This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

A New Outlook on Routing in Cognitive Radio Networks: Minimum-Maintenance-Cost Routing

Ilario Filippini, Eylem Ekici, *Senior Member, IEEE*, and Matteo Cesana

***Abstract*—Cognitive radio networks (CRNs) are composed of frequency-agile radio devices that allow licensed (primary) and unlicensed (secondary) users to coexist, where secondary users opportunistically access channels without interfering with the operation of primary ones. From the perspective of secondary users, spectrum availability is a time-varying network resource over which multihop end-to-end connections must be maintained. In this paper, a theoretical outlook on the problem of routing secondary user flows in a CRN is provided. The investigation aims to characterize optimal sequences of routes over which a secondary flow is maintained. The optimality is defined according to a novel metric that considers the maintenance cost of a route as channels, and/or links must be switched due to the primary user activity. Different from the traditional notion of route stability, the proposed approach considers subsequent path selections, as well. The problem is formulated as an integer programming optimization model. Properties of the problem are also formally introduced and leveraged to design a heuristic algorithm when information on primary user activity is not complete. Numerical results are presented to assess the optimality gap of the heuristic routing algorithm in realistic CRN scenarios.**

***Index Terms*—Cognitive radio networks (CRNs), optimization, stable routing.**

I. Introduction

With the abundant use of wireless devices, wireless spectrum, especially the ISM band, becomes increasingly scarce. Recent studies by the Federal Communications Commission (FCC) highlighted that many spectrum bands allocated through static assignment policies are used only in bounded geographical areas or over limited periods of time, and that the average utilization of such bands varies between 15%–85% [1].

Spectrum scarcity and the underutilization of statically allocated spectrum motivate the investigation of allocation policies to allow unlicensed users to utilize the licensed spectrum bands in a dynamic and noninterfering manner. Advancements in the field of software-defined radios allow the development of spectrum-agile devices that can be programmed to operate on a wide spectrum range and tuned to any frequency band in that range with negligible delay [2]. Resulting so-called *cognitive radios* (CRs) sense the spectrum, dynamically identify currently unused spectrum blocks for data communications, and intelligently access the unoccupied blocks called *spectrum opportunities (SOPs)* [3].

We consider in this paper a general scenario of *cognitive radio networks* (CRNs) [4] where two kinds of users share a common spectrum portion: Primary (or licensed) users (PUs) have priority in spectrum utilization, and secondary users (SUs) must access the spectrum in a nonintrusive manner. PUs use traditional wireless communication systems with static spectrum allocation. SUs are equipped with CRs and exploit SOPs to sustain their communication activities without interfering with PUs.

The realization of CRNs requires research in many areas, including spectrum sensing techniques [5], [6], spectrum management policies [7], nondisruptive spectrum transition procedures, and cognitive multiple access strategies [8], [9]. Although most of the research has focused on single-hop scenarios of CRNs, the analysis of end-to-end communication has very recently attracted the attention of the research community, being of profound importance for networking applications [10]. In their most general form, CRNs are wireless multihop networks where information may be forwarded through several relay points to its destination. Yet, CRNs possess properties that set them apart from traditional multihop wireless networks (ad hoc and wireless mesh networks). Different from ad hoc networks, where topological changes are caused by node mobility, topological changes in CRNs occur primarily due to changes in PU activity, which may affect several links at a same time. Furthermore, the channel availability in CRNs is also significantly different than in traditional wireless multichannel multihop networks. Nodes in CRNs potentially have partially overlapping or nonoverlapping sets of available channels. The available channel set at an SU is of time-varying nature and changes in correlated or uncorrelated manner with respect to the sets of other nodes. Consequently, the development of network-layer solutions in CRNs is a nontrivial task that requires novel algorithms able to cope with the spectrum availability throughout the entire path and the necessity of rerouting in case specific portions of the currently active path are "impaired" by the presence of an activating PU.

In this paper, we address the problem of routing in CRNs by offering a novel point of view on the definition of *route stability*. Traditionally, a stable route is defined as a route that must be repaired with a small probability, which ignores the cost of route repair or *route maintenance*. The maintenance cost represents the effort needed or penalty paid to maintain end-to-end

Manuscript received May 06, 2010; revised January 23, 2011; June 25, 2011; May 04, 2012; and September 19, 2012; accepted October 30, 2012; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor S. Subramaniam.

I. Filippini and M. Cesana are with the Department of Electronics and Information, Politecnico di Milano, 20133 Milano, Italy. (e-mail: filippini@elet.polimi.it; cesana@elet.polimi.it).

E. Ekici is with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210 USA. (e-mail: ekici@ece.osu.edu).

Digital Object Identifier 10.1109/TNET.2012.2236569

1063-6692/$31.00 © 2013 IEEE

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

connectivity. Such a cost may account for the service interruption time while switching routes, the amount of resources used to find new available links and routes, and/or the signaling required to include new links into the new route.

To alleviate the problems associated with the traditional definition of route stability, which may lead to selection of routes that require large repair costs, we aim to find a *sequence of paths* throughout the lifetime of a connection that *minimizes the maintenance cost*. To this end, we associate a cost to each new link activation and define a stable route as a route that is maintained during its lifetime with a small total cost. This definition intrinsically considers a tradeoff between the reliability of a given route and the effort needed to repair that route during its entire lifetime.

The salient contributions of the work presented in this paper can be listed as follows.

- A new approach to route selection in multihop CRNs is presented that accounts for route maintenance costs.
- Fundamental properties of the problem and the optimal solutions are formally shown.
- Efficient algorithms are derived to compute the optimal solutions under full knowledge of PU activity.
- Heuristic algorithms are introduced to solve the problem under partial knowledge of PU activity.
- Numerical results are presented to assess the quality of minimum-maintenance-cost route selection.

II. Related Work

Previously published work in the field is primarily concerned with the design of novel routing protocols in cognitive environments and with the definition of proper routing metrics. The most diffused approach resorts to some type of on-demand routing protocol, leveraging different metrics to assess the quality of a given path.

In [11] and [12], authors consider metrics based on interference and transmission power. In [11], since each node in the network uses different bands for transmission and each band has a certain interference area according to its transmission power, the objective is to minimize the network-wide product between bandwidth and size of the interference area. In [12], the authors analyze which is the best forwarding strategy for SU transmissions constrained by the interference level that PUs can tolerate.

Delay minimization is the leading principle of works in [13]–[15]. In [13], authors develop a metric that takes into account the switching delay between frequency bands and the medium access delay of a given band. Work in [13] is extend by [14] including queueing delay as well. In [15], the effective transmission time also considers the end-to-end delay of prioritized flows.

Path selection in [16] and [17] is based on the maximum achievable throughput. The authors in [16] express the throughput in terms of spectrum availability, the time percentage a link can be used without being interrupted by PUs. Meanwhile in [17], the authors define a link utility based on the product between the link backlog and the total available bandwidth among its bands according to the Shannon's formula. The utility is then maximized.

A final set of works, [18]–[20], aims at maximizing path stability. In [18], the authors propose a metric that combines link availability, capacity, packet error performance, and switching delay. A tree structure is created for each spectrum band and used for routing and route recovery. In [19], the authors study link availability in terms of path connectivity based on the Laplacian spectrum of the graph. Assigning weights to routes proportionally to the algebraic connectivity of the Laplacian matrix of the SU network graph, they aim at routing packets across paths that avoid network areas that do not guarantee stability and high connectivity. Finally, in [20], mobile SUs are considered as well.

Roughly speaking, the main purpose of the aforementioned work is finding "high-quality" routes in a dynamic network setting. Although some of the proposed routing metrics consider future link availability statistics, the definition of the "best path" does not account directly the cost associated with switching to another path when necessary. In contrast, we provide a mathematical framework to design routes in CRNs with minimum route maintenance cost.

Since we make wide use of mathematical programming formulations throughout the paper, it is worth mentioning here also the related literature leveraging optimization approaches in the field of CRNs. Miao and Tsang propose in [21] a mixed integer linear programming formulation for the problem of achieving throughput optimal routing and scheduling for secondary transmissions. Shi and Hou propose in [22] a nonlinear mathematical programming formulation for the same problem, which is solved through a distributed algorithm. A centralized heuristic based on a graph-layering approach is introduced in [23].

To the best of our knowledge, ours is the first effort in the literature to analyze the interactions of successive route selections for a given flow and to propose a routing algorithm to minimize the route maintenance cost in a multihop CRN.

III. Problem Definition and Notation

In the reference network scenario, see Fig. 1, SUs (*circles*) coexist in a given geographical area with PUs (*squares*). PUs have licensed access to a given spectrum portion, whereas SUs can opportunistically access the spectrum whenever vacated by PUs. From the SUs' perspective, the spectrum availability is a random process that depends on the activity of PUs. In the figure, the circles around PUs represent blocked areas[1] for SUs whenever the corresponding PU is actively using its licensed spectrum portion. In this scenario, SUs are interested in establishing and maintaining multihop paths toward destination SUs. Since PUs may intermittently activate/deactivate, rerouting and route repair are needed to maintain multihop connectivity among SUs. Rerouting and route repair involve a route maintenance cost that accounts for the signaling overhead to coordinate with other SUs, the corresponding consumed power, etc.

[1]We use this concept just for the sake of presentation in order to better define which SU links are influenced by PU activities. Clearly, it depends on the considered communication technology. The generality of our formulation allows to use any transmission technique: The model only requires to define which SU link can/cannot be active according to PU activities.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

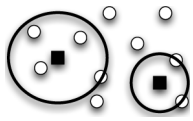


Fig. 1. Reference CRN scenario.

The aforementioned network scenario can be abstracted as follows. Given a two-dimensional Euclidean space $\mathcal{S}$, a set of SU nodes $\mathcal{N}$ is deployed in $\mathcal{S}$. Nodes in $\mathcal{N}$ are equipped with a single transceiver and communicate through wireless links. Along each link, transmissions can be carried out on a given channel m that belongs to the set of available channels $\mathcal{M}$.[2] The set of available links is denoted by $\mathcal{L}$, the availability of link (i,j) from node i to node j only depends on the SU deployment. Link (i,j) using channel m is denoted by $l(i,j)^m$, that is, $l(i,j)^m = \langle (i,j), m \rangle$.

An *epoch* in the epoch set $\mathcal{E} \triangleq \{1..E\}$ is defined as the time period during which there is no change in PU activities. Progress in time is given by the ordered sequence of epochs, from 1 to E.

