Spectrum Sharing for Multi-Hop Networking with Cognitive Radios

Y. Thomas Hou, Senior Member, IEEE, Yi Shi, Student Member, IEEE, and Hanif D. Sherali

Abstract—Cognitive Radio (CR) capitalizes advances in signal processing and radio technology and is capable of reconfiguring RF and switching to desired frequency bands. It is a frequency-agile data communication device that is vastly more powerful than recently proposed multi-channel multi-radio (MC-MR) technology. In this paper, we investigate the important problem of multi-hop networking with CR nodes. For such a network, each node has a pool of frequency bands (typically of unequal size) that can be used for communication. The potential difference in the bandwidth among the available frequency bands prompts the need to further divide these bands into sub-bands for optimal spectrum sharing. We characterize the behavior and constraints for such a multi-hop CR network from multiple layers, including modeling of spectrum sharing and sub-band division, scheduling and interference constraints, and flow routing. We develop a mathematical formulation with the objective of minimizing the required network-wide radio spectrum resource for a set of user sessions. Since the formulated model is a mixed-integer non-linear program (MINLP), which is NP-hard in general, we develop a lower bound for the objective by relaxing the integer variables and using a linearization technique. Subsequently, we design a near-optimal algorithm to solve this MINLP problem. This algorithm is based on a novel sequential fixing procedure, where the integer variables are determined iteratively via a sequence of linear programs. Simulation results show that solutions obtained by this algorithm are very close to the lower bounds obtained via the proposed relaxation, thus suggesting that the solution produced by the algorithm is near-optimal.

Index Terms—Cognitive Radio (CR), spectrum sharing, multi-hop networking, interference modeling, cross-layer optimization.

I. INTRODUCTION

RECENT studies sponsored by the FCC have shown that traditional fixed allocation policy is becoming inadequate in addressing today’s rapidly evolving wireless communications. Studies show that many allocated spectrum blocks are not used in certain geographical areas and are idle most of the time. These frequency bands are called the spectrum “white space” (or “hole”). Measurements conducted by the Shared Spectrum Company [18] find that even in the most crowded area near downtown Washington, DC, where both government and commercial spectrum use is intensive, 62% of the spectrum remain white space (a bandwidth is considered white space if it is wider than 1 MHz and remains unoccupied for at least 10 minutes). Another measurement, also conducted by the Shared Spectrum Company [19], shows that even during the 2004 Republican National Convention in New York City (perhaps the most heavily-congested area in the U.S. at that time), there was still significant white space available in the public sector spectrum. These studies have prompted the FCC to explore new innovative policies to encourage dynamic access to the under-utilized spectrum [7]. Wireless devices are allowed to sense and explore a wide range of the frequency spectrum and identify currently unused spectrum bands for data communication. This approach is also called dynamic spectrum access (DSA).

The enabling physical layer technology to realize DSA is cognitive radio (CR), which is a frequency-agile data communication device that has a rich control and monitoring (spectrum sensing) interface [12], [21]. It capitalizes advances in signal processing and radio technology, as well as recent advancements in spectrum policy [25]. A frequency-agile radio module is capable of sensing the available bands [3], [9], [10], [20], [26], [30], reconfiguring RF, and switching to newly-selected frequency bands. Thus, a CR can be programmed to tune and operate on specific frequency bands over a wide spectrum range [25]. An even more profound advance in CR technology is that there is no requirement that selected frequencies/channels be contiguous: the radio can send packets over non-contiguous frequency bands. From an application perspective, CR allows a single radio to provide a wide variety of functions, acting as a cell phone, broadcast receiver, GPS receiver, wireless data terminal, etc.

In this paper, we focus on the multi-hop networking problem for a CR-based wireless network. For such a network, each node senses a set of spectrum bands that it can use. Due to the unequal size of spectrum bands, it is necessary to further divide each band into sub-bands (likely of unequal size) to schedule transmission and reception. There are many fundamental problems that can be posed for such a wireless network in the context of rates and capacity. In this paper, we consider the following problem. Suppose there is a set of user sessions in the network that is characterized by a set of source-destination pairs each having a certain rate requirement. Then, how can we perform spectrum allocation, scheduling and interference avoidance, and multi-hop multi-path routing such that the required network-wide radio spectrum resource is minimized?

To formulate the problem mathematically, we characterize behaviors and constraints from multiple layers for a general multi-hop CR network. Special attention is given to modeling of spectrum sharing and unequal (non-uniform) sub-band
division, scheduling and interference modeling, and multi-path routing. We formulate an optimization problem with the objective of minimizing the required network-wide radio spectrum resource for a set of source-destination pair rate requirements. Since such a problem formulation is a mixed-integer non-linear program (MINLP), which is NP-hard in general [8], we aim to derive a near-optimal solution.

We present a near-optimal algorithm for the formulated MINLP problem. First, we develop a lower bound for the objective by relaxing the integer variables and employing a linearization technique. This lower bound will be used as a measure for the quality of any solution. Then we present a novel sequential fixing (SF) solution procedure where the determination of integer variables is performed iteratively through a sequence of linear programs (LPs). Upon fixing all the integer variables, other variables in the optimization problem can be solved using an LP. Since the solution obtained by the proposed SF algorithm represents an upper bound for the objective, we compare it to the lower bound developed earlier. Simulations show that the results obtained by the SF algorithm are very close to the lower bound, thus suggesting that (1) the lower bound is very tight; and (2) the solution obtained by the SF algorithm is even closer to the optimum and thus is near-optimal. The significance of this theoretical work is to provide a performance benchmark which can be used to evaluate protocols and distributed algorithms for real implementation.

