Spectrum Assignment and Sharing for Delay Minimization in Multi-Hop Multi-Flow CRNs

Wei Li*, Xiuzhen Cheng*, Tao Jing†, Yong Cui‡, Kai Xing§, Wendong Wang¶

*Department of Computer Science, The George Washington University, Washington DC, USA
†School of Electronics and Information Engineering, Beijing Jiaotong University, Beijing, P. R. China
‡Department of Computer Science, Tsinghua University, Beijing, China
§Department of Computer Science and Technology, University of Science and Technology of China, Hefei, China
¶State Key Lab of Network and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China

Email: {weili, cheng}@gwu.edu, tjing@bjtu.edu.cn, cuiyong@tsinghua.edu.cn, kxing@ustc.edu.cn, wdwang@bupt.edu.cn

Abstract

This paper investigates the problem of spectrum assignment and sharing to minimize the total delay of multiple concurrent flows in multi-hop cognitive radio networks. We first analyze the expected per-hop delay, which incorporates the sensing delay and transmission delay characterizing the PU activities and spectrum capacities. Then we formulate a minimum delay optimization problem with interference constraints, and propose an approximation algorithm termed MCC to solve the problem. According to our theoretical analysis, MCC has a bounded performance ratio and a low computational complexity. Finally, we exploit the minimum potential delay fairness in spectrum sharing to mitigate the inter-flow contentions. Extensive simulation study has been performed to validate our design and to compare the performance of our algorithms with that of the state-of-the-art.

Index Terms

Cognitive radio networks; delay minimization; spectrum assignment; spectrum sharing

I. INTRODUCTION

With the dramatic increase of the spectrum demand, FCC’s traditional static spectrum assignment policy leads to the problem of spectrum scarcity. Nevertheless, a large portion of the licensed spectrum is still not in use or is severely underutilized, resulting in an inefficient spectrum utilization [1]–[4]. To resolve this conflict, Cognitive Radio Networking (CRN) is proposed as an effective technology to enhance the spectrum utilization. In a cognitive radio network, dynamic spectrum access plays the most crucial role for effectively sharing the limited spectrum between the primary and secondary network, which has been extensively studied from different aspects [3], [5]–[8]. In this paper, we focus on both spectrum assignment and spectrum sharing for delay minimization in multi-hop multi-flow CRNs. This is a challenging problem due to the following reasons. First, delay analysis should characterize the specific properties of CRNs such as the PU activities and spectrum capacities. Although previous work [9]–[13] has performed extensive study on the delay performance, none of them considers the impact of the imperfection of spectrum sensing on the performance of SUs. Moreover, the spectrum sensing time is deemed as a type of delay in CRNs [14]–[17] but unfortunately it is largely ignored by the existing research, as elaborated in our Section II Related Work. Second, the inter-flow contentions need to be handled to coordinate spectrum access, i.e., spectrum sharing. In other words, we should study spectrum bandwidth distribution to improve the end-to-end performance of multiple concurrent flows sharing the same link, which has not been addressed to our best knowledge. Third, both the intra-link and the inter-link interference should be taken into account, which makes spectrum assignment in multi-hop network scenarios more difficult than that in the single-hop scenario. Finally, spectrum assignment is NP-hard and thus approximation algorithms should be sought for efficiency purpose.

In spectrum assignment, our objective is to minimize the total expected delay. We first establish a delay model to characterize the PU activities and the limited spectrum capacities. The delay of each single-hop link consists of the sensing delay and the transmission delay. Based on our delay model, we formulate the problem of spectrum
assignment for delay minimization in multi-hop multi-flow CRNs as a 0-1 integer programming and solve it by exploiting an approximation algorithm termed \textbf{Minimum Cost Coloring (MCC)}. Our theoretical analysis reveals that MCC has a bounded performance ratio. Additionally, we propose an approach to fairly distribute spectrum bandwidths among different flows to mitigate the inter-flow contentions via spectrum sharing. In particular, we exploit the \textit{minimum potential delay fairness} [18], which captures the long-term throughput that a user should expect to receive from a fully saturated network. The optimal solution to our bandwidth allocation problem with minimum potential delay fairness can be efficiently obtained by Lagrange Multipliers.

The multi-fold contributions of this paper are summarized as follows.

- We propose a novel delay model that considers the impact of spectrum sensing, which can accurately evaluate the transmission delay in cognitive radio networks. This delay model takes into account both the sensing delay and the transmission delay, characterizing the PU behaviors and the limited spectrum capacities.
- We formulate a delay optimization problem constrained on the inter-flow and intra-flow interference for multi-hop multi-flow CRNs based on the delay model.
- We propose a spectrum assignment algorithm MCC for the delay minimization problem based on graph theory to obtain an approximate solution with a bounded performance ratio and a low computational complexity.
- To mitigate the inter-flow contentions, we establish an optimization problem of spectrum sharing to achieve the \textit{minimum potential delay fairness}. For this problem, Lagrange Multipliers are employed to find the optimal solution.
- We perform an extensive comparison-based simulation study to validate our delay model and verify the performance of our algorithms in terms of the average throughput, the average end-to-end delay, the average packet-drop ratio, and the Jain’s fairness index of the network. The simulation results confirm the effectiveness of our proposed algorithms.

The rest of the paper is organized as follows. The related work is summarized in Section II. The network model and problem formulation are presented in Section III. Section IV details our approximation algorithm MCC and provides a thorough theoretical performance analysis. Spectrum sharing to mitigate the inter-flow spectrum contentions is investigated in Section V. After reporting our performance evaluation results in Section VI, we conclude this paper in Section VII.

