of these scenarios as described next.

Let $\mathcal{E} = \{\mathcal{E}_k : 1 \leq k \leq |\mathcal{E}|\}$ denote the set of all feasible simultaneous packet transmission scenarios of the link set $\mathcal{L} = \{1, 2, ..., L\}$. Note that $|\mathcal{E}| \leq \prod_{i=1}^{L}(G_i + 1)$ with equality for the case where no two links share any common node.

Let Q denote an $L \times |\mathcal{E}|$ matrix such that the element in the l-th row and k-th column of Q denoted by q_{lk} is the number of packets transmitted by link l in the k-th simultaneous packet transmission scenario in $\mathcal{E}$, i.e. $\mathcal{E}_k$.

The minimum length scheduling problem is then formulated as an Integer Programming (IP) problem [55] as

minimize

$$t^T c \tag{5.24}$$

subject to

$$Qc \geqslant G \tag{5.25}$$

variables

$$c^{(k)} \in \{0, 1\}, \quad k \in [1, |\mathcal{E}|] \tag{5.26}$$

where t is $|\mathcal{E}| \times 1$ vector whose k-th element $t^{(k)}$ is the minimum time slot length corresponding to the k-th simultaneous packet transmission scenario $\mathcal{E}_k$, G is $L \times 1$ vector whose k-th element is the packet requirement of the k-th link, i.e. G_k. The variable of the IP problem is the $|\mathcal{E}| \times 1$ vector c whose k-th element denoted by $c^{(k)}$ takes value 1 if the transmission scenario $\mathcal{E}_k$ exists in the optimal schedule and 0 otherwise. Equation (5.25) represents the packet requirements of the links in the network.

There arise two fundamental difficulties in solving this IP problem. First, it requires an exponential effort to determine the time slot length required for each possible transmission scenario. Second, even if the time slot lengths are determined, there are exponential number of integer variables in the foregoing IP formulation which makes the problem intractable. Since there are $\prod_{i=1}^{L}(G_i + 1)$ possible transmission scenarios, the complexity of determining the time slot length corresponding to all transmission scenarios is $O(\prod_{i=1}^{L}(G_i + 1) f_p)$ where f_p is the complexity of the power control algorithm used depending on the transmission rate model and equal to $O(|L|^3)$, $O(|L|^3)$

and $O(M^{|L|}|L|^3)$ for constant, continuous and discrete rate transmission models as derived in Sections 5.4.1, 5.4.2 and 5.4.3 respectively. The complexity of solving the IP problem with $\prod_{i=1}^{L}(G_i+1)$ variables given the corresponding time slot lengths on the other hand is $O(2^{\prod_{i=1}^{L}(G_i+1)})$. The complexity of solving the optimal scheduling problem is therefore $O(\prod_{i=1}^{L}(G_i+1)f_p + 2^{\prod_{i=1}^{L}(G_i+1)})$.

5.6 Branch and Price Method Based Heuristic Algorithm for Scheduling

Branch and Price Method Based Heuristic Algorithm (BPH) developed to overcome the intractability problem of the IP formulation in Section 5.5 is based on using Column Generation Method based heuristic algorithm to solve the LP relaxation of the IP problem at each node of the Branch-and-Bound tree.

BPH starts with the exponential IP formulation given in Equations (5.24)-(5.26). The LP relaxation of the exponential IP formulation is solved using the Column Generation Method (CGM). If the solution is fractional, one of the fractional variables is picked and two new IP problems are formed by setting the value of the corresponding variable to 0 and 1, which we call branching forming new nodes in the tree. After branching, each resulting node has an associated IP formulation. Note that at each node, the associated IP formulation does not contain some of transmission scenarios in $\mathcal{E}$ since the associated variables to these transmission scenarios are either set to 0 or 1 through the branch on which the particular node is located. We iteratively solve the LP relaxation of the IP formulations at the nodes in the tree using CGM. If the CGM does not provide a feasible solution at a certain node, we stop branching (fathom) that node. If the solution of the CGM is integral, we update the best integral solution and fathom that node. If the solution is fractional and the solution is worse than the best integral solution, we also fathom that node; otherwise we continue branching by using one of the fractional variables and setting its value to 0 and 1 resulting in two new IP problems, i.e. nodes. This process continues until every node is fathomed.

Column Generation Method (CGM), which is an elegant method generally used for solving large-scale LP problems [55], is used to solve the LP-relaxation of the IP problem associated with the nodes of the Branch and Price tree. CGM decomposes the resulting LP problem into two subproblems: Restricted Master Problem and Pricing Problem.

5.6.1 Restricted Master Problem

The Restricted Master Problem (RMP) is similar to the original problem except that only a small subset of the transmission scenarios $\mathcal{E}^{(s)} \subset \mathcal{E}$ is considered as

minimize

$$(t^{(s)})^T c \tag{5.27}$$

subject to

$$Q^{(s)} c \geqslant G \tag{5.28}$$

variables

$$0 \leq c^{(k)} \leq 1, \quad k \in [1, |\mathcal{E}^s|] \tag{5.29}$$

where $t^{(s)}$ and $Q^{(s)}$ only contain the timeslot lengths in vector t and the columns of the transmission scenario matrix Q corresponding to the subsets in $\mathcal{E}^s$ respectively.

To start CGM, we need to choose an initial set $\mathcal{E}^{(s)}$ that guarantees the feasibility of the RMP. A proper choice of $\mathcal{E}^{(s)}$ contains the transmission scenarios where only one link becomes active at a time transmitting all its packets resulting in $t^{(s)}$ as a vector containing the corresponding minimum time slot lengths and $Q^{(s)}$ as an $L \times L$ diagonal matrix whose diagonal elements are the packet requirements of the links.

We can solve RMP and its dual problem with simplex method in polynomial time to obtain its primal optimal solution c^p and dual optimal solution c^d. Since we consider only an arbitrary subset $\mathcal{E}^{(s)} \subset \mathcal{E}$, c^p is not necessarily the optimal solution to the original problem given in Equations (5.24)-(5.26).

In the original problem, the cost coefficient of each variable $c^{(k)}$ in the objective function is $t^{(k)}$ for $k \in [1, |\mathcal{E}|]$. Thus the reduced cost of a column $q^{(k)}$ in the matrix Q but not in $Q^{(s)}$ is

$$r^{(k)} = t^{(k)} - (c^d)^T q^{(k)} \tag{5.30}$$

where $t^{(k)}$ is the minimum length of the time slot for the transmission scenario $\mathcal{E}_k$ corresponding to the column $q^{(k)}$.

If there exists a column $q^{(k)}$ the reduced cost of which denoted by $r^{(k)}$ is less than 0, the objective function of the RMP can be further reduced by including the column $q^{(k)}$ in the matrix $Q^{(s)}$. Otherwise, the current solution c^p is the optimal solution to the original problem. Instead of calculating the reduced cost of each column $q^{(k)}$ in the matrix Q but not in $Q^{(s)}$, which requires an exponential effort, we can solve the following Pricing Problem which will generate the column with the minimum reduced cost among all columns in Q matrix.

5.6.2 Pricing Problem

The goal of formulating the Pricing Problem (PP) is to find a transmission scenario that can improve the objective of the RMP when included in $Q^{(s)}$ matrix as a column. The objective of the PP formulation is given as

$$t - \sum_{l=1}^{L} c_l^d \alpha_l \tag{5.31}$$

where c_l^d is the l-th element of the dual optimal solution of the RMP c^d; α_l is the number of packets transmitted by link l; t, the length of the time slot corresponding to the transmission scenario represented by the number of packets α_l transmitted by each link l where $l \in [1, L]$. The constraints and variables of the optimization problem are the same as those of the power control problem provided in Section 5.4 for all transmission models with the additional variable α_l, $l \in [1, L]$.

The PP formulation is a Mixed-Integer Programming problem for which there is no known polynomial-time algorithm [55]. We therefore propose the following heuristic algorithm to efficiently and rapidly solve PP.

5.6.3 Heuristic Algorithm for Pricing Problem

The heuristic algorithm called Reduced Cost Minimization Algorithm (RCMA) is based on iteratively including links and corresponding number of packet transmissions one at a time such that the reduced cost of the resulting set is minimized.

Let S and α^S denote the set of links and the number of packet transmissions by the links in the set S respectively such that each link $l \in S$ transmits α_l packets, i.e. $\alpha^S = \{\alpha_{S(1)}, \alpha_{S(2)}, ..., \alpha_{S(|S|)}\}$ where $S(i)$ is the i-th element of the set S. The reduced cost of the transmission scenario represented by the pair (S, α^S) is then defined as

$$r^{(S,\alpha^S)} = t^{(S,\alpha^S)} - \sum_{l \in S} c_l^d \alpha_l \tag{5.32}$$

where $t^{(S,\alpha^S)}$ is the length of the time slot corresponding to the transmission scenario where the nodes in the set S transmit the number of packets in the set α^S.

RCMA algorithm iteratively includes the links and corresponding number of packet transmissions in the initially empty sets S and α^S respectively (Line 1) such that the link and corresponding number of packet transmissions minimize the reduced cost of the resulting transmission scenario after their addition (Lines $3-5$). The algorithm stops when the set S includes all the links in $\mathcal{L}$ (Line 2) or the addition of a link does not decrease the reduced cost of the transmission scenario represented by (S, α^S) (Lines $6-8$).

Algorithm 11 Reduced Cost Minimization Algorithm (RCMA)

1: $S = \emptyset$, $\alpha^S = \emptyset$, $D = \mathcal{L}$;
2: **while** $S \neq \mathcal{L}$ **do**
3:     **if** $\min_{l \in D, \alpha_l \in [1, G_l]} r^{(S+\{l\}, \alpha^S+\{\alpha_l\})} < r^{(S,\alpha^S)}$ **then**
4:         (k, α_k)=arg $\min_{l \in D, \alpha_l \in [1, G_l]} r^{(S+\{l\}, \alpha^S+\{\alpha_l\})}$;
5:         $S = S + \{k\}, D = D - \{k\}, \alpha^S = \alpha^S + \{\alpha_k\}$;
6:     **else**
7:         **break**;
8:     **end if**
9: **end while**

5.6.4 Complexity of BPH

BPH investigates all the nodes in the branch-and-price tree in the worst case. The total number of the nodes in the branch-and-price tree is less than $2^{\prod_{i=1}^{L}(G_i+1)+1}$ since each node is obtained by setting a subset of $\prod_{i=1}^{L}(G_i+1)$ variables to either 0 or 1. On the other hand, at each node of the branch-and-price tree, CGM is used to solve the LP relaxation of the associated IP formulation. The maximum number of iterations in the CGM is equal to the maximum number of columns CGM can generate, which is equal to the total number of transmission scenarios, i.e. $\prod_{i=1}^{L}(G_i+1)$. In each iteration of CGM, the RMP solves an LP problem of at most $\prod_{i=1}^{L}(G_i+1)$ variables resulting in the worst case complexity of $O(L(\prod_{i=1}^{L}(G_i+1))^4 d)$ using Interior Method [104] where d is the maximum number of bits used to represent the coefficients of the objective function and constraints of the optimization problem, i.e. t^k, q_{lk} and G_l, for $l \in [1, L]$ and $k \in \left[1, \prod_{i=1}^{L}(G_i+1)\right]$. The RCMA algorithm solving the PP of the CGM in each iteration has complexity of $O(\sum_{i=1}^{L}(G_i))$ provided that the time slot length corresponding to all transmission scenarios is generated once in the worst case with complexity $O(\prod_{i=1}^{L}(G_i+1)f_p)$. Therefore, the overall complexity of BPH is $O(\prod_{i=1}^{L}(G_i+1)f_p + 2^{\prod_{i=1}^{L}(G_i+1)+1}(L(\prod_{i=1}^{L}(G_i+1))^5 d))$. Note that the worst case complexity of BPH is much higher than that of the optimal IP formulation. However, since the fathoming mechanism prevents the generation of all the nodes in the branch-and-price tree and CGM only generates a subset of the columns, the average runtime of BPH is less than that of the optimal IP formulation as will be demonstrated via extensive simulations in Section 5.8.

5.7 Column Generation Method Based Heuristic Sheduling Algorithms

BPH described in Section 5.6 explores the whole branch-and-bound tree by solving the LP relaxation of the IP problem at each node using the Column Generation Method (CGM) based heuristic algorithm. We now propose heuristic methods of lower complexity that still exploit the CGM solution given in Sections 5.6.1 and 5.6.3.