Finally, we focus on maintenance costs. In general, since the set of all routes from node s to node d, $\mathcal{P}_{sd}(e)$, varies in subsequent epochs, we may be forced by PUs to switch from route $P_i \in \mathcal{P}_{sd}(e)$ at epoch e to a new route $P_j \in \mathcal{P}_{sd}(e+1)$. We associate a cost to this change given by $C[P_i, P_j]$. We say a *link change* occurs when a new link must be established during the transition from epoch e to epoch $e+1$. This means that in order to build route P_j, one or multiple new links must be set up w.r.t. P_i. A *channel change* occurs instead when a link already used in path P_i at epoch e must change only its operating channel in path P_j at epoch $e+1$. Maintenance costs C_{L} and C_{C} are assigned to each added new link in a *link change* and to each channel switch in a *channel change*, respectively. Clearly, if a link does not experience any change from epoch e to epoch $e+1$, then its maintenance cost is 0.

The transition cost—that is, the maintenance cost associated to the switch from route P_i at epoch e to route P_j at epoch $e+1$, $C[P_i, P_j]$—is given by the total cost of changes, that is

$$C[P_i, P_j] = \sum_{l(a,b)^m \in P_j} c[l(a,b)^m, P_i] \tag{1}$$

where $c[l(a,b)^m, P_i]$ is: C_{L}, if $l(a,b)^m$ is a new link w.r.t. P_i; C_{C}, if $l(a,b)^m$ changes the operating channel w.r.t. the one in P_i; 0, if $l(a,b)^m$ is a link of P_i. Note that this cost function allows to simplify the definition of an epoch. Since there is no change when a PU switches from active to idle, an epoch can last until a PU switches from idle to active.

The minimum-maintenance-cost *Route Selection Problem* (RSP) in CRNs can be stated as follows.

(RSP) Given two nodes s and d in $\mathcal{N}$, find a route $P(e)$ from s to d for each epoch $e \in \mathcal{E}$, $P(e) \in \mathcal{P}_{sd}(e)$, such that the sum of the initial route $P(1)$'s setup cost and all the required transition costs is minimized

$$\min_{\{P(e)\}_{e\in\mathcal{E}}} \left\{ C_1[P(1)] + \sum_{e=1}^{|\mathcal{E}|-1} C[P(e), P(e+1)] \right\}$$
$$\text{s.t.} \quad P(e) \in \mathcal{P}_{sd}(e) \qquad \forall\, e \in \mathcal{E}$$

where $C_1[P(1)]$ is the route $P(1)$'s setup cost and $\{P(e)\}_{e\in\mathcal{E}}$ is the sequence of selected routes.

IV. Minimum-Maintenance-Cost Routing

We consider a set $\mathcal{K}$ of traffic demands, where S_k denotes the source node, D_k the destination node, and V_k the traffic load for a given flow k, $\forall k \in \mathcal{K}$. We define two sets of variables. Route selection variables $x_{ij}^{k,c}[e]$ are defined as

$$x_{ij}^{k,c}[e] = \begin{cases} 1, & \text{demand } k \text{ path goes through link } (i,j) \\ & \quad \text{over channel } c \text{ at epoch } e \\ 0, & \text{otherwise} \end{cases}$$

and link cost variables $p_{ij}^{k,c}[e]$ that express the change cost when link (i,j) with channel c is selected for demand k at epoch e.

The model for the (RSP) is the following:

$$\min \quad \sum_{\substack{e\in\mathcal{E}:\\ e\neq 1}} \sum_{k\in\mathcal{K}} \sum_{c\in\mathcal{M}} \sum_{(i,j)\in\mathcal{L}} x_{ij}^{k,c}[e] p_{ij}^{k,c}[e] + \sum_{k\in\mathcal{K}} \sum_{c\in\mathcal{M}} \sum_{(i,j)\in\mathcal{L}} C_{\mathrm{S}} x_{ij}^{k,c}[1] \tag{2}$$

$$\text{s.t.} \quad \sum_{c\in\mathcal{M}} \sum_{(j,i)\in\mathcal{L}} x_{ji}^{k,c}[e] - \sum_{c\in\mathcal{M}} \sum_{(i,j)\in\mathcal{L}} x_{ij}^{k,c}[e] = \begin{cases} -1, & i = S_k \\ 1, & i = D_k \\ 0, & o/w \end{cases} \quad \forall i\in\mathcal{N}, k\in\mathcal{K}, e\in\mathcal{E} \tag{3}$$

$$\sum_{c\in\mathcal{M}} x_{ij}^{k,c}[e] \leq 1 \qquad \forall (i,j)\in\mathcal{L}, e\in\mathcal{E}, k\in\mathcal{K} \tag{4}$$

$$x_{ij}^{k,c}[e] \leq \frac{B_{ij}^{c}[e]}{V_k} \qquad \forall (i,j)\in\mathcal{L}, c\in\mathcal{M}, e\in\mathcal{E}, k\in\mathcal{K} \tag{5}$$

$$p_{ij}^{k,c}[e] = \begin{cases} 0, & x_{ij}^{k,c}[e-1]=1 \\ C_{\mathrm{C}}, & x_{ij}^{k,c}[e-1]=0 \wedge \sum_{c'\in\mathcal{M}} x_{ij}^{k,c'}[e-1]=1 \\ C_{\mathrm{L}}, & \sum_{c'\in\mathcal{M}} x_{ij}^{k,c'}[e-1]=0 \end{cases} \quad \forall (i,j)\in\mathcal{L}, k\in\mathcal{K}, c\in\mathcal{M}, e\in\mathcal{E}: e\neq 1 \tag{6}$$

$$\sum_{h\in\mathcal{K}} \sum_{(l,m)\in IS_{(ij)}} \left(x_{lm}^{h,c}[e] \frac{V_h}{B_{lm}^{c}[e]} \right) - M_{ij}^{k,c}[e]\left(1 - x_{ij}^{k,c}[e]\right) \leq 1 \quad \forall (i,j)\in\mathcal{L}, k\in\mathcal{K}, c\in\mathcal{M}, e\in\mathcal{E} \tag{7}$$

$$x_{ij}^{k,c}[e]\in\{0,1\},\ p_{ij}^{k,c}[e]\geq 0 \qquad \forall (i,j)\in\mathcal{L}, k\in\mathcal{K}, c\in\mathcal{M}, e\in\mathcal{E}. \tag{8}$$

The objective function (2) minimizes the sum of change costs of the links included in the routes selected for every epoch and the setup cost of the initial route. Cost c_{S} is the setup cost of each link of the initial route; it captures the amount of communication resources required to establish the initial route. Therefore, the objective function takes into account the resource consumption both to set the route and to repair it over epochs. Constraints (3) are flow balance equations that define routing paths. Constraints (4) force the flow along a link of each single demand to be served with a single channel in a given epoch. In order to simplify the problem, multiple channels on a link can be used to serve multiple demands in an epoch, but each demand must be served with a single channel on each link. Constraints (5) are availability constraints for each link on each channel. Note that

[2]Note that the definition of a channel is completely arbitrary. It can be a carrier, a set of subcarriers, a slot in a frame, etc.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

the parameter $B_{ij}^{c}[e]$ expresses the capacity of the link $l(i,j)^c$ during epoch e, and it is equal to 0 if the presence of PU at epoch e prevents its use. Constraints (6) define the change cost for each link in each epoch.[3]

If link $l(i,j)^c$ is used in epoch e and:

1) is used at epoch $e-1$ for the same demand, then its change cost at epoch e is 0;
2) is used at epoch $e-1$ for the same demand with a different channel, then it gives a channel change cost;
3) is not used at epoch $e-1$, then it gives a link change cost.

It is worth remarking that, although link change costs and channel change costs are expressed here as scalar values to easy up the presentation, they can be vectors of values (f.i., $C_{\mathrm{L},(ij)c}[e]$) that vary for each epoch, for each link and for each channel in order to provide a scenario that includes maintenance cost dependence on geographical positions, epochs, PU activities, channel interferences, etc. According to the *fluid traffic interference model* [24], constraints (7) limit the total traffic of links belonging to the interference set of each link $l(i,j)$, denoted by $IS_{(ij)}$, where $M_{ij}^{k,c}[e]$ is the smallest constant that is large enough to ensure that the inequality is always satisfied when $x_{ij}^{k,c}[e] = 0$ (regardless of the other variable values). Constraints (8) define variable domains.

The model above is nonlinear due to the objective function and constraints (6). However, it can be linearized by replacing the objective function (2) with the following:

$$\min \sum_{e\in\mathcal{E}:e\neq 1}\sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{M}}\sum_{(i,j)\in\mathcal{L}} p_{ij}^{k,c}[e]+\sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{M}}\sum_{(i,j)\in\mathcal{L}} C_{\mathrm{S}}x_{ij}^{c,k}[1]$$

and constraints (6) with the following two sets of constraints:

$$p_{ij}^{k,c}[e] \geq C_{\mathrm{L}} - C_{\mathrm{L}}\left(\sum_{c'\in\mathcal{M}}\left(x_{ij}^{k,c'}[e-1]\right)+1-x_{ij}^{k,c}[e]\right) \quad \forall (i,j)\in\mathcal{L}, k\in\mathcal{K}, c\in\mathcal{M}, e\in\mathcal{E}: e\neq 1$$

$$p_{ij}^{k,c}[e] \geq C_{\mathrm{C}} - C_{\mathrm{C}}\left(x_{ij}^{k,c}[e-1]+1-x_{ij}^{k,c}[e]\right) \quad \forall (i,j)\in\mathcal{L}, k\in\mathcal{K}, c\in\mathcal{M}, e\in\mathcal{E}: e\neq 1.$$

Note that since variables $p_{ij}^{k,c}$ are minimized, only the tighter constraint will be active forcing a link change cost C_{L} or a channel change cost C_{C}, $(C_{\mathrm{L}} \geq C_{\mathrm{C}})$.

The proposed formulation is rather complete, accounting for all the relevant aspects related to CRNs routing. Nevertheless, in specific but realistic CRN scenarios, some of the assumptions (constraints) given in the formulation may be relaxed. As an example, in some cases connectivity among nodes is the main issue, whereas capacity limitations are much more loose. Think of a network scenario where SUs perform transmissions over short opportunistic, high-capacity links. In this scenario, flows carry small amounts of traffic with respect to link capacities. Constraints (7) and demand index "k" can be dropped in the previous formulation, thus focusing on routing a single traffic demand throughout epochs. This problem where $\mathcal{K} = 1$ and $v_k = 1$ is denoted as *(RSP-1)*. The global routing can be found by the superimposition of the route of each single demand at every epoch determined solving (RSP-1).

In Section V, we first focus on this first simplified scenario in order to analyze its properties and behaviors. The useful insights coming from this study have inspired the work on the complete scenario, which is given in Section IV-B.