II. RELATED WORK

There exist quite a few recent works focusing on the delay analysis in CRNs [9]–[13]. In [9], the authors presented a hybrid protocol model for SUs to exploit the spatial gap among PUs for frequency reuse. This hybrid model suffices to guide the secondary network to obtain the same delay scaling as a standalone network, without harming the transmissions of the primary system. In [10], Liang \textit{et. al} derived the average packet transmission delay for bursty traffic and Poisson traffic by taking into consideration both the periodical switching and the triggered switching for real-time traffic. In [11], Wang \textit{et. al} analyzed SUs’ steady-state queueing delay performance by taking a fluid queue approximation approach. With the consideration of effective channel bandwidth, Chen \textit{et. al} [12] pointed out that the delay and throughput optimality can be simultaneously achieved under the framework of risk sensitive constrained Markov decision process as a measure of queueing delay. In order to examine whether there is indeed an advantage in using dynamic multi-channel MAC in CRNs, Liu \textit{et. al} [13] analyzed the delay performance of an opportunistic multi-channel medium access control (MAC) scheme and compared it to that of the corresponding single channel MAC. Although these delay analysis works do capture one or more unique features of CRNs, none of them considers the impact of spectrum sensing. Nevertheless, the spectrum sensing time can degrade SUs’ transmission performance. Moreover, it is a type of delay resulted from the specific characteristics of opportunistic spectrum access in CRNs [14]–[17].

On the other hand, the increasing number of applications motivates the research on multi-hop CRNs [19]–[23]. Xue \textit{et. al} [19] introduced their optimal control and scheduling algorithms to maximize the throughput of the SUs to meet the collision probability constraints required by the PUs in stable CRNs. In [20], the uneven size of the spectrum bandwidth prompted the need of dividing spectrums into sub-bands for optimal spectrum sharing, and a mixed integer non-linear programming was established to minimize the total required network-wide spectrum resource for a set of user sessions. In [21], Pan \textit{et. al} extended the mathematical formulation in [20] to consider the joint routing and link scheduling problem under the uncertain spectrum supplies. Shiang \textit{et. al} [22] proposed a
distributed resource management algorithm that allows the network nodes to exchange information and to choose a proper available spectrum for delay sensitive transmissions. They also considered the trade-off between the learning efficiency of the spectrum selection and the cost of the required information exchange. Besides dynamic spectrum access, the interference temperature constraints between PUs and SUs were taken into account by [23]. Correspondingly, a probabilistic frequency selection scheme was used to minimize the scheduling delay.

To our best knowledge, [24] is the only work that employs a per-hop transmission delay composed of the propagation delay and the waiting time of an available spectrum band/channel. This work established the scaling law of the multi-hop delay in ad hoc CRNs and analyzed the relationship between the multi-hop delay and the source-destination distance without clearly estimating the multi-hop delay. As a comparison, our delay model takes into account both the sensing delay and the transmission delay that is affected by the source-destination distance and the spectrum bandwidth, to evaluate the expected delay of multi-hop communications. On the other hand, most existing works [19]–[23] assume that spectrum sensing is perfect, while in this paper we consider a more practical scenario with imperfect spectrum sensing. Furthermore, our design objective for spectrum assignment is to minimize the total expected delay, which is different from those in [19]–[23]. Additionally, we study the problem of minimum potential delay fairness to mitigate the inter-flow contentions for spectrum sharing in multi-hop multi-flow CRNs, which has not yet been addressed.

III. NETWORK MODEL

A. Opportunistic Spectrum Access

We consider a multi-hop CRN with \( M \) SUs in a set \( U \) and \( N \) available licensed channels in a set \( C \). Each SU is allowed to access one available licensed channel only when the channel is not occupied by the PUs. Because the SUs reside at different physical locations, the set of available channels for each SU might be different. Let \( C_i \subseteq C \) be the set of available channels for SU \( i \), and \( C_{ij} = C_i \cap C_j \) \((i \neq j)\) be the set of overlapping available channels of SU \( i \) and SU \( j \). Let \( U_i^A \) denote the set of single-hop neighbors of SU \( i \).

In the CRN, a set of multiple multi-hop concurrent data flows, denoted by \( F \), is considered. If flow \( f \in F \) crosses a link \((i, j)\) from SU \( i \) to SU \( j \), we claim that \((i, j) \in f \). In such a case, we assume that \( C_{ij} \neq \emptyset \). Since we focus on spectrum management to minimize the total delay of the flows, all the routing paths are assumed to be known.

B. Delay Analysis

Assume that all SUs have the ability to detect the traffic statistics of the PUs on each channel and learn about PU activities, which can be modeled as an ON-OFF Poisson process [17], [25]. With such a model, PU activities on channel \( k \) have two states, i.e., ON and OFF. An ON state, which is also called a busy state, indicates that a PU is active in the channel, while an OFF state, which is also named an idle state, implies that a PU is inactive. Let \( \mu_k \) and \( \eta_k \) be the arrival rate and the departure rate of the PU traffic on channel \( k \), respectively. Then the busy probability and idle probability of channel \( k \) can be computed by \( P_{on}^k = \frac{\mu_k}{\mu_k + \eta_k} \) and \( P_{off}^k = \frac{\eta_k}{\mu_k + \eta_k} \), respectively [17], [25].

Fig. 1. The structure of a time slot.

In this paper, we consider a time-slotted network model shown in Fig. 1. A time slot \( T \) has a fixed length, which consists of a sensing period and a data transmission period, denoted by \( T_s^k \) and \( T_t^k \) for channel \( k \), respectively. Note that \( T_s^k \) and \( T_t^k \) are determined by the corresponding channel information, such as the PU activity, the channel bandwidth, and so on. Since the length of a time slot is fixed and SUs cannot transmit during the sensing period, it is necessary to enhance the sensing efficiency, which is defined as \( T_s^k = \frac{T - T_t^k}{T} \), in order to obtain a longer data transmission time [16], [17]. On the other hand, spectrum sensing is not perfect in many real-world scenarios due to the unavoidable false alarms and the detection errors. Thus we consider a cooperative scheme whose detection
error rate can converge to 0 as the number of SUs increases [16], [17]. According to [17], the sensing time can be computed by the following equation:

$$T^s_k = \frac{1}{B_k \gamma^2} [Q^{-1}(\frac{P}{\gamma}) + (\gamma + 1)Q^{-1}(\frac{P_{off} \overline{P}_f}{P_{on}})]^2,$$  

(1)

where $B_k$ is the bandwidth of channel $k$, $\gamma$ is the received signal-to-noise ratio (SNR) at the secondary receiver, $\overline{P}_f$ is the average false alarm rate, and $Q(\cdot)$ is the probability distribution function of the standard Gaussian.