The remainder of this paper is organized as follows. In Section II, we review related work on CR and state-of-the-art on cross-layer optimization for MC-MR networks. In Section III, we characterize the behavior of CR networks from multiple layers and formulate them as mathematical constraints. We also elaborate on the optimal radio resource sharing problem and formulate it as an MINLP problem. In Section IV, we develop a lower bound for this MINLP problem by relaxing integer variables and using linearization. In Section V, we describe the proposed SF algorithm. Section VI presents simulation results and demonstrates the near-optimal performance of the SF algorithm. Section VII concludes this paper.

II. RELATED WORK

CR is based on software defined radio (SDR) [25]. Since its inception, SDR development has witnessed rapid advances. Standards bodies such as IEEE 802 Standards Committee, the SDR Forum, the Object Management Group have been instrumental in promoting open standards for SDR commercialization. Among others, the Software Communications Architecture core framework is the result from standardization efforts on SDR. The IEEE 802.22 working group is in the process of developing a standard for a CR-based interface for use by license-exempt devices on a non-interfering basis in spectrum that is allocated to the TV Broadcast Service. CR employs all the technologies that are available to SDR, plus the additional capability of spectrum sensing and cognition (learning and adaptation).

In CR research community, there have been extensive activities devoted to effective sharing of spectrum or spectrum allocation. For a multi-user single-hop communication in a network environment, a number of approaches have been proposed. For example, in [4], [22], game theory was applied to study spectrum sharing, while in [13], [14], pricing mechanism was used. In [6], Etkin et al. studied a utility maximization problem and solved it under certain condition. In [23], Peng et al. studied the spectrum assignment problem with the aim of maximizing the total utility. In these efforts, routing is not part of the problem.

For the multi-hop networking problem with CRs, there is limited amount of work to date available in the literature. In [32], Zhao et al. designed a distributed coordination approach for spectrum sharing. They showed that this approach offers throughput improvement over a dedicated channel approach. In [29], Ugarte and McDonald studied the network capacity problem for multi-hop CR-based networks and found an upper bound, although it is not clear how tight this bound is. In [31], Xin et al. studied how to assign frequency bands at each node to form a topology such that a certain performance metric can be optimized. A layered graph was proposed to model frequency bands available at each node and to facilitate topology formation and achieve optimization objective. The authors considered the so-called fixed channel approach whereby the radio is assumed to operate on only one channel at a specific time. In [28], Steenstrup studied three different frequency assignment problems: common broadcast frequencies, non-interfering frequencies for simultaneous transmissions, and frequencies for direct source-destination communications. Each is viewed as a graph-coloring problem, and both centralized and distributed algorithms were presented. Within these limited efforts, there remains a lack of results on fundamental theoretical performance limits for multi-hop CR networks.

A closely related line of research is the so-called multi-channel multi-radio (MC-MR) networks (e.g., [1], [5], [15], [16], [24]). It is important to understand that a CR is vastly more powerful and flexible than MC-MR technology. First, the MC-MR platform employs a traditional hardware-based radio technology (i.e., signal processing, modulation, etc., are all implemented in the hardware), and thus each radio can only operate on a single channel at a time and there is no switching of channels on the packet level. As a result, the number of concurrent channels that can be used at a wireless node is limited by the number of hardware-based radios. In contrast, the radio technology in CR is software-based; a CR is capable of switching frequency bands on the packet level. As a result, the number of concurrent frequency bands that can be shared by a single CR is typically much larger than that which can be supported by MC-MR. Second, due to the nature of hardware-based radio technology in MC-MR, a common assumption in MC-MR is that there is a set of “common channels” available for every node in the network; each channel typically has the same bandwidth. However, such an assumption is hardly true for CR networks, in which each node may not have an identical set of frequency bands and each band is likely to be of unequal size. Due to this difference, CR is required to work on a set of frequency bands that are scattered over widely-separated slices of the frequency spectrum with different bandwidths. In summary, these important differences between MC-MR and CR warrant that the algorithmic design
for a CR network is substantially more complex than that for MC-MR. In some sense, an MC-MR-based wireless network can be considered as a special case of a CR-based wireless network. Thus, algorithms designed for CR networks can be tailored to address MC-MR networks, while the converse is not true.

III. CR NETWORK MODEL AND PROBLEM FORMULATION

Table I lists all the relevant notation used in this paper. We consider an ad hoc network consisting of a set of \( N \) nodes. Among these nodes, there are a set of \( L \) uni-cast communication sessions. Denote \( s(l) \) and \( d(l) \) the source and destination nodes of session \( l \in L \), and \( r(l) \) the rate requirement (in b/s) of session \( l \).

A. Modeling of Multi-layer Characteristics

**Modeling of Spectrum Sharing and Sub-band Division.** This mathematical modeling feature and constraints are unique to CR networks and do not exist in MC-MR networks. In a multi-hop CR network, the available spectrum bands at one node may be different from another node in the network. Given a set of available frequency bands at a node, the size (or bandwidth) of each band may differ drastically. For example, among the least-utilized spectrum bands found in [19], the bandwidth between [1240, 1300] MHz (allocated to amateur radio) is 60 MHz, while bandwidth between [1525, 1710] MHz (allocated to mobile satellites, GPS systems, and meteorological applications) is 185 MHz. Such large difference in bandwidths among the available bands suggests the need for further division of the larger bands into smaller sub-bands for more flexible and efficient frequency allocation. Since equal sub-band division of the available spectrum band is likely to yield sub-optimal performance, an unequal division is desirable.

