Maximizing Capacity in Multi-Hop Cognitive Radio Networks Under the SINR Model

Yi Shi, *Member, IEEE,* Y. Thomas Hou, *Senior Member, IEEE,* Sastry Kompella, *Member, IEEE,*
and Hanif D. Sherali

Abstract—Cognitive radio networks (CRNs) have the potential to utilize spectrum efficiently and are positioned to be the core technology for the next-generation multi-hop wireless networks. An important problem for such networks is its capacity. We study this problem for CRNs in the SINR (signal-to-interference-and-noise-ratio) model, which is considered to be a better characterization of interference (but also more difficult to analyze) than disk graph model. The main difficulties of this problem are two-fold. First, SINR is a non-convex function of transmission powers; an optimization problem in the SINR model is usually a non-convex program and NP-hard in general. Second, in the SINR model, scheduling feasibility and the maximum allowed flow rate on each link are determined by SINR at the physical layer. To maximize capacity, it is essential to follow a cross-layer approach; but joint optimization at physical (power control), link (scheduling), and network (flow routing) layers with the SINR function is inherently difficult. In this paper, we give a mathematical characterization of the joint relationship among these layers. We devise a solution procedure that provides a $(1 - \varepsilon)$ optimal solution to this complex problem, where ε is the required accuracy. Our theoretical result offers a performance benchmark for any other algorithms developed for practical implementation. Using numerical results, we demonstrate the efficacy of the solution procedure and offer quantitative understanding on the interaction of power control, scheduling, and flow routing in a CRN.

Index Terms—Theory, multi-hop cognitive radio network, non-linear optimization, SINR model, cross-layer, capacity.

✦

1 INTRODUCTION

COGNITIVE radio networks (CRNs) have great potential to improve spectrum efficiency and will be the core technology for the next-generation multi-hop wireless networks. Within such a network, each node is equipped with a cognitive radio (CR) for wireless communications, which is capable of sensing the available frequency bands (i.e., those bands that are currently not used by primary users) and reconfiguring RF to switch to newly-selected frequency bands.

From wireless networking perspective, CR offers a whole new set of research problems in algorithm design and protocol implementation. To appreciate such opportunity, we compare CR with a closely related wireless technology called *multi-channel multi-radio* (MC-MR) [2], [20], [21]. First, MC-MR employs traditional *hardware-based* radio technology and thus each radio can only operate on a single channel at a time. In contrast, the radio technology in CR is *software-based*; a soft radio is capable of switching frequency bands on a per-packet basis. As a result, the number of concurrent frequency bands utilized by a CR is typically much larger than by MC-MR. Second, a common assumption for MC-MR is

that there is a set of "common channels" available at every node in the network. Such assumption is hardly true for CRNs, where each node may have a different set of available frequency bands. These important differences between MC-MR and CR suggest that algorithm design for CRNs is substantially more complex than that for MC-MR networks. In some sense, an MC-MR network can be considered as a special case of a CRN.

The capacity of a CRN is usually considered a key problem in fundamental understanding. In this paper, we aim to study this problem in the SINR (signal-to-interference-and-noise-ratio) model. In this model, concurrent transmissions are allowed and interference (due to transmissions by non-intended transmitter) is treated as noise. A transmission is successful if and only if SINR at the receiver is greater than or equal to a threshold. Moreover, the achieved transmission capacity is also a function of SINR (via Shannon's formula). The current understanding is that the SINR model is better than the so-called "disk graph model" (or "protocol model" [15]) for interference characterization.

Although the SINR model is thought to be more realistic than the protocol model, a number of problems arise when trying to carry out mathematical analysis in it. First, SINR at a receiver not only depends on the transmission power at the corresponding transmitter, but also depends on the transmission powers at other transmitters. Mathematically, SINR is a non-convex function of multiple variables. Many optimization problems in the SINR model are non-convex problems and NP-hard. Second, since both scheduling feasibility and the maximum allowed flow rate on each link are determined by SINR, an optimal solution to maximize capacity must be developed with joint consideration of network, link,

- *Y. Shi and Y.T. Hou are with the Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061. E-mail: {yshi,thou}@vt.edu.*
- *S. Kompella is with Information Technology Division, U.S. Naval Research Laboratory, Washington, DC, 20375. E-mail: sastry.kompella@nrl.navy.mil.*
- *H.D. Sherali is with the Grado Department of Industrial and Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061. E-mail: hanifs@vt.edu.*

Report Documentation Page

1. REPORT DATE **27 MAR 2010**	2. REPORT TYPE	3. DATES COVERED **00-00-2010 to 00-00-2010**

4. TITLE AND SUBTITLE **Maximizing Capacity in Multi-Hop Cognitive Radio Networks Under the SINR Model**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Research Laboratory,Information Technology Division,Washington,DC,20375**	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

14. ABSTRACT
Cognitive radio networks (CRNs) have the potential to utilize spectrum efficiently and are positioned to be the core technology for the next-generation multi-hop wireless networks. An important problem for such networks is its capacity. We study this problem for CRNs in the SINR (signal-to-interference-and-noise-ratio) model, which is considered to be a better characterization of interference (but also more difficult to analyze) than disk graph model. The main difficulties of this problem are two-fold. First, SINR is a non-convex function of transmission powers; an optimization problem in the SINR model is usually a non-convex program and NP-hard in general. Second, in the SINR model, scheduling feasibility and the maximum allowed flow rate on each link are determined by SINR at the physical layer. To maximize capacity, it is essential to follow a cross-layer approach; but joint optimization at physical (power control), link (scheduling), and network (flow routing) layers with the SINR function is inherently difficult. In this paper, we give a mathematical characterization of the joint relationship among these layers. We devise a solution procedure that provides a (1 ? ") optimal solution to this complex problem, where " is the required accuracy. Our theoretical result offers a performance benchmark for any other algorithms developed for practical implementation. Using numerical results, we demonstrate the efficacy of the solution procedure and offer quantitative understanding on the interaction of power control, scheduling, and flow routing in a CRN.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**14**	

and physical layers. Due to these difficulties, theoretical results on CRNs in the SINR model remain limited.

In this paper, we investigate the network capacity problem for multi-hop CRNs in the SINR model. For a given instance where each node has access to a set of available bands (likely heterogeneous), we study the network capacity problem via a cross-layer optimization approach. In particular, we consider how to maximize the rates of a set of user communication sessions, with joint consideration at physical layer (via power control), link layer (via frequency band scheduling), and network layer (via flow routing). We give a mathematical characterization of these layers and formulate a mixed integer nonlinear program (MINLP) problem. For this optimization problem, we first identify core optimization variables and core optimization space based on the physical significance of the variables. We devise an algorithm based on the branch-and-bound framework to obtain a $(1-\varepsilon)$ optimal solution.

Although branch-and-bound framework is standard, many components within this framework are not specified. We design the following problem-specific components. (1) We propose a reformulation-linearization technique (RLT) to develop a tight linear relaxation so as to obtain a tight upper bound. (2) For the lower bound, we design a local search algorithm by analyzing and removing infeasibility in the linear relaxation solution. (3) For problem partitioning, we propose to choose a partition variable based on its physical significance. With these problem-specific designs, the overall solution can find a $(1-\varepsilon)$ optimal solution much faster than brute-force exhaustive search.

Our theoretical result offers a solution for a given network instance and can serve as a performance benchmark for a multi-hop CRN. Such a result is not yet available in the literature. If the available bands at each node in the network varies on a relatively larger time scale than the execution of our solution, then our solution can be used to compute a $(1-\varepsilon)$ optimal result for each time period during which the available bands at each node is static. If the available bands at each node in the network varies on a relatively smaller time scale, then the performance benchmark can be obtained offline by our solution. One can use our result as a reference performance benchmark to measure the quality of some other proposed algorithms (possibly heuristic, despite distributed) that are developed for actual deployment.

The remainder of this paper is organized as follows. In Section 2, we review related work on cross-layer optimization. Section 3 gives a mathematical characterization of power control, scheduling, and routing in the SINR model for multi-hop CRNs. In Section 4, we reformulate the optimization problem and obtain a clean and compact formulation. Section 5 analyzes the core optimization space and describes the main algorithm to obtain $(1-\varepsilon)$ optimal solution. In Section 6, we develop tight upper and lower bounds inside the main algorithm. We discuss how to interpret and apply our theoretical result in Section 7. Section 8 presents numerical results and Section 9 concludes this paper.

2 RELATED WORK

There have been active research efforts on cross-layer optimization for wireless networks. Many of these efforts are based on the disk graph model (see, e.g., [2], [18], [20], [24] for wireless networks with traditional radios and [26], [28], [33], [34] for CRNs in particular). Under this model, the impact of interference from neighboring nodes is solely determined by whether or not a node falls within the interference range of other transmitting nodes. The controversy surrounding (or arguments against) the protocol model is that a binary decision of whether or not interference exists (based on interference range) does not accurately capture physical layer characteristics. As a result, the accuracy (and validity) of results developed under protocol model remains unclear.

In contrast, the SINR model is widely regarded as a better model for interference characterization. Although such model is preferred, there are many difficulties in carrying out analysis with this model due to the computational complexity SINR involves, particularly when we study cross-layer optimization in a multi-hop environment. As a result, many previous efforts were done on single-hop networks, e.g., [3], [10], [13], [14], [17]. For multi-hop networks, various simplifications have been employed. For example, in [4], the authors assumed synchronized power control, where transmission power at each node in the network is adjustable but is "synchronized" (identical). Needless to say, such synchronization in power control cannot offer optimal network performance. There are also some efforts studying cross-layer problems involving two layers instead of three layers (physical, link, and network, as we do in this paper). For example, in [6], Bhatia and Kodialam optimized power control and routing, but assumed some frequency hopping mechanism is in place for scheduling, which helps simplify joint consideration of scheduling. In [9], Elbatt and Ephremides optimized joint power control and scheduling, but assumed routing was given a priori. In [27], Shu and Krunz studied how to maximize the total rate on all links in a CRN, with the consideration of power control and channel assignment. For cross-layer optimization in the SINR model involving three layers (physical, link, and network), nearly all existing efforts (e.g., [7], [8]) followed a "layer-decoupled" approach to simplify analysis. Under such an approach, the final solution is obtained by determining algorithm/mechanism for one layer at a time and then piecing up them together instead of solving a joint optimization problem. Due to de-coupling in the solution procedure, these approaches are heuristic at best and cannot offer any performance guarantee.

On another line of research, various efforts have been made to study asymptotic behavior (or scaling laws) of wireless networks (see, e.g., [1], [5], [15], [16], [19], [21], [22], [29], [30], [31], [32]). These efforts differ from ours in this paper, which focuses on designing optimal cross-layer algorithms for a *finite-sized* network.

3 MATHEMATICAL MODELS

We consider a multi-hop network with a set of CR nodes $\mathcal{N}$. Each node $i \in \mathcal{N}$ senses its environment and finds a

set of available frequency bands $\mathcal{M}_i$ for the given time instance (i.e., those bands that are currently not used by primary users), which may not be the same as the available frequency bands at other nodes. We assume that the bandwidth of each frequency band (channel) is W. Denote $\mathcal{M}$ the union of all frequency bands among all the nodes in the network, i.e., $\mathcal{M} = \bigcup_{i \in \mathcal{N}} \mathcal{M}_i$. Denote $\mathcal{M}_{ij} = \mathcal{M}_i \bigcap \mathcal{M}_j$, which is the set of frequency bands that is common between nodes i and j and thus can be used for communication between these two nodes. In the rest of this section, we present mathematical characterization of each layer and obtain a formulation for the capacity problem in this study.

3.1 Power Control, Scheduling, and Their Relationship in the SINR Model

Power control on each transmitting node at the physical layer affects SINR at a receiving node. These SINR values in turn will affect scheduling decision at link layer. That is, if a node is scheduled to receive, then its SINR must be at least α (minimum requirement). Therefore, power control and scheduling are tightly coupled via SINR and cannot be modeled separately. In this section, we formulate power control, scheduling, and their relationship in the SINR model.