A. Simplified Scenario: (RSP-1)

In this section, we show that (RSP-1) is solvable in polynomial time. We take here a constructive approach: We first show that (RSP-1) in the case of a single channel, (SC-RSP-1), $(|\mathcal{M}| = 1)$ belongs to P, and propose a polynomial-time solution approach. Then, we show that the general (RSP-1) can be reduced to an equivalent single-channel formulation through a polynomial reduction.

1) Single Channel RSP-1 (SC-RSP-1): It can be written as

$$\min \quad \sum_{e\in\mathcal{E}:e\neq 1}\sum_{(i,j)\in\mathcal{L}} C_{\mathrm{L}}h_{ij}[e] + \sum_{(i,j)\in\mathcal{L}} C_{\mathrm{S}}x_{ij}[1] \tag{9}$$

$$\text{s.t.} \quad \sum_{(j,i)\in\mathcal{L}} x_{ji}[e] - \sum_{(i,j)\in\mathcal{L}} x_{ij}[e] = \begin{cases} -1, & i = S \\ 1, & i = D \\ 0, & \text{otherwise} \end{cases} \quad \forall i\in\mathcal{N}, e\in\mathcal{E} \tag{10}$$

$$x_{ij}[e] \leq B_{ij}[e] \qquad \forall (i,j)\in\mathcal{L}, e\in\mathcal{E} \tag{11}$$

$$h_{ij}[e] \geq x_{ij}[e] - x_{ij}[e-1] \quad \forall (i,j)\in\mathcal{L}, e\in\mathcal{E}: e\neq 1 \tag{12}$$

$$x_{ij}[e]\in\{0,1\},\ h_{ij}[e]\geq 0 \qquad \forall (i,j)\in\mathcal{L}, e\in\mathcal{E}. \tag{13}$$

Variable x and constraint meanings are similar to ones in the previous MILP formulation (2)–(8). Variables h, replacing variables p, are indicator variables for a link change at link $l(i,j)$ at epoch e.

Indices k and c are dropped due to the use of a single channel and a single demand, as well as constraints (4) and (7).

Note that constraints (12) are equivalent to constraints (6) thanks to objective function (9). We build the proof of polynomial-time characteristic of (SC-RSP-1) (9)–(13) on the following theorems.

Theorem IV.1: The matrix of coefficients of (SC-RSP-1) is totally unimodular.

Proof Sketch: Denoting the sets of variables $x_{ij}[e]$ and $h_{ij}[e]$ as $\mathbf{v} = [\mathbf{x}, \mathbf{h}]^T$, the set of constraints (10)–(13) can be written as $\mathbf{A}\mathbf{v} \leq \mathbf{b}$, where $\mathbf{b}$ is called vector of known coefficients and $\mathbf{A}$ matrix of coefficients. The proof is based on two observations. First, the rows of $\mathbf{A}$ are composed only of elements in $\{-1, 0, +1\}$. Second, matrix $\mathbf{A}$ satisfies the following: Each collection Q of rows can be partitioned into two subsets (Q_1, Q_2) such that the sum of the rows in one subset minus the sum of the rows in the other subset is a vector with entries only in the set $\{-1, 0, +1\}$. A direct inspection of all the possible collections of rows in matrix A verifies such property. This is a sufficient condition for $\mathbf{A}$ to be totally unimodular [25]. □

Knowing that its matrix of coefficients is totally unimodular, we can leverage the following theorem to assess the complexity of (SC-RSP-1).

[3]Note that a branch with a disconnection cost can be easily added to constraints (6) using route selection variables. However, since timeout-based keep-alive link probing is widely used in routing protocols for CRNs, we assume no additional resources are needed to remove a link. Since the timeout is at routing layer, the problem of medium access control is completely decoupled to it, and so are spectrum sensing mechanisms to access the channel.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

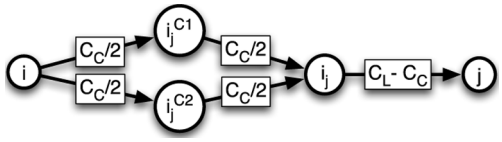


Fig. 2. Equivalent structure of a multichannel link.

Theorem IV.2: Let $\mathbf{A}$ be an integer matrix. The matrix $\mathbf{A}$ is totally unimodular if and only if the polyhedron $P(\mathbf{b}) = \{\mathbf{x} \in \mathbb{R}_+^n | \ \mathbf{Ax} \leq \mathbf{b}\}$ is integral, $\forall \mathbf{b} \in \mathbb{Z}^m$ for which $P(\mathbf{b}) \neq \emptyset$ [25].

This second theorem states that vertices of the solution space polyhedron of problems with totally unimodular coefficient matrices are integer, even if the problem variables $\mathbf{x}$ are fractional. This allows to solve the continuous relaxation of the MILP guaranteeing an integer optimum solution. Moreover, since the vertices of the solution space polyhedron are integer regardless of the variables' integrality, optimum solutions of the integer formulation and the linear relaxation are guaranteed to be equal. Consequently, we can obtain the optimal solution of (SC-RSP-1) by solving a linear program (LP). Since LP solution methods are known to take polynomial time [26], we can state that (SC-RSP-1) $\in P$.

2) Multiple Channel RSP-1: In this section, we describe a polynomial-time reduction algorithm to reduce any formulation of the general (RSP-1) to an equivalent formulation of (SC-RSP-1). This proves that the general (RSP-1) is also polynomial-time since its solution time complexity is given by the sum of the complexities of (SC-RSP-1) and of the reduction procedure, both polynomial with respect to instance sizes.

The driving idea of the reduction procedure is to create an equivalent SU connectivity graph epoch by epoch such that channel-switching operations in the original graph are equivalent to link-switching operations in the equivalent one. To this extent, for each link $l(i,j)$ in the original graph, we introduce the following elements in the equivalent graph:

1) node i itself and an additional node i_j;
2) nodes i_j^c, one for each available channel c on $l(i,j)$;
3) for each channel c, a link from node i to node i_j^c and one from node i_j^c to node i_j are added. In addition, a link from node i_j to node j is added to the graph.

The link switching costs, $C_{\mathrm{L}(lm)}$, in the equivalent graph are defined as follows:

- $\frac{C_\mathrm{C}}{2}$, for links $l(l,m)$ from node i to nodes i_j^c;
- $\frac{C_\mathrm{C}}{2}$, for links $l(l,m)$ from nodes i_j^c to nodes i_j;
- $C_\mathrm{L}(C_\mathrm{S}) - C_\mathrm{C}$, for the link $l(l,m)$ from node i_j to node j.

The new multiple-channel link representation is shown in Fig. 2. The objective function (9) is replaced by

$$\min \sum_{e \in \mathcal{E}: e \neq 1} \sum_{(i,j) \in \mathcal{L}} C_{\mathrm{L}(ij)} h_{ij}[e] + \sum_{(i,j) \in \mathcal{L}} C_{\mathrm{S}(ij)} x_{ij}[1]. \quad (14)$$

Note that objective function (14) does not modify the matrix of coefficients, which is still totally unimodular.

The modeling of channel c unavailability between nodes i and j during epoch e is obtained by imposing $B_{i,i_j^c}[e] = B_{i_j^c,i_j}[e] = 0$. Note that the unavailability of multiple-channel link $l(i,j)$ is equivalent to the unavailability of every channel between nodes i and j. Imposing $B_{i,i_j^c}[e] = B_{i_j^c,i_j}[e] = 0, \ \forall c \in \mathcal{M}$, prevents flows from entering link $l(i_j,j)$, thus the route cannot go directly from node i to node j.

As for the maintenance cost from epochs $e-1$ to e, if channel c becomes unavailable between nodes i and j and the new selected route uses channel d, we have $x_{i,i_j^d}[e] = x_{i_j^d,i_j}[e] = 1$ and $x_{i,i_j^d}[e-1] = x_{i_j^d,i_j}[e-1] = 0$, while $x_{i_j,j}[e] = x_{i_j,j}[e-1] = 1$. Therefore, the objective function (14) adds $\frac{C_\mathrm{C}}{2} + \frac{C_\mathrm{C}}{2} = C_\mathrm{C}$ for a channel change. In case the link must be switched completely, the newly established link is characterized by $x_{i,i_j^c}[e] = x_{i_j^c,j}[e] = x_{i_j,j}[e] = 1$, thus the additional cost in the objective function (14) is $\frac{C_\mathrm{C}}{2} + \frac{C_\mathrm{C}}{2} + C_\mathrm{L} - C_\mathrm{C} = C_\mathrm{L}$, which is consistent with the original problem formulation.

Finally, since integer solutions provide a unique non-split path from source to destination in the modified single channel scenario, per-link per-demand single-channel selection constraints (4) are implicitly satisfied. With the procedure described above, (RSP-1) can be solved with the following polynomial-time algorithm.

Algorithm 1: PolyAlg (RSP-1)

1: $\mathrm{Reduce(RSP\text{-}1)} \rightarrow \mathrm{SC\text{-}RSP\text{-}1_{eq}}$
2: $\mathrm{LP\text{-}SOLVE(SC\text{-}RSP\text{-}1_{eq})}$

In Step 1, the instance of (RSP-1) is reduced to an equivalent single-channel formulation leveraging the reduction procedure described above. Note that the reduction steps are polynomial in number, i.e., $O(|\mathcal{L}||\mathcal{M}|)$. The second and final step solves the equivalent single-channel formulation leveraging LP solvers [26].

B. Complete Scenario: (RSP)

In this section, we show that the Route Selection Problem in the complete scenario, (RSP), is a NP-hard problem. We present a heuristic algorithm leveraging the Lagrangian Relaxation of (RSP). After discussing the properties of the selected relaxation, we conclude by describing our algorithm.

In the complete scenario, capacity constrainst are considered, and multiple source–destination demands must be routed under the *fluid traffic interference model* described by constraints (7). It is easy to show the following theorem.

Theorem IV.3: (RSP) is NP-hard.

Proof Sketch: We can easily reduce the Two-commodity Integral Flow in Directed Graphs Problem (D2CIF) to a recognition version of the Route Selection Problem in the complete scenario (RSP-r). (D2CIF) is the problem of finding a flow routing for two commodities with integer flow requests in a capacitated directed graph. Therefore, it is equivalent to find a feasible routing of an (RSP) instance with two integer demands and only one epoch in the particular case where sets $IS_{(ij)}$ contain only the link (i,j). Since (D2CIF) is NP-complete [27], (RSP-r), which is at least as difficult, is NP-complete. This means that the optimization problem (RSP) is NP-hard. □

When dealing with large-scale instances, we may want to rely on a heuristic algorithm to obtain a solution of (RSP) close to

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

the optimum in much shorter time than solving the MILP formulation (2)–(8) to the optimality. We have designed such an algorithm from the lessons learned in the simplified scenario (RSP-1).