More formally, we model the union of the available spectrum among all the nodes in the network as a set of \( M \) unequally sized bands (see Fig. 1). Denote \( M \) the set of these bands and \( M_i \subseteq M \) the set of available bands (or white-space) at node \( i \in N \), which is likely to be different from that at another node, say \( j \in N \), i.e., possibly \( M_i \neq M_j \). For example, at node \( i \), \( M_i \) may consist of bands I, III, and V, while at node \( j \), \( M_j \) may consist of bands I, IV, and VI. Denote \( W^{(m)} \) the bandwidth of band \( m \in M \). For more flexible and efficient bandwidth allocation and to overcome the disparity in the bandwidth size among the spectrum bands, we assume that band \( m \) can be further divided into up to \( K^{(m)} \) sub-bands, each of which may be of unequal bandwidth. Denote \( u^{(m,k)} \) the fraction of bandwidth for the \( k \)-th sub-band in band \( m \), which is part of our cross-layer optimization variables. Then we have

\[
\sum_{k=1}^{K^{(m)}} u^{(m,k)} = 1. 
\]

Note that some \( u^{(m,k)} \)'s can be 0 in the final optimization solution, in which case we will have fewer number of sub-bands than \( K^{(m)} \). As an example, Fig. 1 shows \( M \) bands in the network and for a specific band \( m \), it displays a further division into \( K^{(m)} \) sub-bands. Then the \( M \) bands in the network are effectively divided into \( \sum_{m=1}^{M} K^{(m)} \) sub-bands, each of which may be of different size.

**Transmission Range and Interference Range.** We assume that the power spectral density from the transmitter of a CR node is \( Q \). In this paper, we assume that all nodes use the same power density for transmission. The more complex issue of power control will be deferred for future research. A widely-used model for power propagation gain is [11]

\[
g_{ij} = \beta \cdot d_{ij}^{-n}, \tag{1}
\]

where \( \beta \) is an antenna related constant, \( n \) is the path loss index, and \( d_{ij} \) is the distance between nodes \( i \) and \( j \).\(^1\) We assume that a data transmission is successful only if the received

---

\(^1\)In this paper, we consider a uniform gain model and assume the same gain model on all frequency bands. The case of a non-uniform gain model or a band-dependent gain behavior can be extended without much technical difficulty.
power spectral density at the receiver exceeds a threshold \( Q_T \). Likewise, we assume interference will become non-negligible only if it produces a power spectral density over a threshold of \( Q_I \) at a receiver. Based on the threshold \( Q_T \), the transmission range for a node is thus \( R_T = (\beta Q_I/Q_T)^{1/n} \), which comes from \( \beta (R_T)^{-n} Q = Q_T \). Similarly, based on the interference threshold \( Q_I (\ll Q_T) \), the interference range for a node is \( R_I = (\beta Q_I/Q_I)^{1/n} \). Since \( Q_I < Q_T \), we have \( R_I > R_T \). Both, the transmission range \( R_T \) and the interference range \( R_I \), will be used in the modeling of the interference constraints as follows.

**Scheduling and Interference Constraints.** Scheduling can be done either in time domain or frequency domain. In this paper, we consider frequency domain sub-band assignment, i.e., how to assign sub-bands at a node for transmission and reception. A feasible scheduling on frequency bands must ensure that there is no interference at the same node and among the nodes.

Suppose that band \( m \) is available at both node \( i \) and node \( j \), i.e., \( m \in M_i \cap M_j \). To simplify the notation, let \( M_{ij} = M_i \cap M_j \). Denote

\[
x_{ij}(m,k) = \begin{cases} 1 & \text{if node } i \text{ transmits data to node } j \text{ on sub-band } (m,k), \\ 0 & \text{otherwise.}
\end{cases}
\]

For a node \( i \in N \) and a band \( m \in M_i \), denote \( T_i^m \) the set of nodes that can use band \( m \) and are within the transmission range to node \( i \), i.e.,

\[
T_i^m = \{ j : d_{ij} \leq R_T, j \neq i, m \in M_j \}.
\]

Note that node \( i \) cannot transmit to multiple nodes on the same frequency sub-band. We therefore have

\[
\sum_{q \in T_i^m} x_{ij}(m,k) \leq 1. \tag{2}
\]

Also, node \( i \) cannot use the same frequency sub-band for transmission and reception, due to “self-interference” at the physical layer. That is, if \( x_{ij}(m,k) = 1 \), then for any \( q \in T_j^m \), \( x_{jq}(m,k) \) must be 0. In other words, we have

\[
x_{ij}(m,k) + \sum_{q \in T_j^m} x_{jq}(m,k) \leq 1. \tag{3}
\]

Note that in (3), we are referring to a specific node \( j \) to which node \( i \) is transmitting. If \( x_{ij}(m,k) = 1 \), then \( \sum_{q \in T_j^m} x_{jq}(m,k) = 0 \), i.e., node \( j \) cannot use the same frequency sub-band \( (m,k) \) for transmission. On the other hand, if \( x_{ij}(m,k) = 0 \), then \( \sum_{q \in T_j^m} x_{jq}(m,k) \leq 1 \), i.e., node \( j \) may use frequency sub-band \( (m,k) \) for transmission, but can only use it for one receiving node \( q \in T_j^m \) (same as in (2)).