On the other hand, the expected packet transmission rate $r_k(i, j)$ of link $(i, j)$ on channel $k$ is dependent on the channel capacity $c^e_k(i, j)$ given by [16], [17],

$$r_k(i, j) = \frac{T^t_k}{T} c^e_k(i, j) = \frac{T^t_k}{B_k} \log_2(1 + \frac{P_{off}}{N_0}),$$  

(2)

with $p$ denoting the transmit power, $N_0$ denoting the noise experienced by the receiver, and $g_{ij} = d^{-\alpha}_{ij}$ ($d_{ij}$ is the physical distance between $i$ and $j$, $\alpha$ is the shadowing factor ranging from 2.0 to 5.0 [26]) denoting the channel gain.

For simplicity, we assume that all the packets have the same size $L$. Thus, the corresponding expected packet transmission time $t_k(i, j)$ can be estimated by

$$E[t_k(i, j)] = t_k(i, j) = \frac{L}{r_k(i, j)}.$$  

(3)

In our model, the expected service time of a packet, denoted by $\tau_k(i, j)$, is defined to be the duration from the instant when the packet reaches the head of the queue in the transmitter to the instant when it successfully departs from the queue. This expected service time includes two parts, i.e., the sensing time and the transmission time. The average spectrum sensing time attributed to a packet can be evaluated by $\frac{T^s_k}{\lambda T}$, where $\lambda T$ is the expected number of packets arriving in the queue during a time slot. Therefore we have $\tau_k(i, j) = t_k(i, j) + \frac{T^s_k}{\lambda T}$.

Previous work [25] indicates that a packet service system can be modeled as a standard $M/G/1$ queue. Thus the average queuing delay $T^q_k(i, j)$ of the link $(i, j)$ on channel $k$ is given by Queuing Theory:

$$E[T^q_k(i, j)] = \frac{\lambda E[\tau_k(i, j)]^2}{2(1 - \lambda E[\tau_k(i, j)])} = \frac{\lambda E[(t_k(i, j) + \frac{T^s_k}{\lambda T})^2]}{2(1 - \lambda E[t_k(i, j) + \frac{T^s_k}{\lambda T}])},$$  

(4)

where $\lambda$ is the average packet arrival rate of the SUs.

Based on the above analysis, the total delay of each link consists of three parts, i.e., the sensing delay (or sensing time), the queuing delay, and the transmission time. Therefore, the total expected delay of link $(i, j)$ on channel $k$, $T^D_k(i, j)$, is obtained by combining (1), (3) and (4),

$$E[T^D_k(i, j)] = E[T^s_k] + (1 - \frac{T^s_k}{T})E[T^q_k(i, j)] + E[t_k(i, j)].$$  

(5)

The component $(1 - \frac{T^s_k}{T})$ in the second term of (5) is derived from the observation that packets might enter the queue during the sensing period. Thus the overlapping time between sensing and queuing must be removed. From (5), we notice that the expected per-hop delay is influenced by the channel capacity and the PU’s activity in CRNs.

C. Interference Model

There are two main types of interferences in CRNs, i.e., the interference between PUs and SUs, and the interference among SUs. In our model, we assume that all unused channels can be detected based on existing spectrum sensing technologies before channel selection. Thus accessing unused channels does not generate interference between PUs and SUs.

For simplicity, we assume that all SUs transmit at the same power level $p$. But the proposed model and algorithm can be easily extended to the case when transmit powers vary. Thus, all SUs have the same transmission range $R_T = \sqrt{\frac{P}{\gamma_{th}N_0}}$ and interference range $R_I = \beta R_T$, where $\gamma_{th}$ is the SNR threshold for successful communications, and $\beta \geq 1$ is a constant characterizing the relationship between $R_I$ and $R_T$. In this paper, we set $\beta = 2$ if the transmitter and receiver are on the same channel [27].

In multi-hop CRNs, we take both the intra-link interference and the inter-link interference into consideration.
1) Intra-link Interference: Let $X_{N \times (M \times M)} = \{x_k(i, j)\}$ be a binary matrix indicating the channel selection, where $x_k(i, j) = 1$ if and only if channel $k$ is assigned to link $(i, j)$. Note that each node can access an available channel, but cannot communicate with multiple nodes on the same channel at the same time. Thus we have,

$$\sum_{k \in C_i} \sum_{j \in U^A_i} x_k(i, j) = 1. \quad (6)$$

On the other hand, each node cannot use the same channel for sending and receiving data simultaneously due to “self-interference”, which can be expressed as follows:

$$x_k(i, j) + \sum_{u \in U^A_j} x_k(j, u) \leq 1. \quad (7)$$

2) Inter-link Interference: In our model, we assume that a media access control protocol with acknowledgement is adopted. Therefore, to reduce the interference among different links, two links are not allowed to access the same channel if any end node of one link resides in the interference range of any end node of the other link, i.e.,

$$x_k(i, j) + \sum_{u, v \in U^I_i \cup U^I_j} x_k(u, v) \leq 1, \quad (8)$$

where $U^I_i = \{u | u \in U, d_{ui} \leq R_I\}$. Eq. (8) indicates that if link $(i, j)$ selects channel $k$, any link $(u, v)$ that has one end node in the interference range of SU $i$ or SU $j$ is not allowed to use channel $k$.