Scheduling for transmission at each node in the network can be done either in frequency domain or time domain. In this paper, we consider scheduling in frequency domain in the form of assigning frequency bands (channels). Since time domain based formulation is similar to that for frequency domain, our approach can be extended to time domain based formulation and will yield similar results.

To maximize the capacity, there may still be concurrent transmissions within the same channel (and thus interference). Denote scheduling variables x_{ij}^m as follows.

$$x_{ij}^m = \begin{cases} 1 & \text{If node } i \text{ transmits data to node } j \text{ on band } m, \\ 0 & \text{otherwise.} \end{cases}$$

Due to interference, a node i can use a band m for transmitting to only a single node j or receiving from only a single node k. That is,

$$\sum_{k \in \mathcal{N}, k \neq i}^{m \in \mathcal{M}_k} x_{ki}^m + \sum_{j \in \mathcal{N}, j \neq i}^{m \in \mathcal{M}_j} x_{ij}^m \leq 1 \quad (i \in \mathcal{N}, m \in \mathcal{M}_i). \quad (1)$$

For power control, we assume that the transmission power at a node can be tuned to a finite number of levels between 0 and $P_{\max}$. To model this discrete power control, we introduce an integer parameter Q that represents the total number of power levels to which a transmitter can be adjusted, i.e., transmission power can be $0, \frac{1}{Q}P_{\max}, \frac{2}{Q}P_{\max}, \cdots, P_{\max}$. Denote $q_{ij}^m \in \{0, 1, 2, \cdots, Q\}$ the integer levels for transmission power. Clearly, when node i does not transmit data to node j on band m, q_{ij}^m is 0. Thus, power control and scheduling is coupled with each other via the following relationship.

$$q_{ij}^m \begin{cases} \in [1, Q] & \text{If } x_{ij}^m = 1, \\ = 0 & \text{otherwise.} \end{cases} \quad (i, j \in \mathcal{N}, i \neq j, m \in \mathcal{M}_{ij}) \quad (2)$$

Consider a transmission from node i to node j on band m. When there is interference from concurrent transmissions on the same band, SINR at node j, denoted as s_{ij}^m, is

$$s_{ij}^m = \frac{g_{ij} \frac{q_{ij}^m}{Q} P_{\max}}{\eta W + \sum_{k \in \mathcal{N}, k \neq i}^{m \in \mathcal{M}_k} \sum_{h \in \mathcal{N}, h \neq k}^{m \in \mathcal{M}_h} g_{kj} \frac{q_{kh}^m}{Q} P_{\max}}$$
$$(i, j \in \mathcal{N}, i \neq j, m \in \mathcal{M}_{ij}), \quad (3)$$

where η is the ambient Gaussian noise density and g_{ij} is the propagation gain from node i to node j.

Note that in theory, for any small SINR, the corresponding capacity is still positive (by Shannon's capacity formula). But in practice, if SINR is too small, then the achieved capacity will also be very small. In this case, such a weak link will not be very useful to carry traffic flow. Thus, we may use a threshold to remove such weak links from considerations. In this regard, we introduce a threshold for SINR, i.e., a transmission from node i to node j on band m is considered *successful* if and only if $s_{ij}^m \geq \alpha$. Then we have the following coupling relationship for scheduling (x_{ij}^m) and SINR (s_{ij}^m).

$$x_{ij}^m = 1 \Longleftrightarrow s_{ij}^m \geq \alpha \quad (i, j \in \mathcal{N}, i \neq j, m \in \mathcal{M}_{ij}). \quad (4)$$

3.2 Routing and Link Capacity

We assume there is a set of $\mathcal{L}$ active user communication (unicast) sessions in the network. Denote $s(l)$ and $d(l)$ the source and destination nodes of session $l \in \mathcal{L}$ and $r(l)$ the minimum rate requirement (in b/s) of session l. In our study, we aim to maximize a common scaling factor K for all session rates. That is, we aim to determine the maximum K such that a rate of $K \cdot r(l)$ can be transmitted from $s(l)$ to $d(l)$ for *each* session $l \in \mathcal{L}$ in the network.

To route each data flow from its source node to its corresponding destination node, multi-hop relaying may be necessary, due to limited transmission power at each node. Further, for optimality and flexibility, it is desirable to allow flow splitting and multi-path routing. This is because a single-path flow routing for a session is overly restrictive and is unlikely to offer optimal solution.

Mathematically, this can be modeled as follows. Denote $f_{ij}(l)$ the data rate on link $i \to j$ that is attributed to session l. If node i is the source node of session l, i.e., $i = s(l)$, then

$$\sum_{j \in \mathcal{N}, j \neq i}^{\mathcal{M}_{ij} \neq \emptyset} f_{ij}(l) = r(l)K \quad (l \in \mathcal{L}, i = s(l)). \quad (5)$$

If node i is an intermediate relay node for flow attributed to session l, i.e., $i \neq s(l)$ and $i \neq d(l)$, then

$$\sum_{j \in \mathcal{N}, j \neq i, s(l)}^{\mathcal{M}_{ij} \neq \emptyset} f_{ij}(l) = \sum_{k \in \mathcal{N}, k \neq i, d(l)}^{\mathcal{M}_{ki} \neq \emptyset} f_{ki}(l)$$
$$(l \in \mathcal{L}, i \in \mathcal{N}, i \neq s(l), d(l)). \quad (6)$$

If node i is the destination node of session l, i.e., $i = d(l)$, then

$$\sum_{k \in \mathcal{N}, k \neq i}^{\mathcal{M}_{ki} \neq \emptyset} f_{ki}(l) = r(l)K \quad (l \in \mathcal{L}, i = d(l)). \quad (7)$$

In addition to the above flow balance equations at each node $i \in \mathcal{N}$ for session $l \in \mathcal{L}$, the aggregated flow rates on each radio link cannot exceed this link's capacity. For a link $i \to j$, we have

$$\sum_{l \in \mathcal{L}}^{s(l) \neq j, d(l) \neq i} f_{ij}(l) \leq \sum_{m \in \mathcal{M}_{ij}} W \log_2(1 + s_{ij}^m)$$
$$(i, j \in \mathcal{N}, i \neq j, \mathcal{M}_{ij} \neq \emptyset) . \quad (8)$$

This constraint shows the coupling relationship between flow routing and SINR.

3.3 The Capacity Problem

In this paper, we study a capacity problem for multi-hop CRNs. For capacity problem, the simplest objective is the sum of throughput achieved by all sessions. However, such an objective may lead to biased rate allocation among sessions [23]. Another objective is maxmin, i.e., maximizing the minimum throughput among all sessions. Maxmin has been used in a number of works (e.g., [15], [21]) to ensure fairness. In this work, we consider how to maximize a common scaling factor K for all sessions under some given minimum rate requirements $r(l)$, i.e., what is the maximum factor K such that a rate of $K \cdot r(l)$ can be transmitted from $s(l)$ to $d(l)$ for *each* session $l \in \mathcal{L}$ in the network. Note that this objective is a more general form of maxmin in the sense that when $r(l) = 1$ for each session $l \in \mathcal{L}$, this objective becomes the maxmin throughput objective. On the other hand, when $r(l) \neq 1$, i.e., the minimum required rate for each session l may be different from session to session, the objective function becomes to maximize each session's rate proportional to its minimum rate requirement.

Putting together the objective and all the constraints for power control, scheduling, and flow routing, we have the following formulation.

$$\begin{aligned}
\text{Max} \quad & K \\
\text{s.t.} \quad & \text{constraints (1)--(8)}
\end{aligned}$$

$$x_{ij}^m \in \{0,1\}, q_{ij}^m \in \{0, 1, 2, \cdots, Q\}, t_i^m, s_{ij}^m \geq 0$$
$$(i, j \in \mathcal{N}, i \neq j, m \in \mathcal{M}_{ij})$$
$$K, f_{ij}(l) \geq 0 \quad (l \in \mathcal{L}, i, j \in \mathcal{N}, i \neq j, i \neq d(l), j \neq s(l), \mathcal{M}_{ij} \neq \emptyset) .$$

Note that we assume $\mathcal{M}_i$, the set of current available bands at node i, is given for a particular time instance. The solution to the above optimization problem will offer the best performance for this given instance. Since $\mathcal{M}_i$ may change over time, the optimal solution may also change over time.

4 REFORMULATION

The formulation in Section 3.3 is still in a "raw" form and is the first step in setting up our cross-layer optimization problem. Much work needs to be done to put it into a more "clean" and compact form that is amenable to mathematical operation. In this section, we analyze each constraint carefully and perform this necessary and important reformulation step.

- The constraint described in (2) is not suitable for mathematical programming. We reformulate it with the following linear constraint.

$$x_{ij}^m \leq q_{ij}^m \leq Qx_{ij}^m \quad (i, j \in \mathcal{N}, i \neq j, m \in \mathcal{M}_{ij}) . \quad (9)$$

It is easy to verify that this constraint is equivalent to (2).

A closer look at (9), in conjunction with (3) and (4), suggests that (9) can be further simplified to remove redundancy. That is, we find that (3) and (4) yield $x_{ij}^m \leq q_{ij}^m$ and thus the left-hand-side (LHS) of (9) can be removed. To show this, we consider the two cases for x_{ij}^m. For the case when $x_{ij}^m = 0$, $x_{ij}^m \leq q_{ij}^m$ holds by the definition of q_{ij}^m. For the case when $x_{ij}^m = 1$, by (4), s_{ij}^m must be positive. Then by (3), q_{ij}^m must also be positive. By the definition of q_{ij}^m, $q_{ij}^m \geq 1$ when it is positive. Therefore, $q_{ij}^m \geq x_{ij}^m$. So (9) (or (2)) can be replaced by the following constraint.

$$q_{ij}^m \leq Qx_{ij}^m \quad (i, j \in \mathcal{N}, i \neq j, m \in \mathcal{M}_{ij}) . \quad (10)$$

- Constraint (3) is in the form of a fraction. In a mathematical programming, product is a better form. We can re-write it as

$$\begin{aligned}
s_{ij}^m &= \frac{g_{ij} \frac{q_{ij}^m}{Q} P_{\max}}{\eta W + \sum_{k \in \mathcal{N}, k \neq i}^{m \in \mathcal{M}_k} \sum_{h \in \mathcal{N}, h \neq k}^{m \in \mathcal{M}_h} g_{kj} \frac{q_{kh}^m}{Q} P_{\max}} \\
&= \frac{g_{ij} q_{ij}^m}{\frac{\eta W Q}{P_{\max}} + \sum_{k \in \mathcal{N}, k \neq i}^{m \in \mathcal{M}_k} \sum_{h \in \mathcal{N}, h \neq k}^{m \in \mathcal{M}_h} g_{kj} q_{kh}^m} \\
&\qquad (i, j \in \mathcal{N}, i \neq j, m \in \mathcal{M}_{ij}) .
\end{aligned}$$

This is equivalent to

$$\frac{\eta W Q}{P_{\max}} s_{ij}^m + \sum_{k \in \mathcal{N}, k \neq i}^{m \in \mathcal{M}_k} \sum_{h \in \mathcal{N}, h \neq k}^{m \in \mathcal{M}_h} g_{kj} q_{kh}^m s_{ij}^m - g_{ij} q_{ij}^m = 0$$
$$(i, j \in \mathcal{N}, i \neq j, m \in \mathcal{M}_{ij}) . \quad (11)$$

Note that in (11), q_{kh}^m and s_{ij}^m are variables while all other symbols are constants. Thus, we have a double sum of nonlinear terms $q_{kh}^m s_{ij}^m$ in (11). To reduce the number of nonlinear terms, denote

$$t_k^m = \sum_{h \in \mathcal{N}, h \neq k}^{m \in \mathcal{M}_h} q_{kh}^m \quad (k \in \mathcal{N}, m \in \mathcal{M}_k) . \quad (12)$$

Then (11) can be re-written as

$$\frac{\eta W Q}{P_{\max}} s_{ij}^m + \sum_{k \in \mathcal{N}, k \neq i}^{m \in \mathcal{M}_k} g_{kj} t_k^m s_{ij}^m - g_{ij} q_{ij}^m = 0$$
$$(i, j \in \mathcal{N}, i \neq j, m \in \mathcal{M}_{ij}) , \quad (13)$$

which now only involves a single sum of nonlinear terms $t_k^m s_{ij}^m$. Note that by introducing new variables t_k^m, we decrease the number of nonlinear terms from $O(|\mathcal{N}|^4 \cdot |\mathcal{M}|)$ in (11) to $O(|\mathcal{N}|^3 \cdot |\mathcal{M}|)$ in (13).