In addition to the above constraints at the same node, there are also scheduling constraints due to potential interference among the nodes in the network. In particular, for a frequency sub-band \( (m,k) \), if node \( i \) uses this sub-band for transmitting data to a node \( j \in T_i^m \), then any other node that can produce interference on node \( j \) should not use this sub-band.\(^2\)

\(^2\)Note that the so-called “hidden terminal” problem is a special case under this constraint.

\[\text{this constraint, we denote } P^m_j \text{ the set of nodes that can produce interference at node } j \text{ on band } m, \text{i.e.,}\]

\[P^m_j = \{ p : d_{pj} \leq R_T, p \neq j, T_p^m \neq 0 \}.\]

The physical meaning of \( T_p^m \neq 0 \) in the above definition is that node \( p \) may use band \( m \) for a valid transmission to a node in \( T_p^m \) and then may interfere node \( j \). Then we have

\[x_{ij}(m,k) + \sum_{q \in T_p^m} x_{jq}(m,k) \leq 1 \quad (p \in P^m_j, p \neq i). \tag{4}\]

In (4), if \( x_{ij}(m,k) = 1 \), i.e., node \( i \) uses frequency sub-band \((m,k)\) to transmit to node \( j \), then any node \( p \) that can produce interference on node \( j \) should not transmit on this sub-band, i.e., \( \sum_{q \in T_p^m} x_{pq}(m,k) = 0 \). On the other hand, if \( x_{ij}(m,k) = 0 \), (4) degenerates into (2), i.e., node \( p \) may transmit on sub-band \((m,k)\) to one node \( q \in T_p^m \), i.e., \( \sum_{q \in T_p^m} x_{pq}(m,k) \leq 1 \).

It is important to understand that in the interference constraint (4), if \( x_{ij}(m,k) = 0 \), two nodes that can produce interference at node \( j \) but are far apart and outside each other’s interference range can use the same sub-band \((m,k)\) for transmission. We use an example to illustrate this point. In Fig. 2, suppose node 1 is transmitting to node 2 on sub-band \((m,k)\), then any node that can produce interference at node 2 (i.e., node 3 or 5) cannot use the same sub-band for transmission. On the other hand, if node 1 is not using sub-band \((m,k)\) to transmit to node 2, then node 3 may use this sub-band to transmit (to node 4) as stated in (4). Likewise, node 5 may also use this sub-band to transmit (to node 6) as stated in (4). That is, both nodes 3 and 5 may use the same sub-band for transmission.

We now use a compact form to include both (3) and (4).

Denote

\[T_j^m = \begin{cases} P^m_j \cup \{j\} & \text{if } T_j^m \neq 0, \\ P^m_j & \text{otherwise.} \end{cases}\]

\[\begin{array}{c}
\text{Fig. 2. An example illustrating interference among links.}
\end{array}\]
Thus, both (3) and (4) can be described by the following constraint.

\[ \sum_{q \in T_{ij}^m} x_{pq}^{(m,k)} \leq 1 \quad (p \in T_j^m, p \neq i) \]

**Routing.** At the network level, a source node may need a number of relay nodes to route the data stream toward its destination node. Clearly, a route having only a single path may be overly restrictive and is not able to take advantage of load balancing. A set of paths (or multi-path) is more flexible to route the traffic from a source node to its destination. Mathematically, this can be modeled as follows.

Denote \( f_{ij}(l) \) the data rate on link \((i, j)\), which is attributed to session \( l \), where \( i \in \mathcal{N}, j \in \bigcup_{m \in M_l} T_i^m \), and \( l \in \mathcal{L} \). To simplify the notation, let \( T_i = \bigcup_{m \in M_i} T_i^m \). If node \( i \) is the source node of session \( l \), i.e., \( i = s(l) \), then

\[ \sum_{j \in T_i} f_{ij}(l) = r(l) . \]

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

\[ \sum_{j \in T_i, j \neq s(l)} f_{ij}(l) = \sum_{p \in T_i, p \neq d(l)} f_{pi}(l) . \]

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

\[ \sum_{p \in T_i} f_{pi}(l) = r(l) . \]

It can be easily verified that if (5) and (6) are satisfied, then (7) must be satisfied. As a result, it is sufficient to list only (5) and (6) in the formulation.

In addition to the above flow balance equations at each node \( i \) for each session \( l \), the aggregate flow rates on each radio link cannot exceed this link’s capacity. To model this mathematically, we need to first find the capacity on link \((i, j)\) in sub-band \((m, k)\). If node \( i \) sends data to node \( j \) on sub-band \((m, k)\), i.e., \( x_{ij}^{(m,k)} = 1 \), then the capacity on link \((i, j)\) in sub-band \((m, k)\) is

\[ c_{ij}^{(m,k)} = u^{(m,k)} W^{(m)} \log_2 \left( 1 + \frac{g_{ij} Q}{\eta} \right) , \]

where \( \eta \) is the ambient Gaussian noise density. Note that the denominator inside the log function contains only \( \eta \). This is due to one of our interference constraints stated earlier, i.e., when node \( i \) is transmitting to node \( j \) on sub-band \((m, k)\), then all the other neighbors of node \( j \) within its interference range are prohibited from using this sub-band. This interference constraint significantly helps to simplify the calculation of the link capacity \( c_{ij}^{(m,k)} \). When \( x_{ij}^{(m,k)} = 0 \), we have \( c_{ij}^{(m,k)} = 0 \). Thus, \( c_{ij}^{(m,k)} \) can be written in the following compact form.