As shown in Section IV-A, (RSP) is polynomial in case capacity constraints are dropped and a single traffic demand must be routed. Building on this observation, the heuristic leverages a relaxed problem formulation where capacity constraints are added to the objective function through Lagrangian multipliers. The remaining set of constraints, (3)–(6) and (8), describes a routing scenario where all the demands are routed together without capacity constraints. The corresponding matrix of coefficients can be arranged in a block diagonal form, where each block consists of the (RSP-1) matrix of coefficients associated to the single source–destination demand. Therefore, it is straightforward to prove the following theorem.

Theorem IV.4: The matrix of coefficients corresponding to constraints (3)–(6) and (8), $\hat{\mathbf{A}}$, is totally unimodular.

Proof Sketch: The proof directly derives from the proof of Theorem IV.1. We show there that the matrix of coefficients of (SC-RSP-1) satisfies a property on its rows. Furthermore, the extension of (SC-RSP-1) to deal with multiple channels, i.e., (RSP-1), does not affect its constraints. Therefore, the property is still valid for the coefficients of (RSP-1). Since $\hat{\mathbf{A}}$ is block-diagonal, only zeros will be added to the rows of every collection Q; as a result, the property valid for (RSP-1) is also valid for $\hat{\mathbf{A}}$. This is a sufficient condition for $\hat{\mathbf{A}}$ to be totally unimodular [25]. □

We can now derive the Lagrangian Relaxation of (RSP), $(\mathrm{RSP})_{\mathrm{L}}(\bar{u})$, for any value of the vector $\bar{u} = \left\{u^k_{i,j}[e] \geq 0, (i,j) \in \mathcal{L}, k \in \mathcal{K}, e \in \mathcal{E}\right\}$. The linearized objective function is shown in (15) at the bottom of the page, while the set of constraints of $(\mathrm{RSP})_{\mathrm{L}}(\bar{u})$ consists of constraints (3)–(6) and (8) of (RSP). Thanks to Theorem IV.4 and Theorem IV.2, we can state that $(\mathrm{RSP})_{\mathrm{L}}(\bar{u})$ belongs to P, for every value of $\bar{u}$, and we can solve it, quickly and efficiently, by using LP methods.

Since $(\mathrm{RSP})_{\mathrm{L}}(\bar{u})$ is easy to solve, we have designed our heuristic algorithm for (RSP) on the idea of the classical *subgradient method* [28], where the key operation is to iteratively solve $(\mathrm{RSP})_{\mathrm{L}}(\bar{u})$ while adjusting the value of $\bar{u}$. At each iteration, capacity constraints are checked and, if satisfied, the solution is compared to the best solution in terms of route maintenance cost met so far. The one with the lowest cost is stored, and the algorithm moves to the next iteration. When the algorithm reaches the end of the last iteration, the stored solution is the result of the algorithm. However, it may happen that the algorithm cannot find a feasible solution because either the scenario does not have a solution, or the scenario has an exact (RSP) solution, but the heuristic algorithm cannot find it. In such cases, the considered solution is the one computed in the last iteration. Even though it is not feasible, i.e., capacity constraints are violated, it is the solution that most likely has a small capacity violation because of convergence of the subgradient method to a solution where the penalty added to the objective function is minimum [see (15)]. This solution, although not allowing to route the entire traffic demand, can provide a good throughput. We have verified it by pushing as much flow as possible through the routes indicated by the solution, taking into account capacity constraints. The *routing efficiency*, $\eta \in [0,1]$, can then be defined as the fraction of the total demands admissible under the computed routing scheme. Note that $\eta = 1$ for feasible solutions. The details of the procedure are given in Algorithm 2.

Algorithm 2: LagrAlg (RSP)

```
Compute an upper bound z_UB
ū_1 ← 0̄, Cost_min ← ∞, Sol_best ← ∅
for it = 1 to it_max do
    z(ū_it) ← Solve (RSP)_L(ū_it)
    if (RSP)_L(ū_it) satisfies capacity constraints (7) then
        Compute (RSP)_L(ū_it) maintenance cost (see Section III)
        if (RSP)_L(ū_it) cost is less than Cost_min then
            Sol_best ← Solution of (RSP)_L(ū_it)
        end if
    end if
    ū_it+1 ← UpdateMultiplier(z(ū_it), z_UB, it)
end for
if Sol_best ≠ ∅ then
    Return Sol_best
else
    Return Solution of (RSP)_L(ū_it_max)
end if
```

The upper bound z_{UB} is computed solving $(\mathrm{RSP})_{\mathrm{L}}(\bar{u})$ with $\bar{u} = \left\{u^k_{i,j}[e] = 0.001\right\}$ and, if capacity constraints (7) are violated, a rerouting phase reroutes each single demand in order to make the solution feasible. The rerouting is based on shortest paths with a metric such that the weight of link (i,j) is equal to inverse of the residual capacity in its interference

$$(\mathrm{RSP})_{\mathrm{L}}(\bar{u}) \qquad z(\bar{u}) = \min_{\bar{x}\bar{p}} \left\{ \sum_{e\in\mathcal{E}:e\neq 1} \sum_{k\in\mathcal{K}} \sum_{c\in\mathcal{M}} \sum_{(i,j)\in\mathcal{L}} p^{k,c}_{ij}[e] + \sum_{k\in\mathcal{K}} \sum_{c\in\mathcal{M}} \sum_{(i,j)\in\mathcal{L}} C_{\mathrm{S}} x^{k,c}_{ij}[1] \right\}$$
$$+ \sum_{(i,j)\in\mathcal{L}} \sum_{k\in\mathcal{K}} \sum_{e\in\mathcal{E}} u^k_{ij}[e] \left(\sum_{h\in\mathcal{K}} \sum_{(l,m)\in IS_{ij}} \left(x^{h,c}_{lm}[e] V_h \right) - M^{k,c}_{ij}[e] \left(1 - x^{k,c}_{ij}[e]\right) - B^c_{ij}[e] \right) \quad (15)$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

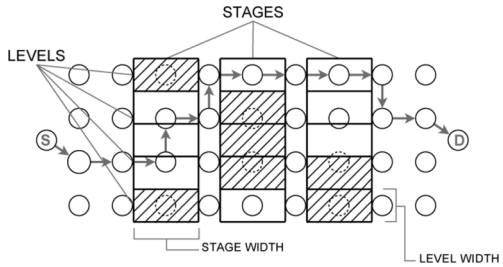


Fig. 3. CRN scenario under evaluation.

set $IS_{(ij)}$, thus penalizing links in highly loaded areas. The $UpdateMultiplier$ procedure is based on the one of the classical subgradient method [28] where the initial scaling coefficient is set equal to 2 and halved after two nonimproving steps. The value of $it_{\max}$ has been experimentally set to 5.

C. Results

1) Simplified Scenario: To test the effect of parameter values on solution characteristics, we solve the (RSP-1) in the reference cognitive network scenario depicted in Fig. 3. SUs are deployed on a grid, and their communication range is equal to the grid step. The transmitting SU and the receiving one are located on the left and right edges of the grid, respectively. A variable number of PUs are active in the scenario and induce *interference areas*. Since we want to stress the routing design algorithm, we further assume that *interference areas* of PUs form vertical barriers between source and destination.

The barriers can be described as sets of rectangular interference areas organized in rows (*levels*) and columns (*stages*). In case multiple channels are available, we have a superimposition of elementary PUs (one for each channel) within the same rectangular interference area.

Unless differently specified, the reference scenario considers 104 SUs (on an 8×13 grid), a two-stages barrier covering three columns of the grid per-stage, and five levels with a total height equal to the grid height; optimal routes are evaluated over 15 epochs with a PU activity probability of 0.3 on each channel, in each epoch. The PU activity patterns randomly generated in each instance are used to determine link availabilities. We set $C_{\mathrm{C}} = 1$ and $C_{\mathrm{L}} = C_{\mathrm{S}} = 10$. All results are averaged on 25 realizations of the PU activity process.

As the first step of analysis, the route maintenance cost is observed for different PU activity probability and interference area shapes and dimensions. Fig. 4(a) shows minimum routing maintenance cost as the PU activity probability increases. As expected, as PU activity increases, links among SUs are more likely to be unavailable, thus the frequency of switching link and/or channel increases. Interestingly, when multiple channels are available, route maintenance cost decreases. Indeed, if transmissions can be carried out over many channels, the probability of finding at least one available channel between two nodes in an epoch is higher. As a result, expensive link changes occur less frequently, and routes can be repaired just by changing the operating channel of their links. We can further observe that the gain of having multiple channels is higher when PUs are more active. Finally, note how the addition of just one channel to the single-channel case can substantially reduce the route maintenance costs.

As for the shape of the interference areas, Fig. 4(b) shows the effect of the barrier width on route maintenance. The stage width is expressed in the number of grid columns affected by PU activity. Intuitively, maintenance costs increase with the stage width since a larger number of SU links are affected by the interference areas. The effect of PU barrier locations is shown in Fig. 4(c). Results are obtained by splitting a single wide barrier into several geographically disjoint stages (3, 4, 6).[4] Results show that the more dispersed the PU activity is, the higher is the route maintenance cost. Crossing a wide interference area with an almost synchronous behavior induces a smaller maintenance cost than crossing several smaller but uncorrelated interference areas.

Fig. 4(d) shows how the route length varies in different scenarios. Here, the average number of hops during epochs and the length of the initial route are shown as number of stages varies. In the single-channel case, since the maintenance cost among epochs has a big impact on the overall cost, the initial route tends to be longer, preferring links that require less changes in the future. In case of multiple channels, the initial setup is the dominant cost, so the first epoch route is shorter. The availability of multiple channels allows to perform channel changes instead of link changes. A similar behavior is exhibited by the average route length during the epochs. This is readily confirmed in Fig. 4(e), which reports the number of hops of the initial route when varying the ratio between link change and channel change costs. For the case of four channels, the curve also reports the cost share of the route at the first epoch with respect to the overall route maintenance cost. The impact of the route setup cost in the first epoch grows as the link setup cost increases. Route maintenance, exploiting channel changes, becomes increasingly negligible. In addition, note that while the initial route length is actually constant with two and three available channels, with more available channels, it shortens when the maintenance cost ratio is higher.