D. Problem Formulation

To sum up, mathematically the spectrum assignment for expected delay minimization can be formulated by the following optimization problem, with $\{x_k(i, j)\}$ being the set of variables to be determined:

$$\min \sum_{f \in F} \sum_{(i, j) \in f} \sum_{j \in U^A_i} \sum_{k \in C_{ij}} x_k(i, j) E[T^D_k(i, j)] \quad (9a)$$

s.t.

$$\sum_{k \in C_i} \sum_{j \in U^A_i} x_k(i, j) = 1, i \in U, \quad (9b)$$

$$x_k(i, j) + \sum_{u \in U^A_j} x_k(j, u) \leq 1, i, j \in U, k \in C_{ij}, \quad (9c)$$

$$x_k(i, j) + \sum_{u, v \in U^I_i \cup U^I_j} x_k(u, v) \leq 1, i, j \in U, k \in C_{ij}, \quad (9d)$$

$$x_k(i, j) \in \{0, 1\}, i, j \in U, k \in C. \quad (9e)$$

The above optimization programming is referred as the Spectrum Assignment for Minimum Expected Delay (SA-MED). The constraint (9b) shows that one SU is allowed to access only one available channel at each time; the constraints (9c) and (9d) indicate the intra-link and inter-link interference-free conditions, respectively; the constraint (9e) specifies the range of the variable $x_k(i, j)$. Note that the NP-hardness of this problem can be easily derived from previous work [28].

IV. THE COLORING ALGORITHM FOR SA-MED

In this section, we propose a novel centralized algorithm that takes into account the expected delay, the conflict constraints, and the available spectrum of each link, termed Minimum Cost Coloring (MCC), to resolve the NP-hard optimization problem SA-MED. MCC is executed by placing a central spectrum manager to collect the necessary information and determine the spectrum assignment for SUs, and then inform the SUs of the assignment through a common control channel. In our approach, we first reduce SA-MED to a variant of the graph coloring problem by mapping channels into colors, and then assign channels to links.
A. Link Conflict Graph

Denote the link conflict graph by $G(V, E, C)$. We use Fig. 2 to illustrate the construction of $G$. Fig. 2(a) is a flow topology of a CRN, in which each node is a SU, each solid edge is a single-hop link, each dashed edge indicates that the physical distance between two SUs is less than the interference range, and the label of an edge represents the available spectrum band of the link. Note that this flow topology is a sub-graph supporting the set of all flows, excluding the links without any flow. In the conflict graph $G$, $C$ is the set of available channels in the network, and $v \in V$ corresponds to a link in the flow topology. An edge exists between two vertices $u$ and $v$ if the corresponding link in the flow topology has at least one common available channel and the two incident vertices might interfere with each other if they are assigned the same channel. The set of colors beside an edge in $G$ is the set of common available channels. An edge associated with a color set is called a colored edge.

Thus solving the problem SA-MED is equivalent to coloring each vertex in $G$ by using an available channel to minimize the total expected delay. The coloring algorithm is constrained by that if there is a $k$-colored edge between any two vertices in $G$, they are not allowed to use color $k$ simultaneously. There are two main characteristics in this conflict graph:

- **Link-based Vertex.** In order to reflect both the intra-link interference and the inter-link interference in multi-hop CRNs, we use a vertex to represent a single-hop link of the flow topology.
- **Multi-colored Edge.** We extend the graph to a multi-color conflict graph by considering the impact of PU activities on available channels and SUs’ interference.

B. Coloring via Labeling

In order to measure the interference between any two vertices $u, v \in V$, we define a 0-1 binary interference factor $I_k(u, v)$, which is equal to 1 if and only if there is a $k$-colored edge between $u$ and $v$. For each vertex $v \in V$, its $k$-color-specific degree, denoted by $N_{vk}^I$, is the number of conflict edges, i.e.,

$$N_{vk}^I = \sum_{u \in U_v} I_k(u, v).$$  \hspace{1cm} (10)

Eq. (10) indicates the number of neighbors that cannot simultaneously use color $k$ if $k$ is assigned to $v$, measuring the influence on neighbors when $k$ is assigned to vertex $v$.

Furthermore, we introduce a new concept called “color-cost”, indicating the expected delay of available channels. Let $\text{cost}_{vk}$ be the cost of using color $k$ at vertex $v$ in its neighborhood:

$$\text{cost}_{vk} = E[T_{vk}^D](N_{vk}^I + 1),$$  \hspace{1cm} (11)

where $E[T_{vk}^D] = E[T_k^D(i, j)]$, with $(i, j)$ being the link in the flow topology that maps to the vertex $v$ in $G$.

Since the set of available channels of each vertex $v \in V$ is not empty, $v$ is associated to a vector of color-costs with $|C_v|$ elements. For the minimum expected delay optimization, we label and color the vertex $v$ by $\text{label}_v = \min_{k \in C_v} \{\text{cost}_{vk}\}$ and $\text{color}_v = \arg \min_{k \in C_v} \{\text{cost}_{vk}\}$, respectively. The methods of labeling and coloring consider the trade-off between spectrum utilization in terms of selecting the color with the minimum cost and minimum interference to neighbors in terms of the $k$-color-specific degree.
The whole process of our coloring algorithm is summarized in Alg. 1. At each stage, MCC calculates the label values for all non-colored vertices. Then, MCC selects a vertex \( v^* \) with the minimum label value, and colors it by the color \( k^* \) with the minimum cost. After coloring, the colored vertex and the corresponding associated edges are removed from the graph. Moreover, \( k^* \) is deleted from the sets of available channels of the neighbors of \( v^* \) and the corresponding incident edges of \( v^* \). Obviously, the interference constraints keep on changing as other vertices are processed, and both the label and the cost values are updated accordingly. This coloring process terminates after each vertex receives a color.

**Algorithm 1 Minimum Cost Coloring (MCC)**

1: while \( V \neq \emptyset \) do
2: For each vertex \( v \in V \), calculate \( label_v \)
3: Find \( v^* = \arg \min_{v \in V} \{ label_v \} \), assign \( k^* = color_{v^*} = \arg \min_{k \in C_{v^*}} \{ cost_{v^*k} \} \) to \( v^* \)
4: Delete \( v^* \) and all its incident edges
5: Remove \( k^* \) from the lists of \( v^* \) ’s neighbors and the corresponding incident edges
6: For each stand-alone vertex (without any edge), assign a color with the lowest cost, and delete the vertex
7: end while

**C. Theoretical Analysis**

Although the delay minimization problem proposed in Section III-D is NP-hard, we show that the MCC algorithm can obtain the approximate solution in polynomial time with a bounded performance ratio.

**Theorem 1:** In a multi-hop CRN with \( M \) SUs and \( N \) available channels, the time complexity of MCC is \( O(M^4N) \).