\[ c_{ij}^{(m,k)} = x_{ij}^{(m,k)} \cdot u^{(m,k)} W^{(m)} \log_2 \left( 1 + \frac{g_{ij} Q}{\eta} \right) . \]

Now, returning to our earlier requirement that the aggregate data rates on each link \((i, j)\) cannot exceed the link’s capacity, we have,

\[ \sum_{l \in \mathcal{L}, s(l) \neq j, d(l) \neq i} f_{ij}(l) \leq \sum_{m \in M_i} \sum_{k=1}^{K^{(m)}} c_{ij}^{(m,k)} \]

\[ = \sum_{m \in M_i} \sum_{k=1}^{K^{(m)}} x_{ij}^{(m,k)} \cdot u^{(m,k)} W^{(m)} \log_2 \left( 1 + \frac{g_{ij} Q}{\eta} \right) . \]

**B. Problem Formulation**

For the multi-hop CR networks that we are investigating, various performance objectives can be used. In this paper, we use the total required radio resource to support the user sessions as our performance objective. The radio resource can be measured in terms of the total bandwidth used by all nodes in the network, which is the simplified form of the so-called space-bandwidth product proposed in [17] with fixed transmission power spectral density. It is not hard to see that the solution procedure in this paper can be applied when other performance objectives are used.

To recap, we are given a set of source-destination pairs (user sessions) in the network, each with a certain rate requirement. Each node in the network has a set of available frequency bands that it can use for communication. We want to find an optimal solution to divide the set of available frequency bands at each node, the scheduling of sub-bands for transmission and reception, and multi-hop routing for each flow such that the total radio bandwidth used in the network is minimized (or the solution declares that there is no feasible solution). Mathematically, we have the following optimization problem.

Min \[ \sum_{i \in \mathcal{N}} \sum_{m \in M_i} \sum_{j \in T_i^m} \sum_{k=1}^{K^{(m)}} W^{(m)} x_{ij}^{(m,k)} u^{(m,k)} \]

s.t

1. \[ \sum_{k=1}^{K^{(m)}} x_{ij}^{(m,k)} = u^{(m,k)} \quad (m \in M) \]
2. \[ \sum_{q \in T_i^m} x_{pq}^{(m,k)} \leq 1 \quad (i \in \mathcal{N}, m \in M_i, 1 \leq k \leq K^{(m)}) \]
3. \[ x_{ij}^{(m,k)} + \sum_{q \in T_i^m} x_{pq}^{(m,k)} \leq 1 \quad (i \in \mathcal{N}, m \in M_i, j \in T_j^m, 1 \leq k \leq K^{(m)}, p \in T_j^m, p \neq i) \]
4. \[ \sum_{l \in \mathcal{L}, s(l) \neq j, d(l) \neq i} f_{ij}(l) - \sum_{m \in M_i} \sum_{k=1}^{K^{(m)}} W^{(m)} \log_2 \left( 1 + \frac{g_{ij} Q}{\eta} \right) x_{ij}^{(m,k)} u^{(m,k)} \leq 0 \quad (i \in \mathcal{N}, j \in T_i) \]
5. \[ \sum_{j \in T_i, j \neq s(l)} f_{ij}(l) = r(l) \quad (l \in \mathcal{L}, i = s(l)) \]
6. \[ \sum_{j \in T_i, j \neq d(l)} f_{ij}(l) = 0 \quad (l \in \mathcal{L}, i \neq d(l)) \]
7. \[ x_{ij}^{(m,k)} = 0 \text{ or } 1, u^{(m,k)} \geq 0 \quad (i \in \mathcal{N}, m \in M_i, j \in T_j^m, 1 \leq k \leq K^{(m)}) \]
8. \[ f_{ij}(l) \geq 0 \quad (l \in \mathcal{L}, i \neq d(l), j \in T_i, j \neq s(l)) \]

where \( W^{(m)}, g_{ij}, Q, \eta, \) and \( r(l) \) are all constants, and \( x_{ij}^{(m,k)}, u^{(m,k)}, \) and \( f_{ij}(l) \) are all optimization variables.

The above optimization problem is a mixed-integer non-linear programming (MINLP) problem, which is NP-hard in
general [8]. Although existing software (e.g., BARON [2]) can solve very small-sized network instances (e.g., several nodes), the time complexity becomes prohibitively high for large-sized networks.

Our approach to solve this problem is as follows. In Section IV, we first explore a lower bound for the objective, which can be obtained by relaxing the integer variables and using a linearization technique. Using this lower bound as a performance benchmark, in Section V, we develop a highly effective algorithm based on a novel sequential fixing (SF) procedure. Using simulation results, we show that the SF algorithm has a performance very close to the lower bound. Since the optimal objective value lies between the lower bound and the solution obtained by the SF algorithm, the solution by the SF algorithm must be even closer to the true optimum.