Fig. 4(f) shows the minimum route maintenance costs as the number of epochs increases. In case of a single available channel, the route maintenance cost increases by 120% going from 5 to 25 epochs. On the other hand, cost increases are less relevant when multiple channels are available: 83% and 32% increase in case of two and three channels, respectively. As more channels are available, maintenance costs among epochs are less influent, and the main component of the overall routing cost is due to the initial route setup.

2) Complete Scenario: Since the complete scenario is more difficult to solve, we resort to instances where the end-to-end routing is easier to maintain. This allows us to obtain a reasonable ratio between feasible solutions and generated instances in our tests. The new reference scenario still considers 104 SUs (on an 8×13 grid), but with a one-stage barrier covering four columns of the grid. The parameter $B_{ij}^{c}[e]$ has been fixed equal to 54 for every available link (i, j) and for every available channel c at each epoch e, while the other parameters have been left unchanged.

We focus here on the effects of the traffic load. Other parameters impact the solution in a similar way as the analysis of (RSP-1). Fig. 5 shows the route maintenance cost when the

[4]To provide a fair comparison, the splitting is carried out in such a way that the number of links affected by the PU activity is kept constant.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

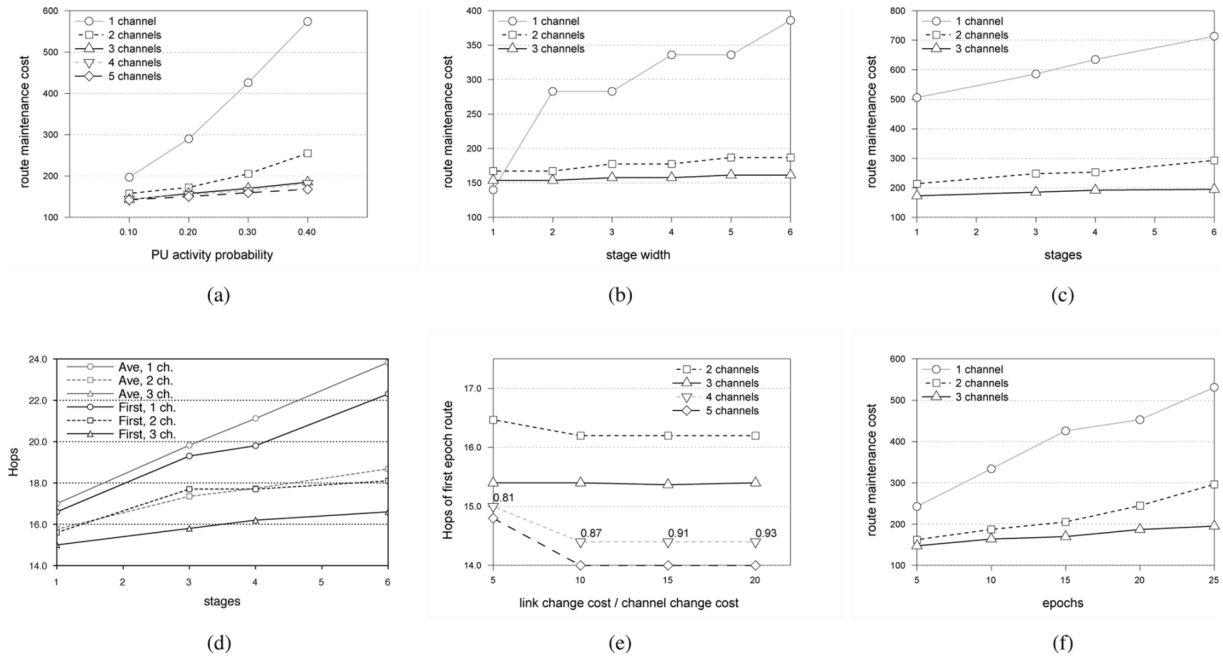


Fig. 4. Quality of minimum-maintenance-cost routing. (a) Route maintenance cost versus PU activity.(b) Route maintenance cost versus stage width. (c) Route maintenance cost versus stages. (d) First epoch route length versus stages. (e) First epoch route length (and cost share) versus available channels. (f) Route maintenance cost versus epochs.

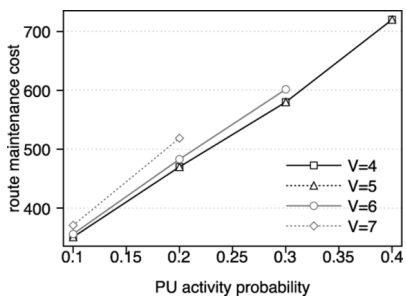


Fig. 5. Route maintenance cost for different per-demand traffic load V when the PU activity probability varies.

PU activity probability varies. The scenario consists of two parallel demands to be routed from the left side to the right side of the grid, each demand requiring a traffic amount equal to $V = 4, 5, 6, 7$. Two kinds of effects can be appreciated. When the load is low ($V \leq 5$), any load increase does not affect the maintenance cost. The network is far from the saturation point, and the obtained solutions are equal to the ones computed without capacity constraints, i.e., solving (RSP-1) for each demand. When the load becomes tangible ($V > 5$), increasing the traffic load causes an increase of the route maintenance cost. Due to interflow interference, solutions show routes far apart in order not to violate capacity constraints. This leads the selection toward links that are not good in terms of route maintenance cost, but sufficiently apart to build a feasible solution. Moreover, note that the increase of traffic load makes infeasible almost every instance with high PU activity probability due to the stringent capacity constraints. Finally, the slope of the curves increases when V increases. This means that the higher the traffic load, the higher the impact on the solution of the PU activity. Roughly speaking, when a network is highly loaded, the effort needed to simultaneously repair broken paths and face capacity limitations increases.

TABLE I
PERFORMANCE OF ALGORITHM 2 WITH TWO DEMANDS AND ONE CHANNEL

Instances		Model	Alg. 2	Sol. $\eta = 1$		Sol. $\eta < 1$	
V	a	time [s]	time [s]	% tot.	gap	% tot.	η
4	0.[1-4]	42.4	5.6	100	0.00	0.0	-
6	0.1	289.7	9.2	100	0.05	0	-
	0.2	2966.5	23.2	75	0.10	25	0.9
	0.3	4569.3	24.5	50	0.05	50	0.9

TABLE II
PERFORMANCE OF ALGORITHM 2

Instances			Model	Alg. 2	Sol. $\eta = 1$		Sol. $\eta < 1$	
D_0	C_0	V	time [s]	time [s]	% tot.	gap	% tot.	η
2	1	4	42.4	5.6	100	0.00	0	-
2	1	6	2608.5	18.9	75	0.07	25	0.9
3	1	2	546.4	17.0	100	0.00	0	-
3	1	4	3046.7	42.7	92	0.06	8	0.9
2	2	6	975.3	563.9	100	0.01	0	-
2	3	6	1854.0	410.2	100	0.00	0	-

Tables I and II compare the performance achieved by Algorithm 2 against the optimum. The optimality gap is computed as $(\frac{C_{\mathrm{AL}} - C_{\mathrm{OPT}}}{C_{\mathrm{OPT}}})$, where C_{AL} is the cost obtained using Algorithm 2 and C_{OPT} is the optimal cost obtained solving the model.

Table I shows the results in case two traffic demands with traffic load $V = 4, 6$ must be routed in a network under several values of the PU activity probability a. The time to solve to optimality the MILP model (*Model time*) and the running time of Algorithm 2 (*Alg. 2 time*) are reported. Moreover, columns of type Sol. $\eta = 1$ and Sol. $\eta < 1$ refer to the runs of Algorithm 2 terminated with a feasible and an unfeasible solution, respectively. The percentage of the total number of instances for which Algorithm 2 has/has not found a solution is reported under the *% tot.* column, while the column *gap* indicates the average cost gap of

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

feasible solutions. Finally, in case a feasible solution cannot be found (Sol. $\eta < 1$), the average *routing efficiency* is also reported.

We can see that the time saving of Algorithm 2 is considerable; it can reach two orders of magnitude, while the optimality gap is contained within 10% for feasible solutions. When the solution is not feasible, the algorithm reaches a *routing efficiency* of 0.9, which is remarkable. Moreover, note that when load is low, thus the solution easier, the PU activity probability does not actually affect the short solving times, and Algorithm 2 always finds feasible solutions. Vice versa, when the load is higher, solving instances becomes hard. Increasing the PU activity probability results in longer solving times, but this effect is less evident using Algorithm 2. In these cases, it finds a larger percentage of unfeasible solutions, but with a good *routing efficiency*.

Table II shows performance results of CRN instances, where the number of demands to be routed (D_0), available channels (C_0), and per-demand traffic amount (V) vary. Results are obtained averaging over the available values of PU activity probability, while the columns are the same as in Table I. The solution time is mainly determined by the traffic saturation level of the scenario. Higher traffic demands determine instances that are harder to solve; the number of demands is less influent for the model. The performance of Algorithm 2 is good even with heavy traffic. Analyzing the effect of introducing more available channels, we can see that the algorithm still performs well, but with a reduced time saving. The reason is that, despite the higher number of variables, instances having more available channels become easier to solve directly with the model.

V. Routing Heuristic

The algorithms presented in Sections IV-A and IV-B provide an efficient way to solve (RSP), even for large-size instances, in case the information on current and future PU activity is fully available at the SUs. The availability at the SU of the exact and up-to-date current spectrum occupation pattern is consistent with the recent directions on spectrum access regulations; indeed, the Federal Communication Commission has recently promoted the opportunistic use of white spaces in the spectrum below 900 MHz and in the 3-GHz bandwidth through the use of centrally maintained spectrum databases indicating over time and space the channel availabilities [29]. Before sending or receiving data, cognitive opportunistic devices will be required to access these databases to determine available channels.

However, the assumption on full and perfect knowledge of future PUs activity adopted in the model is obviously not realistic. To this extent, we drop in this section this assumption, and we design a routing algorithm operating with a more limited knowledge of the PU activity, based on the statistical characterization of the spectrum opportunities. Again, we follow a constructive approach: We start off by highlighting some properties of the minimum-maintenance-cost routing (Section V-A), which we leverage in the design of novel routing metrics (Section V-B) to be used in an operational algorithm (Section V-C). The optimality gap and the behavior of the proposed algorithm is also assessed through numerical evaluations.

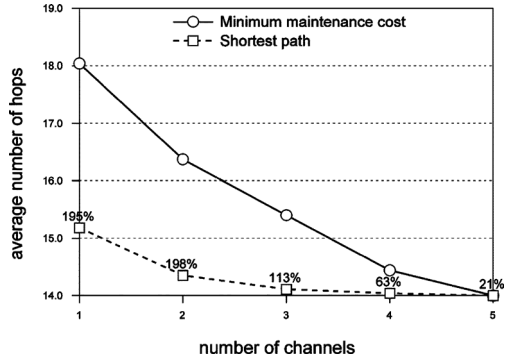


Fig. 6. Average number of hops versus number of channels.