5.7.1 Restricted Master Heuristic

Restricted Master Heuristic (RMH) first solves the LP relaxation of the optimal IP formulation generating the columns of the Q matrix required for the optimal fractional solution then formulates the IP problem with these columns. The algorithm starts with the exponential IP formulation given in Equations (5.24)-(5.26). The LP relaxation of the exponential IP formulation is solved using the Column Generation Method (CGM). In CGM, the set of transmission scenarios $\mathcal{E}^{(s)}$ is initialized to guarantee the feasibility of the RMP and extended by including the transmission scenario that minimizes the reduced cost by the use of the PP at each iteration of CGM. The resulting set $\mathcal{E}^{(s)}$, the corresponding time slot length vector $t^{(s)}$ and transmission scenario matrix $Q^{(s)}$ are then input to the original IP formulation given in Equations (5.24)-(5.26) by setting $Q = Q^{(s)}$ and $t = t^{(s)}$.

RMH requires solving CGM in $\prod_{i=1}^{L}(G_i + 1)$ iterations with at most $\prod_{i=1}^{L}(G_i + 1)$ variables and the optimal IP formulation with $\prod_{i=1}^{L}(G_i + 1)$ variables in the worst case. The overall complexity of RMH is therefore $O(\prod_{i=1}^{L}(G_i + 1)f_p + L(\prod_{i=1}^{L}(G_i + 1))^5 d + 2^{\prod_{i=1}^{L}(G_i+1)})$. Similar to BPH, the worst case complexity of RMH is much higher than that of the optimal IP formulation. However, since in most cases CGM terminates in much smaller number of iterations than the total number of transmission scenarios given by $\prod_{i=1}^{L}(G_i + 1)$, the average runtime of the RMH is much smaller than that of the optimal IP formulation as will be demonstrated via extensive simulations in Section 5.8.

5.7.2 Rounding Heuristic

Rounding Heuristic (RH) explores only one branch of the branch-and-price tree. The algorithm starts by solving the LP-relaxation of the exponential IP formulation given in Equations (5.24)-(5.26) using the CGM method. Among the fractional values in the solution, the largest one is chosen setting its value to 1, in contrast to exploring both values of 0 and 1 as in the Branch-and-Price Method. The packet requirements of the links are then updated by considering that the transmission scenario corresponding to the variable is set to 1. In each iteration, the largest fractional value in the solution of the CGM for the LP relaxation of the resulting problem is set to 1 updating the packet requirements of the links accordingly. The algorithm stops when the packet requirements of all the links are met.

RH requires solving CGM on the nodes in one branch of the branch-and-price tree. As the algorithm proceeds on each node of the investigated branch, one of the transmission scenarios is included by setting the corresponding variable to 1, resulting in the transmission of at least one packet. The maximum number of the nodes on the branch is therefore $\sum_{i=1}^{L}(G_i)$. Therefore, the worst case complexity of RH is $O(\prod_{i=1}^{L}(G_i+1)f_p+\sum_{i=1}^{L}(G_i)(L(\prod_{i=1}^{L}(G_i+1))^5 d))$, which is much smaller than that of BPH.

The worst case complexities of the proposed algorithms and the optimal IP formulation (OPT) are summarized in Table 5.1.

OPT	$O(\prod_{i=1}^{L}(G_i+1)f_p+2^{\prod_{i=1}^{L}(G_i+1)})$
BPH	$O(\prod_{i=1}^{L}(G_i+1)f_p+2^{\prod_{i=1}^{L}(G_i+1)+1}(L(\prod_{i=1}^{L}(G_i+1))^5 d))$
RMH	$O(\prod_{i=1}^{L}(G_i+1)f_p+L(\prod_{i=1}^{L}(G_i+1))^5 d+2^{\prod_{i=1}^{L}(G_i+1)})$
RH	$O(\prod_{i=1}^{L}(G_i+1)f_p+\sum_{i=1}^{L}(G_i)(L(\prod_{i=1}^{L}(G_i+1))^5 d))$

Table 5.1: Complexity Comparison of the Algorithms

5.8 Performance Evaluation

The goal of this section is to evaluate the performance of the proposed scheduling algorithms including BPH, RMH and RH together with the power control and rate adaptation algorithms for constant rate, continuous rate and discrete rate transmission models, and compare their performance to that of the Traditional Scheduling (TS) algorithm for constant and discrete rate transmission models. TS algorithm is described as follows: The CGM based scheduling algorithm proposed in [72] is implemented to obtain the optimal schedules neglecting the packet based transmissions. Since the resulting optimal schedules divide the packet transmissions into several data chunks each of which is transmitted in different time slots, the schedules are adapted for packet based transmissions by updating the time slot lengths such that each link transmits an integer number of packets, which is the smallest number greater than or equal to the ratio of the number of bits assigned to that link to the packet length of the link, if the link did not satisfy its packet traffic requirement until that time slot. The performance of the TS algorithm is not included for continuous rate transmission model since no algorithm has been proposed to minimize the schedule length by using the continuous rate transmission model in the literature and CGM cannot be applied to the continuous rate transmission model since the number of columns corresponding to the feasible transmission rates of the links is infinite.

We used MATLAB on a computer with a 2.5 GHz CPU and 4 GB RAM to run the simulations. Simulation results are obtained based on 1000 independent random network topologies where various numbers of links with $1m$ fixed length are uniformly distributed over a square area of 3×3 meters. The attenuation of the links are determined using the path loss model given by $PL(d) = PL(d_0) - 10n \log_{10}\left(\frac{d}{d_0}\right) + Z$, where d is the distance between the transmitter and receiver, d_0 is the reference distance, $PL(d)$ is the path loss at distance d, n is the path loss exponent and Z is a Gaussian random variable with zero mean and σ_z^2 variance. The parameters used in the simulations are $n = 4$, $\sigma_z^2 = 2$ dB^2, $PL(d_0) = 30$ dB, $d_0 = 1$ m, $N_0 = 10^{-8}$ W/Hz, $p_{max} = 10$ mW, $W = 100$ MHz, $R_l = 100$ bits, G_l is randomly determined from the set $\{1, 2, 3\}$ for each $l \in [1, L]$, $\beta_l = 10$ dB for all $l \in [1, L]$ in constant rate transmission model, $\epsilon = 0.001$ in continuous rate transmission model and $\gamma = (-\infty, 10, 20, 30)$ dB in discrete rate transmission model. The transmission rate corresponding to these SINR values are calculated by using Equation (5.2).

Figures 5.3 a), b) and c) show 95% confidence interval bars depicted around the mean of the approximation ratio of the proposed scheduling algorithms including BPH, RMH and RH algorithms and the TS algorithm for constant rate, continuous rate and discrete rate transmission models respectively. The approximation ratio is defined as the ratio of the schedule

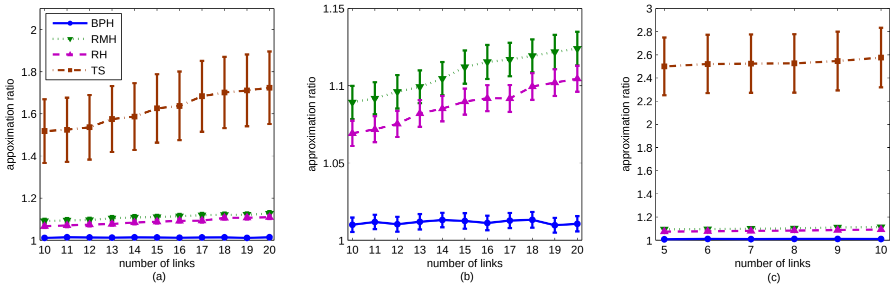

Figure 5.3: Approximation ratio of BPH, RMH, RH and TS scheduling algorithms for a) constant rate, b) continuous rate, c) discrete rate transmission models.

length obtained by the heuristic algorithm to that obtained by OPT. Note that the number of the links is limited to 20 for constant and continuous rate models, and to 10 for discrete rate models since solving the OPT is intractable for networks with higher number of links. For constant and discrete rate transmission models, the proposed algorithms outperform TS algorithm. We observe similar behavior for all the proposed algorithms under all transmission models: The approximation ratio of BPH is very close to 1, i.e. optimal solution, and robust to the increase in the number of the links whereas the approximation ratio of RMH and RH is worse than that of BPH and increases as the number of the links increases. Moreover, the performance of RH is slightly better than RMH since RH continues to search for the best set of transmission scenarios in each iteration whereas RMH uses only the set of transmission scenarios determined at the output of the first CGM.

Figures 5.4 a), b) and c) illustrate the average runtime performance of BPH, RMH, RH and OPT scheduling algorithms for constant rate, continuous rate and discrete rate transmission models respectively. For constant and continuous rate transmission models, the average runtime of the RMH and RH scheduling algorithms increases almost linearly in the number of the links due to the greedy structure of these algorithms and linear increase in the runtime of the corresponding power control algorithms. On the other hand, the runtime of the BPH and OPT algorithms is exponential in the number of the links. However, the exponent of the runtime of the OPT and

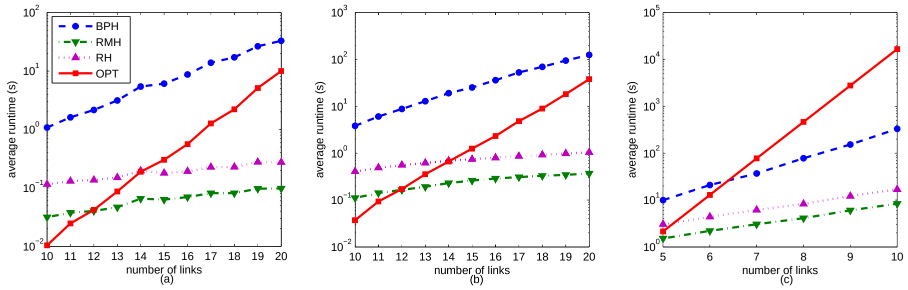

Figure 5.4: Average runtime of BPH, RMH, RH and OPT scheduling algorithms for a) constant rate, b) continuous rate, c) discrete rate transmission models.

BPH algorithms are around 2 and 1.3 respectively, which creates a significant runtime difference as the number of the links grows larger. Although the runtime of the OPT algorithm is lower than that of the BPH algorithm for up to 20 links, when projected, it is expected that the BPH algorithm would outperform OPT algorithm for 22 links and the runtime of the BPH would be negligible compared to the OPT algorithm for 30 links. Note that it is impossible to solve the OPT algorithm for higher than 20 links using MATLAB while the BPH algorithm provides solutions for larger number of the links. The reason for this intractability difference is that the OPT algorithm solves the scheduling problem as a pure IP problem whereas the BPH algorithm searches for the optimal solution over a BPH tree by exploiting the fathoming and the efficiency of the CGM based method together with the proposed heuristic method to solve the pricing problem. Furthermore, the runtime behavior of the scheduling algorithms for the discrete rate transmission model is very different from that of constant and continuous rate models. The average runtime of the OPT algorithm increases dramatically as the number of links increases and becomes 100 times greater than that of the BPH algorithm for 10 links. This is a direct result of the exponential complexity of the FTFA power control algorithm: The OPT algorithm uses FTFA to determine the time slot length for all possible subsets of the links whereas the BPH algorithm uses FTFA for smaller number of links since the average number of concurrently transmitting links in the optimal solution is smaller than the number of the links in the network.

Figure 5.5 shows the length of the schedule for constant rate, continuous rate and discrete rate transmission models normalized by that of the continuous rate model. As expected, the length of the schedule is lowest and highest for continuous rate and constant rate models respectively. The schedule lengths of the discrete rate and constant rate transmission models get closer to that of the continuous rate model as the number of the links increases mainly due to exponentially increasing number of possible link rate vectors in the number of the links leading to the selection of a better link rate vector.

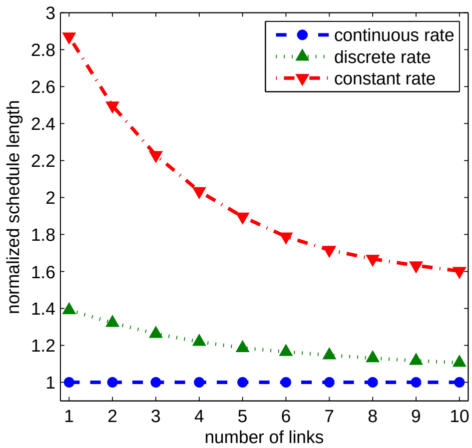

Figure 5.5: Effect of transmission model on the schedule length.