IV. A LOWER BOUND FOR THE OBJECTIVE FUNCTION

The complexity of the problem formulated in Section III-B arises from the binary $x_{ij}^{(m,k)}$ variables and the product of variables $x_{ij}^{(m,k)}u_{ij}^{(m,k)}$. To pursue a lower bound for the objective, we first multiplying (9) and (10) by the corresponding $u_{ij}^{(m,k)}$, so that $x_{ij}^{(m,k)}$ appears throughout as a product with $u_{ij}^{(m,k)}$.

We then relax the integrality (binary) requirement on $x_{ij}^{(m,k)}$ with $0 \leq x_{ij}^{(m,k)} \leq 1$ and replace $x_{ij}^{(m,k)}u_{ij}^{(m,k)}$ with a single variable, say $s_{ij}^{(m,k)}$, i.e., $s_{ij}^{(m,k)} = x_{ij}^{(m,k)}u_{ij}^{(m,k)} \leq u_{ij}^{(m,k)}$. Such a relaxation leads to the following lower-bounding problem formulation.

$$\min \sum_{i \in \mathcal{N}} \sum_{m \in \mathcal{M}} \sum_{j \in \mathcal{T}^m} \sum_{k=1}^{K(m)} W^{(m)} s_{ij}^{(m,k)}$$

subject to

$$\sum_{k=1}^{K(m)} s_{ij}^{(m,k)} = 1 \quad (m \in \mathcal{M})$$

$$\sum_{q \in \mathcal{T}^m} s_{iq}^{(m,k)} - u_{ij}^{(m,k)} \leq 0 \quad (i \in \mathcal{N}, m \in \mathcal{M}, 1 \leq k \leq K(m))$$

$$s_{ij}^{(m,k)} + \sum_{q \in \mathcal{T}^m} s_{pq}^{(m,k)} - u_{ij}^{(m,k)} \leq 0 \quad (i \in \mathcal{N}, m \in \mathcal{M}, 1 \leq k \leq K(m), p \in \mathcal{T}^m, p \neq i)$$

$$\sum_{j \in \mathcal{L}} f_{ij}(l) - \sum_{m \in \mathcal{M}} \sum_{k=1}^{K(m)} W^{(m)} \log_2 \left(1 + \frac{g_{ij}Q}{\eta}ight) s_{ij}^{(m,k)} \leq 0 \quad (i \in \mathcal{N}, j \in \mathcal{T}_i)$$

$$\sum_{j \neq i, j \in \mathcal{T}_i} f_{ij}(l) - \sum_{p \neq i, p \in \mathcal{T}_i} f_{pi}(l) = 0 \quad (l \in \mathcal{L}, i \in \mathcal{N}, i \neq s(l), d(l))$$

$$u_{ij}^{(m,k)}, s_{ij}^{(m,k)} \geq 0 \quad (i \in \mathcal{N}, m \in \mathcal{M}, 1 \leq k \leq K(m))$$

$$f_{ij}(l) \geq 0 \quad (l \in \mathcal{L}, i \in \mathcal{N}, i \neq s(l), d(l), j \in \mathcal{T}_j, j \neq s(l))$$

This new (relaxed) formulation is a standard linear program (LP), the solution of which can be obtained in polynomial time. Due to the relaxation (and thus enlarged optimization space), the solution value to this LP problem yields a lower bound for the objective of the original problem in Section III-B. Note that there may not exist a feasible solution that achieves this lower bound.

Nevertheless, this lower bound offers a benchmark to measure the quality of a feasible solution, which we will develop in the next section. It turns out that this lower bound is extremely tight (see results in Section V). This can be explained by the convex hull results presented by Sherali et al. [27].

V. A NEAR-OPTIMAL ALGORITHM BASED ON SEQUENTIAL FIXING

A. Basic Algorithm

We now take a closer look at the original MINLP problem formulation in Section III-B. Observe that once the binary values for all $x$ variables are determined, i.e., whether or not a node will indeed use a particular sub-band to send data to another node, then this MINLP reduces to an LP, which can be solved in polynomial time. In other words, the key obstacle in solving this MINLP problem lies in the determination of the binary values for the $x_{ij}^{(m,k)}$ variables. To this end, we propose a two-step solution procedure: i) fix the binary values for $x_{ij}^{(m,k)}$ iteratively through a sequence of LPs; ii) once all the $x_{ij}^{(m,k)}$ variables are fixed, find a solution (to determine how to divide sub-bands and flow routing) corresponding to this set of $x_{ij}^{(m,k)}$ values. Such a two-step approach will yield a sub-optimal (upper bound) solution to the original MINLP problem. The quality of this algorithm can be assessed by comparing its solution to the lower bound that we developed in the previous section.

As said, the key to the two-step approach resides in the determination of the binary values for all the $x_{ij}^{(m,k)}$ variables. Our main idea is to fix (set) the values of the $x_{ij}^{(m,k)}$ variables sequentially through solving a series of relaxed LP problems, with each iteration setting at least one binary value for some $x_{ij}^{(m,k)}$. Specifically, during the first iteration, we relax all binary variables $x_{ij}^{(m,k)}$ to 0 ≤ $x_{ij}^{(m,k)}$ ≤ 1 as in Section IV to obtain an LP. Upon solving this LP, we have a solution with each $x_{ij}^{(m,k)} = s_{ij}^{(m,k)}/u_{ij}^{(m,k)}$ being a value between 0 and 1. Among all the $x$-values, suppose some $x_{ij}^{(m,k)}$ has the largest value. Then we fix (set) this particular $x_{ij}^{(m,k)}$ to 1. As a result of this fixing, by (9), we also need to fix $x_{ij}^{(m,k)} = 0$ for $q \neq j$. Further, by (10), we can fix $x_{pq}^{(m,k)}$ to 0 for $p \neq i$ and $q \neq j$. Technically, in an implementation, we can fix all the $x$-variables that have a value of 1 and perform an additional fixing for the largest fractional variable as above.