A. Properties and Observations

The fundamental property of minimum-maintenance-cost routing is given in the form of theorem. It states that if two PU activity processes are equivalent from an epoch e_0 to the last one, and the corresponding optimal route selection sequences include the same route at the same epoch e_0, then sums of transition costs given by the two sequences from e_0 to the end are *equal* and *minimum*, even if path selections, except at epoch e_0, are not necessarily the same. We define two PU activity processes equivalent from epoch f to epoch l when their induced sets $\mathcal{P}_{sd}(e)$ are equivalent for $e \in [f, l]$. Leveraging the following theorem, we show in Corollary V.2 that if the route selection sequence reaches a particular route in epoch e_0, then routes at future epochs do not depend on choices made before e_0. A sketch of the proof of Theorem V.1 and the proof of Corollary V.2 are given in the Appendix.

Theorem V.1: Given a scenario, two PU activity processes Υ, and Υ', and indicating with $\hat{P}_{\Upsilon}(e)$ the route selected at epoch e under PU activity process Υ, if the following conditions hold:

- $\Upsilon \equiv \Upsilon'$ from epoch e_0 to epoch $|E|$;
- $\left\{\hat{P}_{\Upsilon}(e)\right\}_{e\in\mathcal{E}}$ and $\left\{\hat{P}_{\Upsilon'}(e)\right\}_{e\in\mathcal{E}}$ are min-cost route selection sequences under corresponding PU activity processes, where $\hat{P}_{\Upsilon}(e)$ and $\hat{P}_{\Upsilon'}(e)$ are not necessarily the same;
- $\hat{P}_{\Upsilon}(e_0) = \hat{P}_{\Upsilon'}(e_0)$;

then the sums of transition costs of $\left\{\hat{P}_{\Upsilon}(e)\right\}_{e\in\{e_0..|\mathcal{E}|\}}$ and $\left\{\hat{P}_{\Upsilon'}(e)\right\}_{e\in\{e_0..|\mathcal{E}|\}}$ are equal and minimum.

Corollary V.2. (Memoryless Property): If the route selection reaches a particular route in epoch e_0, routes at future epochs do not depend on choices made before e_0.

Corollary V.2 introduces a nice property of the minimum-maintenance-cost routing, which is proved to be memoryless in time under specific assumptions.

We highlight here further properties of minimum-maintenance-cost routing as observed from numerical results. Such properties, given here in the form of observations, are leveraged to develop a heuristic routing algorithm in Section VI.

- *Observation I:* Using minimum-hop routing in individual epochs results in extremely high route maintenance costs even though the average path length of optimal routes converges to minimum-hop path lengths with increased number of channels per link. Fig. 6 reports the average number of hops obtained when solving (RSP-1) compared

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

against the average number of hops obtained solving shortest-path problem in each epoch for different numbers of available channels. The shortest-path curve is labeled with the percentage increase in route maintenance cost with respect to (RSP-1). As clearly demonstrated by the figure, routes created by (RSP-1) feature a larger number of hops, but on the other hand, the cost for maintaining the shortest path is considerably and expectedly higher than the one to maintain RSP-generated routes (up to 200% increase).

- *Observation II:* A good indicator to assess the quality of a link is the average uninterrupted link lifetime under the minimum-cost route selection problem. Previous work in the field [16] considers the average link availability to assess the stability of a given link. Here, we are more interested in the expected *time-to-switch*. Time-to-switch is the time between the selection of a link until a forced switch due to a failure of that link. If a link goes up and down frequently, its use incurs a high maintenance cost even if its availability is high on average.

B. Routing Metric Design

From Section V-A, any consistent routing metric for setting minimum-maintenance-cost path cannot be based on the hop count only (*Observation I*), but must also account for the continuous lifetime of the links (*Observation II*). Moreover, the memoryless property (Corollary V.2) allows us to neglect the past history in the definition of the metric, thus focusing on the current epoch and the future ones only.

Leveraging these properties and observations, we propose hereafter two new routing metrics for (RSP). We start from a metric expressly devised to the simplified scenario (RSP-1), which is then modified to obtain a second metric where capacity constraints are taken into account.

Ideally, the link "quality" depends on two factors: the cost of switching from the current link to another link l, C_l^{sw} (*switching cost*), and the expected cost to repair link l in the future, C_l^{rep} (*repair cost*). The former represents the "short-term" investment to maintain the route, the latter the expected "long-term" one.

Each link l can be weighted by the following metric:

$$w_l = \frac{C_l^{\mathrm{sw}} + \alpha C_l^{\mathrm{rep}}}{E\left[\mathrm{TTS}_l\right]}. \tag{16}$$

Parameter α allows tuning different cost contributions, and $E\left[\mathrm{TTS}_l\right]$ represents the average time to switch for the link l. Following Observation II, the longer the continuous lifetime of a link is, the lower maintenance cost incurs. Therefore, the denominator is used to give lower weights to links available for longer time periods.

The *switching cost* for link l operating on channel c from epoch e to epoch $e+1$ can be defined as

$$C_l^{\mathrm{sw}} = \mathrm{dist} \cdot \mathrm{ch.cost} \tag{17}$$

where the value of dist is computed considering the minimum number of hops between link l and the selected route at epoch e, and the value of $\mathrm{ch.cost}$ depends on whether link l appears in the current route with the same channel c $(\mathrm{ch.cost} = 0)$, with a different channel $(\mathrm{ch.cost} = C_{\mathrm{C}})$, or not $(\mathrm{ch.cost} = C_{\mathrm{L}})$.

The *repair cost* is defined as

$$C_l^{\mathrm{rep}} = \sum_{n=1}^{\infty} \left(n \cdot \frac{\mathrm{ch.cost}}{\langle E\left[\mathrm{TTL}\right]\rangle_n} \cdot P_n^l \cdot \prod_{j=0}^{n-1} \left(1 - P_j^l\right) \right). \tag{18}$$

The term $\langle E\left[\mathrm{TTL}\right]\rangle_n$ denotes the expected time-to-switch averaged over the n-hop neighborhood. The sum gives the expected cost to pay to repair link l when it becomes unavailable. Differently from the modeling approach that assumes full knowledge of the PU activity, the proposed metrics only leverage a statistical knowledge of the PU behavior. To this extent, P_n^l denotes the probability to find an available link n hops away from l, which is available as a backup when l fails

$$P_n^l = \langle Pr\{\text{link } i \text{ is UP}|\text{link } l \text{ is DOWN}\}\rangle_{i \in N_n^l}$$

where N_n^l denotes the set of links at exactly n-hops from link l. Note that this equation provides an approximated value of the exact probability since an exact prediction on where the new route will go through after link l's failure is not possible.

We call $X_u(t)$ the ergodic random binary process describing the activity of a PU u: $X_u(t) = 1$ if PU u is active at time t. X_u is a time sample of $X_u(t)$. Given a link k: $\mathcal{U}_k$ is the set of PUs that prevent k from transmitting when they become active. Since $Pr\{\text{link } i \text{ is UP}|\text{link } l \text{ is DOWN}\} = \frac{Pr\{i \text{ is UP} \cap l \text{ is DOWN}\}}{Pr\{l \text{ is DOWN}\}}$, we compute the following probabilities for link in N_n^l:

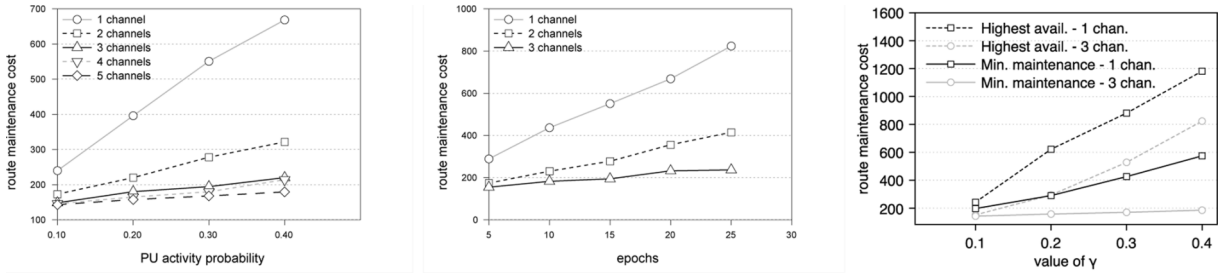
$$\begin{aligned}
&\Pr\{l \text{ is DOWN}\} \\
&\quad = 1 - \Pr\{X_i = 0, \forall i \in \mathcal{U}_l\} \\
&\quad = 1 - E[\textstyle\prod_{i \in \mathcal{U}_l}(1 - X_i)] \\
&\Pr\{i \text{ is UP} \cap l \text{ is DOWN}\} \\
&\quad = \Pr\{\cap_{j \in \mathcal{U}_i} X_j = 0 \cap \cup_{k \in \mathcal{U}_l} X_k = 1\} \\
&\quad = E[\Pi_{j \in \mathcal{U}_i}(1 - X_j)(1 - \textstyle\prod_{k \in \mathcal{U}_l} X_k)].
\end{aligned}$$

If the correlation among PUs with overlapping interference areas is known, then the value of $E\left[\prod_i X_i\right]$ can be computed exactly. When no correlation information is available, one possible approximation is splitting the expected value of the X_i product into the product of their expected values $E\left[X_i\right]$, i.e., the product of average activities of each single PU ($E\left[X_i\right] = \Pr\left\{X_i = 1\right\}$). This is equivalent to considering PU activity processes independent of each other.

Finally, the average *time-to-switch* is determined considering the mean length of the uninterrupted availability periods of link l. Since PU activity processes are usually ergodic, the estimate of $E\left[\mathrm{TTS}\right]$ value can be computed averaging over past observations.