- Similar to (2), the constraint described in (4) is not suitable for mathematical programming. We now

$$
\begin{array}{lll}
\text{Max} & K & \\[4pt]
\text{s.t.} & \sum_{k\in\mathcal{N},k\neq i}^{m\in\mathcal{M}_k} x_{ki}^m + \sum_{j\in\mathcal{N},j\neq i}^{m\in\mathcal{M}_j} x_{ij}^m \leq 1 & (i\in\mathcal{N},m\in\mathcal{M}_i)\\[10pt]
& q_{ij}^m - Q x_{ij}^m \leq 0 & (i,j\in\mathcal{N},i\neq j,m\in\mathcal{M}_{ij})\\[10pt]
& \sum_{j\in\mathcal{N},j\neq i}^{m\in\mathcal{M}_j} q_{ij}^m - t_i^m = 0 & (i\in\mathcal{N},m\in\mathcal{M}_i)\\[10pt]
& \frac{\eta W Q}{P_{\max}} s_{ij}^m + \sum_{k\in\mathcal{N},k\neq i}^{m\in\mathcal{M}_k} g_{kj} t_k^m s_{ij}^m - g_{ij} q_{ij}^m = 0 & (i,j\in\mathcal{N},i\neq j,m\in\mathcal{M}_{ij})\\[10pt]
& \alpha x_{ij}^m - s_{ij}^m \leq 0 & (i,j\in\mathcal{N},i\neq j,m\in\mathcal{M}_{ij})\\[10pt]
& \sum_{j\in\mathcal{N},j\neq i}^{\mathcal{M}_{ij}\neq\emptyset} f_{ij}(l) - r(l)K = 0 & (l\in\mathcal{L},i=s(l))\\[10pt]
& \sum_{j\in\mathcal{N},j\neq i,s(l)}^{\mathcal{M}_{ij}\neq\emptyset} f_{ij}(l) - \sum_{k\in\mathcal{N},k\neq i,d(l)}^{\mathcal{M}_{ki}\neq\emptyset} f_{ki}(l) = 0 & (l\in\mathcal{L},i\in\mathcal{N},i\neq s(l),d(l))\\[10pt]
& \sum_{l\in\mathcal{L}}^{s(l)\neq j,d(l)\neq i} f_{ij}(l) - \sum_{m\in\mathcal{M}_{ij}} W\log_2(1+s_{ij}^m) \leq 0 & (i,j\in\mathcal{N},i\neq j,\mathcal{M}_{ij}\neq\emptyset)\\[10pt]
& x_{ij}^m \in\{0,1\}, q_{ij}^m\in\{0,1,2,\cdots,Q\}, t_i^m, s_{ij}^m \geq 0 & (i,j\in\mathcal{N},i\neq j,m\in\mathcal{M}_{ij})\\[10pt]
& K, f_{ij}(l) \geq 0 & (l\in\mathcal{L},i,j\in\mathcal{N},i\neq j,i\neq d(l),j\neq s(l),\mathcal{M}_{ij}\neq\emptyset)
\end{array}
$$

Fig. 1. Problem formulation.

show that (4) can be eliminated once we have (10), (13) and the following new constraint.

$$ s_{ij}^m \geq \alpha x_{ij}^m \quad (i,j\in\mathcal{N},i\neq j,m\in\mathcal{M}_{ij}) . \tag{14} $$

We now verify that (10), (13), and (14) can lead to the following two relationships described in (4), i.e.,

$$ x_{ij}^m = 1 \Rightarrow s_{ij}^m \geq \alpha \quad (i,j\in\mathcal{N},i\neq j,m\in\mathcal{M}_{ij}) \tag{15} $$
$$ s_{ij}^m \geq \alpha \Rightarrow x_{ij}^m = 1 \quad (i,j\in\mathcal{N},i\neq j,m\in\mathcal{M}_{ij}) \tag{16} $$

First we consider the case when $x_{ij}^m = 1$. By (14), we have $s_{ij}^m \geq \alpha$, i.e., (15) holds. Now consider the case when $s_{ij}^m \geq \alpha$. By (13), since t_k^m's are non-negative variables and $\eta, W, Q, P_{\max}, g_{kj}, g_{ij}$ are all positive constants, q_{ij}^m must be positive. Then by (10), x_{ij}^m must be positive when q_{ij}^m is positive. Since x_{ij}^m can only take 0 or 1 by its definition, x_{ij}^m must be 1. That is, (16) holds.

- Finally, we can prove that (7) is redundant once we have (5) and (6). Thus, we will leave (5) and (6) in the formulation and remove constraint (7).

With these careful reformulations, we now have a cleaner and more compact problem formulation, which is shown in Fig. 1.

5 SOLUTION OVERVIEW

For the optimization problem in Fig. 1, $Q, \eta, W, \alpha, P_{\max}$, g_{ij}, and $r(l)$ are constants and $K, x_{ij}^m, q_{ij}^m, t_i^m, s_{ij}^m$, and $f_{ij}(l)$ are optimization variables. This formulation is a mixed integer non-linear program (MINLP), which is NP-hard in general [11] and cannot be solved by CPLEX. In Section 5.1, we first analyze the intricate relationship among the variables and identify the core variables among all the variables. We show that the dependent variables can be derived once these core variables are fixed. We call the optimization space for the core variables as the core optimization space. In Section 5.2, we present the main algorithm on how to determine an optimal solution in the core optimization space.

5.1 Core Variables and Their Optimization Space

For the complex MINLP problem, its variables include $x_{ij}^m, q_{ij}^m, t_i^m, s_{ij}^m, f_{ij}(l)$, and K, which collectively contribute a seemingly huge optimization space. However, a closer investigation of these variables show that they are inter-dependent. In particular, we find that x_{ij}^m and q_{ij}^m variables are "core" variables and other variables can all be derived by core variables. As a result, we can focus our study on an optimization space by the core variables, which is a much smaller space.

We now show how to derive dependent variables from core variables. For t_i^m and s_{ij}^m variables, they can be derived by (12) and (13), respectively, once x_{ij}^m and q_{ij}^m are given. For $f_{ij}(l)$ variables and K variable, their optimal values are determined by a linear program (LP). That is, once we have the values for all x_{ij}^m and q_{ij}^m variables and we have computed the values for all t_i^m and s_{ij}^m variables by (12) and (13), respectively, the optimization problem reduces to a network flow problem, which is an LP.

With this new understanding of optimization variables, we can now focus our efforts on the core variables x_{ij}^m and q_{ij}^m.

5.2 Main Algorithm

In this section, we describe our main algorithm, which is based on the branch-and-bound framework [25]. Those readers who are familiar with branch-and-bound can skip this section and go to Section 6, where we design problem specific algorithms for each component of branch-and-bound.

Under branch-and-bound, we aim to provide a $(1-\varepsilon)$ optimal solution, where ε is a small positive constant reflecting our desired accuracy in the final solution. In case we set $\varepsilon = 0$, an optimal solution can be obtained.

We can start by developing upper and lower bounds (UB and LB) for the objective function (see Fig. 2(a) for an example). Branch-and-bound requires to develop
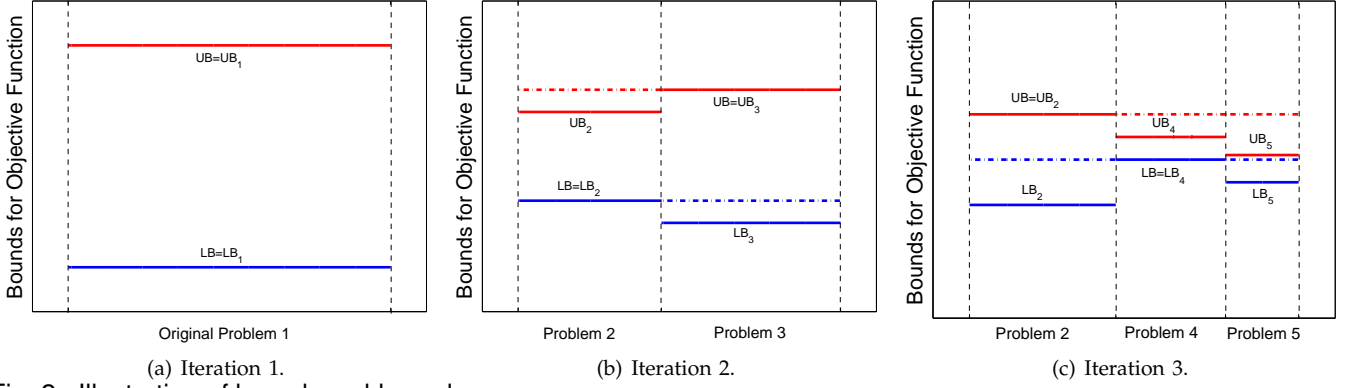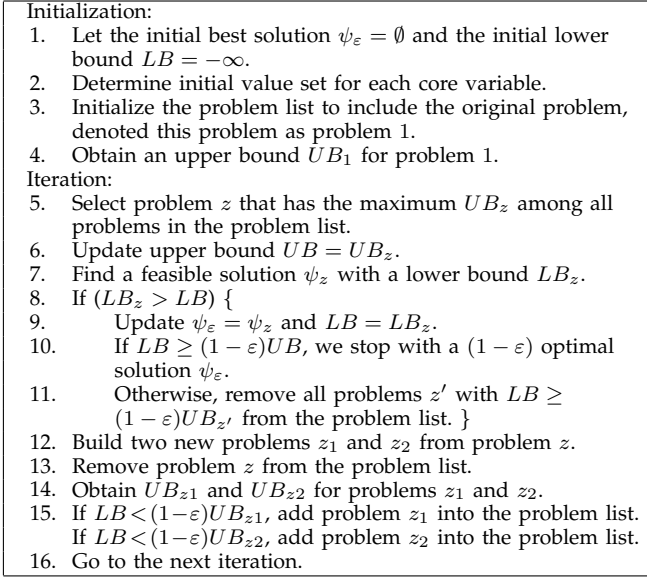
Fig. 2. Illustration of branch-and-bound.

Fig. 3. Main algorithm.

problem specific algorithms for finding upper and lower bounds, which will be described in Section 6. After we obtain the two bounds, we compare the gap between them. If the upper and lower bounds are close to each other, i.e., $LB \geq (1 - \varepsilon)UB$, then the feasible solution corresponding to the current lower bound LB is $(1 - \varepsilon)$ optimal and we are done.

Otherwise, we need to further narrow the gap between UB and LB. To do this, branch-and-bound partitions the original problem 1 into two new problems 2 and 3 (see Fig. 2(b)). This is accomplished by choosing an appropriate core variable x_{ij}^m or q_{ij}^m and dividing its value set into two smaller sets. The choice of specific core variables is important as its affects complexity. We will show how to do this in Section 6.