Now, having fixed some $x$-variables in the first iteration, we update the problem to obtain a new LP for the second iteration as follows. For those $x_{ij}^{(m,k)}$-variables that are already fixed at 1, since $s_{ij}^{(m,k)} = x_{ij}^{(m,k)}u_{ij}^{(m,k)} = u_{ij}^{(m,k)}$, we can replace the corresponding $s_{ij}^{(m,k)}$ by $u_{ij}^{(m,k)}$. For those $x_{ij}^{(m,k)}$ and $x_{pq}^{(m,k)}$ that are fixed to 0, we can set $s_{iq}^{(m,k)} = 0$ and $s_{pq}^{(m,k)} = 0$. As a result, all the terms in the LP involving these $s$-variables can be removed and the corresponding constraint in (11) and (12) can also be removed.
In the second iteration, we solve this new LP and then fix some additional \( x \)-variables based on the same process (now the ordering of the \( x \)-values is done only for the remaining un-fixed \( x \)-variables). The iteration continues and eventually we fix all \( x \)-variables to either 0 and 1.

Upon fixing all the \( x \)-values, the original MINLP reduces to an LP problem, which can be solved in polynomial time. Unlike the solutions obtained in Section IV, the final solution obtained here is a feasible solution since all \( x \)-values are binary. The complete Sequential Fixing (SF) algorithm is given in Fig. 3.

**B. An Iteration-Speedup Technique**

In the SF algorithm, we need to solve a sequence of LPs. The complexity of SF is polynomial. By exploiting the space and frequency dimensions involved in radio resource allocation, we may decrease the number of LPs by fixing more \( x \)-variables during each iteration in Fig. 3. As a result, the complexity can be further decreased. From a space dimension viewpoint, a sub-band usage will only have an impact within the interference range and the same sub-band can be used by other links outside this range. Thus, for the same sub-band \((m, k)\), we may fix multiple links that have non-overlapping interference ranges within a single iteration of the SF algorithm. From the frequency dimension viewpoint, the transmission in one sub-band will not interfere with the transmission in a different sub-band. Thus, for the same link \((i, j)\), we may fix multiple sub-bands within a single iteration of the SF algorithm. Specifically, we can use a threshold \( \alpha > 0.5 \) in this fixing process and fix all the \( x \)-variables that exceed \( \alpha \) to 1 in a single iteration. Note that in (9) and (10), it is required that at most one binary variable \( x_{ij}^{(m,k)} = 1 \) while in the relaxed problem, there is at most one fraction \( s_{ij}^{(m,k)}/u_{ij}^{(m,k)} > 0.5 \). Thus, \( \alpha > 0.5 \) ensures that both the constraints (9) and (10) (interference constraints at each node and among the nodes) will hold during the SF procedure.\(^3\) In the case that none of the \( x \)-variables exceed \( \alpha \), we will fall back to the basic algorithm in Fig. 3 and simply choose the largest valued \( x \)-variable.

**VI. SIMULATION RESULTS**

In this section, we present simulation results for our SF algorithm and compare it to the lower bound obtained in Section IV. We consider \( |\mathcal{N}| = 20, 30 \) or 40 nodes in a 500 x 500 area (in meters). Among these nodes, there are \( |\mathcal{C}| = 5 \) active sessions, each with a random rate within [10, 100] Mb/s.

We assume that there are \( M = 5 \) bands that can be used for the entire network (see Table II). Bands I and II are among the least-utilized (less than 2%) spectrum bands found in [19] and bands III, IV, and V are unlicensed ISM bands used for 802.11. Recall that available bands at each CR node is a subset of these five bands based on its location and the available bands at any two nodes in the network may not be identical.

In the simulation, this is done by randomly selecting a subset of bands from the pool of five bands for each node. Further, we assume bands I to V can be divided into 3, 5, 2, 4, and 4 sub-bands although other desirable divisions can be used. Note that the size of each sub-band may be unequal and is part of the optimization problem.

We assume that the transmission range at each node is 100 m and that the interference range is 150 m, although other settings can be used. The path loss index \( n \) is assumed to be 4 and \( \beta = 62.5 \). The threshold \( Q_T \) is assumed to be 10\( \eta \). Thus, we have \( Q_I = (\frac{100}{100})^n Q_T \) and the transmission power spectral density \( Q = (100)^n Q_T / \beta = 1.6 \cdot 10^7 \eta \).

Note that it is possible that there is no feasible solution for a specific data set. This could be attributed to dis-connectivity in the network (due to random network topology), resource bottleneck in a hot area, etc. Thus, we only report results based on those data sets that have feasible solutions.

We first present simulation results for 100 data sets for 20-node networks that can produce feasible solutions. For each data set, the network topology, source/destination pair and bit rate of each session, and available frequency bands at each node are randomly generated. We use the SF algorithm to determine the cost, which is the total required bandwidth in the objective function. As discussed, we compare this result with the lower bound developed in Section IV. The running time for each simulation is less than 10 seconds on a Pentium 3.4 GHz machine.