5.9 Conclusion and Discussion

We formulate the joint optimization of power control, rate adaptation and scheduling with the goal of generating minimum length schedule given traffic demands on the links in practical packet based wireless ad hoc networks. In contrast to the traditional approach where the optimization problem ignores the packet level details assigning the transmission of any amount of data to the time slots at different rates, we assume that only an integer number of packets can be transmitted in a time slot. We first formulate the optimal power control and rate adaptation problem with the goal of minimizing the length of the time slot given the number of packets transmitted by the subset of the links and propose both optimal and fast heuristic algorithms. Given the minimum time slot lengths corresponding to all the subset of the links concurrently transmitting any possible number of packets, the joint optimization of power control, rate adaptation and scheduling problem can then be reduced to a pure scheduling problem. The scheduling problem is an IP problem in which the number of variables is exponential in the number of the links in the network. In order to solve this large-scale IP, we propose Branch and Price method and two Column Generation method based heuristic algorithms, and compare their performance to that of the Traditional Scheduling algorithm that ignores the constraints of the packet transmissions via extensive simulations. The proposed algorithms outperform the traditional scheduling algorithm. Moreover, Branch and Price method based heuristic algorithm (BPH) provides performance very close to the optimal solution and robust to the increasing number of the links with much smaller runtime than that of the optimal algorithm enabling the usage of the method in large networks. The Column Generation method based heuristic algorithms on the other hand provide performance slightly worse than the BPH but degrading more as the number of the links increases with runtime much smaller than the BPH.

In the future, we are planning to extend the proposed framework for different objectives such as throughput maximization and energy minimization to provide realistic solutions for various packet based wireless network applications.

Chapter 6

Power Control and Scheduling for Ultra-wideband Wireless Networks

Ultra-wideband (UWB) is a wireless technology often defined to be a transmission from an antenna for which the emitted signal bandwidth exceeds the lesser of 500MHz and 20% of the center frequency. This large bandwidth provides resistance to multi-path fading, power loss due to the lack of line-of-sight and intentional/ unintentional interference therefore achieves robust performance at high data rate and very low transmit power. UWB applications range from multimedia services such as voice and video conversations, video streaming and high-rate data transfer to high reliability time-critical services such as medical monitoring, industrial control and maintenance, intra-vehicular sensor networks. In this study, we focus on the time-critical applications of UWB where the objective is to minimize the length of the schedule, i.e. delay minimization, given the traffic demand of the links in the network in contrast to the commonly studied multimedia applications where the aim is to maximize the throughput while providing a certain level of fairness to the links [105, 106].

In this chapter, we study the optimal power control, rate adaptation and scheduling for ultra-wideband (UWB) wireless networks for two different objectives which are delay minimization and delay constrained energy minimization.

6.1 Introduction

We investigate two different optimization problems for rate-adaptive UWB wireless networks.

First, we study the optimal power control, rate adaptation and scheduling problem for delay minimization subject to data traffic, reliability and transmit power constraints on the links. The main contributions of the study are as follows. We formulate the problem of finding the minimum-length schedule in UWB wireless networks using Linear Programming (LP) problem where the number of variables is exponential in the number of the links. Then, we propose Column Generation Method (CGM) to decompose the original large scale problem into two sub-problems called Restricted Master Problem (RMP) and Pricing Problem (PP). We propose a heuristic method that replaces the intractable optimal PP formulation by a utility maximization algorithm and adapts the exclusion region concept commonly used in UWB systems to the initialization of the RMP. Finally, we demonstrate the speed and accuracy of the proposed algorithm through simulations.

Second, we study the optimal power control, rate adaptation and scheduling for the objective of minimizing energy subject to delay, data traffic, reliability and transmit power constraints on the links. The main contributions are as follows. We determine optimal power and rate allocation for energy minimization with delay constraint in UWB wireless networks. We formulate the delay constrained energy minimization problem as a very large scale Linear Programming (LP) problem with exponential number of variables in the number of the links. Then, we propose Price Minimization based CGM (PM-CGM) scheduling algorithm that decomposes the exponential size LP problem into RMP and PP subproblems. Finally, we exploit the findings obtained in the delay minimization problem to find an initial delay-feasible schedule for initialization of RMP and propose a heuristic method replacing the intractable optimal PP formulation. Simulations illustrate that PM-

CGM decreases the runtime considerably with respect to the optimal LP formulation while achieving very close to optimal solutions.

6.2 Related Work

The scheduling algorithms designed for UWB wireless networks mostly aim at maximizing throughput along with considering fairness among the wireless nodes since they usually consider multimedia applications requiring high-rate data transfer [105, 106, 107, 108]. However, the objective of maximizing throughput cannot be considered to be applicable for time-critical UWB applications where the objective is either delay minimization or delay constrained energy minimization given the data traffic, transmit power and SNIR requirements on the links.

Scheduling for delay constrained systems have been investigated for interference-free and interference-controlling communication schemes. For interference-free scheduling, the studies aim at minimizing energy consumption by determining the optimal packet transmission times and durations where all packets have a common deadline [109, 110] or individual deadlines [111, 112]. The main finding of these studies which is minimization of energy consumption by transmitting the packets in longest possible duration however is not applicable to short range transmissions since the energy consumption due to circuitry during data transmission dominates the energy consumption due to actual data transmission which is shown to be valid for UWB transmissions in [113]. On the other hand, the scheduling algorithms designed for interference-controlling communication schemes aim at determining the best assignment of simultaneous transmissions considering optimal power allocation [114, 115], but not exploiting the rate adaptation for energy minimization. Delay constrained energy minimization has been investigated in a joint framework of optimizing transmission rates, powers and scheduling only for narrowband long-range wireless networks [116].

Determining the best assignment of simultaneous transmissions for optimal scheduling for delay constrained energy minimization necessitates solving very large scale problems since the number of possible concurrently active link sets is exponential in the number of links. Such a large scale problem can be solved using Column Generation Method (CGM) [55] which decomposes it into master and pricing sub-problems and solves rapidly in an iterative scheme. In the context of scheduling in wireless networks, CGM has been investigated for throughput maximization in [117, 118] and delay minimization in [119] where the authors have not considered rate adaptation. The heuristics presented in [119] for fixed-rate scheduling cannot be used for UWB wireless networks. In fact, in the existence of SNIR requirements on the links, fixed-rate allocation based problem formulation is an instance of rate-adaptation based problem formulation where the rate variables are eliminated by assigning them fixed predetermined values.

6.3 Fast Scheduling for Delay Minimization

This section investigates the delay minimization problem for UWB wireless networks.

6.3.1 Joint Optimization Problem

The optimal scheduling, rate adaptation and power control problem for the objective of delay minimization subject to traffic demand, transmit power and SNIR constraints in UWB wireless networks is formulated as follows:

minimize

$$\sum_{n=1}^{N} t^{(n)} \tag{6.1}$$

subject to

$$\sum_{n=1}^{N} t^{(n)} x_l^{(n)} \geq R_l, \quad l \in [1, L] \tag{6.2}$$

$$p_l^{(n)} \leqslant p_{max}, \quad l \in [1, L], n \in [1, N] \tag{6.3}$$

$$x_l^{(n)} \leq K \frac{p_l^{(n)} h_{ll}}{\beta_l \left(N_0 + \sum_{k=1, k \neq l}^{L} p_k^{(n)} h_{kl} \gamma\right)}, \quad l \in [1, L], n \in [1, N] \tag{6.4}$$

$$a_{lk} + 1_{\left\{p_l^{(n)} > 0\right\}} + 1_{\left\{p_k^{(n)} > 0\right\}} \leq 2, \quad l, k \in [1, L], n \in [1, N] \tag{6.5}$$

variables

$$t^{(n)} \geq 0, p_l^{(n)} \geq 0, x_l^{(n)} \geq 0, \quad l \in [1, L], n \in [1, N] \tag{6.6}$$

where L is the number of links, N is the number of time slots, R_l is the data requirement of link l, p_{max} is the maximum allowed power determined by the UWB regulations, K is a system constant mapping SNIR to the achievable rate, N_0 is the background noise energy plus interference from non-UWB systems, h_{ll} is the power gain of link l, h_{kl} is the power gain from the transmitter of link k to the receiver of link l, γ is a parameter depending on the shape and repetition time of the UWB pulse, β_l is the threshold for the ratio of SNIR to the rate for link l determined based on various considerations such as desired bit error rate and modulation schemes, and a_{lk} is a constant that takes the value 1 if links l and k share a common node and 0 otherwise. The variables of the optimization problem are $t^{(n)}$, the length of the n-th time slot; $x_l^{(n)}$, the transmission rate of link l in time slot n; $p_l^{(n)}$, the transmit power of link l in time slot n.

Equation (6.2) represents the data constraint of the links in the network. Equation (6.3) states the upper bound for the transmit power of the links. Equation (6.4) states the upper bound on the transmission rate of the links [120]. This formulation for the maximum achievable transmission rate is based on the UWB characteristics which are adaptability of the transmission rate to the SNIR level at the receiver and linear relation between transmission rate and SNIR level due to very large bandwidth. SNIR calculation is valid for both pulse based time hopping and direct sequence UWB communications and independent of the channel model, i.e. the power gains h_{lk} can be calculated by using the UWB channel model of any application [105, 120, 121]. Finally, Equation (6.5) states that any node in the network cannot transmit to and receive from more than one node simultaneously.

Note that if we fix the transmission rate variables such that $x_l^{(n)}/K = 1$, the optimization problem turns into minimum airtime scheduling problem with airtime traffic demands and SNIR requirements on the links as given in [119]: Equation (6.2) turns into airtime traffic demands on the links and Equations (6.3-6.5) turn into the requirements for a feasible matching of the links.

6.3.2 Optimal Scheduling Formulation

Given the power allocation of the links in a time slot, the achievable rate region for a link given in Equation (6.4) is independent of the rates of the concurrently transmitting links. Since higher rate results in smaller transmission delay, the optimal transmission rate is the maximum achievable rate for delay minimization. Moreover, it has been shown in [122] that any feasible average rate is achievable by $0/p_{max}$ power allocation to the links meaning that any feasible delay of a network can be achieved by $0/p_{max}$ power allocation so no power control is needed for delay minimization in UWB wireless networks. The scheduling problem is therefore independent of the power and rate allocation allowing us to reduce the joint formulation given in Section 6.3.1 to a pure scheduling problem.

Let $\mathcal{E} = \{\mathcal{E}_k : 1 \leq k \leq |\mathcal{E}|\}$ denote the set of all possible subsets of the link set $\mathcal{L} = \{1, 2, ..., L\}$ that can transmit concurrently considering the constraint given in Equation (6.5). Note that $|\mathcal{E}| \leq 2^L$ with equality for the case where no two links share any common node.

Let X denote an $L \times |\mathcal{E}|$ matrix such that the element in the l-th row and k-th column of X denoted by $x_l^{(k)}$ is the optimal transmission rate of link l if link l is in the subset $\mathcal{E}_k$ and 0 otherwise.

The optimal scheduling problem for delay minimization is then formulated as a Linear Programming (LP) problem as follows:

minimize

$$\mathbf{1}^T t \tag{6.7}$$

subject to

$$Xt \geq R \tag{6.8}$$

variables

$$t \geq \mathbf{0} \tag{6.9}$$

where $\mathbf{1}$ is $|\mathcal{E}| \times 1$ all-ones vector, $\mathbf{0}$ is $|\mathcal{E}| \times 1$ all-zeros vector, R is the vector containing the data requirements of the links, i.e. $R = [R_1, R_2, ..., R_L]$. The variable of the LP problem is the vector $t = [t^{(1)}, t^{(2)}, ..., t^{(|\mathcal{E}|)}]$, where $t^{(k)}$ is the time slot length allocated to the subset $\mathcal{E}_k$. Equation (6.8) represents the data requirements of the links in the network.

It requires an exponential effort to both form the X matrix and solve this LP problem since the number of variables is exponential in the number of the links. This approach is intractable for large number of links even with the use of efficient optimization problem solvers such as Fast Lipschitz [123]. On the other hand, by virtue of Caratheodory theorem, the number of non-zero variables in the optimal solution of this LP problem is at most L suggesting the use of the column generation method where the columns of X are generated on-the-fly iteratively as needed.

6.3.3 Exclusive Region and Utility Maximization Based Column Generation Method (EXUM-CGM)

The Column Generation Method (CGM) decomposes large scale original LP problem into two subproblems, a Restricted Master Problem (RMP) and a Pricing Problem (PP), and solves them iteratively.

6.3.3.1 Restricted Master Problem

The Restricted Master Problem is similar to the original problem except that only a small subset of feasible link sets $\mathcal{E}^s \subset \mathcal{E}$ is considered:

minimize

$$\mathbf{1}^T t \tag{6.10}$$

subject to

$$X^s t \geq R \tag{6.11}$$

variables

$$t \geq \mathbf{0} \tag{6.12}$$

where X^s denotes the transmission rate matrix corresponding to the subset $\mathcal{E}^s$ and the length of all-ones vector $\mathbf{1}$, all-zeros vector $\mathbf{0}$ and variable t is $|\mathcal{E}^s| \times 1$.