The very same routing metric for the simplified scenario (RSP-1) can be extended to the complete capacity-constrained scenario (RSP). The rationale is to ideally allocate the cost–time ratio we have computed so far to the effective capacity of the link, that is, the actual bandwidth that can be used considering interference effects. The new metric divides the maintenance cost associated to the link by the time we can continuously use that link and the transmission rate we can actually achieve

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



(a) (b) (c)

Fig. 7. Route maintenance cost when adopting Algorithm 3. (a) Route maintenance cost versus PU activity. (b) Route maintenance cost versus epochs. (c) Route maintenance cost varying the metric.

using that link. The link l's weight according to the modified metric is given by

$$w_l = \frac{C_l^{\mathrm{sw}} + \alpha C_l^{\mathrm{rep}}}{E\left[\mathrm{TTS}_l\right] \cdot \left(\max\left\{B_l - \mathrm{load}_l, 0\right\}\right)^{\gamma}} \tag{19}$$

where the new terms B_l, load_l, and γ have been introduced. B_l indicates the nominal capacity of link l, which depends on the used channel and technology. load_l represents the load of link l's interference set, that is, the sum of flows through links belonging to link l's interference set. The term $\max\left\{B_l - \mathrm{load}_l, 0\right\}$ is the residual capacity of link l that can be actually used by traffic flows. Finally, parameter γ allows to tune the weight of the residual capacity in order to privilege solutions with different levels of link saturation. We can think of a practical meaning of w_l when $\gamma = 1$: It represents the per-bit maintenance cost of the traffic transmitted during the period link l is expected be available without interruptions.

C. Routing Algorithm Structure and Results

We use the routing metrics provided in Section V-B in a heuristic to solve the minimum-maintenance-cost routing problem when exact information on PU activity is not available. The main idea of the algorithm is to assign weights to each link in every epoch based on the metric given in (16) or (19) and iteratively find the shortest source–destination route from epoch to epoch. The algorithm is composed of the following steps.

Algorithm 3: Routing Algorithm

1: Initialize
2: **for** $e \in \mathcal{E}$ **do**
3: $\quad \mathrm{AssignWeigths}(e, e-1)$
4: $\quad \mathrm{GetShortestPath}(e)$
5: **end for**

In Step 1, the algorithm computes the shortest path between the intended source and destination, assuming all PUs are inactive (initialization phase). After that, for each epoch, the algorithm assigns weights to SU links according to metric in (16) or (19), and then applies the Bellman–Ford algorithm to find the shortest path at epoch e.

TABLE III
OPTIMALITY GAP OF ALGORITHM 3

PU act.	Optimality Gap				
	1 ch.	2 ch.	3 ch.	4 ch.	5 ch.
0.1	0.22	0.10	0.04	0.02	0.01
0.2	0.34	0.28	0.14	0.06	0.05
0.3	0.32	0.35	0.14	0.09	0.05
0.4	0.36	0.36	0.19	0.17	0.07

In the following, we report results on the quality of solutions of (RSP-1) obtained with metric in (16) and of (RSP) with metric in (19).

1) Simplified Scenario (RSP-1): Fig. 7(a) and (b) reports the route maintenance cost obtained running Algorithm 3 when varying PU activity and the number of epochs, respectively. As observed in these figures, the maintenance cost obtained through the heuristics has a similar trend and behavior as the minimum maintenance cost obtained in Section IV-C [see Fig. 4(a) and (f)]. Table III reports the optimality gap between results obtained using the heuristic algorithm (without using exact future PU activity knowledge) and optimal results (using the optimization algorithm) as the PU activity increases.

The error trend is related to the number of changes faced by selected routes, and the gap increases when the probability of a change in the selected route is higher due to a higher PU activity probability.

Finally, it is worth assessing the quality of the routing metric proposed in (16) with respect to classical metrics of route stability based on the average link availability [16]. To this extent, Fig. 7(c) compares the solution of Algorithm 3 when these two metrics are used to weigh the links among the SUs epoch by epoch. As shown in the figure (and anticipated in Observation 2), routing according to the average link availability induces considerably higher route maintenance cost (up to four times).

2) Complete Scenario (RSP): Fig. 8(a)–(c) reports the behavior of solutions in terms of routing efficiency when the parameter γ varies. Each figure includes four plots, which have been obtained with a PU activity probability a equal to 0.1, 0.2, 0.3, and 0.4. The three figures show the results when one, two, and three channels are available, respectively. We can see that the effect of increasing γ is to generate solutions that achieve higher network efficiencies. When more channels are available, it is easier to find solutions where all the demands can be routed satisfying capacity constraints. Therefore, increasing γ leads to smaller improvements, in particular when the PU activity probability is low. Obtained solutions can provide more and more

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

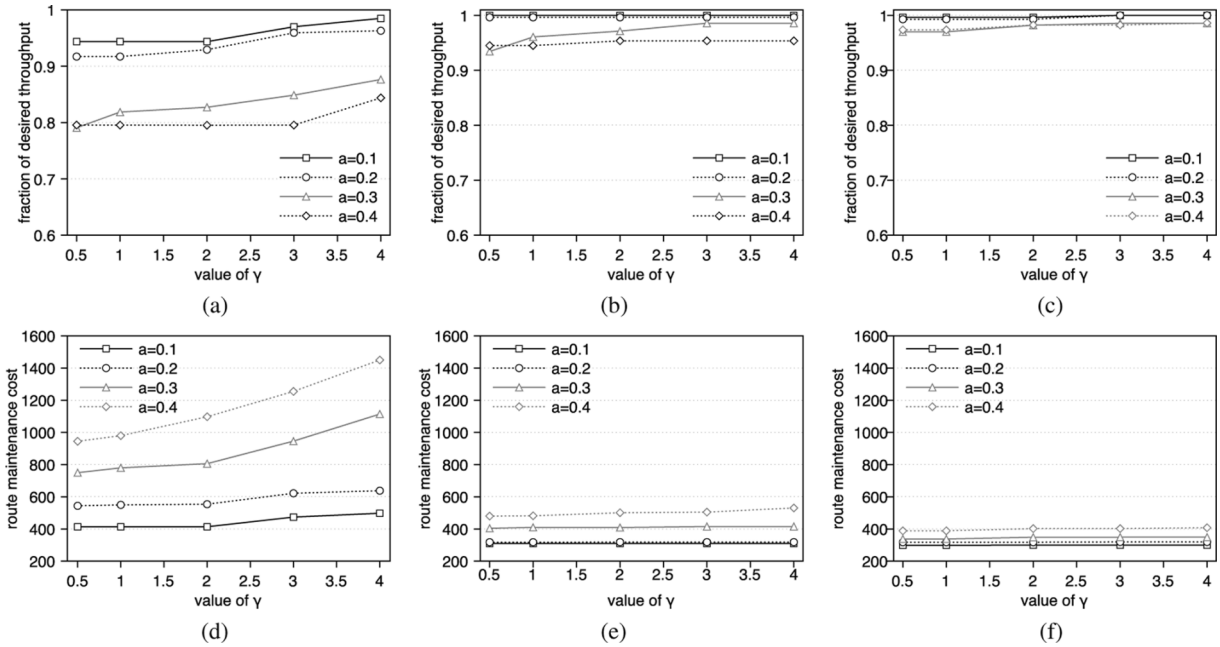


Fig. 8. Achievable routing efficiency and route maintenance cost for different numbers of available channels when parameter γ varies. (a) Routing efficiency with one channel. (b) Routing efficiency with two channels. (c) Routing efficiency with three channels. (d) Maintenance cost with one channel. (e) Maintenance cost with two channels. (f) Maintenance cost with three channels.

frequently the desired throughput within capacity constraints, thus results are closer to 1.

Fig. 8(d)–(f) shows the route maintenance cost of solutions analyzed in Fig. 8(a)–(c). Note that the price to have solutions with higher routing efficiency is a larger route maintenance cost. As expected, costs increase when the values of γ increase. Again, we can see that the availability of more channels reduces the effect of varying γ. Since the solutions can easily satisfy capacity constraints, we have cost curves almost flat with low route maintenance costs.

Finally, we want to assess the quality of obtained results with respect to the optimal solution given by the model in Section IV. Like in Table III, we have the worst performance when only one channel is available. However, in this scenario, the model finds feasible solutions for $a = 0.3, 0.4$ only for less than 10% of the instances. This explains the two corresponding curves in Fig. 8(a), which are far from 1, even with high values of γ. For $a = 0.1, 0.2$, the model provides feasible solutions with an average cost of 371 and 519, respectively. The results of Algorithm 3 instead vary from 414 to 498 with $a = 0.1$ and from 544 to 638 when $a = 0.2$, according to the value of γ. Again, we want to remark the fact that the heuristic algorithm computes routes using only PU activity statistics, while optimal results are computed using a full-knowledge optimization algorithm.

VI. Test Cases

We show here how the proposed routing framework can be mapped to a realistic CRN routing scenario. We are not interested in designing a full-fledged routing protocol for CRNs, but rather we consider a simplified reference routing protocol and show how its characteristics can be used to derive the parameters of the model.

Suppose we have a simplified system where SUs agree on the communication channel to use through a common control channel (CCC). CCC is an out-of-band signaling channel where SUs fall back after each transmission; it can be implemented, for instance, using an additional interface tuned on a dedicated band. The routing algorithm is on-demand. The source SU broadcasts a Route Request (RREQ) message including a Request ID (RID) and an intended destination, which is rebroadcasted at each intermediate SU after storing the association $(\mathrm{last\_hop}, \mathrm{RID})$. Upon reception of the RREQ message, the destination selects the best route according to a routing metric and sends back a route reply (RREP) message along the selected path, leveraging stored associations.

In case of link failure due to PU access to the channel, the system reacts either by changing the working channel or by rerouting. In the former case, the two extremes of the failed link fall back to the CCC and, if it is possible, renegotiate a new channel. If there is no available channel, rerouting is mandatory. Suppose that in case of rerouting, the system behaves as follows. The upstream SU t broadcasts on CCC a special RREQ message toward the destination in order to repair the broken path. The special RREQ message travels within the network until it reaches an SU u that already belongs to the broken path. Node u replies with an RREP message on CCC along the reverse discovered path between t and u. During this phase, traversed SUs must agree on the used channel and tune their interfaces on it before retransmitting the RREP message. Finally, the broken path is patched with a new subpath from node t to node u.

We can now compute the value of parameters C_{C} and C_{L} of the reference routing protocol. Maintenance cost can be associated, as an example, to delays for route repair. Channel change cost can be consequently defined as $C_{\mathrm{C}} = T_{\mathrm{agree}} + T_{\mathrm{switch}}$, where T_{agree} is the time needed to reach a channel agreement

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

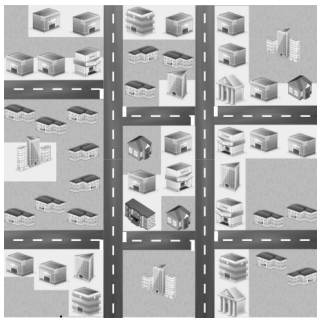


Fig. 9. Neighborhood scenario.

between two SUs on the CCC (f.i., due to an RTS-CTS and CST-CSR exchange), and T_{switch} is the interface switching time to the new channel. Typical values for the two delay components can be set as $T_{\mathrm{agree}} = 775\ \mu$s, $T_{\mathrm{switch}} = 100\ \mu$s [30]. The link change cost instead can be computed as $C_{\mathrm{L}} = T_{\mathrm{RREQ}} + T_{\mathrm{RREP}} + T_{\mathrm{agree}} + T_{\mathrm{switch}}$, where T_{RREQ} and T_{RREP} are the time needed to transmit a RREQ message and a RREP message, respectively. Assuming a rate of 1 Mb/s on CCC, $T_{\mathrm{RREQ}} = 192\ \mu$s and $T_{\mathrm{RREP}} = 160\ \mu$s according to [31]. Note that, as already stated in Section IV, these time values can change among nodes and during epochs, and C_{C} and C_{L} values can vary accordingly. For instance, SUs may use different rates to transmit RREQ and RREP messages, or interfaces may have different switching times. Finally, note that C_{L} and C_{C} can be set considering many other different route maintenance costs, like additional energy consumption or increased network load, even with different levels of details.