Figure 4 shows the normalized costs obtained by the SF algorithm with respect to the lower bound costs for 100 data sets. The average normalized cost among the 100 simulations is 1.04 and the standard derivation is 0.07. There are two observations that can be made from this figure. First, since the ratio of the solution obtained by SF (upper bound of optimal solution) to the lower bound solution is close to 1 (in many cases, they coincide with each other), the lower bound must be very tight. Second, since the optimal solution (unknown) is between the solution obtained by the SF algorithm and the lower bound, the SF solution must be even closer to the optimum.
To get a sense of how the actual (rather than normalized) numerical results appear in the simulations, we list the first 40 sets of results in Table III. Note that in many cases, the result obtained by the SF algorithm is identical to the respective lower bound obtained via relaxation. This indicates that the solution found by SF is optimal.

Simulation results for 100 random data sets for 30-node and 40-node networks that produce feasible solutions are displayed in Figs. 5 and 6, respectively. For 30-node networks, the average normalized cost among the 100 simulations is 1.10 and the standard derivation is 0.16. For 40-node networks, the average normalized cost among the 100 simulations is 1.18 and the standard derivation is 0.16. Thus, the SF solutions are also close to the optimal solutions.

VII. CONCLUSIONS

In this paper, we conducted a systematic study on the important problem of multi-hop networking with CR nodes. The nature of the problem calls for a characterization and modeling of multi-layer behaviors and constraints. We characterized behaviors and constraints for a multi-hop CR network from multiple layers, including the modeling of spectrum sharing and sub-band division, scheduling and interference constraints, and flow routing. We formulated an optimization problem with the objective of minimizing the required network-wide radio spectrum resource for a set of user sessions. Since the problem formulation is an MINLP, we developed a lower bound to estimate the objective function. Subsequently, we developed a novel sequential fixing algorithm to the cross-layer optimization problem. Simulation results showed that results obtained by this algorithm are very close to the lower bound, thus confirming that they are near-optimal.
ACKNOWLEDGEMENTS

The authors thank the anonymous reviewers for their constructive comments. The work of Y.T. Hou and Y. Shi has been supported in part by NSF Grant CNS-0721570. The work of H.D. Sherali has been supported in part by NSF under Grant CMMI-0552676.

REFERENCES

Y. Thomas Hou (S’91–M’98–SM’04) received the B.E. degree from the City College of New York in 1991, the M.S. degree from Columbia University in 1993, and the Ph.D. degree from Polytechnic University, Brooklyn, New York, in 1998, all in Electrical Engineering. Since August 2002, he has been with Virginia Tech, the Bradley Department of Electrical and Computer Engineering, Blacksburg, VA, where he is now an Associate Professor. His current research interests are radio resource (spectrum) management and networking for cognitive radio wireless networks, optimization and algorithm design for wireless ad hoc and sensor networks, and video communications over dynamic ad hoc networks. From 1997 to 2002, Dr. Hou was a Researcher at Fujitsu Laboratories of America, Sunnyvale, CA, where he worked on scalable architectures, protocols, and implementations for differentiated services Internet; service overlay networking; video streaming; network bandwidth allocation policies and distributed flow control algorithms.

Prof. Hou 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 algorithm design and optimizations for wireless ad hoc and sensor networks. He has published extensively in leading journals and top tier conferences and received several best paper awards from IEEE. He holds two U.S. patents and has three more pending.

Prof. Hou is active in professional services and is currently serving as an Editor of IEEE Transactions on Wireless Communications, ACM/Springer Wireless Networks (WINET), and Elsevier Ad Hoc Networks Journal. He was Co-Chair of Technical Program Committee (TPC) of the Second International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM 2007), Orlando, FL, August 1-3, 2007. He was the Founding Chair of the First IEEE Workshop on Networking Technologies for Software Defined Radio (SDR) Networks (held in conjunction with IEEE SECON 2006), Reston, VA, Sept. 25, 2006.

Prof. Hou was Co-Chair (with Tony Ephremides) of NSF Workshop on Bridging the Gap between Wireless Networking Technologies and Advances at the Physical Layer, August 27-28, 2007, Reston, VA. He will serve as Co-Chair of Technical Program Committee (TPC) of IEEE INFOCOM 2009, to be held in Rio de Janeiro, Brazil.

Yi Shi (S’02) received his B.S. degree from University of Science and Technology of China, Hefei, China, in 1998, a M.S. degree from Institute of Software, Chinese Academy of Science, Beijing, China, in 2001, a second M.S. degree from Virginia Tech, Blacksburg, VA, in 2003, all in computer science, and a Ph.D. degree in computer engineering from Virginia Tech, in 2007. He is currently a Senior Research Associate in the Department of Electrical and Computer Engineering at Virginia Tech.

His current research focuses on algorithms and optimization for cognitive radio wireless networks, MIMO and cooperative communication networks, sensor networks, and ad hoc networks. His work has appeared in some highly selective international conferences (ACM MobiCom, ACM MobiHoc, and IEEE INFOCOM) and IEEE journals.

While an undergraduate, he was a recipient of Meritorious Award in International Mathematical Contest in Modeling in 1997 and 1998, respectively. He was a recipient of Chinese Government Award for Outstanding Students Abroad in 2006. He was a TPC member of IEEE Workshop on Networking Technologies for Software Defined Radio (SDR) Networks (held in conjunction with IEEE SECON 2006), Reston, VA, Sept. 25, 2006.

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 238 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 National Academy of Engineering.