To start CGM, we need an initial X^s matrix that guarantees the feasibility of the RMP. An intelligent choice of X^s can reduce the number of iterations required in CGM since the solution converges to the optimal solution in each iteration of CGM and the initial X^s matrix determines the distance of the initial solution to the optimal point. We therefore propose the heuristic algorithm called Exclusion Region Based Algorithm (ERBA) based on the exclusion region concept commonly used in UWB network scheduling algorithms to obtain a good initial subset $\mathcal{E}^s$. The algorithm determines the first subset in $\mathcal{E}^s$ by choosing the maximal subset of nodes in $\mathcal{L}$ that are outside the exclusion region of each other then proceed by removing the link that satisfies its data requirement and including maximal subset of additional links satisfying exclusion region constraint to determine the next subset to be included in $\mathcal{E}^s$ in each step. The exclusion region is defined as a static size circle around the receiver of a link such that the nodes inside that circle cannot transmit simultaneously with the transmitter of that link. Size of the exclusion region is determined using the formulations in [106].

Algorithm 12 Exclusion Region Based Algorithm

$\mathcal{E}^s = \emptyset$; $\mathcal{E}_k = \emptyset$; $D = \mathcal{L}$; $k = 1$;
while $D \neq \emptyset$ **do**
  while there are links in D outside the exclusion regions of the links in $\mathcal{E}_k$ **do**
    add an arbitrary link outside the exclusion regions of the links in $\mathcal{E}_k$;
  end while
  $\mathcal{E}^s = \mathcal{E}^s + \mathcal{E}_k$;
  drop the link that satisfies its data requirement from D and $\mathcal{E}_k$;
  $\mathcal{E}_{k+1} = \mathcal{E}_k$; $k = k + 1$;
end while

We can solve RMP and its dual problem with simplex method in polynomial time to obtain its primal optimal solution t^p and dual optimal solution t^d respectively. Since we consider only an arbitrary subset $\mathcal{E}^s \subset \mathcal{E}$, t^p is not necessarily the optimal solution to the original problem. In the original problem, the cost coefficient of every variable in the objective is 1. Then, the reduced cost of a column $x^{(k)}$ in the matrix X but not in X^s is

$$c^{(k)} = 1 - \left(t^d\right)^T x^{(k)} \tag{6.13}$$

If the reduced cost $c^{(k)}$ of any column $x^{(k)}$ is less than 0, the objective function of the RMP can be further reduced by adding the column $x^{(k)}$ to the matrix X^s. Otherwise, the current solution t^p is the optimal solution to the original problem. Instead of calculating the reduced cost of each column $x^{(k)}$ in the matrix X but not in X^s, which requires an exponential effort, we can solve the following Pricing Problem which will generate the column with the minimum reduced cost among all columns in X matrix.

6.3.3.2 Pricing Problem

Pricing Problem is formulated as follows:

maximize

$$\sum_{l=1}^{L} t_l^d K \frac{b_l p_{max} h_{ll}}{\beta_l \left(N_0 + \sum_{k \neq l} b_k p_{max} h_{kl} \gamma \right)} \tag{6.14}$$

subject to

$$a_{lk} + b_l + b_k \leq 2, \quad l, k \in [1, L] \tag{6.15}$$

variables

$$b_l \in \{0, 1\}, \quad l \in [1, L] \tag{6.16}$$

where t_l^d is the l-th element of the vector t^d representing the dual optimal solution of the RMP in the current iteration. The variable of the optimization problem is b_l, which takes value 1 if link l is active and 0 otherwise.

The objective of the optimization problem is to minimize the reduced cost given in Equation (6.13) so maximize $\left(t^d\right)^T x^{(k)}$ over all columns $x^{(k)}$ of the matrix X leading to maximizing $\sum_{l=1}^{L} t_l^d x_l$ over all possible rate columns $[x_1, x_2, ..., x_L]$ in the matrix X. Once the active links are determined through variable b_l for each link l, the rate used in the corresponding column of the matrix X is the maximum achievable rate given in Equation (6.4) with the assigned power allocation p_{max} to the active links as explained in Section 6.3.2, i.e. $x_l = K \frac{b_l p_{max} h_{ll}}{\beta_l \left(N_0 + \sum_{k \neq l} b_k p_{max} h_{kl} \gamma \right)}$, resulting in the objective function given in Equation (6.14). Equation (6.15) represents the constraint on the concurrent transmissions previously given in Equation (6.5) in terms of integer variables b_l.

The PP formulation is a non-linear integer programming problem for which there is no known polynomial-time algorithm. We therefore propose the following heuristic algorithm called Utility Maximization Algorithm (UMA) to efficiently and rapidly solve PP. We define the utility u_S for a set S of links as

$$u_S = \sum_{l \in S} t_l^d K \frac{p_{max} h_{ll}}{\beta_l \left(N_0 + \sum_{k \neq l, k \in S} p_{max} h_{kl} \gamma \right)} \tag{6.17}$$

UMA algorithm iteratively adds links to the initially empty set G such that in each iteration, the link picked is the one that maximizes the utility of the set after its addition. The algorithm stops when the set G includes all the links in the link set $\mathcal{L}$ or the addition of a link cannot increase the utility of the set G. If the utility of the resulting set G returned by the UMA algorithm is higher than 1, the corresponding rate vector is passed to the RMP. Otherwise, the CGM algorithm stops.

Algorithm 13 Utility Maximization Algorithm

$G = \emptyset$; $D = \mathcal{L}$;
while $G \neq \mathcal{L}$ **do**
  if $\max_{i \in D} u_{G+i} > u_G$ **then**
    $k = \arg\max_{i \in D} u_{G+i}$;
    $G = G + k$; $D = D - k$;
  else
    break;
  end if
end while

6.3.4 Performance Evaluation

We used MATLAB on a computer with a 2.5GHz CPU and 4GB RAM to run the simulations. The attenuation of the links are determined using the path loss model given by $PL(d) = PL(d_0) - 10\alpha \log_{10}\left(\frac{d}{d_0}\right) + Z$, where d is the distance between the transmitter and receiver, d_0 is the reference distance, $PL(d_0)$ is the path loss at the reference distance, α is the path loss exponent and Z is a Gaussian random variable with zero mean and σ_z^2 variance. The parameters of the model used in the simulations are $\alpha = 4$, $\sigma_z^2 = 2$, $PL(d_0) =$

$30dB$, $d_0 = 1m$, $N_0 = 10^{-8}W/Hz$, $p_{max} = 10mW$, $\gamma = 10^{-3}$, $K = 10^6$, $\beta_l = 10dB$ for all $l \in \mathcal{L}$. Simulation results are obtained based on 1000 independent random network topologies where various numbers of links with $1m$ fixed length are uniformly distributed over square areas of different sizes indicating different network densities. The reason for choosing fixed length links is to perform a worst case performance comparison with alternative scheduling algorithms.

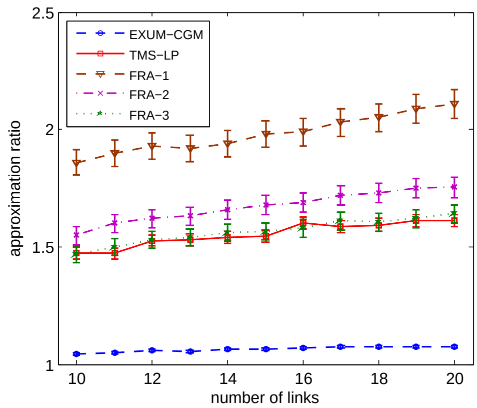

Figure 6.1: Approximation ratio of EXUM-CGM, TMS-LP and FRA algorithms for different number of links.

Fig. 6.1 shows 95% confidence interval bars depicted around the mean of the approximation ratio of EXUM-CGM, Throughput Maximization Set based LP (TMS-LP) and Fixed Rate Allocation (FRA) algorithms for different number of links. The approximation ratio is the ratio of the delay obtained by the heuristic algorithm to the delay obtained by the optimal exponential LP formulation (EXP-LP). In TMS-LP, first the links are assigned to a number of time slots proportional to their traffic weights such that a throughput maximizing subset of links is allocated to each time slot using the exclusion region concept in [105] then the resulting matrix X^s containing only these predetermined throughput maximizing subsets is input to the LP problem given in Equations (6.10) and (6.11) to determine the schedule. In FRA, first all subsets of the links feasible at a predetermined fixed transmission rate are determined then the resulting rate matrix X each column containing the fixed rate in all active links in each feasible subset is input to the LP problem given in Equations (6.7) and (6.8) to determine the schedule. We simulated FRA algorithm for three predetermined rate allocations considering each network separately. Let x_1 and x_2 be the minimum and maximum fixed transmission rate to get a feasible schedule using FRA algorithm. In FRA-1, FRA-2 and FRA-3, fixed transmission rates are calculated as $(2x_1 + x_2)/3$, $(x_1 + x_2)/2$ and $(x_1 + 2x_2)/3$ respectively. The approximation ratio of EXUM-CGM is closer to 1 with a smaller variance than the other algorithms degrading only slightly as the number of links increases. On the other hand, the approximation ratio of TMS-LP algorithm is around 1.5 with higher variance than EXUM-CGM and increases significantly as the number of links increases demonstrating the unsuitability of throughput maximization based identification of the concurrently transmitting subsets for delay minimization problem. The approximation ratios of FRA algorithms are even worse than TMS-LP with higher variances demonstrating the superiority of adaptive rate over fixed rate allocation.

Fig. 6.2 compares the average runtime of EXUM-CGM and EXP-LP. The runtime of EXUM-CGM increases linearly in the number of the links compared to the exponential increase of that of EXP-LP. Fig. 6.3 on the other

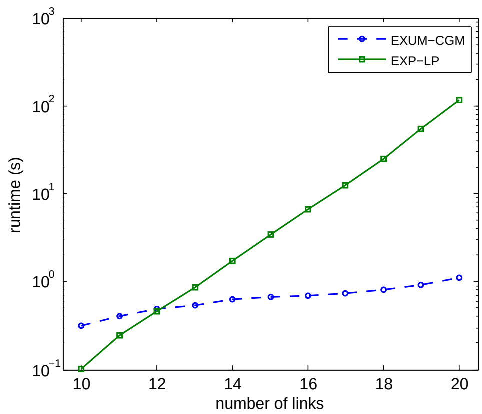

Figure 6.2: Average runtime of EXUM-CGM and EXP-LP for different number of links.

hand shows the robustness of EXUM-CGM to varying network density. For a wide range of network density, EXUM-CGM performs close to optimal and again superior than the other algorithms.

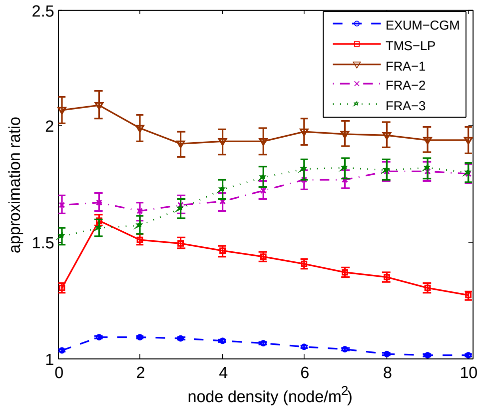

Figure 6.3: Approximation ratio of EXUM-CGM, TMS-LP and FRA algorithms for different network densities.

6.4 Delay Constrained Energy Minimization

This section investigates the delay constrained energy minimization problem for UWB wireless networks.

6.4.1 Joint Optimization Problem

The optimal power control, rate adaptation and scheduling problem for energy minimization in UWB wireless networks subject to total delay, link traffic demand, transmit power and SNIR constraints is formulated as follows:

minimize

$$\sum_{n=1}^{N}\sum_{l=1}^{L} p_l^{(n)} t^{(n)} + \sum_{n=1}^{N}\sum_{l=1}^{L} 1_{\left\{p_l^{(n)}>0\right\}} (p_{tx} + p_{rx}) t^{(n)} \tag{6.18}$$

subject to

$$\sum_{n=1}^{N} t^{(n)} \leq D_{max} \tag{6.19}$$

$$\sum_{n=1}^{N} t^{(n)} x_l^{(n)} \geq R_l, \quad l \in [1, L] \tag{6.20}$$

$$p_l^{(n)} \leqslant \ p_{max}, \quad l \in [1, L], n \in [1, N] \tag{6.21}$$

$$x_l^{(n)} \leq K \frac{p_l^{(n)} h_{ll}}{\beta_l \left(N_0 + \sum_{k=1, k \neq l}^{L} p_k^{(n)} h_{kl} \gamma \right)}, \quad l \in [1, L], n \in [1, N] \tag{6.22}$$

$$a_{lk} + 1_{\left\{p_l^{(n)} > 0\right\}} + 1_{\left\{p_k^{(n)} > 0\right\}} \leq 2, \quad l, k \in [1, L], n \in [1, N] \tag{6.23}$$

where N is the number of time slots, L is the number of links, D_{max} is the maximum delay requirement; i.e., maximum schedule length, R_l is the data requirement of link l, p_{max} is the maximum transmit power due to UWB regulations, p_{tx} and p_{rx} are the constant powers dissipated at the transmitter and receiver respectively during data transmission in active mode, K is a constant representing the linear mapping between SNIR level at the receiver and the achievable transmission rate, N_0 is the background noise, h_{kl} is the power gain from the transmitter of link k to the receiver of link l, γ is a UWB parameter depending on the pulse repetition time and shape, β_l is the SNIR threshold value for link l depending on the desired bit/packet error rate and modulation schemes, and a_{lk} is a constant with value 1 if two links l and k share a common wireless node and 0 otherwise. The variables are $p_l^{(n)}$, the transmission power of link l in time slot n; $x_l^{(n)}$, the transmission rate of link l in time slot n and $t^{(n)}$, the length of time slot n.