After partitioning, the core optimization space is divided into two sub-spaces for problems 2 and 3, respectively. We again obtain upper bounds UB_2 and UB_3 and lower bounds LB_2 and LB_3 for problems 2 and 3, respectively. Since the optimization space of problems 2 and 3 are both smaller than that of problem 1, we can have tighter upper bounds, i.e., $\max\{UB_2, UB_3\} \leq UB_1$, which in turn yield better lower bounds with $\max\{LB_2, LB_3\} \geq LB_1$ for our maximization problem (see Fig. 2(b)). Then the upper bound of the original

problem is updated as $UB = \max\{UB_2, UB_3\}$ and the lower bound is updated as $LB = \max\{LB_2, LB_3\}$. As a result, we now have a smaller gap between UB and LB. Then we either have a $(1 - \varepsilon)$ optimal solution (if $LB \geq (1 - \varepsilon)UB$) or continue to choose a problem with the maximum upper bound (Problem 3 in Fig. 2(b)) and perform partitioning for this problem. By choosing a problem with the maximum upper bound for partition, we can ensure that UB is decreased after each partition.

An important technique in branch-and-bound is that we can remove some problems from further consideration and thus reduce complexity. In particular, if we find a problem z with $LB \geq (1 - \varepsilon)UB_z$ (see problem 4 in Fig. 2(c)), we conclude that this problem can be removed from further consideration without loss of $(1 - \varepsilon)$ optimality.

Figure 3 shows the main algorithm. Since our core optimization space is finite (with finite number of core variables x_{ij}^m and q_{ij}^m, and each core variable has a finite integer value set), the branch-and-bound algorithm is guaranteed to converge (even for $\varepsilon = 0$) [25]. Here, branch-and-bound is much faster than brute-force exhaustive search because non-improving problems are being removed during the process to avoid wasting precious cycle time in future computation. As a result, for all the network instances studied in Section 8, brute-force exhaustive search cannot find an optimal solution, while our algorithm can find near-optimal solutions for all network instances (with various network sizes and user sessions).

6 DETERMINING BOUNDS AND PARTITIONING PROBLEMS

The main algorithm was presented in Fig. 3. Several components (i.e., determining upper and lower bounds, partitioning problems) in this main algorithm are yet to be developed. These algorithms should exploit problem specific properties to optimize performance. In this section, we design algorithms for these components.

6.1 Determining Upper Bound

To find an upper bound for a problem in branch-and-bound (see lines 4 and 14 in Fig. 3), we propose to construct a linear relaxation. That is to say, we linearize

Max $\qquad K$

s.t. $\quad \sum_{k\in\mathcal{N},k\neq i}^{m\in\mathcal{M}_k} x_{ki}^m + \sum_{j\in\mathcal{N},j\neq i}^{m\in\mathcal{M}_j} x_{ij}^m \leq 1$ $\qquad\qquad (i\in\mathcal{N}, m\in\mathcal{M}_i)$

$\qquad q_{ij}^m - Qx_{ij}^m \leq 0$ $\qquad\qquad (i,j\in\mathcal{N}, i\neq j, m\in\mathcal{M}_{ij})$

$\qquad \sum_{j\in\mathcal{N},j\neq i}^{m\in\mathcal{M}_j} q_{ij}^m - t_i^m = 0$ $\qquad\qquad (i\in\mathcal{N}, m\in\mathcal{M}_i)$

$\qquad \frac{\eta WQ}{P_{\max}} s_{ij}^m + \sum_{k\in\mathcal{N},k\neq i}^{m\in\mathcal{M}_k} g_{kj} u_{ijk}^m - g_{ij} q_{ij}^m = 0$ $\qquad (i,j\in\mathcal{N}, i\neq j, m\in\mathcal{M}_{ij})$

$\qquad$ Linear constraints for u_{ijk}^m $\qquad\qquad (i,j,k\in\mathcal{N}, i\neq j, m\in\mathcal{M}_{ij}, m\in\mathcal{M}_k)$

$\qquad \alpha x_{ij}^m - s_{ij}^m \leq 0$ $\qquad\qquad (i,j\in\mathcal{N}, i\neq j, m\in\mathcal{M}_{ij})$

$\qquad \sum_{j\in\mathcal{N},j\neq i}^{\mathcal{M}_{ij}\neq\emptyset} f_{ij}(l) - r(l)K = 0$ $\qquad\qquad (l\in\mathcal{L}, i = s(l))$

$\qquad \sum_{j\in\mathcal{N},j\neq i,s(l)}^{\mathcal{M}_{ij}\neq\emptyset} f_{ij}(l) - \sum_{k\in\mathcal{N},k\neq i,d(l)}^{\mathcal{M}_{ki}\neq\emptyset} f_{ki}(l) = 0$ $\qquad (l\in\mathcal{L}, i\in\mathcal{N}, i\neq s(l),d(l))$

$\qquad \sum_{l\in\mathcal{L}}^{s(l)\neq j,d(l)\neq i} f_{ij}(l) - \sum_{m\in\mathcal{M}_{ij}} \frac{W}{\ln 2} c_{ij}^m \leq 0$ $\qquad (i,j\in\mathcal{N}, i\neq j, \mathcal{M}_{ij}\neq\emptyset)$

$\qquad$ Linear constraints for c_{ij}^m $\qquad\qquad (i,j\in\mathcal{N}, i\neq j, m\in\mathcal{M}_{ij})$

$\qquad t_i^m, s_{ij}^m, c_{ij}^m, u_{ijk}^m \geq 0$ $\qquad\qquad (i,j,k\in\mathcal{N}, i\neq j, m\in\mathcal{M}_{ij}, m\in\mathcal{M}_k)$

$\qquad K, f_{ij}(l) \geq 0$ $\qquad\qquad (l\in\mathcal{L}, i,j\in\mathcal{N}, i\neq j, i\neq d(l), j\neq s(l), \mathcal{M}_{ij}\neq\emptyset)$

$\qquad (\mathbf{x},\mathbf{q})\in\Omega_z$

Fig. 5. Linear relaxation.

Fig. 4. A convex hull for $c_{ij}^m = \ln(1+s_{ij}^m)$.

all constraints in Fig. 1 so that the relaxed problem can be solved by an LP. This solution provides an upper bound.

Note that in Fig. 1, $t_k^m s_{ij}^m$ and $\log_2(1+s_{ij}^m)$ are nonlinear terms. For product $t_k^m s_{ij}^m$, we apply a novel method based on *Reformulation-Linearization Technique* (RLT) [25]. This is done by introducing a new variable $u_{ijk}^m = t_k^m s_{ij}^m$ and add linear constraints for the new variable. Suppose t_k^m and s_{ij}^m are bounded by $(t_k^m)_L \leq t_k^m \leq (t_k^m)_U$ and $(s_{ij}^m)_L \leq s_{ij}^m \leq (s_{ij}^m)_U$, respectively. Thus, we have

$$[t_k^m - (t_k^m)_L] \cdot [s_{ij}^m - (s_{ij}^m)_L] \geq 0 ,$$
$$[t_k^m - (t_K^m)_L] \cdot [(s_{ij}^m)_U - s_{ij}^m] \geq 0 ,$$
$$[(t_k^m)_U - t_k^m] \cdot [s_{ij}^m - (s_{ij}^m)_L] \geq 0 ,$$
$$[(t_k^m)_U - t_k^m] \cdot [(s_{ij}^m)_U - s_{ij}^m] \geq 0 .$$

Substituting $u_{ijk}^m = t_k^m s_{ij}^m$, we have the following linear constraints for u_{ijk}^m.

$$(t_k^m)_L \cdot s_{ij}^m + (s_{ij}^m)_L \cdot t_k^m - u_{ijk}^m \leq (t_k^m)_L \cdot (s_{ij}^m)_L ,$$

$$(t_k^m)_U \cdot s_{ij}^m + (s_{ij}^m)_L \cdot t_k^m - u_{ijk}^m \geq (t_k^m)_U \cdot (s_{ij}^m)_L ,$$
$$(t_k^m)_L \cdot s_{ij}^m + (s_{ij}^m)_U \cdot t_k^m - u_{ijk}^m \geq (t_k^m)_L \cdot (s_{ij}^m)_U ,$$
$$(t_k^m)_U \cdot s_{ij}^m + (s_{ij}^m)_U \cdot t_k^m - u_{ijk}^m \leq (t_k^m)_U \cdot (s_{ij}^m)_U .$$

For the nonlinear term $\log_2(1+s_{ij}^m) = \frac{1}{\ln 2}\ln(1+s_{ij}^m)$, we propose to employ three tangential supports for $\ln(1+s_{ij}^m)$, which is a convex hull linear relaxation (see Fig. 4). Suppose that we have the bounds for s_{ij}^m, i.e., $(s_{ij}^m)_L \leq s_{ij}^m \leq (s_{ij}^m)_U$. We introduce a variable $c_{ij}^m = \ln(1+s_{ij}^m)$ and consider how to get a linear relaxation for c_{ij}^m. The curve of $c_{ij}^m = \ln(1+s_{ij}^m)$ can be bounded by four segments (or a convex hull), where segments I, II, and III are tangential supports and segment IV is the chord (see Fig. 4). In particular, the three tangent segments are tangential at points $(1+(s_{ij}^m)_L, \ln(1+(s_{ij}^m)_L))$, $(1+\beta, \ln(1+\beta))$, and $(1+(s_{ij}^m)_U, \ln(1+(s_{ij}^m)_U))$, where

$$\beta = \frac{[1+(s_{ij}^m)_L]\cdot[1+(s_{ij}^m)_U]\cdot[\ln(1+(s_{ij}^m)_U)-\ln(1+(s_{ij}^m)_L)]}{(s_{ij}^m)_U - (s_{ij}^m)_L} - 1$$

is the horizontal location for the point that is intersected by extending segments I and III; segment IV is the segment that joins points $(1+(s_{ij}^m)_L, \ln(1+(s_{ij}^m)_L))$ and $(1+(s_{ij}^m)_U, \ln(1+(s_{ij}^m)_U))$. The convex region defined by the four segments can be described by the following four *linear* constraints.

$$[1+(s_{ij}^m)_L]\cdot c_{ij}^m - s_{ij}^m \leq [1+(s_{ij}^m)_L]\cdot[\ln(1+(s_{ij}^m)_L)-1]+1 ,$$
$$(1+\beta)\cdot c_{ij}^m - s_{ij}^m \leq (1+\beta)\cdot[\ln(1+\beta)-1]+1 ,$$
$$[1+(s_{ij}^m)_U]\cdot c_{ij}^m - s_{ij}^m \leq [1+(s_{ij}^m)_U]\cdot[\ln(1+(s_{ij}^m)_U)-1]+1 ,$$
$$[(s_{ij}^m)_U-(s_{ij}^m)_L]\cdot c_{ij}^m + [\ln(1+(s_{ij}^m)_L)-\ln(1+(s_{ij}^m)_U)]\cdot s_{ij}^m$$
$$\geq (s_{ij}^m)_U\cdot\ln(1+(s_{ij}^m)_L) - (s_{ij}^m)_L\cdot\ln(1+(s_{ij}^m)_U) .$$

As a result, the nonlinear ln (or log) term is relaxed into linear constraints.

After relaxing all nonlinear terms for a problem, say problem z, we have a relaxed problem $\hat{z}$ in Fig. 5, which is an LP. In problem $\hat{z}$, $\mathbf{x}$ and $\mathbf{q}$ are the vectors for all x_{ij}^m and q_{ij}^m variables, respectively, $(x_{ij}^m)_L, (x_{ij}^m)_U, (q_{ij}^m)_L$, and $(q_{ij}^m)_U$ are constant bounds, and $\Omega_z = \{(\mathbf{x}, \mathbf{q}) : (x_{ij}^m)_L \leq x_{ij}^m \leq (x_{ij}^m)_U, (q_{ij}^m)_L \leq q_{ij}^m \leq (q_{ij}^m)_U\}$ is the core optimization space of $(\mathbf{x}, \mathbf{q})$.