**Proof:** Let \( |V| \) be the number of vertices in graph \( G = (V, E, C) \). The while loop in MCC repeats the coloring process at most \( |V| \) time. During each iteration of MCC, the computation of the values of all labels takes at most \( |V| \cdot N \) time. Therefore, the time complexity of MCC is \( O(|V|^2N) \). Since \( |V| \leq \frac{1}{2}M(M - 1) \), the time complexity of MCC becomes \( O(M^4N) \).

Next, we perform a theoretical analysis on the upper bound of the total cost using MCC. The cost bound of using one available spectrum band is estimated as,

\[
Bound = \sum_{v \in V} \min_{k \in C_v} \{ E[T_{vk}^D](N_{vk}^I + 1) \}.
\]

Correspondingly, the total coloring cost of using MCC is,

\[
Cost = \sum_{f \in F} \sum_{(i,j) \in f} \sum_{j \in U^k} \sum_{k \in C_i} x_k(i, j)E[T_k^D(i, j)].
\]

In order to analyze the upper bound and the performance ratio of MCC, we introduce the following notations.

- \( S^{(q)} = \{(v, k, q)\} \) : the set of chosen vertex-color pairs at the \( q \)-th coloring stage.
- \( l^{(q)}_{vk} \) : the availability of color \( k \) after the \( q \)-th coloring stage, i.e., \( l^{(q)}_{vk} = 1 \) if and only if color \( k \) is available for the vertex \( v \) after the \( q \)-th coloring stage.
- \( F^{(q)} = \{(v, k)|l^{(q)}_{vk} = 1\} \) : the set of available vertex-color pairs after the \( q \)-th coloring stage.
- \( u^{(q)}_{vk} \) : the indicator of the disabled vertex-color pairs due to the \( q \)-th coloring stage, i.e., \( u^{(q)}_{vk} = 1 \) if and only if \( l^{(q-1)}_{vk} = 1 \) and \( l^{(q)}_{vk} = 0 \).
- \( U^{(q)} = \{(v, k)|u^{(q)}_{vk} = 1\} \) : the set of disabled vertex-color pairs due to the \( q \)-th coloring stage.
- \( D^{(q)}_{vk} \) : the vertex \( v \)’s \( k \)-color-specific degree after the \( q \)-th coloring stage, i.e., \( D^{(q)}_{vk} = \sum_{w \in V} I_k(w, v)l^{(q)}_{wk}l^{(q)}_{vk} \). Let \( D^{(0)}_{vk} = N_{vk}^I \).
- \( M(cost) \) : the set of vertex-color pairs with the minimum color-cost, i.e., \( M(cost) = \{(v, k)|v \in V, k = \arg \min_{k \in C_v} \{ cost_{vk} \} = \arg \min_{k \in C_v} \{ E[T_{vk}^D](N_{vk}^I + 1) \} \}. \)
**Lemma 1:** \( S(q), \mathcal{F}(q), \) and \( \mathcal{U}(q) \) meet the following relationships:

\[
\mathcal{F}^{(q-1)} = \mathcal{F}(q) \cup S(q) \cup \mathcal{U}(q), \\
\mathcal{F}(q) \cap S(q) = \mathcal{F}(q) \cap \mathcal{U}(q) = S(q) \cap \mathcal{U}(q) = \emptyset, \\
\mathcal{F}(0) = \bigcup_{q} S(q) \bigcup_{q} \mathcal{U}(q).
\]

**Lemma 2:** \( D_{vk}^{(q)} \leq D_{vk}^{(s)} \leq D_{vk}^{(0)}, \forall q \geq s \geq 0. \)

*Proof:* Since \( i_{vk}^{(q)} \in \{0, 1\} \) and \( i_{vk}^{(q)} \leq i_{vk}^{(q-1)} \), the degree of the vertex \( v, D_{vk}^{(q)} \), is a non-increasing function of \( q \). Thus we have \( D_{vk}^{(q)} \leq D_{vk}^{(s)} \leq D_{vk}^{(0)}, \forall q \geq s \geq 0. \)

**Lemma 3:** \( E[T_{vk}^{D_{vk}^{(q-1)}}](D_{vk}^{(q-1)} + 1) \leq E[T_{vk}^{D_{vk}^{(q)}}](D_{vk}^{(q)} + 1) \leq E[T_{vk}^{D_{vk}^{(0)}}](D_{vk}^{(0)} + 1), \forall q > 0, (v, k) \in \mathcal{F}^{(q-1)}, (v, q, k) \in S(q). \)

*Proof:* According to the definition of \( S(q) \) and the labeling and coloring methods in Section IV-B, we have,

\[
(v, k) = \arg\min_{(v, k) \in \mathcal{F}^{(q-1)}} \{ E[T_{vk}^{D_{vk}^{(q-1)}}](D_{vk}^{(q-1)} + 1) \}.
\]

Furthermore, from Lemma 1 and 2, we have,

\[
E[T_{vk}^{D_{vk}^{(q-1)}}](D_{vk}^{(q-1)} + 1) \leq E[T_{vk}^{D_{vk}^{(q)}}](D_{vk}^{(q)} + 1) \leq E[T_{vk}^{D_{vk}^{(0)}}](D_{vk}^{(0)} + 1).
\]

**Lemma 4:**

\[
\sum_{(v, k) \in \mathcal{M}(\text{cost}) \cap S(q)} E[T_{vk}^{D_{vk}^{(0)}}](D_{vk}^{(0)} + 1) \\
\geq (1 - |S(q) \setminus \mathcal{M}(\text{cost})|) E[T_{vk}^{D_{vk}^{(0)}}](D_{vk}^{(q-1)} + 1).
\]