The objective of the optimization problem is to minimize the total energy consumption in the network. Equation (6.19) states the maximum delay constraint for the schedule to be constructed. Equations (6.20) and (6.21) represent the traffic demand and maximum transmission power constraints of the links respectively. Equation (6.22) is the achievable rate formulation for the links based on the rate adaptivity characteristic of UWB such that the transmitter of a link can adapt its transmission rate linearly to the SNIR level achieved at the receiver due to very large bandwidth. Equation (6.23) states that a particular node in the network can be either the transmitter or receiver end of, at most, one active link in a time slot.

Since it is hard to solve this general non-convex programming problem in this joint optimization framework, it is beneficial to solve optimally for each set of variables independently.

6.4.2 Optimal Rate and Power Allocation

For a fixed power allocation, using maximum achievable rate satisfying Equation (6.22) is optimal since increasing the transmission rate decreases the time duration required for the transmission of a fixed amount of data, which consequently decreases both the delay and energy consumption of a link. More important issue for the foregoing joint optimization problem is optimal power allocation since power allocation of a specific link interests all other links by means of interference it creates.

It is shown in [105] that any feasible rate allocation or average power consumption can be achieved with $0/p_{max}$ allocation. However, it is not obvious that the $0/p_{max}$ allocation is the optimum choice in delay constrained energy minimization problem. Suppose that we have achieved a schedule for minimizing energy while satisfying the delay constraint with an arbitrary power allocation. According to [105], it is possible to achieve the same delay with a new schedule in which the power allocation is limited to $0/p_{max}$ allocation. It is also possible to achieve the same energy consumption with a new schedule in which the power allocation is limited to $0/p_{max}$ allocation. However, schedules constructed with $0/p_{max}$ allocation for the aforementioned purposes are not necessarily the same, i.e., while we are constructing a new schedule based on $0/p_{max}$ power allocation in order to achieve the same delay, energy consumption may increase and vice versa. For this reason, we state the optimality of power allocation for delay constrained energy minimization.

Theorem 9. *In the optimal solution for the delay constrained energy minimization problem described in Section 6.4.1, each link is either active with maximum transmission power or inactive in a time slot.*

Proof: Suppose that an arbitrary schedule satisfying the constraints in the delay constrained energy minimization problem is constructed for a network. Let a be the duration of a time slot in which links $\{1, 2, .., m\}$ are active with corresponding power and rate allocations $\{p_1, p_2, .., p_m\}$ and $\{x_1, x_2, .., x_m\}$

respectively. Let f_k denote the number of data bits transmitted by link $k \in [1, m]$ in this time slot. Let E be the total energy consumption in this time slot. We will show that we can separate this time slot into two time slots such that when any link with an arbitrary power level is assigned to transmit power p_{max} and 0 in the first and second slot respectively while keeping the transmit power of the remaining links the same, the time duration of the first and second time slots denoted by $a^{(1)}$ and $a^{(2)}$ respectively required for the transmission of total f_k data bits by every link $l \in [1, m]$ results in total delay less than a, i.e. $a^{(1)} + a^{(2)} \leqslant a$, and energy consumption less than E.

First, we will show that $a^{(1)} + a^{(2)} \leqslant a$. Assume that in the first time slot the transmit power p_i of an arbitrary link i is assigned to p_{max} while keeping the transmit powers of the remaining links the same. The data rate of link i is $x_i^{(1)} = (p_{max}/p_i)\, x_i$ whereas the data rate of link j where $j \neq i$ is given by

$$x_j^{(1)} = \frac{N_0 + U_j + p_i C}{N_0 + U_j + p_{max} C} x_j \tag{6.24}$$

where U_j denotes the interference of the links except link i on link j and $C = h_{ij}\gamma$. Here, it is obvious that only link i satisfies its data requirement in the first time slot since $x_i^{(1)} \geq x_i$ and $x_j^{(1)} \leq x_j$ where $j \neq i$ therefore $a^{(1)} = (p_i/p_{max})\, a$. In order to satisfy the data requirements of the links other than link i, these links are allocated in the second time slot. The data rate of link j where $j \neq i$ in the second time slot is given by

$$x_j^{(2)} = \frac{N_0 + U_j + p_i C}{N_0 + U_j} x_j \tag{6.25}$$

For an arbitrary link j to satisfy its data requirement, the equation given as

$$a x_j = a^{(1)} x_j^{(1)} + a_j^{(2)} x_j^{(2)} \tag{6.26}$$

should be satisfied where $a_j^{(2)}$ denotes the actual amount of time required for link j to meet its data requirement. Substituting $a^{(1)}$ with $(p_i/p_{max})\, a$, $a_j^{(2)}$ is given by

$$a_j^{(2)} = a\{\frac{N_0 + U_j}{N_0 + U_j + p_i C} - \frac{p_i}{p_{max}} \frac{N_0 + U_j}{N_0 + U_j + p_{max} C}\} \tag{6.27}$$

Since we want all the links to satisfy their data requirements, $a^{(2)} = \max_{j \in [1,m], j \neq i} a_j^{(2)}$. Suppose that this maximum is achieved for link k. Then, the total length of

$a^{(1)}$ and $a^{(2)}$, say $a^{'}$, is given by

$$a^{'} = a\{1 - \frac{p_i C}{N_0 + U_k + p_i C} + \frac{p_i C}{N_0 + U_k + p_{max} C}\} \tag{6.28}$$

It is clear that $a^{'} \leq a$ for every pair i,k and equality holds only if $p_i = p_{max}$. We can conclude that any link with an arbitrary power level can be allocated with 0/ p_{max} transmit power allocation resulting in a new schedule with less delay.

Now, we will investigate the energy consumption. Energy consumption in time slot of duration a is given by

$$E = a \sum_{j=1}^{m} \{p_j + p_{tx} + p_{rx}\} \tag{6.29}$$

Let $E^{(1)}$ and $E^{(2)}$ denote the energy consumption in the first and second time slots respectively. $E^{(1)}$ and $E^{(2)}$ are formulated as follows:

$$E^{(1)} = a^{(1)} \sum_{j=1}^{m} \{p_j + p_{tx} + p_{rx}\} + a^{(1)}\{p_{max} - p_i\} \tag{6.30}$$

$$E^{(2)} = a^{(2)} \sum_{j=1}^{m} \{p_j + p_{tx} + p_{rx}\} - a^{(2)} p_i \tag{6.31}$$

The total energy consumption, say $E^{'}$, is then given by

$$E^{'} = a^{'} \sum_{j=1}^{m} \{p_j + p_{tx} + p_{rx}\} + a^{(1)} p_{max} - \{a^{(1)} + a^{(2)}\} p_i \tag{6.32}$$

Substituting p_{max} with $p_i a / a^{(1)}$, we get

$$E^{'} = a^{'} \sum_{j=1, j \neq i}^{m} \{p_j + p_{tx} + p_{rx}\} + a p_i + a^{'}\{p_{tx} + p_{rx}\} \tag{6.33}$$

Since $a^{'} \leq a$, it follows that $E^{'} \leq E$ and equality holds only if $p_i = p_{max}$. $\square$

6.4.3 Optimal Scheduling

Given the optimal rate and power allocation provided in Section 6.4.2, the joint optimization problem presented in Section 6.4.1 can be reduced to a pure scheduling problem. However, since we do not know the sets of concurrently transmitting links prior to scheduling, we need to consider all possible subsets of links for concurrent transmission as described next.

Let $\mathcal{E} = \{\mathcal{E}_k : 1 \leq k \leq |\mathcal{E}|\}$ denote the set of all feasible subsets of the link set $\mathcal{L} = \{1, 2, ..., L\}$. Note that $|\mathcal{E}|$ is equal to 2^L if any two links do not share a common node. Let $\mathbf{X}$ be an $L \times |\mathcal{E}|$ transmission rate matrix, where the element x_{lk} of $\mathbf{X}$ is the optimal transmission rate of link l in link set $\mathcal{E}_k$ given as

$$x_{lk} = \frac{p_{max} h_{ll}}{\beta_l} \left(N_0 + \sum_{i \in \{\mathcal{E}_k / l\}} p_{max} h_{il} T_f \gamma \right)^{-1} \tag{6.34}$$

Notice that each column of the matrix $\mathbf{X}$ gives the optimal transmission rates of the concurrently active links in a subset. Furthermore, let $\mathbf{e}$ denote a $1 \times |\mathcal{E}|$ vector whose k-th element is equal to the number of active links in link set $\mathcal{E}_k$.

Then, optimal scheduling problem for energy minimization with delay constraint can be formulated as a Linear Programming (LP) problem as:

minimize

$$(p_{max} + p_{tx} + p_{rx})\mathbf{et} \tag{6.35}$$

subject to

$$\mathbf{Xt} \geq \mathbf{R} \tag{6.36}$$

$$-\mathbf{1t} \geq -D_{max} \tag{6.37}$$

where $\mathbf{1}$ is a $1 \times |\mathcal{E}|$ all-ones vector and $\mathbf{R}$ is a $L \times 1$ vector containing the traffic demand requirements of the links, i.e. $\mathbf{R} = [R_1, R_2, ..., R_L]$. The variable of the problem is $|\mathcal{E}| \times 1$ vector $\mathbf{t}$ whose k-th element is the time slot duration allocated for the subset $\mathcal{E}_k$. Equations (6.36) and (6.37) represent the traffic demand requirements of the links and delay requirement of the schedule respectively.

There are two difficulties in solving this optimization problem although it is an LP problem. An exponential time effort is required to form $\mathbf{X}$ matrix and solve the corresponding LP since the number of link subsets so the number of variables is exponential in the number of links. These difficulties can be overcome by reducing the size of $\mathcal{E}$ thus dealing with a small size rate matrix $\mathbf{X}$.

6.4.4 Price Minimization Based Column Generation Method (PM-CGM)

The intractability of the exponential size LP formulation given in Section 6.4.3 can be overcome by using Column Generation Method (CGM). In CGM, the large scale original LP problem is decomposed into a Restricted Master Problem (RMP) and a Pricing Problem (PP) and the original problem can be solved in an iterative way. We start with an initial feasible RMP which is a restricted; i.e., small-scale, version of the exponential LP formulation. Then, we pass the dual solution of RMP to PP. If the optimal solution of the PP can be used to improve the solution of the RMP, we pass the vector corresponding to the optimal solution to RMP by adding it to the constraint matrix of RMP as a column. Then RMP is solved again and its dual is passed to PP to make PP generate another column. This iterative behaviour continues until the optimal solution of the PP cannot be used to improve the solution of the RMP. This column generation idea is the center of PM-CGM algorithm.

6.4.4.1 Restricted Master Problem (RMP)

We reduce the original problem to the following restricted problem given as

minimize

$$\boldsymbol{e}^s\mathbf{t} \tag{6.38}$$

subject to

$$\boldsymbol{X}^s\mathbf{t} \geq \mathbf{R} \tag{6.39}$$

$$-\mathbf{1}\mathbf{t} \geq -D_{max} \tag{6.40}$$

The difference between the original problem and the foregoing restricted problem is that only a subset of $\mathcal{E}^s \subset \mathcal{E}$ is considered in the restricted problem and $\boldsymbol{X}^s$ and $\boldsymbol{e}^s$ denote the transmission rate matrix and the vector of numbers of active links corresponding to the subset $\mathcal{E}^s$ respectively. Note that we have removed the constant multiplier $(p_{max} + p_{tx} + p_{rx})$ in the objective since it does not affect the optimal solution. For the first iteration of PM-CGM, we need an initial transmission rate matrix $\boldsymbol{X}^s$ that guarantees a feasible solution for the RMP; i.e, both data requirements of the links and total delay requirement for the schedule can be satisfied. However, this is not a straightforward problem. For now, we assume that such an initial matrix is available. We will discuss this in Section 6.4.4.3.