A relaxed problem $\hat{z}$ can be solved by an LP in polynomial time. Denote its solution as $LP(\hat{z})$. This gives us an upper bound to problem z.

6.2 Determining Lower Bound

To find a lower bound for problem z, it is sufficient to find a feasible solution to this problem. By feasible solution, we mean that it satisfies all constraints for problem z in Fig. 1, despite that the objective value corresponding to the feasible solution may not be optimal (maximum). Although any feasible solution to problem z can serve as a lower bound, we strive to find one feasible solution that can offer a tight lower bound. We can find such a feasible solution (denoted as ψ_z) based on $LP(\hat{z})$ by searching its neighborhood, which we call *local search.*

The local search algorithm begins with an initial feasible solution. Such a solution may be far away from the optimum and may not provide a tight lower bound. Thus, we will iteratively improve the current solution to achieve a better lower bound. Until we can no longer improve (increase) the lower bound, we are done.

To obtain an initial feasible solution, we set $x_{ij}^m = (x_{ij}^m)_L$ for scheduling and $q_{ij}^m = (q_{ij}^m)_L$ for power control. Then we can compute SINR value s_{ij}^m by (3). When an SINR value is larger than or equal to α, the achieved capacity is $W \log_2(1 + s_{ij}^m)$. Otherwise (i.e., SINR $< \alpha$), the transmission is considered unsuccessful. Note that although the flow rates $f_{ij}(l)$ in the relaxed solution $LP(\hat{z})$ guarantee flow balance at each node, such flow rates may exceed the capacities on some links under the initial x_{ij}^m and q_{ij}^m values. To find feasible flow rates under current x_{ij}^m and q_{ij}^m, we compare the achievable link capacity (under current x_{ij}^m and q_{ij}^m values) to the aggregated flow rates $f_{ij}(l)$ on each link $i \rightarrow j$ by computing the ratio between the two (denoted as λ_{ij}) as follows.

$$\lambda_{ij} = \frac{\sum_{m \in \mathcal{M}_{ij}} W \log_2(1 + s_{ij}^m)}{\sum_{l \in \mathcal{L}}^{s(l) \neq j, d(l) \neq i} f_{ij}(l)} . \qquad (17)$$

If $\lambda_{ij} < 1$ for some link $i \rightarrow j$, then the aggregated flow rates exceed the link capacity and the link capacity constraint on $i \rightarrow j$ is violated. In this case, we need to scale down the flow rates on link $i \rightarrow j$ (to satisfy link capacity constraint) and the flow rates on all other links (to maintain flow balance in the network) by a value $\lambda \leq \lambda_{ij}$. On the other hand, we want to have a λ as large as possible so as to maximize the scaling factor (our objective). Such a value is the bottleneck value λ_{ij} among all links (denoted as $\lambda_{\min} = \min\{\lambda_{ij} : i, j \in \mathcal{N}, i \neq j, \mathcal{M}_{ij} \neq \emptyset\}$). We now have a complete solution $\lambda_{\min} \cdot f_{ij}(l), (x_{ij}^m)_L, (q_{ij}^m)_L$ for routing, scheduling, and power control, respectively. The achieved objective is

---

Initialization:
1. Set $x_{ij}^m = (x_{ij}^m)_L$ and $q_{ij}^m = (q_{ij}^m)_L$.
2. Compute the ratio λ_{ij} by (17) for each link $i \rightarrow j$ and denote $\lambda_{\min} = \min\{\lambda_{ij} : i, j \in \mathcal{N}, i \neq j, \mathcal{M}_{ij} \neq \emptyset\}$.

Iteration:
3. Suppose $\lambda_{ij} = \lambda_{\min}$.
4. If we can increase q_{ij}^m on a used band {
5.   Suppose band m has the largest q_{ij}^m values in solution $LP(\hat{z})$ among these bands.
6.   Increase q_{ij}^m such that $q_{ij}^m \leq (q_{ij}^m)_U$ and for any other
7.   link $k \rightarrow h$, their newly updated $\lambda_{kh} > \lambda_{\min}$. }
8. else, if we can increase q_{ij}^m on an available and unused band {
9.   Suppose band m has the largest q_{ij}^m values in solution $LP(\hat{z})$ among these bands.
10.   Increase q_{ij}^m such that $q_{ij}^m \leq (q_{ij}^m)_U$ and for any other
11.   link $k \rightarrow h$, their newly updated $\lambda_{kh} > \lambda_{\min}$.
12.   If q_{ij}^m increases, then $x_{ij}^m = 1$. }
13. else the iteration terminates.

Fig. 6. Pseudocode of proposed local search algorithm.

$\lambda_{\min} \cdot K$, where K is the objective value in the relaxed solution $LP(\hat{z})$.

In the next iteration, we aim to improve the current solution. Note that if we can increase $\lambda_{\min}$, then the current solution is improved. Suppose link $i \rightarrow j$ is the link with $\lambda_{ij} = \lambda_{\min}$. To increase λ_{ij}, we try to increase transmission power q_{ij}^m on some band m under the constraint $q_{ij}^m \leq (q_{ij}^m)_U$. Based on the constraints in Fig. 1, we may update the values of other variables to maintain feasibility. For example, by the first constraint in Fig. 1, we need to increase x_{ij}^m from 0 to 1 if q_{ij}^m is increased form 0 to a positive value. Moreover, as a consequence of increased q_{ij}^m, the interference to other transmissions on band m is increased and thus the achieved capacities for other links are decreased. Thus, q_{ij}^m can be successfully increased only if for any other link $k \rightarrow h$, its updated λ_{kh} will not fall below the current $\lambda_{\min}$. If the current solution can be improved (with a larger $\lambda_{\min}$), then we continue to the next iteration of improvement. Otherwise, the local search algorithm terminates. The pseudocode of our local search algorithm is given in Fig. 6.

6.3 Partitioning Problem

Our proposed partitioning approach differs from that in standard branch-and-bound procedure. In standard branch-and-bound procedure, partitioning (see line 12 in Fig. 3) is done by choosing a variable with the largest relaxation error and uses its value in the relaxed solution $LP(\hat{z})$ to divide its value set into two smaller sets. The reason of this approach (with the largest relaxation error) is that such a variable is likely to lead to a larger gap between upper and lower bounds. Thus, we should partition its value set such that the relaxation error will became smaller. This division (on value set) also divides the optimization space for problem z into two smaller spaces, which result in two new problems z_1 and z_2, respectively.

Such standard partitioning technique, however, does not explore any problem specific property on choosing partition variables. We find that if we weigh the significance of each variable when choosing a partitioning

variable, the complexity of the overall algorithm can be decreased significantly. For our problem, x variables are more important than q variables. That is, we should first decide whether or not a band is used for transmission. Only if a band is used, then we further decide the transmission power on this band. Thus, we first choose one x variable for partitioning and after x variables are all determined (either 0 or 1), we consider q variables for partitioning.

When choosing a specific x_{ij}^m variable for partitioning, we pick the one with the maximum relaxation error, which is defined as $\min\{\hat{x}_{ij}^m, 1 - \hat{x}_{ij}^m\}$, where $\hat{x}_{ij}^m$ is the relaxed solution in $LP(\hat{z})$ and is not necessarily an integer. Once partitioned, the value set for this x can have element 0 or 1. This variable can be replaced by a constant 0 or 1 in the two new subproblems.

We further observe that by fixing the value of x_{ij}^m, some other variables may also be fixed. We note that by (1), a node cannot transmit and receive on the same band. Based on this observation, if the value of x_{ij}^m is set to 1, then we have $x_{ki}^m = 0$ for $k \in \mathcal{N}, k \neq i, m \in \mathcal{M}_k$ and $x_{jh}^m = 0$ for $h \in \mathcal{N}, h \neq j, m \in \mathcal{M}_h$. On the other hand, if the value of x_{ij}^m is set to 0, then we have $q_{ij}^m = 0$ based on (2).

After all x variables are fixed, we can start partition on q variables. Again, we choose one q variable with the maximum relaxation error for partitioning. The relaxation error of a variable q_{ij}^m is defined as $\min\{q_{ij}^m - \lfloor q_{ij}^m \rfloor, \lfloor q_{ij}^m \rfloor + 1 - q_{ij}^m\}$, where q_{ij}^m is the relaxed solution in $LP(\hat{z})$ and is not necessarily integer. The value set of q_{ij}^m in problem z is $\{(q_{ij}^m)_L, (q_{ij}^m)_L + 1, \cdots, (q_{ij}^m)_U\}$. Then its new value set in the two subproblems will be $\{(q_{ij}^m)_L, (q_{ij}^m)_L + 1, \cdots, \lfloor q_{ij}^m \rfloor\}$ and $\{\lfloor q_{ij}^m \rfloor + 1, \lfloor q_{ij}^m \rfloor + 2, \cdots, (q_{ij}^m)_U\}$, respectively.

7 DISCUSSION

We now discuss how to interpret and apply our theoretical results. Note that in previous sections, the problem setting under study and the proposed solution procedure are concerned with an instance of CRN where the available bands at each node is given (static). Our theoretical result offers a performance benchmark for a multi-hop CRN. Such a result is not yet available in the literature.

We make no claim that our solution will be deployed in its current form in the real world. Instead, we position our result to be used as a reference optimal result/performance benchmark for any other algorithms that will be developed for deployment. To study the performance of other proposed algorithm, we can record its performance over a time period T. For performance benchmarking/comparison, we can run our solution offline by breaking up the time period T into smaller time intervals where the available bands can be assumed static within each small time interval $t_1, t_2, \cdots$. For each time interval t_i, our algorithm can provide a solution π_i. Then a complete solution Π for the entire time period T is obtained by piecing up together all solutions π_i. Solution Π is either optimal (by setting $\varepsilon = 0$) or near-

optimal (if $\varepsilon > 0$).[1] Then we can compare how far the performance of the other proposed algorithm is from the optimum.

8 NUMERICAL RESULTS

In this section, we present numerical results for the proposed solution. Our goals are to demonstrate the efficacy of the solution procedure and to offer quantitative understanding on the joint optimization at different layers in the SINR model.

8.1 Simulation Setting

We consider 20-, 30-, and 50-node CRNs with each node randomly located in a 50x50 area (see Tables 1, 3, and 8). For the ease of exposition, we normalize all units for distance, bandwidth, rate, and power based on (8) with appropriate dimensions. We assume there are $|\mathcal{M}| = 10$, 20, and 30 frequency bands in the network for the 20-, 30-, and 50-node CRNs and each band has a bandwidth of $W = 50$. At each node, only a subset of these bands is available. For the 20- and 30-node CRNs, we assume there are 5 user communication sessions (see Tables 2 and 4) and for the 50-node CRN, the number of user communication sessions is 10 (see Table 9). The source node and destination node for each session are randomly selected and the minimum rate requirement of each session is randomly generated within $[1, 10]$.

We assume $g_{ij} = d_{ij}^{-4}$ and the SINR threshold $\alpha = 3$ [12]. We assume that under the maximum transmission power, a node at distance 20 can receive data when there is no interference. Thus, we have $\frac{(20)^{-4} P_{\max}}{\eta W} = \alpha$, i.e., the maximum transmission power is $P_{\max} = \alpha \cdot (20)^4 \cdot \eta W = 4.8 \cdot 10^5 \eta W$. We assume that power control can be done in $Q = 10$ levels. For our proposed solution, we set ε to 0.1, which guarantees that the solution is at least 90% optimal.