*Proof:* Since the vertex-color pair \((v, k)\) is selected at the \( q \)th coloring stage, we have the following relationship according to Lemma 3,

\[
\sum_{(v, k) \in \mathcal{M}(\text{cost}) \cap S(q)} E[T_{vk}^{D_{vk}^{(0)}}](D_{vk}^{(0)} + 1) \\
\geq \sum_{(v, k) \in \mathcal{M}(\text{cost}) \cap S(q)} E[T_{vk}^{D_{vk}^{(0)}}](D_{vk}^{(q-1)} + 1) \\
= \sum_{(v, k) \in \mathcal{M}(\text{cost}) \cap S(q)} E[T_{vk}^{D_{vk}^{(0)}}](D_{vk}^{(q-1)} + 1) \\
= (1 - |S(q) \setminus \mathcal{M}(\text{cost})|) E[T_{vk}^{D_{vk}^{(0)}}](D_{vk}^{(q-1)} + 1).
\]

**Theorem 2:** The total coloring cost of \( \text{MCC} \) is bounded, i.e., \( \text{Cost} \leq \text{Bound} \).

*Proof:* Since \( \mathcal{M}(\text{cost}) \subseteq \mathcal{F}(0) \) and \( \mathcal{F}(0) = \bigcup_{q} S(q) \bigcup_{q} \mathcal{U}(q) \), we have,

\[
\mathcal{M}(\text{cost}) = \bigcup_{q} S(q) \bigcup_{q} \mathcal{M}(\text{cost}) \bigcup_{q} \mathcal{U}(q) \bigcap \mathcal{M}(\text{cost}).
\]
Thus, after executing MCC, the cost bound meets the following condition:

\[
\text{Bound} = \sum_{v \in V, k \in C_v} \min \left\{ E[T_{vk}]D_{vk}^{(0)}(D_{vk}^{(0)} + 1) \right\}
\]

\[
= \sum_{(v,k) \in M(\text{cost})} E[T_{vk}]D_{vk}^{(0)} + 1
\]

\[
= \sum_{q} \sum_{(v,k) \in M(\text{cost}) \cap S(q)} E[T_{vk}]D_{vk}^{(0)} + 1
\]

\[
+ \sum_{q} \sum_{(v,k) \in M(\text{cost}) \cap \bar{S}(q)} E[T_{vk}]D_{vk}^{(0)} + 1
\]

\[
\geq \sum_{v \in V} \min_{k \in C} \sum_{q} E[T_{vk}]D_{vk}^{(0)} + 1 \left( D_{vk}^{(0)} + 1 \right)
\]

\[
= \sum_{q} E[T_{v_qk_q}] \left( D_{v_qk_q}^{(q-1)} + 1 \right) \left( 1 - |S^{(q)} \setminus M(\text{cost})| \right)
\]

\[
+ \sum_{q} E[T_{v_qk_q}] \left( D_{v_qk_q}^{(q-1)} + 1 \right) \left( D_{v_qk_q}^{(q-1)} + 1 \right)
\]

\[
+ |S^{(q)} \setminus M(\text{cost})|
\]

\[
= \sum_{q} E[T_{v_qk_q}] \left( D_{v_qk_q}^{(q-1)} + 1 \right)^2
\]

Note that \( \sum_{j \in U, k \in C} x_k(i, j) = 1 \), then \( \sum_{q} E[T_{v_qk_q}] = \text{Cost} \). Therefore, \( \text{Cost} \leq \text{Bound} \). \( \blacksquare \)

Note that channel assignment for total cost minimization is equivalent to the minimum-weighted maximum-independent set problem, if there is only one available channel in the network. Additionally, we can draw the following conclusion according to [28], [29].

**Theorem 3:** Let \( \text{Cost}^*(G) \) be the optimal solution, and \( \text{Cost}(G) \) be the solution obtained from MCC on graph \( G \). Then, the performance ratio is bounded by

\[
\frac{1}{\Delta + 1} \leq \frac{\text{Cost}^*(G)}{\text{Cost}(G)} \leq \frac{1}{2},
\]

where \( \Delta \) is the maximum degree of the vertices in \( G \).

**Proof:** From Theorem 2 we have,

\[
\text{Cost}^*(G) = \min \left\{ \sum_{v \in V, k \in C_v} E[T_{vk}] \right\}
\]

\[
\geq \sum_{v \in V, k \in C_v} \min \left\{ E[T_{vk}] \right\}
\]

\[
= \sum_{v \in V, k \in C_v} \min \left\{ \frac{E[T_{vk}]D_{vk}^{(0)}D_{vk}^{(0)} + 1}{D_{vk}^{(0)}} \right\}
\]

\[
\geq \frac{1}{\Delta + 1} \sum_{v \in V, k \in C_v} \min \left\{ E[T_{vk}]D_{vk}^{(0)}D_{vk}^{(0)} + 1 \right\}
\]

\[
= \frac{1}{\Delta + 1} \text{Bound}
\]

\[
\geq \frac{1}{\Delta + 1} \text{Cost}(G).
\]

That is, \( \frac{1}{\Delta + 1} \leq \frac{\text{Cost}^*(G)}{\text{Cost}(G)} \).
Let \( 1 \leq h \leq N_{vk} \), then \( \frac{cost_{vk}}{N_{vk} + 1} \leq \frac{cost_{vk}}{h + 1} \). According to the definition of \( cost_{vk} \) in (11), we have,

\[
Cost^*(G) \leq \sum_{v \in V, k \in C_v} E[T_{vk}] = \sum_{v \in V, k \in C_v} \frac{cost_{vk}}{N_{vk} + 1} \leq \sum_{v \in V, k \in C_v} \frac{cost_{vk}}{h + 1} = \frac{1}{h + 1} \sum_{v \in V, k \in C_v} cost_{vk} \leq \frac{1}{2} \sum_{v \in V, k \in C_v} cost_{vk} = \frac{1}{2} \frac{1}{2} Cost(G).
\]

That is, \( \frac{Cost^*(G)}{Cost(G)} \leq \frac{1}{2} \).

**V. SPECTRUM BANDWIDTH SHARING**

Even with a channel that does not cause interference to PUs, a flow might still not be able to obtain a full spectrum capacity because of the inter-flow contentions. In this section, we discuss the problem of spectrum sharing among different flows in multi-hop CRNs, with an objective of improving the end-to-end performance, which is restricted by the bottleneck links [18], [30].