RMP can be solved using simplex method in polynomial time and primal optimal solution $\boldsymbol{t}^p$ and dual optimal solution $\boldsymbol{t}^d$ corresponding to the dual of RMP can be obtained. Note that $\boldsymbol{t}^p$ may not be the optimal solution of the original problem; i.e., $\boldsymbol{t}^p$ is the optimal solution only when $\mathcal{E}^s$ contains the optimal sets. The reduced cost of a column in the original constraint matrix of the problem given by Equations (6.35-6.37) but not in the restricted constraint matrix of the problem given by Equations (6.38-6.40) is

$$c_k = e_k - (\sum_{i=1}^{L} t_i^d x_{ik}) + t_{L+1}^d \tag{6.41}$$

where t_l^d is the l-th element of the optimal dual solution of the RMP. Negativity of the reduced cost of any column in the original constraint matrix implies that the optimal solution of the RMP can be improved; i.e. decreased, by adding the column with the negative reduced cost to the restricted constraint matrix. If there is not a column with negative reduced cost, the solution $\boldsymbol{t}^p$ of the RMP is the optimal solution to the original large scale problem. In order to find if there is such a column, one needs to solve the pricing problem discussed next.

6.4.4.2 Pricing Problem (PP)

Pricing Problem that is used to generate a column with the minimum reduced cost is formulated as

minimize

$$\sum_{l=1}^{L} b_l - \sum_{l=1}^{L} t_l^d K \frac{b_l p_{max} h_{ll}}{\beta_l \left(N_0 + \sum_{k=1,k\neq l}^{L} b_k p_{max} h_{kl} \gamma \right)} + t_{L+1}^d \tag{6.42}$$

subject to

$$a_{lk} + b_l + b_k \leq 2, \quad l, k \in [1, L] \tag{6.43}$$

$$b_l \in \{0, 1\}, \quad l \in [1, L] \tag{6.44}$$

The variables of the optimization problem are b_l for each link l, which acts as an indicator for activity of link l; i.e., takes value 1 if link l is active and 0 otherwise.

PP minimizes the reduced cost given by Equation (6.41) over all columns $\boldsymbol{x}^{(k)}$ of the original transmission rate matrix $\boldsymbol{X}$ subject to Equation (6.43) representing the availability of concurrent transmissions originally given in Equation (6.23) using integer variables b_l. After determining b_l for each link l, $\boldsymbol{x}^{(k)}$ is the column vector of maximum achievable rates given by Equation (6.22) with the assigned power allocation $b_l p_{max}$, i.e. $x_l = K \frac{b_l p_{max} h_{ll}}{\beta_l \left(N_0 + \sum_{k=1, k \neq l}^{L} b_k p_{max} h_{kl} \gamma\right)}$. Note that, one must append a (-1) to the column $\boldsymbol{x}^{(k)}$ before passing to the RMP to completely characterize a column in the constraint matrix due to the delay constraint represented by Equation (6.40). Moreover, cost coefficient of the variable corresponding to the vector passed to the RMP is $\sum_{l=1}^{L} b_l$.

The foregoing PP formulation is a non-linear integer programming problem which is NP-hard and cannot be solved in polynomial time. In order to deal with this complexity problem, we propose Price Minimization Algorithm (PMA) to solve PP fast and efficiently. We define the price $P_{\mathcal{S}}$ for a set $\mathcal{S}$ of links as

$$P_{\mathcal{S}} = |\mathcal{S}| - \sum_{l \in S} t_l^d K \frac{p_{max} h_{ll}}{\beta_l \left(N_0 + \sum_{k \neq l, k \in S} p_{max} h_{kl} \gamma\right)} + t_{L+1}^d \tag{6.45}$$

In each iteration, PMA algorithm picks one link to add to set $\mathcal{S}$ such that this link minimizes the price of the set after it is added to the set (Lines 4-5). The algorithm terminates either when the set $\mathcal{S}$ includes all the links in the link set $\mathcal{L}$ (Line 2) or the price of the set $\mathcal{S}$ cannot further decrease by the addition of a link (Line 3). If the price of the output set $\mathcal{S}$ returned by the PMA is less than 0, the vector of transmission rates corresponding to set $\mathcal{S}$, with a (-1) appended as the last element, is passed to the RMP

Price Minimization Algorithm (PMA)

$\mathcal{S} = \emptyset$; $\mathcal{S}' = \mathcal{L}$;
while $\mathcal{S} \neq \mathcal{L}$ **do**
  if $\min_{i \in \mathcal{S}'} P_{\mathcal{S}+\{i\}} < P_{\mathcal{S}}$ **then**
    $k = \mathrm{argmin}_{i \in \mathcal{S}'} P_{\mathcal{S}+\{i\}}$;
    $\mathcal{S} = \mathcal{S} + \{k\}$;
    $\mathcal{S}' = \mathcal{S}' - \{k\}$;
  else
    break;
  end if
end while

since it decreases the objective of the RMP; i.e., the energy consumption. Otherwise, the PM-CGM algorithm terminates.

6.4.4.3 Determination of a Feasible Initial Matrix for RMP

PM-CGM algorithm requires an initial transmission rate matrix $\boldsymbol{X}^s$ that guarantees a feasible solution for the RMP as stated in Section 6.4.4.1. In order to produce an $\boldsymbol{X}^s$ matrix that satisfies both data and delay requirements, we propose solving the delay minimization problem given as

minimize

$$\mathbf{1t} \tag{6.46}$$

subject to

$$\mathbf{Xt} \geq \mathbf{R} \tag{6.47}$$

Feasibility condition of the energy minimization problem with delay constraint is equivalent to the condition that the optimal solution of the above delay minimization problem is less than the delay requirement D_{max}. Hence, if we solve this delay minimization problem, the columns in $\mathbf{X}$ corresponding to the positive variables in the optimal solution $\mathbf{t}$ constitutes a feasible initial matrix $\boldsymbol{X}^s$ if the optimal delay is less than D_{max}. However, since this is again a large scale LP programming problem with exponential number of variables in the number of links, it is intractable. In Section 6.3, we have proposed a column generation based fast scheduling algorithm for the delay minimization problem with data constraints on the links. The solution procedure is directly applicable here. Note that we can stop at the particular iteration of the column generation that returns a schedule with delay less than D_{max} since it suffices to create a schedule with length less than or equal to the delay constraint D_{max}.

6.4.5 Performance Evaluation

Simulations are performed on a computer with a CPU of 2.5GHz processing speed and 4GB RAM in MATLAB. The path loss model used is $PL(d) = PL(d_0) - 10\alpha \log_{10}\left(\frac{d}{d_0}\right) + Z$ where d is the distance between the transmitter and receiver, d_0 is the reference distance with value 1m, $PL(d)$ is the path loss at distance d, α is the path loss exponent and Z is a Gaussian random variable with zero mean and σ_z^2 variance. The simulation parameters are given in Table-6.1.

α	4	σ_z^2	2	p_{max}	10mW
β_l	10 dB	$PL(d_0)$	30 dB	p_{tx}	30mW
K	10^6	N_0	10^{-8} W/Hz	p_{rx}	100mW

Table 6.1: Simulation Parameters

For a better characterization of the algorithm performance, we have performed the simulations on 1000 independent random network topologies for each simulation scenario. As an alternative scheduling algorithm, we use Fixed Rate Allocation algorithm, denoted as FRA. In FRA algorithm, among all possible sets of links, the feasible ones are determined based on a fixed predetermined rate $\mathbf{r}$. The fixed transmission rate $\mathbf{r}$ is determined as the average of the minimum and the maximum transmission rates for which a feasible schedule can be constructed. The aim of using FRA algorithm for comparison is to illustrate the superiority of rate-adaptivity of PM-CGM over fixed rate allocation schemes.

Figure-6.4 illustrates the approximation ratio performance of PM-CGM and FRA algorithms for different number of links with 95% confidence intervals depicted around the means where approximation ratio is the ratio of the energy consumption achieved by the algorithm to the minimum energy consumption obtained by solving the exponential LP (EXP-LP) formulation given in Equations (6.35-6.37). The approximation ratio of PM-CGM scheduling algorithm is very close to 1; i.e. optimal solution, and outperforms FRA algorithm significantly for different number of links. In addition, the approximation ratio of PM-CGM is robust to the increase in the number of the links whereas the approximation ratio of the FRA algorithm increases as the number of links increases. The small lengths of the confidence interval bars also demonstrate the robustness of PM-CGM to different topologies.

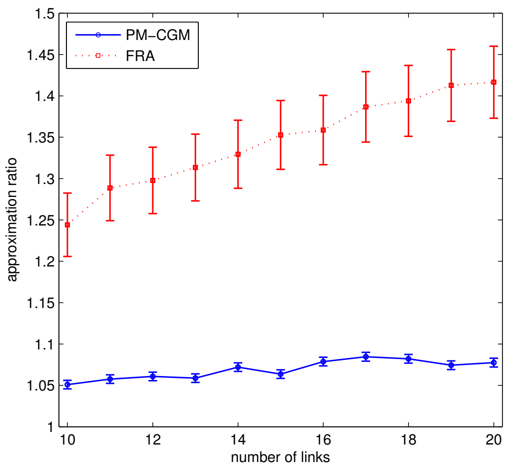

Figure 6.4: Approximation ratio of PM-CGM and FRA algorithms for different number of links with 95% confidence intervals depicted around the means.

Besides performing very close to optimal, PM-CGM decreases the running time required to solve EXP-LP to a greater extent as Figure-6.5 illustrates. The runtime of PM-CGM increases linearly whereas the runtime of EXP-LP increases exponentially with the number of links. The average runtime of PM-CGM is only 1% of the average runtime of EXP-LP for a network of 20 links. Note that the reason of performing simulations for up to 20 links is the intractability of EXP-LP for more than 20 links; i.e., it is not possible to solve EXP-LP in MATLAB.

Figure-6.6 illustrates the robustness of the performance of PM-CGM to varying network densities. The PM-CGM is closer to the optimal at low and

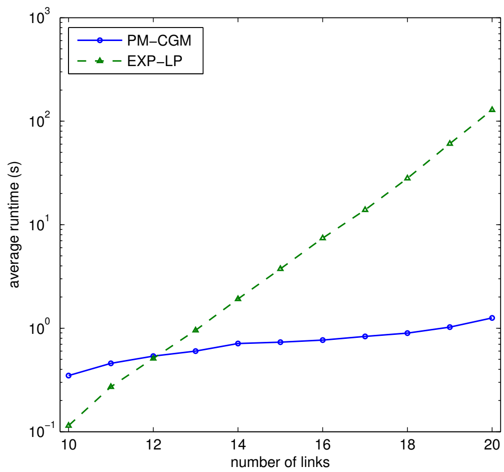

Figure 6.5: Average running time comparison of PM-CGM and EXP-LP for different number of links.

high network densities that correspond to the cases where the links transmit all simultaneously and one by one respectively since optimality of concurrent transmissions is more obvious in these cases compared to medium network densities.

6.5 Conclusion and Discussion

In this chapter, we study the optimal power control, rate adaptation and scheduling problem for delay minimization and delay constrained energy minimization in UWB wireless networks subject to link traffic demand, transmit power and SNIR requirements. Upon stating the optimal rate and power allocation analytically; i.e., each link is either active with maximum power and corresponding maximum achievable rate or inactive at a time at the optimal solution, the scheduling problems are formulated as large scale LP problems of exponential size in the number of the links. We then use the elegant column generation method in the optimization theory to solve these large scale LP problems and propose fast scheduling algorithms performing very close to optimal and decreasing the runtime required to solve the exponential LP problem significantly. Promisiming future directions are the adaptation

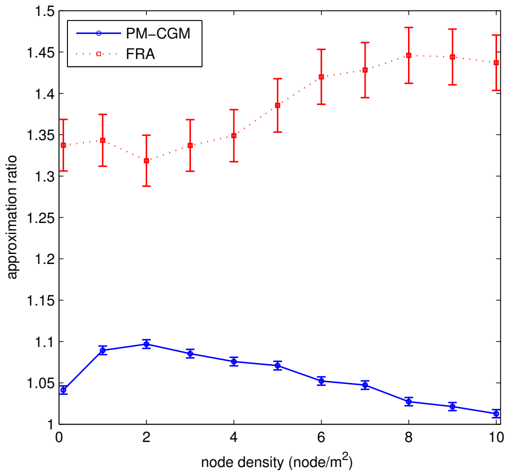

Figure 6.6: Approximation ratio of PM-CGM and FRA algorithms for different node densities with 95% confidence intervals depicted around the means.

of CGM for the optimization problems with different objectives and constraints and to propose a general framework for the application of CGM to any wireless network scheduling problem with any constraint depending on the specific application requirements.