We note that the brute force approach cannot solve the problem even for the 20-node CRN. The solution space of the capacity problem in Section 3.3 includes all possible sets of values for $(x_{ij}^m, q_{ij}^m, K, f_{ij}(l))$. Thus, the number of solutions examined in the brute force approach is clearly more than the number of all possible sets of values for q_{ij}^m variables. We now analyze the number of all possible sets of values for q_{ij}^m variables for the 20-node CRN. The number of q_{ij}^m variables is about $20 \cdot (20 - 1) \cdot 5 = 1900$, where 5 is an approximation of the average number of available bands on a link. Each q_{ij}^m variable has $(Q + 1)$ or 11 possible values. Thus, the number of all possible sets of values for q_{ij}^m variables is about 11^{1900}. Therefore, for the 20-node CRN, the number of solutions examined in the brute force approach is at least 11^{1900}. Suppose

---

1. We discuss the case of $\varepsilon = 0$. The discussion for $\varepsilon > 0$ is similar and thus is omitted. First of all, since available bands are static in any time interval t_i, solution π_i is optimal as we discussed in Section 5.2. That is, there is no solution $\hat{\pi}_i$ that can achieve better performance than π_i in time interval t_i. Since this claim holds for each time interval t_i, there is no solution $\hat{\Pi}$ that can achieve better performance than Π for any time interval. Therefore, solution Π is optimal.

TABLE 1
Location and available frequency bands at each node for a 20-node network.

Node	Location	Available Bands	Node	Location	Available Bands
1	$(0.1, 9.9)$	1, 2, 3, 4, 7, 8, 9, 10	11	$(28.1, 25.6)$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
2	$(29.2, 31.7)$	1, 2, 3, 4, 5, 7, 8, 10	12	$(32.3, 38)$	1, 8, 9, 10
3	$(3, 31.1)$	1, 4, 5, 6	13	$(47.2, 2.6)$	3, 5, 10
4	$(11.8, 40.1)$	1, 2, 3, 4, 6, 9, 10	14	$(44.7, 15)$	2, 3, 6, 7, 8
5	$(15.8, 9.7)$	1, 2, 3, 5, 6, 8, 9	15	$(44.7, 24)$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
6	$(16.3, 19.5)$	3, 5, 6, 8, 9	16	$(47.9, 43.8)$	1, 3
7	$(0.6, 27.4)$	1, 4, 8, 9, 10	17	$(46.4, 16.8)$	1, 7, 9
8	$(22.6, 40.9)$	1, 2, 3, 5, 7, 9, 10	18	$(11.5, 12.2)$	2, 5, 6, 10
9	$(35.3, 10.3)$	2, 9	19	$(28.2, 14.8)$	4, 5, 6, 7, 8, 9, 10
10	$(31.9, 19.6)$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	20	$(2.5, 14.5)$	1, 7, 10

TABLE 2
Source node, destination node, and minimum rate requirement of each session in the 20-node network.

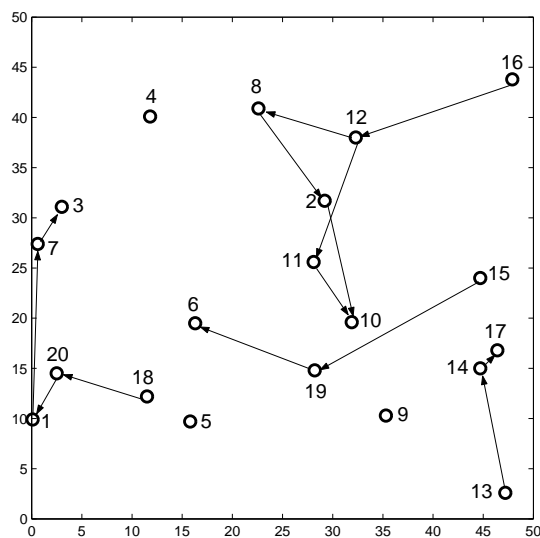| Session l	Source Node $s(l)$	Dest. Node $d(l)$	Min. Rate Req. $r(l)$
1	16	10	9
2	18	3	1
3	12	11	4
4	13	17	3
5	15	6	2

Fig. 7. The routing topology for the 20-node 5-session network.

one solution can be examined in 10^{-6} second. Then the running time is $11^{1900} \cdot 10^{-6} > 11^{1894}$ seconds, which is $11^{1894}/(365 \cdot 24 \cdot 60 \cdot 60) > 11^{1894}/10^8 > 11^{1886}$ years. Therefore, the brute force approach cannot be used to solve our problem even for the 20-node CRN.

8.2 Results

For the 20-node network with 5 sessions, our solution achieves a scaling factor $K = 13.24$. Then, based on the minimum rate requirement $r(l)$ in Table 2, the flow rates $K \cdot r(l)$ for the five sessions are $119.16, 13.24, 52.96, 39.72, 26.48$, respectively. Figure 7 shows the routing topology in the final solution. The flow rates for each session on the links along its path

are as follows.
Session $l = 1$: $f_{16,12}(1) = 119.16$, $f_{12,8}(1) = 103.30$, $f_{12,11}(1) = 15.86$, $f_{8,2}(1) = 103.30$, $f_{11,10}(1) = 15.86$, $f_{2,10}(1) = 103.30$;
Session $l = 2$: $f_{18,20}(2) = 13.24$, $f_{20,1}(2) = 13.24$, $f_{1,7}(2) = 13.24$, $f_{7,3}(2) = 13.24$;
Session $l = 3$: $f_{12,11}(3) = 52.96$;
Session $l = 4$: $f_{13,14}(4) = 39.72$, $f_{14,17}(4) = 39.72$;
Session $l = 5$: $f_{15,19}(5) = 26.48$, $f_{19,6}(5) = 26.48$.
It is easy to verify that flow balance holds at source and destination nodes for each session, as well as at all relay nodes. Note that flow splitting and multi-path routing are used for session 1, which has the largest rate requirement.

Our solution also solves the scheduling variables x_{ij}^m as follows. Note that it is sufficient to list only non-zero x_{ij}^m variables.
Band $m = 1$: $x_{7,3}^1 = 1$, $x_{16,12}^1 = 1$;
Band $m = 2$: $x_{8,2}^2 = 1$;
Band $m = 3$: $x_{13,14}^3 = 1$;
Band $m = 4$: $x_{1,7}^4 = 1$, $x_{2,10}^4 = 1$;
Band $m = 5$: $x_{11,10}^5 = 1$;
Band $m = 6$: $x_{15,19}^6 = 1$;
Band $m = 7$: $x_{14,17}^7 = 1$, $x_{20,1}^7 = 1$;
Band $m = 8$: $x_{12,11}^8 = 1$;
Band $m = 9$: $x_{12,8}^9 = 1$, $x_{19,6}^9 = 1$;
Band $m = 10$: $x_{18,20}^{10} = 1$.
The transmission power levels on their respective frequency bands are as follows.
Band $m = 1$: $q_{7,3}^1 = 1$, $q_{16,12}^1 = 7$;
Band $m = 2$: $q_{8,2}^2 = 2$;
Band $m = 3$: $q_{13,14}^3 = 2$;
Band $m = 4$: $q_{1,7}^4 = 7$, $q_{2,10}^4 = 2$;
Band $m = 5$: $q_{11,10}^5 = 1$;
Band $m = 6$: $q_{15,19}^6 = 9$;
Band $m = 7$: $q_{14,17}^7 = 1$, $q_{20,1}^7 = 1$;
Band $m = 8$: $q_{12,11}^8 = 3$;
Band $m = 9$: $q_{12,8}^9 = 1$, $q_{19,6}^9 = 3$;
Band $m = 10$: $q_{18,20}^{10} = 1$.
Note that the same frequency band may be used by concurrent transmissions. For example, since $x_{7,3}^1 = 1$ and $x_{16,12}^1 = 1$, we have that both nodes 7 and 16 are transmitting on band 1. Such concurrent transmissions

TABLE 3
Location and available frequency bands at each node for a 30-node network.

Node	Location	Available Bands	Node	Location	Available Bands
1	$(7, 0.7)$	1, 2, 6, 7, 16, 17, 19, 20	16	$(30.3, 28.1)$	7, 8, 11, 16, 17, 19, 20
2	$(5, 4)$	3, 5, 9, 12, 14, 15	17	$(32, 41.1)$	7, 11, 16, 17, 19, 20
3	$(6.8, 14)$	1, 2, 6, 7, 8, 11, 16, 17, 19, 20	18	$(14.1, 33.7)$	3, 4, 5
4	$(15.7, 3.3)$	1, 2, 7, 16, 20	19	$(23, 46.4)$	3, 12, 15
5	$(9.5, 17)$	3, 4, 5, 9, 12	20	$(30.3, 9.3)$	5, 9
6	$(19.4, 17.1)$	1, 2, 6, 7, 8, 16, 19, 20	21	$(17.6, 29.2)$	1, 2, 6, 7, 8, 11, 16, 17, 19, 20
7	$(34.7, 14.6)$	3, 4, 5, 9, 12, 14	22	$(27.1, 27.8)$	9, 12, 14, 15
8	$(4.9, 25.9)$	3, 4, 12	23	$(26.9, 45.9)$	3, 4, 5, 9, 10, 12, 13, 14, 15, 17
9	$(46.6, 42.1)$	10, 18	24	$(43.3, 32.4)$	1, 2, 11, 16, 17, 20
10	$(8.3, 38.3)$	3, 4, 5, 9, 14	25	$(45.4, 8.2)$	3, 4, 5, 9, 12, 14
11	$(26.7, 11.1)$	1, 6, 7, 8, 11, 16, 17, 19, 20	26	$(43.4, 35)$	3, 5, 9, 15
12	$(36.4, 47.3)$	10, 13, 18	27	$(41.3, 45.1)$	1, 16, 20
13	$(24.3, 21.2)$	1, 2, 6, 8, 11, 19	28	$(14.4, 30.3)$	1, 2, 6, 7, 8, 11, 16, 17, 20
14	$(23.1, 0.8)$	3, 5, 9, 14	29	$(41.6, 41.7)$	3, 4, 5, 9, 10, 12, 14, 15, 18
15	$(21.4, 19.2)$	4, 9, 12, 14	30	$(25.9, 12)$	1, 2, 6, 7, 8, 11, 16, 17, 19, 20

TABLE 4
Source node, destination node, and minimum rate requirement of each session in the 30-node network.

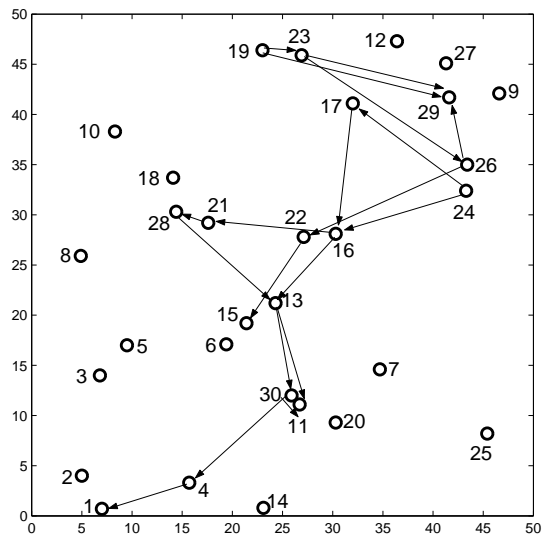| Session l	Source Node $s(l)$	Dest. Node $d(l)$	Min. Rate Req. $r(l)$
1	16	28	4
2	24	11	7
3	13	1	1
4	19	29	8
5	26	15	1

Fig. 8. The routing topology for the 30-node 5-session network.

are allowed as long as the SINR at each receiving node is no less than α. For example, SINR at receiving node 12 on band 1 is $s_{16,12}^1 = \frac{g_{16,12} \cdot q_{16,12}^1}{\frac{\eta W Q}{P_{\max}} + g_{7,12} \cdot q_{7,3}^1} = \frac{(1.303 \cdot 10^{-5}) \cdot 7}{2.083 \cdot 10^{-5} + (8.011 \cdot 10^{-7}) \cdot 1} = 4.22$, which is larger than $\alpha = 3$. Thus, transmission $16 \to 12$ on band 1 is successful. Following the same token, we can compute all SINR s_{ij}^m values as follows.