**A. Sharing Model and Formulation**

We seek to distribute spectrum bandwidths to different flows to achieve fairness. Instead of adopting the two classical fairness criteria, i.e., max-min fairness and proportional fairness, we employ an alternative utility function called **minimum potential delay fairness** that could achieve the bandwidth-sharing objective by minimizing the sum of the reciprocal of the bit rates [18]:

\[
\min \sum_{f \in F} \frac{1}{r^f},
\]

where \( r^f \) is the effective bit rate of the end-to-end flow \( f \). Note that (18) can capture the long-term throughput that a user could expect from a network [18]. According to [27], the **minimum potential delay fairness** belongs to the so called \( \lambda \)-fairness class with \( \lambda = 2 \). The max-min fairness and the proportional fairness also belongs to the \( \lambda \)-fairness class, with \( \lambda = \infty \) and \( \lambda = 1 \), respectively.

For a given effective flow rate, the allocated bandwidth \( B^f_k(i, j) \) for flow \( f \) in link \((i, j)\) over spectrum \( k \) is derived according to Shannon’s Theorem:

\[
B^f_k(i, j) = \begin{cases} 
\frac{r^f}{\log_2(1 + P_{ij} N_0)}, & \text{if } x_k(i, j) = 1, \\
0, & \text{otherwise}. 
\end{cases}
\]

In addition, a feasible bandwidth allocation must satisfy the link capacity constraint. That is, the sum of the effective flow bit rates on the link from SU \( i \) to SU \( j \) cannot exceed the capacity \( c^f_k(i, j) \) of spectrum \( k \) on link \((i, j)\), i.e.,

\[
\sum_{f \in F} x_k(i, j)r^f \leq c^f_k(i, j).
\]
Thus, we can establish the following optimization problem of Bandwidth Sharing for Minimum Potential Delay fairness (BS-MPD):

\[
\min_{\mathbf{r}} \sum_{f \in F} \frac{1}{r_f} \quad (21a)
\]

subject to

\[
\sum_{f \in F} x_k(i, j)r_f \leq c_k(i, j), f \in F, (i, j) \in f, k \in C, \quad (21b)
\]

\[
r_f(i, j) \geq 0, f \in F, (i, j) \in f. \quad (21c)
\]

The constraint (21b) requires that the sum of all effective flow bit rates on each link does not exceed the spectrum capacity; the constraint (21c) shows the ranges of the variables. The problem BS-MPD aims at finding a fair end-to-end flow rate allocation \(\{r_f\}\) for a given spectrum assignment \(\{x_k(i, j)\}\). Then, we can distribute the spectrum bandwidths to all flows according to (19).

### B. Lagrange Multipliers for BS-MPD

Note that the objective function of the problem BS-MPD (21a) is concave and nonincreasing, and all the constraints are linear. Thus, we can find a unique global optimal solution in its feasible set within a polynomial time. According to the theory of convex optimization, we can use Lagrange Multipliers to obtain the optimal solution. Let \(\Phi(\mathbf{r}) = \sum_{f \in F} \frac{1}{r_f}\) and \(\Psi(\mathbf{r}) = \mathbf{X}\mathbf{r} - \mathbf{C}\alpha\), where \(\mathbf{r}, \mathbf{X}\), and \(\mathbf{C}\alpha\) are the flow bit rate vector, the spectrum assignment matrix, and the spectrum capacity vector, respectively. Then the following Lagrange function can be obtained,

\[
H(\mathbf{r}, \theta) = \Phi(\mathbf{r}) - \theta \Psi(\mathbf{r}), \quad (22)
\]

in which \(\theta\) is the vector of Lagrange Multipliers.

Accordingly, we can establish the following set of equations,

\[
\begin{cases}
\frac{\partial H}{\partial r_f} = \frac{\partial \Phi}{\partial r_f} - \theta \frac{\partial \Psi}{\partial r_f} = 0, & f \in F, \\
\Psi = 0. & \end{cases} \quad (23)
\]

**Theorem 4:** The optimal solution of the problem BS-MPD corresponds to the value of \(\mathbf{r}\) obtained from (23).

### VI. Simulation

#### A. Methodology

In this section, we examine the accuracy of our delay model via a comprehensive comparison analysis and evaluate the performance of our proposed algorithms through an extensive simulation study. Note that our MCC is proposed for spectrum assignment. The output of MCC is taken as the input to BS-MPD for spectrum sharing to alleviate the inter-flow spectrum contentions. We denote this version of BS-MPD by MCC+BS. We also compare our algorithms with two others proposed for multi-hop CRNs: the Probabilistic Frequency Selection (PFS) algorithm [23] for minimizing the scheduling delay in terms of the number of time slots experienced per packet, and the Minimum-Delay Channel Selection (MDCS) algorithm [22] for selecting the channel with the minimum delay based on the local network information.

All these algorithms are examined according to the following performance metrics: (1) the average throughput; (2) the average end-to-end delay; (3) the average packet-drop ratio; and (4) Jain’s fairness index, i.e.,

\[
J = \frac{\sum_{f \in F} y_f^2}{|F| \sum_{f \in F} (y_f^2)^2},
\]

where \(y_f\) is the throughput of the flow \(f\). A larger value of \(J \in [0, 1]\) indicates a better fairness in resource management [31]. For the average end-to-end delay, we provide an upper bound of MCC termed UppMCC, which is computed from (12) based on the link conflict graph, for the comparison purpose.