Chapter 7

Conclusion

In this thesis, we have investigated the resource allocation in delay and energy constrained wireless networks. We have provided novel optimization frameworks for the joint optimization of power control, rate adaptation and scheduling in WNCSs. We first investigated the joint optimization of controller and communication systems encompassing efficient abstractions of both systems for WNCSs, which may lead to broader adoption and real-world deployment, and proposed energy efficient resource allocation methods. Second, we have introduced a novel approach to provide maximum level of adaptivity accommodating packet losses and changes in network topology while exploiting periodic nature of the sensor node transmissions in WNCSs. Third, we have revisited the minimum length scheduling problem in wireless ad hoc networks introducing a different constraint from the literature which is packetized transmissions meaning that packet transmission of a node cannot be interrupted and divided into multiple chunks to be transmitted in different time slots. We have proposed optimal and heuristic solutions for this new problem. Lastly, we have proposed optimal power control and rate allocation strategies and fast scheduling algorithms for delay minimization and delay constrained energy minimization problems in UWB wireless networks.

We believe that our studies on WNCSs are very promising for the emerging field of machine to machine communications over next generation cellular systems which is commonly referred as 5G. This is due to the similarity in the application level requirements in both systems and the traffic characteristics of the sensors in WNCSs and machine type devices in M2M. In future, we plan to incorporate our findings and solution approaches in WNCSs to M2M communication to be a part of the efforts on the realization of the fully M2M capable 5G communication systems.

References

[1] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 138–162, Jan. 2007.

[2] A. Willig, "Recent and emerging topics in wireless industrial communication," *IEEE Transactions on Industrial Informatics*, vol. 4, no. 2, pp. 102–124, May 2008.

[3] J. R. Moyne and D. M. Tilbury, "The emergence of industrial control networks for manufacturing control, diagnostics, and safety data," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 29–47, Jan. 2007.

[4] A. Aswani, N. Master, J. Taneja, D. Culler, and C. Tomlin, "Reducing transient and steady state electricity consumption in HVAC using learning-based model-predictive control," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 240–253, Jan. 2012.

[5] *Minimum Aviation System Performance Standard for Automatic Dependent Surveillance Broadcast (ADS-B)*, RTCA, 2002, DO-242A.

[6] H. Li, L. Lai, and H. V. Poor, "Multicast routing for decentralized control of cyber physical systems with an application in smart grid," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 6, pp. 1097–1107, July 2012.

[7] *ISA-100.11a-2009 Wireless systems for industrial automation: Process control and related applications*, ISA, 2009.

[8] *WirelessHART data sheet*, HART Communication Foundation, 2007, http://www.hartcomm2.org/hartprotocol/wirelesshart/wirelesshartmain.html.

[9] R. Steigman, and J. Endresen, "Introduction to WISA and WPS, WISA-wireless interface for sensors and actuators and WPS-wireless

proximity switches," *White paper*, 2004, http://www.eit.uni-kl.de/litz/WISA.pdf.

[10] H. Shariatmadari, R. Ratasuk, S. Iraji, A. Laya, T. Taleb, R. Jantti, and A. Ghosh, "Machine-type communications: current status and future perspectives toward 5g systems," *Communications Magazine, IEEE*, vol. 53, no. 9, pp. 10–17, September 2015.

[11] A. Biral, M. Centenaro, A. Zanella, L. Vangelista, and M. Zorzi, "The challenges of {M2M} massive access in wireless cellular networks," *Digital Communications and Networks*, vol. 1, no. 1, pp. 1 – 19, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S235286481500005X

[12] Y. Sadi, S. C. Ergen, and P. Park, "Minimum energy data transmission for wireless networked control systems," *Wireless Communications, IEEE Transactions on*, vol. 13, no. 4, pp. 2163–2175, April 2014.

[13] Y. Sadi and S. Coleri Ergen, "Joint optimization of communication and controller components of wireless networked control systems," in *Communications (ICC), 2015 IEEE International Conference on*, June 2015, pp. 6487–6493.

[14] Y. Sadi and S. Ergen, "Joint optimization of wireless network energy consumption and control system performance in wireless networked control systems," *Wireless Communications, IEEE Transactions on*, submitted.

[15] Y. Sadi and S. Coleri Ergen, "Energy and delay constrained maximum adaptive schedule for wireless networked control systems," *Wireless Communications, IEEE Transactions on*, vol. 14, no. 7, pp. 3738–3751, July 2015.

[16] Y. Sadi and S. Ergen, "Minimum length scheduling with packet traffic demands in wireless ad hoc networks," *Wireless Communications, IEEE Transactions on*, vol. 13, no. 7, pp. 3738–3751, July 2014.

[17] ——, "Fast scheduling for delay minimization in uwb wireless networks," *Communications Letters, IEEE*, vol. 16, no. 9, pp. 1400 –1403, september 2012.

[18] Y. Sadi and S. Coleri Ergen, "Delay constrained energy minimization in uwb wireless networks," in *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*, April 2013, pp. 597–602.

[19] P. Varaiya, “Smart cars on smart roads: Problems of control,” *IEEE Transactions on Automatic Control*, vol. 38, no. 2, pp. 195–207, Feb. 1993.

[20] R. A. Gupta and M. Chow, “Networked control system: Overview and research trends,” *IEEE Transactions on Industrial Electronics*, vol. 57, no. 7, pp. 2527–2535, July 2010.

[21] N. Pereira, B. Andersson, and E. Tovar, “Widom: A dominance protocol for wireless medium access,” *IEEE Transactions on Industrial Informatics*, vol. 3, no. 2, pp. 120–130, May 2007.

[22] H. Hoang, G. Buttazzo, M. Jonsson, and S. Karlsson, “Computing the minimum EDF feasible deadline in periodic systems,” in *IEEE Real-Time Systems Symposium*, Aug. 2006, pp. 125–134.

[23] Y. Wu, G. Buttazzo, E. Bini, and A. Cervin, “Parameter selection for real-time controllers in resource-constrained systems,” *IEEE Transactions on Industrial Informatics*, vol. 6, no. 4, pp. 610–620, Nov. 2010.

[24] M. Baldi, R. Giacomelli, and G. Marchetto, “Time-driven access and forwarding for industrial wireless multihop networks,” *IEEE Transactions on Industrial Informatics*, vol. 5, no. 2, pp. 99–112, May 2009.

[25] B. Demirel, Z. Zou, P. Soldati, and M. Johansson, “Modular co-design of controllers and transmission schedules in wirelesshart,” in *IEEE Conference on Decision and Control and European Control Conference*, December 2011.

[26] P. Park, J. Araujo, and K. H. Johansson, “Wireless networked control system co-design,” in *IEEE International Conference on Networking, Sensing and Control (ICNSC)*, April 2011, pp. 486–491.

[27] J. Bai, E. P. Eyisi, F. Qiu, Y. Xue, and X. D. Koutsoukos, “Optimal cross-layer design of sampling rate adaptation and network scheduling for wireless networked control systems,” in *ACM/IEEE Third International Conference on Cyber-Physical Systems (ICCPS)*, April 2012, pp. 107–116.

[28] X. Liu and A. Goldsmith, “Wireless network design for distributed control,” in *IEEE Conference on Decision and Control*, Dec. 2004, pp. 2823–2829.

[29] E. Uysal-Biyikoglu, B. Prabhakar, and A. E. Gamal, "Energy-efficient packet transmission over a wireless link," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 487–499, Aug. 2002.

[30] A. Fu, E. Modiano, and J. N. Tsitsiklis, "Optimal transmission scheduling over a fading channel with energy and deadline constraints," *IEEE Transactions on Wireless Communications*, vol. 5, no. 3, pp. 630–641, March 2006.

[31] M. A. Antepli, E. Uysal-Biyikoglu, and H. Erkan, "Optimal packet scheduling on an energy harvesting broadcast link," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1721–1731, Sept. 2011.

[32] O. Ozel, J. Yang, and S. Ulukus, "Optimal broadcast scheduling for an energy harvesting rechargeable transmitter with a finite capacity battery," *IEEE Transactions on Wireless Communications*, vol. 11, no. 6, pp. 2193–2203, June 2012.

[33] W. Chen, M. J. Neely, and U. Mitra, "Energy-efficient transmissions with individual packet delay constraints," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2090–2109, May 2008.

[34] X. Zhong and C. Wu, "Energy-efficient wireless packet scheduling with quality of service control," *IEEE Transactions on Mobile Computing*, vol. 6, no. 10, pp. 1158–1170, Oct. 2007.

[35] M. A. Zafer and E. Modiano, "A calculus approach to energy-efficient data transmission with quality-of-service constraints," *IEEE/ACM Transactions on Networking*, vol. 17, no. 3, pp. 898–911, June 2009.

[36] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2349–2360, Sept. 2005.

[37] T. Arampatzis, J. Lygeros, and S. Manesis, "A survey of applications of wireless sensors and wireless sensor networks," in *Proc. of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation*, 2005, pp. 719–724.

[38] S. C. Ergen and P. Varaiya, "Pedamacs: Power efficient and delay aware medium access protocol for sensor networks," *IEEE Transactions on Mobile Computing*, vol. 5, no. 7, pp. 920–930, July 2006.

[39] G. Walsh, O. Beldiman, and L. Bushnell, "Asymptotic behavior of nonlinear networked control systems," *IEEE Transactions on Automatic Control*, vol. 46, no. 7, pp. 1093–1097, July 2001.

[40] G. Walsh, H. Ye, and L. Bushnell, "Stability analysis of networked control systems," *IEEE Transactions on Control Systems Technology*, vol. 10, no. 3, pp. 438–446, May 2002.

[41] D. Carnevale, A. R. Teel, and D. Nesic, "A lyapunov proof of an improved maximum allowable transfer interval for networked control systems," *IEEE Transactions on Automatic Control*, vol. 52, no. 5, pp. 892–897, May 2007.

[42] W. P. M. H. Heemels, A. R. Teel, N. van de Wouw, and D. Nesic, "Networked control systems with communication constraints: Tradeoffs between transmission intervals, delays and performance," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1781–1796, Aug. 2010.

[43] G. Karagiannis, O. Altintas, E. Ekici, G. Heijenk, B. Jarupan, K. Lin, and T. Weil, "Vehicular networking: A survey and tutorial on requirements, architectures, challenges, standards and solutions," *IEEE Communications Surveys Tutorials*, vol. 13, no. 4, pp. 584 –616, July 2011.

[44] *Industrial communication networks - Wireless communication network and communication profiles - WirelessHART*, IEC, iEC 62591.

[45] *IEEE 802.15.4 standard: Wireless Medium Access Control and Physical Layer Specifications for Low-Rate Wireless Personal Area Networks*, IEEE, 2006, http://www.ieee802.org/15/pub/TG4.html.

[46] I. 802.15.4e, "Ieee 802.15 wpan task group 4e (tg4e)." [Online]. Available: http://ieee802.org/15/pub/TG4e.html

[47] P. Park, C. Fischione, A. Bonivento, K. Johansson, and A. Sangiovanni-Vincent, "Breath: An adaptive protocol for industrial control applications using wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 6, pp. 821–838, June 2011.

[48] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York, NY: Wiley-Interscience, 1991.

[49] J. Proakis, *Digital Communications.* New York, NY: McGraw-Hill, 2000.

[50] M. I. of Technology. Project MAC. Engineering Robotics Group and M. Dertouzos, *Control robotics: The procedural control of physical processes*, 1973.

[51] S. Baruah and J. Goossens, "Scheduling real-time tasks: Algorithms and complexity," *Handbook of scheduling: Algorithms, models, and performance analysis*, vol. 3, 2004.

[52] F. Eisenbrand and T. Rothvoß, "Edf-schedulability of synchronous periodic task systems is conp-hard," in *Proc. of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms.* Society for Industrial and Applied Mathematics, 2010, pp. 1029–1034.

[53] F. Zhang and A. Burns, "Schedulability analysis for real-time systems with edf scheduling," *IEEE Transactions on Computers*, vol. 58, no. 9, pp. 1250–1258, Sept. 2009.

[54] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge University Press, 2004.

[55] D. Bertsimas and T. Tsitsiklis, *Introduction to Linear Optimization.* Athena Scientific, 1997.

[56] O. Khader, A. Willig, and A. Wolisz, "A simulation model for the performance evaluation of wirelessHART TDMA protocol," Technical report, Telecommunication Networks Group, Technical UniversityBerlin, TKN Technical Report Series TKN-11-001, Berlin, Germany, Tech. Rep., 2011.

[57] A. F. Molisch, K. Balakrishnan, C.-C. Chong, S. Emami, A. Fort, J. Karedal, J. Kunisch, H. Schantz, U. Schuster, and K. Siwiak, "IEEE 802.15. 4a channel model-final report," *IEEE P802*, vol. 15, no. 04, p. 0662, 2004.