Band $m = 1$: $s_{7,3}^1 = 118.47$, $s_{16,12}^1 = 4.22$;
Band $m = 2$: $s_{8,2}^2 = 5.84$;

Band $m = 3$: $s_{13,14}^3 = 3.75$;
Band $m = 4$: $s_{1,7}^4 = 3.33$, $s_{2,10}^4 = 3.14$;
Band $m = 5$: $s_{11,10}^5 = 18.87$;
Band $m = 6$: $s_{15,19}^6 = 3.39$;
Band $m = 7$: $s_{14,17}^7 = 1261.14$, $s_{20,1}^7 = 65.46$;
Band $m = 8$: $s_{12,11}^8 = 4.90$;
Band $m = 9$: $s_{12,8}^9 = 3.56$, $q_{19,6}^9 = 4.74$;
Band $m = 10$: $s_{18,20}^{10} = 6.45$.

As expected, we see that the achieved SINR at each receiving node on each band is larger than $\alpha = 3$ in our solution.

For each link, we can further verify that flow rates on this link do not exceed its capacity. For example, for link $16 \to 12$, there is a flow rate $f_{16,12}(1) = 119.16$ on this link. The achieved capacity is $W \log_2(1 + s_{16,12}^1) = W \log_2(1 + 4.22) = 119.21$.

The above results are for 20-node network. The results for 30- and 50-node networks are similar and we abbreviate our discussion. For the 30-node 5-session network, the available bands at each node and the location of each node are shown in Table 3. The source, destination, and minimum rate requirement of each session are shown in Table 4. Our solution gives a routing topology in Fig. 8. The achieved scaling factor is 31.18. Then, based on the minimum rate requirement $r(l)$ in Table 4, the flow rates $K \cdot r(l)$ for the five sessions are $124.72, 218.26, 31.18, 249.44, 31.18$, respectively. The flow rates for each session on links along its paths are shown in Table 5. Our solution solves the scheduling variables x_{ij}^m and power control variables q_{ij}^m. Again, we list only non-zero x_{ij}^m variables in Table 6 and non-zero q_{ij}^m variables in Table 7.

For the 50-node 10-session network, the available bands at each node and the location of each node are shown in Table 8. The source, destination, and minimum rate requirement of each session are shown in Table 9. Our solution gives a routing topology in Fig. 9. The achieved scaling factor is 13.36. The detailed solution on power control, scheduling, and routing are similar to the 20- and 30-node networks and are omitted to conserve space.

TABLE 5
Flow rates for each session in the 30-node 5-session network.

Session l	Session rate $K \cdot r(l)$	Flow rates of session l on each link $f_{ij}(l)$
1	124.72	$f_{16,21}(1) = 124.72$, $f_{21,28}(1) = 124.72$;
2	218.26	$f_{24,16}(2) = 113.91$, $f_{24,17}(2) = 104.36$, $f_{17,16}(2) = 104.36$, $f_{16,21}(2) = 24.22$, $f_{16,13}(2) = 194.05$, $f_{21,28}(2) = 24.22$, $f_{28,13}(2) = 24.22$, $f_{13,11}(2) = 160.59$, $f_{13,30}(2) = 57.68$, $f_{30,11}(2) = 57.68$;
3	31.18	$f_{13,30}(3) = 31.18$, $f_{30,4}(3) = 31.18$, $f_{4,1}(3) = 31.18$;
4	249.44	$f_{19,23}(4) = 167.49$, $f_{19,29}(4) = 81.95$, $f_{23,26}(4) = 81.26$, $f_{23,29}(4) = 86.23$, $f_{26,29}(4) = 81.26$;
5	31.18	$f_{26,22}(5) = 31.18$, $f_{22,15}(5) = 31.18$.

TABLE 6
Scheduling for the 30-node network.

Band	Scheduling	Band	Scheduling	Band	Scheduling	Band	Scheduling
1	$x_{4,1}^1 = 1$, $x_{21,28}^1 = 1$;	6	$x_{13,11}^6 = 1$;	11	$x_{17,16}^{11} = 1$;	16	$x_{24,17}^{16} = 1$, $x_{30,11}^{16} = 1$;
2	$x_{28,13}^2 = 1$;	7	$x_{16,21}^7 = 1$;	12	$x_{19,23}^{12} = 1$;	17	$x_{24,16}^{17} = 1$;
3	$x_{19,29}^3 = 1$;	8	$x_{16,13}^8 = 1$;	13	$x_{ij}^{13} = 0$;	18	$x_{ij}^{18} = 0$;
4	$x_{23,29}^4 = 1$;	9	$x_{26,22}^9 = 1$;	14	$x_{22,15}^{14} = 1$;	19	$x_{13,30}^{19} = 1$;
5	$x_{26,29}^5 = 1$;	10	$x_{ij}^{10} = 0$;	15	$x_{23,26}^{15} = 1$;	20	$x_{30,4}^{20} = 1$.

TABLE 7
Transmission power levels for the 30-node network.

Band	Transmission power	Band	Transmission power	Band	Transmission power	Band	Transmission power
1	$q_{4,1}^1 = 1$, $q_{21,28}^1 = 1$;	6	$q_{13,11}^6 = 2$;	11	$q_{17,16}^{11} = 2$;	16	$q_{24,17}^{16} = 3$, $q_{30,11}^{16} = 1$;
2	$q_{28,13}^2 = 3$;	7	$q_{16,21}^7 = 4$;	12	$q_{19,23}^{12} = 1$;	17	$q_{24,16}^{17} = 7$;
3	$q_{19,29}^3 = 9$;	8	$q_{16,13}^8 = 2$;	13	$q_{ij}^{13} = 0$;	18	$q_{ij}^{18} = 0$;
4	$q_{23,29}^4 = 4$;	9	$q_{26,22}^9 = 7$;	14	$q_{22,15}^{14} = 1$;	19	$q_{13,30}^{19} = 1$;
5	$q_{26,29}^5 = 1$;	10	$q_{ij}^{10} = 0$;	15	$q_{23,26}^{15} = 10$;	20	$q_{30,4}^{20} = 4$.

TABLE 8
Location and available frequency bands at each node for a 50-node network.

Node	Location	Available Bands	Node	Location	Available Bands
1	$(11.1, 21.7)$	2, 3, 4, 8, 25	26	$(25.2, 27.2)$	10, 14, 20, 24, 26
2	$(0.1, 4)$	6, 7, 10, 13, 14, 20, 23, 24, 26, 28	27	$(22.5, 42.2)$	5, 9, 12, 16, 18, 27, 29, 30
3	$(7.2, 16.6)$	6, 10, 14, 20, 23, 24, 26	28	$(30, 31.5)$	6, 13, 24, 26, 28
4	$(11, 32.2)$	6, 7, 10, 13, 14, 20, 23, 24, 26, 28	29	$(35, 22.1)$	6, 10
5	$(16.3, 3.6)$	10, 13, 14, 20, 23	30	$(25.7, 6.2)$	5, 9, 12, 17, 18, 22, 27, 29, 30
6	$(14.5, 24.7)$	8, 11, 25	31	$(34.1, 12.4)$	9, 12, 16, 17, 30
7	$(14.9, 13.7)$	5, 9, 12, 16, 17, 18, 22, 27, 29, 30	32	$(26.4, 30)$	5, 9, 12, 16, 17, 18, 22, 27, 29, 30
8	$(19.5, 14.9)$	7, 24, 28	33	$(14.1, 40.7)$	1, 2, 25
9	$(26.6, 13.4)$	1, 19, 21, 25	34	$(34.4, 46.5)$	9, 17, 18, 30
10	$(22.5, 29.3)$	1, 3, 4, 8, 11, 15, 19	35	$(19, 22.5)$	1, 6, 7, 10, 13, 14, 20, 23, 24, 28
11	$(24.6, 40.5)$	3, 8, 25	36	$(39.9, 25.1)$	6, 13, 14, 20, 23, 24, 26, 28
12	$(38.4, 13.1)$	2, 8, 11, 15	37	$(20.3, 18.2)$	1, 2, 3, 4, 8, 11, 15, 19, 21, 27
13	$(4, 3.9)$	9, 12, 16, 22, 27, 29, 30	38	$(10, 20.5)$	6, 7, 10, 13, 14, 20, 23, 24, 26, 28
14	$(6.1, 18.6)$	9, 12, 16, 17, 18, 22, 27, 30	39	$(20.5, 21.4)$	1, 2, 3, 4, 8, 11, 15, 19, 21, 25
15	$(38.5, 22.6)$	2, 4, 11, 15, 19, 21, 25	40	$(37.1, 28.6)$	7, 10, 13, 14, 20, 23, 24, 26
16	$(1.2, 24.3)$	5, 9, 12, 17, 22, 29, 30	41	$(44.1, 16.1)$	1, 15, 21
17	$(4.9, 42.3)$	5, 27	42	$(41.1, 6)$	9, 29
18	$(18.5, 1.4)$	5, 9, 12, 17, 18, 27, 30	43	$(43, 18.8)$	5, 9, 12, 16, 18, 22
19	$(16.9, 29.1)$	3, 4, 10, 11, 12, 15	44	$(45.4, 24.2)$	9, 12, 16, 17, 18, 30
20	$(33.5, 10.4)$	7, 13, 14, 20, 23, 24, 26, 28	45	$(36.2, 41.2)$	5, 9, 17, 27, 29, 30
21	$(25.6, 12.8)$	6, 7, 20, 23, 24, 28	46	$(27.5, 32.3)$	12, 16, 17, 18, 29, 30
22	$(45.2, 45.5)$	2, 8, 15, 19	47	$(47.8, 13.8)$	22, 27, 29, 30
23	$(43.6, 22.7)$	1, 2, 3, 4, 11, 15, 19, 21	48	$(8.9, 14.8)$	5, 30
24	$(10.6, 40.5)$	4, 15, 19, 21, 25	49	$(6.8, 6.2)$	5, 9, 12, 16, 17, 27, 30
25	$(18.2, 32.7)$	9, 12, 18, 22, 27	50	$(11.7, 35.8)$	1, 2, 3, 4, 8, 11, 15, 19, 21, 25

TABLE 9
Source node, destination node, and minimum rate requirement of each session in the 50-node network.

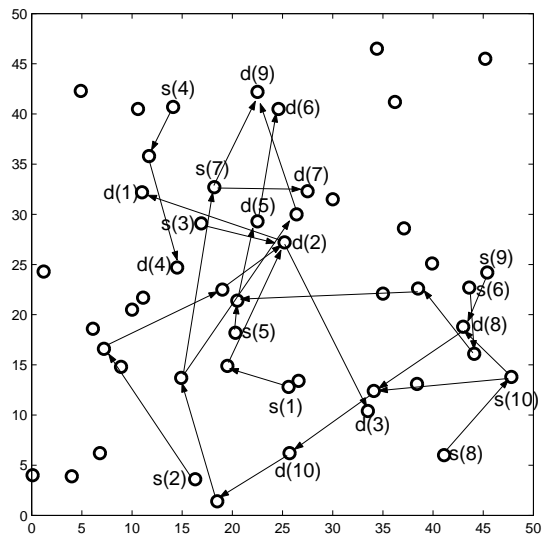| Session l	Source Node $s(l)$	Dest. Node $d(l)$	Min. Rate Req. $r(l)$
1	21	4	4
2	5	26	7
3	19	20	6
4	33	6	10
5	37	10	9
6	23	11	2
7	25	46	3
8	42	43	9
9	44	27	8
10	47	30	1

Fig. 9. The routing topology for the 50-node 10-session network.