We use OMNeT++/MiXiM to simulate a multi-hop CRN, with 1-10 PU channels having bandwidths ranging from 1MHz to 6MHz, and 20 SUs being randomly and uniformly placed in a 30m\(\times\)30m region. There are 5 multi-hop flows whose sources and destinations are selected among all SUs at random. Note that the sources (the destinations) of all flows are different. Then, we exploit the Breath-First-Search scheme to find out the shortest hop-distance path for each flow. At each source SU, the packets arrive with a certain rate \(\lambda\).
TABLE I

<table>
<thead>
<tr>
<th>Settings of the Simulation Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>Number of Channels</td>
</tr>
<tr>
<td>Bandwidth of Channel $k$ ($B_k$)</td>
</tr>
<tr>
<td>Shadowing Factor ($\alpha$)</td>
</tr>
<tr>
<td>Packet Size ($L$)</td>
</tr>
<tr>
<td>Transmission Power ($p$)</td>
</tr>
<tr>
<td>Gaussian Noise ($N_0$)</td>
</tr>
<tr>
<td>Average False Alarm Probability ($P_f$)</td>
</tr>
<tr>
<td>Time slot ($T$)</td>
</tr>
<tr>
<td>PU Traffic Arrival/Departure Rates ($\mu/\eta$)</td>
</tr>
<tr>
<td>Source SUs’ Packet Inter-Arrival Time ($\lambda$)</td>
</tr>
</tbody>
</table>

To examine the impact of the PU activity, we classify the arrival rate and the departure rate of the PU traffic in a spectrum band into three types, i.e., normal-frequency rate with the rate randomly distributed in (0, 1), low-frequency rate with the rate randomly chosen from (0, 0.5), and high-frequency rate with the rate randomly selected from (0.5, 1). Other simulation parameters are set based on IEEE 802.22 [32] and are listed in TABLE I.

B. Validation of the Delay Model

To validate the accuracy of our proposed delay model, we compare the average end-to-end delay obtained from simulation study and the numerical result given by the theoretical analysis when the PU activities, the number of available channels, and the number of flows change, and present the corresponding results in Fig. 3. In the simulation study, the average end-to-end delay is the mean delay of all flows in the network. In the numerical analysis, the average end-to-end delay is the expected value of all flows and is calculated by (5).

From Fig. 3, it can be observed that the curves obtained from the simulation study and that from the theoretical analysis are close to each other at different scenarios, validating the accuracy of our delay model. The gap between the simulation and the numerical analysis is attributed mainly to the following approximations: (i) the computation of the sensing time by the SUs (see (1)), which approximately considers SNR instead of SINR, underestimates the expected sensing delay; (ii) the computation of the expected packet transmission rate (see (2)), which also approximately utilizes SNR, overestimates the actual effective transmission rate, obtaining a slightly smaller delay. However, the effect of these two approximations can be reduced with the increase of the number of available channels.

C. Comparison Study

In this subsection, we report the performance of our proposed algorithms, i.e. MCC and MCC+BS, and compare it with that of PFS and MDCS, by considering the average end-to-end delay, network throughput, and packet-drop ratio.

The average end-to-end delays of all algorithms are presented in Fig. 4. The average delays of both MCC and MCC+BS are lower than those of the other two algorithms due to the following two reasons: (i) our delay model
is more accurate as it considers both the sensing delay and the transmission delay; and (ii) our algorithm MCC has a bounded performance for delay minimization. On the other hand, a much lower delay of MCC is obtained compared with the delay upper bound UppMCC of MCC, which indicates the total delay in the worst case. Note that even in the worst case, the delay of UppMCC is much smaller than those of both PFS and MDCS, indicating the advantage of MCC.

As shown in Fig. 4(a), the lowest average end-to-end delay occurs in the low-high case because of the long-term absence of the PU activities in the available spectrum bands, while the highest delay happens at the high-low case due to a short spectrum available time. In Fig. 4(b), the average end-to-end delay obviously reduces with the increasing number of available channels. In addition, we observe from Fig. 4(c) that as the number of flows increases, the interference among the flows increases, resulting in a larger end-to-end delay.

Fig. 4. Average End-to-End Delay.

Fig. 5 reports the average network throughput at different scenarios. It is observed that our algorithms MCC and MCC+BS outperform others significantly. The main reason is that our model estimates the total delay more accurately by taking into account both the sensing delay and the transmission delay; while both PFS and MDCS only consider the transmission time, ignoring the impact of spectrum sensing and PU activities on CRNs. On the other hand, compared with MCC, MCC+BS obtains a higher long-term throughput because its fair bandwidth distribution could mitigate contentions among different flows with bottleneck links.

We also illustrate the impact of the PU activities, the number of available channels, and the number of flows on the average throughput by Fig. 5(a), Fig. 5(b) and Fig. 5(c), respectively. Obviously, the average throughput of the flows benefits from a longer absent time of the PU activities in the available spectrum bands, a larger number of available channels, and a less interference from other flows.

Fig. 5. Average Throughput.

Furthermore, we analyze the average packet-drop ratios in Fig. 6, which are consistent with the results of the average end-to-end delay.

Finally, the comparison of all algorithms in terms of Jain’s Fairness Index is presented in Fig. 7. It can be observed that a high level of fairness is achieved by MCC+BS as it combines spectrum selection and spectrum sharing to mitigate the interference and the inter-flow contentions. As the number of flows increases, the contentions among the flows become more intensive, leading to a poorer fairness in MCC, PFS and MDCS.
VII. CONCLUSION AND FUTURE RESEARCH

The severe spectrum scarcity and the inefficient spectrum utilization of traditional wireless networks motivate the research on CRNs. In this paper, we investigate how to assign and share available spectrum bands among multiple simultaneous flows to minimize the total transmission latency in multi-hop CRNs. We first provide a comprehensive analysis on the expected delay, which considers the characteristics of the PU activities, the spectrum capacities, and the packet transmission latency. Based on this delay analysis, we formulate the problem of spectrum assignment for delay minimization. To tackle this problem, a novel spectrum assignment algorithm termed MCC with a bounded performance ratio and a low complexity is designed. To mitigate the inter-flow spectrum contentions, we investigate the problem of spectrum sharing by achieving the so-called minimum potential delay fairness. Finally, extensive simulation study is performed and the results verify the effectiveness of our proposed algorithms.

As a future research, we intend to study the approximability of our proposed problem for delay minimization constrained by intra- and inter-flow interference. We also plan to investigate a more comprehensive delay model for cognitive radio networks that can accommodate not only the spectrum sensing delay and transmission delay, but also spectrum negotiation delay and the delay caused by SU mobility.

ACKNOWLEDGMENT

The authors would like to thank the support from the National Science Foundation of the US (CNS-1162057).

REFERENCES