[58] J. Carpenter, S. Funk, P. Holman, A. Srinivasan, J. Anderson, and S. Baruah, "A categorization of real-time multiprocessor scheduling problems and algorithms."

[59] G. C. Buttazzo, "Rate monotonic vs. edf: judgment day," *Real-Time Systems*, vol. 29, no. 1, pp. 5–26, 2005.

[60] A. Burns and S. Baruah, "Sustainability in real-time scheduling," *Journal of Computing Science and Engineering*, vol. 2, no. 1, pp. 74–97, 2008.

[61] A. Frotzsche, U. Wetzker, M. Bauer, M. Rentschler, M. Beyer, S. Elspass, and H. Klessig, "Requirements and current solutions of wireless communication in industrial automation," in *IEEE International Conference on Communications Workshops (ICC)*, June 2014, pp. 67–72.

[62] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke, "A survey on smart grid potential applications and communication requirements," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 1, pp. 28–42, February 2013.

[63] Y. Sadi and S. Ergen, "Optimal power control, rate adaptation, and scheduling for UWB-based intravehicular wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 1, pp. 219–234, 2013.

[64] H. C. Foundation, "Wirelesshart." [Online]. Available: http://www.hartcomm.org

[65] ISA, "Isa-100.11a-2009 wireless systems for industrial automation: Process control and related applications." [Online]. Available: http://www.isa.org

[66] J. Liu, *Real-Time Systems*. Prentice Hall, 2000.

[67] Y. Sadi, S. C. Ergen, and P. Park, "Minimum energy data transmission for wireless networked control systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 4, pp. 2163–2175, April 2014.

[68] Y. Wu, G. Buttazzo, E. Bini, and A. Cervin, "Parameter selection for real-time controllers in resource-constrained systems," *IEEE Transactions on Industrial Informatics*, vol. 6, no. 4, pp. 610–620, November 2010.

[69] W. P. M. H. Heemels, A. R. Teel, N. van de Wouw, and D. Nesic, "Networked control systems with communication constraints: Tradeoffs between transmission intervals, delays and performance," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1781–1796, August 2010.

[70] J. R. Moyne and D. M. Tilbury, "The emergence of industrial control networks for manufacturing control, diagnostics, and safety data," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 29–47, January 2007.

[71] L. Fu, S. C. Liew, and J. Huang, "Fast algorithms for joint power control and scheduling in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 3, pp. 1186 –1197, March 2010.

[72] S. Kompella, J. Wieselthier, A. Ephremides, H. Sherali, and G. Nguyen, "On optimal SINR-based scheduling in multihop wireless networks," *IEEE/ACM Transactions on Networking*, vol. 18, no. 6, pp. 1713 – 1724, Dec. 2010.

[73] L. Lin, X. Lin, and N. Shroff, "Low-complexity and distributed energy minimization in multihop wireless networks," *IEEE/ACM Transactions on Networking*, vol. 18, no. 2, pp. 501 –514, April 2010.

[74] V. V. Vazirani, *Approximation algorithms.* Springer, 2001.

[75] R. A. Horn and C. R. Johnson., *Matrix Analysis.* Cambridge University Press, 1994.

[76] J. Tang, G. Xue, C. Chandler, and W. Zhang, "Link scheduling with power control for throughput enhancement in multihop wireless networks," *Vehicular Technology, IEEE Transactions on*, vol. 55, no. 3, pp. 733 –742, May 2006.

[77] Z. Yang, L. Cai, and W. Lu, "Practical scheduling algorithms for concurrent transmissions in rate-adaptive wireless networks," in *IEEE International Conference on Computer Communications (INFOCOM)*, March 2010, pp. 1–9.

[78] M. Johansson and L. Xiao, "Cross-layer optimization of wireless networks using nonlinear column generation," *Wireless Communications, IEEE Transactions on*, vol. 5, no. 2, pp. 435–445, Feb. 2006.

[79] J. Luo, C. Rosenberg, and A. Girard, "Engineering wireless mesh networks: Joint scheduling, routing, power control, and rate adaptation," *IEEE/ACM Transactions on Networking*, vol. 18, no. 5, pp. 1387 –1400, Oct. 2010.

[80] T. Elbatt and A. Ephremides, "Joint scheduling and power control for wireless ad hoc networks," *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 74–85, Jan. 2004.

[81] U. C. Kozat, I. Koutsopoulos, and L. Tassiulas, "Cross-layer design for power efficiency and qos provisioning in multi-hop wireless networks," *Wireless Communications, IEEE Transactions on*, vol. 5, no. 11, pp. 3306–3315, Nov. 2006.

[82] S. Ramanathan and E. Lloyd, "Scheduling algorithms for multihop radio networks," *Networking, IEEE/ACM Transactions on*, vol. 1, no. 2, pp. 166 –177, Apr. 1993.

[83] W. Wang, Y. Wang, X. Li, W. Song, and O. Frieder, "Efficient interference-aware tdma link scheduling for static wireless networks," in *IEEE Global Communications Conference (GLOBECOM)*, Nov. 2006, pp. 262–273.

[84] J. Mao, Z. Wu, and X. Wu, "A tdma scheduling scheme for many-to-one communications in wireless sensor networks," *Computer Communications*, vol. 30, no. 4, pp. 863–872, Feb. 2007.

[85] S. C. Ergen and P. Varaiya, "Tdma scheduling algorithms for wireless sensor networks," *Wireless Networks*, vol. 16, no. 4, pp. 985 – 997, May 2010.

[86] S. Borbash and A. Ephremides, "Wireless link scheduling with power control and sinr constraints," *Information Theory, IEEE Transactions on*, vol. 52, no. 11, pp. 5106 –5111, Nov. 2006.

[87] A. Behzad and I. Rubin, "Optimum integrated link scheduling and power control for multihop wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 1, pp. 194 – 205, Jan. 2007.

[88] P. Björklund, P. Värbrand, and D. Yuan, "A column generation method for spatial tdma scheduling in ad hoc networks," *Ad hoc networks*, vol. 2, no. 4, pp. 405–418, 2004.

[89] A. Capone, G. Carello, I. Filippini, S. Gualandi, and F. Malucelli, "Routing, scheduling and channel assignment in wireless mesh networks: optimization models and algorithms," *Ad Hoc Networks*, vol. 8, no. 6, pp. 545–563, 2010.

[90] ——, "Solving a resource allocation problem in wireless mesh networks: A comparison between a cp-based and a classical column generation," *Networks*, vol. 55, no. 3, pp. 221–233, 2010.

[91] A. Pantelidou and A. Ephremides, "Minimum schedule lengths with rate control in wireless networks," in *IEEE Military Communications Conference (MILCOM)*, Nov. 2008, pp. 1 –7.

[92] ——, "Minimum-length scheduling and rate control for time-varying wireless networks," in *IEEE Military Communications Conference (MILCOM)*, Oct. 2009, pp. 1 –7.

[93] S. Kompella, H. Sherali, and A. Ephremides, "Optimal scheduling in interference limited fading wireless networks," in *IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2009, pp. 1 –6.

[94] S. Kompella, J. Wieselthier, and A. Ephremides, "Revisiting the optimal scheduling problem," in *42nd Annual Conference on Information Sciences and Systems (CISS)*, March 2008, pp. 492 –497.

[95] S. Kompella, J. E. Wieselthier, and A. Ephremides, "A cross-layer approach to optimal wireless link scheduling with sinr constraints," in *IEEE Military Communications Conference (MILCOM)*, Oct. 2007, pp. 1 –7.

[96] I.-H. Hou and P. R. Kumar, "Real-time communication over unreliable wireless links: A theory and its applications," *Wireless Communications, IEEE*, vol. 19, no. 1, pp. 48–59, Feb. 2012.

[97] R. Bruno, M. Conti, and E. Gregori, "Mesh networks: commodity multihop ad hoc networks," *Communications Magazine, IEEE*, vol. 43, no. 3, pp. 123–131, 2005.

[98] R. Cruz and A. Santhanam, "Optimal routing, link scheduling and power control in multihop wireless networks," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 1, 2003, pp. 702–711.

[99] Y. Li and A. Ephremides, "A joint scheduling, power control, and routing algorithm for ad hoc wireless networks," *Ad Hoc Networks*, vol. 5, no. 7, pp. 959–973, 2007.

[100] M. Johansson and L. Xiao, "Cross-layer optimization of wireless networks using nonlinear column generation," *Wireless Communications, IEEE Transactions on*, vol. 5, no. 2, pp. 435 – 445, feb. 2006.

[101] J. Luo, C. Rosenberg, and A. Girard, "Engineering wireless mesh networks: Joint scheduling, routing, power control, and rate adaptation," *Networking, IEEE/ACM Transactions on*, vol. 18, no. 5, pp. 1387 – 1400, oct. 2010.

[102] E. Uysal-Biyikoglu, B. Prabhakar, and A. El Gamal, "Energy-efficient packet transmission over a wireless link," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 487 – 499, Aug. 2002.

[103] S. Cui, A. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 5, pp. 2349 – 2360, Sept. 2005.

[104] M. S. Bazaraa, J. J. Jarvis, and H. D. Sherali, *Linear programming and network flows*. Wiley-Interscience, 2011.

[105] B. Radunovic and J. L. Boudec, "Optimal power control, scheduling, and routing in uwb networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 7, pp. 1252–1270, September 2004.

[106] K. H. Liu, L. Cai, and X. Shen, "Exclusive-region based scheduling algorithms for uwb wpan," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 933–942, March 2008.

[107] H. Jiang and W. Zhuang, "Effective packet scheduling with fairness adaptation in ultra-wideband wireless networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 2, pp. 680–690, February 2007.

[108] K. H. Liu, L. Cai, and X. Shen, "Multiclass utility-based scheduling for uwb networks," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 2, pp. 1176–1187, March 2008.

[109] E. Uysal-Biyikoglu, B. Prabhakar, and A. E. Gamal, "Energy-efficient packet transmission over a wireless link," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 487–499, August 2002.

[110] F. Zhang and S. Chanson, "Improving communication energy efficiency in wireless networks powered by renewable energy sources," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 6, pp. 2125–2136, November 2005.

[111] W. Chen and U. Mitra, "Energy efficient scheduling with individual packet delay constraints," in *IEEE INFOCOM*, June 2006, pp. 1–6.

[112] X. Zhong and C.-Z. Xu, "Energy-efficient wireless packet scheduling with quality of service control," *IEEE Transactions on Mobile Computing*, vol. 6, no. 10, pp. 1158–1170, October 2007.

[113] T. Wang, W. Heinzelman, and A. Seyedi, "Link energy minimization in ir-uwb based wireless sensor networks," *IEEE Transactions on Wireless Communication*, vol. 9, no. 9, pp. 2800–2811, September 2010.

[114] T. Elbatt and A. Ephremides, "Joint scheduling and power control for wireless ad hoc networks," *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 74–85, January 2004.

[115] L. Fu, S. C. Liew, and J. Huang, "Fast algorithms for joint power control and scheduling in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 3, pp. 1186–1197, March 2010.

[116] R. Madan, S. Cui, S. Lall, and A. J. Goldsmith, "Cross-layer design for lifetime maximization in interference-limited wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 5, no. 11, pp. 3142–3152, November 2006.

[117] M. Johansson and L. Xiao, "Cross-layer optimization of wireless networks using nonlinear column generation," *Wireless Communications, IEEE Transactions on*, vol. 5, no. 2, pp. 435 – 445, feb. 2006.

[118] J. Luo, C. Rosenberg, and A. Girard, "Engineering wireless mesh networks: Joint scheduling, routing, power control, and rate adaptation," *Networking, IEEE/ACM Transactions on*, vol. 18, no. 5, pp. 1387 – 1400, oct. 2010.

[119] L. Fu, S. C. Liew, and J. Huang, "Fast algorithms for joint power control and scheduling in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 3, pp. 1186 –1197, March 2010.

[120] F. Cuomo, C. Martello, A. Baiocchi, and F. Capriotti, "Radio resource sharing for ad hoc networking with uwb," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1722 – 1732, December 2002.

[121] H. Wymeersch, G. Zussman, and M. Z. Win, "Snr analysis for multirate uwb-ir," *IEEE Communication Letters*, vol. 11, no. 1, pp. 49–51, January 2007.

[122] B. Radunovic and J.-Y. Le Boudec, "Power control is not required for wireless networks in the linear regime," in *World of Wireless Mobile and Multimedia Networks, 2005. WoWMoM 2005. Sixth IEEE International Symposium on a*, June 2005, pp. 417 – 427.

[123] C. Fischione, "Fast-lipschitz optimization with wireless sensor networks applications," *IEEE Transactions on Automatic Control*, vol. 56, no. 10, pp. 2319–2331, October 2011.