In summary, we have the following observations for our numerical results. To maximize capacity, it is necessary to consider multi-path routing (i.e., to allow flow splitting) and to consider spatial reuse (i.e., to allow concurrent transmissions to use the same band). Under multi-path multi-hop routing, even the number of user sessions are small, the number of nodes in the network that will be affect by these sessions (i.e., multi-hop relays) can be large.

9 CONCLUSION

In this paper, we investigated a capacity problem for a multi-hop cognitive radio network under the SINR model. We showed the complex inter-relationship among power control, scheduling, and routing in this problem and developed a mathematical formulation for joint optimization. We presented a solution to this problem based on branch-and-bound framework. The novelty of our contribution is the determination of the core optimization space for this problem, which is much smaller than the whole optimization space, and the design of several efficient algorithms within branch-and-bound framework that exploit the physical significance of core variables to obtain tighter upper and lower bounds. The proposed solution is able to guarantee $(1 - \varepsilon)$ optimal solution and can be used as a performance benchmark for other proposed algorithms in the future.
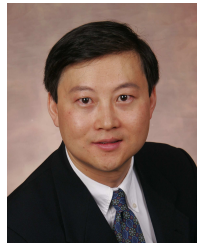
REFERENCES

[1] A. Agarwal and P.R. Kumar, "Capacity bounds for ad hoc and hybrid wireless networks," *ACM SIGCOMM Computer Communications Review,* vol. 34, no. 3, pp. 71–81, July 2004.
[2] M. Alicherry, R. Bhatia, and L. Li, "Joint channel assignment and routing for throughput optimization in multi-radio wireless mesh networks," in *Proc. ACM MobiCom,* pp. 58–72, Cologne, Germany, Aug. 28–Sep. 2, 2005.
[3] M. Andrews and M. Dinitz, "Maximizing capacity in arbitrary wireless networks in the SINR model: Complexity and game theory," in *Proc. IEEE INFOCOM,* pp. 1332–1340, Rio de Janeiro, Brazil, April 19–25, 2009.
[4] A. Behzad and I. Rubin, "Impact of power control on the performance of ad hoc wireless networks," in *Proc. IEEE INFOCOM,* pp. 102–113, Miami, FL, March 13–17, 2005.
[5] V. Bhandari and N.H. Vaidya, "Capacity of multi-channel wireless networks with random (c, f) assignment," in *Proc. ACM MobiHoc,* pp. 229–238, Montreal, QC, Canada, Sep. 9–14, 2007.
[6] R. Bhatia and M. Kodialam, "On power efficient communication over multi-hop wireless networks: joint routing, scheduling and power control," in *Proc. IEEE INFOCOM,* pp. 1457–1466, Hong Kong, China, March 7–11, 2004.
[7] C.C. Chen and D.S. Lee, "A joint design of distributed QoS scheduling and power control for wireless networks," in *Proc. IEEE INFOCOM,* Barcelona, Catalunya, Spain, April 23–29, 2006.
[8] R.L. Cruz and A.V. Santhanam, "Optimal routing, link scheduling and power control in multi-hop wireless networks," in *Proc. IEEE INFOCOM,* pp. 702–711, San Francisco, CA, March 30–April 3, 2003.
[9] T. Elbatt and A. Ephremides, "Joint scheduling and power control for wireless ad-hoc networks," in *Proc. IEEE INFOCOM,* pp. 976–984, New York, NY, June 23–27, 2002.
[10] A. Fanghanel, T. Kesselheim, H. Racke, and B. Vocking, "Oblivious interference scheduling," in *Proc. ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing,* pp. 220–229, Calgary, Alberta, Canada, Aug. 10–12, 2009.
[11] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness, W.H. Freeman and Company,* pp. 245–248, New York, NY, 1979.
[12] A.J. Goldsmith and S.-G. Chua, "Adaptive coded modulation for fading channels," *IEEE Transactions on Communications,* vol. 46, no. 5, pp. 595–602, May 1998.
[13] O. Goussevskaia, Y.A. Oswald, and R. Wattenhofer, "Complexity in geometric SINR," in *Proc. ACM MobiHoc,* pp. 100–109, Montreal, Quebec, Canada, Sep. 9–14, 2007.
[14] O. Goussevskaia and R. Wattenhofer, "Capacity of arbitrary wireless networks," in *Proc. IEEE INFOCOM,* pp. 1872–1880, Rio de Janeiro, Brazil, April 19–25, 2009.
[15] P. Gupta and P.R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory,* vol. 46, no. 2, pp. 388–404, March 2000.

[16] A. K.-Haddad and R. Riedi, "Bounds for the capacity of wireless multihop networks imposed by topology and demand," in *Proc. ACM MobiHoc,* pp. 256–265, Montreal, Quebec, Canada, Sep. 9–14, 2007.

[17] M.M. Halldorsson and R. Wattenhofer, "Wireless communication is in APX," in *Proc. International Colloquium on Automata, Languages and Programming,* pp. 525–536, Rhodes, Greece, July 5–12, 2009.

[18] K. Jain, J. Padhye, V. Padmanabhan, and L. Qiu, "Impact of interference on multi-hop wireless network performance," in *Proc. ACM MobiCom,* pp. 66–80, San Diego, CA, Sep. 14–19, 2003.

[19] S.-W. Jeon, N. Devroye, M. Vu, S.-Y. Chung, and V. Tarokh, "Cognitive networks achieve throughput scaling of a homogeneous network," in *Proc. ICST Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt),* Seoul, Korea, June 26–27, 2009.

[20] M. Kodialam and T. Nandagopal, "Characterizing the capacity region in multi-radio multi-channel wireless mesh networks," in *Proc. ACM MobiCom,* pp. 73–87, Cologne, Germany, Aug. 28–Sep. 2, 2005.

[21] P. Kyasanur and N.H. Vaidya, "Capacity of multi-channel wireless networks: impact of number of channels and interfaces," in *Proc. ACM MobiCom,* pp. 43–57, Cologne, Germany, Aug. 28–Sep. 2, 2005.

[22] S. Li, Y. Liu, and X.-Y. Li, "Capacity of large scale wireless networks under Gaussian channel model," in *Proc. ACM MobiCom,* pp. 140–151, San Francisco, CA, Sep. 14–19, 2008.

[23] B. Radunovic and J.-Y. Le Boudec, "Rate performance objectives of multi-hop wireless networks," in *Proc. IEEE INFOCOM,* pp. 1916–1927, Hong Kong, China, March 7–11, 2004.

[24] K. Ramachandran, E. Belding-Royer, K. Almeroth, and M. Buddhikot, "Interference-Aware Channel Assignment in Multi-Radio Wireless Mesh Networks," in *Proc. IEEE INFOCOM,* Barcelona, Catalunya, Spain, April 23–29, 2006.

[25] H.D. Sherali and W.P. Adams, *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems,* Kluwer Academic Publishers, Dordrecht/Boston/London, Chapter 8, 1999.

[26] Y. Shi and Y.T. Hou, "Optimal power control for multi-hop software defined radio networks," in *Proc. IEEE INFOCOM,* pp. 1694–1702, Anchorage, AL, May 6–12, 2007.

[27] T. Shu and M. Krunz, "Coordinated channel access in cognitive radio networks: A multi-level spectrum opportunity perspective," in *Proc. IEEE INFOCOM,* pp. 2976–2980, Rio de Janeiro, Brazil, April 19–25, 2009.

[28] D. Ugarte and A.B. McDonald, "On the capacity of dynamic spectrum access enabled networks," in *Proc. IEEE DySPAN,* pp. 630–633, Baltimore, MD, Nov. 8–11, 2005.

[29] M. Vu, N. Devroye, M. Sharif, and V. Tarokh, "Scaling laws of cognitive networks," in *Proc. ICST International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom),* pp. 2–8, Orlando, FL, July 31–Aug. 3, 2007.

[30] Z. Wang, P. Giaccone, and E. Leonardi, "A unifying perspective on the capacity of wireless ad hoc networks," in *Proc. IEEE INFOCOM,* pp. 211–215, Phoenix, AZ, April 13–18, 2008.

[31] Y. Xu and W. Wang, "Scheduling partition for order optimal capacity in large-scale wireless networks," in *Proc. ACM MobiCom,* pp. 109–120, Beijing, China, Sep. 20–25, 2009.

[32] C. Yin, L. Gao, and S. Cui, "Scaling laws of overlaid wireless networks: a cognitive radio network vs. a primary network", in *Proc. IEEE GLOBECOM,* pp. 1235–1239, New Orleans, LA, Nov. 30–Dec. 4, 2008.

[33] Y. Yuan, P. Bahl, R. Chandra, T. Moscibroda, and Y. Wu, "Allocating dynamic time-spectrum blocks in cognitive radio networks," in *Proc. ACM MobiHoc,* pp. 130–139, Montreal, Quebec, Canada, Sep. 9–14, 2007.

[34] J. Zhao, H. Zheng, and G. Yang, "Distributed coordination in dynamic spectrum allocation networks," in *Proc. IEEE DySPAN,* pp. 259–268, Baltimore, MD, Nov. 8–11, 2005.

Yi Shi (S'02–M'08) received his Ph.D. degree in computer engineering from Virginia Tech, in 2007. He is currently a Research Scientist in the Department of Electrical and Computer Engineering at Virginia Tech. His research focuses on algorithms and optimization for cognitive radio networks, MIMO and cooperative communication networks, sensor networks, and ad hoc networks. He was a recipient of IEEE INFOCOM 2008 Best Paper Award. He was a recipient of Chinese Government Award for Outstanding Ph.D. Students Abroad in 2006. He served as a TPC member for many major international conferences (including ACM MobiHoc 2009 and IEEE INFOCOM 2009–2011).

Y. Thomas Hou (S'91–M'98–SM'04) received his Ph.D. degree in Electrical Engineering from Polytechnic Institute of New York University in 1998. From 1997 to 2002, Dr. Hou was a Researcher at Fujitsu Laboratories of America, Sunnyvale, CA. Since 2002, he has been with Virginia Polytechnic Institute and State University ("Virginia Tech"), the Bradley Department of Electrical and Computer Engineering, Blacksburg, VA, where he is now an Associate Professor.

Prof. Hou's research interests are cross-layer design and optimization for cognitive radio wireless networks, cooperative communications, MIMO-based ad hoc networks, video communications over dynamic ad hoc networks, and algorithm design for sensor networks. He is a recipient of an Office of Naval Research (ONR) Young Investigator Award (2003) and a National Science Foundation (NSF) CAREER Award (2004) for his research on optimizations and algorithm design for wireless ad hoc and sensor networks. He received five best paper awards from IEEE (including IEEE INFOCOM 2008 Best Paper Award and IEEE ICNP 2002 Best Paper Award) and holds five U.S patents.

Sastry Kompella (S'04–M'07) received his B.E. in Electronics and Communication Engineering from Andhra University, India in May 1996, M.S. in Electrical Engineering from Texas Tech University in May 1998, and the Ph.D. degree in Electrical and Computer Engineering from Virginia Polytechnic Institute and State University, Blacksburg, VA, in 2006. Currently, he is a Researcher in the Information Technology Division at U.S. Naval Research Laboratory, Washington, DC. His research focuses on cross-layer design, optimization, and scheduling in wireless networks.

Hanif D. Sherali is a University Distinguished Professor and the W. Thomas Rice Chaired Professor of Engineering in the Industrial and Systems Engineering Department at Virginia Polytechnic Institute and State University. His areas of research interest are in analyzing problems and designing algorithms for specially structured linear, nonlinear, and integer programs arising in various applications, global optimization methods for non-convex programming problems, location and transportation theory and applications, economic and energy mathematical modeling and analysis. He has published over two hundred refereed articles in various Operations Research journals, has (co-) authored six books in this area, and serves on the editorial board of eight journals. He is an elected member of the U.S National Academy of Engineering.