A. Sample Network Scenarios

In this section, we show how our routing selection scheme can be applied to realistic scenarios. As an example, we consider a wireless community network in the neighborhood depicted in Fig. 9.

We have run Algorithm 3 for two types of potential PU activities and interference areas. The first example considers a single PU that has an interference area that includes the entire neighborhood, which could represent the case of a TV broadcaster. Five channels are available to SUs. The second example considers PUs having an interference area whose diameter is comparable to the SU communication range; three channels are available. This could be, instead, the case of multiple GSM base stations installed within the neighborhood.

Fig. 10 shows selected routing paths during some epochs, where the selected channel is indicated close to each link. Darker interference areas are characterized by a higher number of simultaneously active PU channels. For sake of clarity, the buildings of Fig. 9 are shown as gray squares.

Fig. 10(a) and (b) refers to the TV broadcaster scenario. They show results snapshots from two of the 15 epochs to be optimized. The numbers in the lower right corner indicate which channels PU is using, while traffic demands go from node A to node A', from B to B', and from C to C'. Since PU behaves synchronously within the entire neighborhood, we can see that the route selection reacts to PU transmissions only by changing the working channel of SU links. Routing paths, instead, do not change since the interference due to the PU activity is spatially homogeneous. In addition, note that the algorithm exploits the availability of multiple channels to make the solution feasible in terms of interference constraints.

TABLE IV
ROUTE MAINTENANCE COSTS AND NETWORK EFFICIENCIES FOR EXAMPLE SCENARIOS

Routing Approach	TV Broadcaster		GSM Base Stat.	
	Cost	η	Cost	η
Algorithm 3	191	1.00	551	0.90
Highest availability	196	0.75	1353	0.75
Load-aware highest availability	693	1.00	1484	0.90

Fig. 10(c) and (d) refers to the scenario with GSM base stations at epoch 1 and 11. PU interference areas are depicted as circles, with PU's active channels written in the center. Differently from the previous example, in this scenario, PU interference areas are heterogeneous, which leads the routing selection process to apply rerouting in order to avoid regions with high PU activity.

Table IV reports results on route maintenance cost and routing efficiency, comparing solutions obtained applying Algorithm 3 using (19) against solutions computed to the classical approach of the average link availability. In addition, we have extended the latter metric to consider residual link capacities (load-aware average link availability) in the same way we did with our metric in (19). We can see that Algorithm 3 gives the best result. It generates solutions with lower maintenance cost and higher routing efficiency. Moreover, in order to achieve the same routing efficiency, solutions provided by load-aware average link availabilities require much larger route maintenance costs.

VII. CONCLUSION

In this paper, a theoretical outlook on the problem of routing secondary user flows in a CRN is provided. Defining optimality based on the cost of maintaining a connection as a sequence of paths, we provide several optimal and heuristic algorithms to solve the problem under different assumptions on PU activity knowledge and traffic characteristics. To the best of our knowledge, this is the first attempt to analyze routing multihop CRNs considering route maintenance cost.

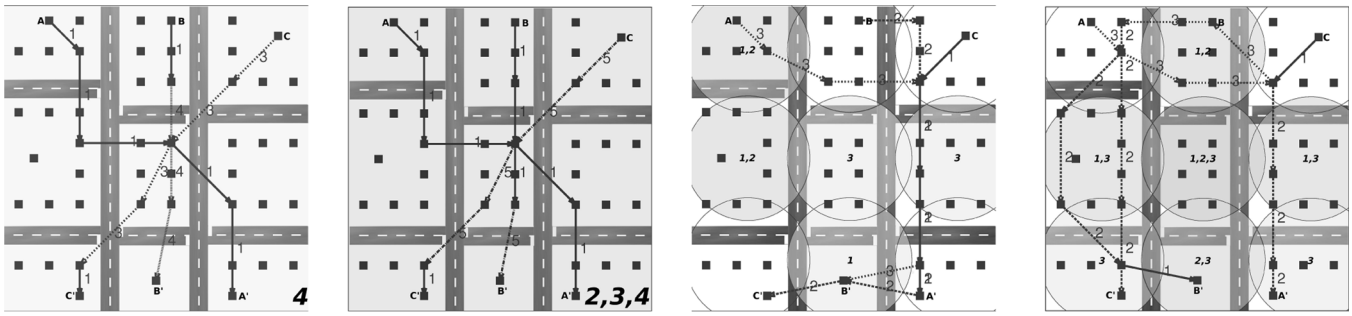
We formally introduce properties of the problem and include computational complexity proofs. The quality and the behavior of generated solutions have been evaluated through numerical results, further sketching guidelines on the applicability and utility of the proposed framework in realistic CRN scenarios.

APPENDIX

Theorem V.1 Proof Sketch: We define a graph $\mathbb{G}(\mathcal{V}, \mathcal{A})$ with a metric $\mathfrak{C} : \mathcal{A} \rightarrow \mathbb{R}^+$ as follows.

- $\mathcal{V} \triangleq \{P_i(e), \forall e \in \mathcal{E}\} \cup \{P'_0, P''_0\}$. $P_i(e)$ represents route P_i at epoch e. P'_0 and P''_0 are two virtual vertices representing, respectively, the initial state in which no route is selected and the end of the time during which nodes s and d must stay connected.
- $\mathcal{A} \triangleq \{(P_i(e), P_j(e+1)) : \forall P_{i,j}(e) \in \mathcal{V} \setminus \{P'_0, P''_0\}\} \cup \{(P'_0, P_i(1)) : \forall P_i \in \mathcal{P}_{sd}(1)\} \cup \{(P_i(|\mathcal{E}|), P''_0) : \forall P_i \in \mathcal{P}_{sd}(|\mathcal{E}|)\}$. Arcs in the

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



(a) (b) (c) (d)

Fig. 10. TV broadcaster and GSM CRN scenarios. (a) TV Tower-Epoch 1. (b) TV Tower-Epoch 4 (c) GSM-Epoch 1. (d) GSM-Epoch 11.

first set express route changes between epochs e and $e+1$. Arcs in the second set express the choice of the initial route. Finally, arcs in the last set join vertices $P_i(|\mathcal{E}|)$ to the end vertex P_0''.

- $\mathfrak{C} : \mathcal{A} \rightarrow \mathbb{R}^+$ is defined as: $\mathfrak{C}\left[(P_i(e), P_j(e+1))\right] = C[P_i, P_j]$, $\mathfrak{C}\left[(P_0', P_i(1))\right] = C\left[P_\emptyset, P_i\right]$, and $\mathfrak{C}\left[(P_i(|\mathcal{E}|), P_0'')\right] = 0$

Note that cost $C[P_\emptyset, P_i]$ is the cost of the selected initial route $P_i = P(1)$. Arcs $(P_i(|\mathcal{E}|), P_0'')$ have length 0 as no route must be arranged after epoch $|\mathcal{E}|$; links of the last route from s to d can simply be destroyed. In addition, $\mathfrak{C}\left[(P_i(e), P_i(e+1))\right] = C[P_i, P_i] = 0$ as route P_i does not change between epochs e and $e+1$, that is, no new links must be activated. A sequence of selected routes during epochs, $\{P(e)\}_{e\in\mathcal{E}}$, can be mapped into a path $\mathfrak{P}^*$ in $\mathbb{G}$, and vice versa. The mapping function $\mathfrak{F}$ is defined as: $P_i = P(1) \Leftrightarrow (P_0', P_i(1)) \in \mathfrak{P}^*$, $P_i = P(|\mathcal{E}|) \Leftrightarrow (P_i(|\mathcal{E}|), P_0'')) \in \mathfrak{P}^*$, and $P_i = P(e) \wedge P_j = P(e+1) \Leftrightarrow (P_i(e), P_j(e+1)) \in \mathfrak{P}^*$. Thus, (RSP) reduces to find a shortest path from vertex P_0' to vertex P_0'' in graph $\mathbb{G}$. Suppose $\mathfrak{P}$ is a shortest path from vertex P_0' to vertex P_0'' in $\mathbb{G}$. The length of such a path is equal to the cost of the corresponding route selection sequence. Since $\mathfrak{P}$ is a shortest path, its length is minimum. This allow us to say that the sum of the initial route $P(1)$'s setup cost and all transition costs between two epochs given by $\mathfrak{P}$ is minimum. A node deployment and the communication range (scenario) define graph $\mathbb{G}$'s vertices and weights, while arcs in $\mathbb{G}$ depend on the PU activity process. Therefore, PU activity processes Υ and Υ' induce graphs $\mathbb{G}_\Upsilon$ and $\mathbb{G}_{\Upsilon'}$, respectively. Since Υ and Υ' are equivalent from epochs e_0 to $|E|$, the subgraph of $\mathbb{G}_\Upsilon$ involving only vertices in $\{P_i(e) : e \geq e_0\} \cup P_0''$ is equal to the same subgraph of $\mathbb{G}_{\Upsilon'}$. This common subgraph is denoted by $\mathbb{G}^{[e_0,|E|]}$. Given two subpaths, $\mathfrak{X}_\Upsilon \in \mathbb{G}_\Upsilon$ and $\mathfrak{X}_{\Upsilon'} \in \mathbb{G}_{\Upsilon'}$, of two shortest paths going through $\mathbb{G}^{[e_0,|E|]}$, for hypotheses, they cross, actually start, in $P_x(e_0)$, where $P_x = P_\Upsilon(e_0) = P_{\Upsilon'}(e_0)$, and end in P_0''. Since they are subpaths of shortest paths on a common subgraph, the length of $\mathfrak{X}_\Upsilon$ from $P_x(e_0)$ to P_0'' must be equal to the length of $\mathfrak{X}_{\Upsilon'}$ from $P_x(e_0)$ to P_0'' and be minimum. Therefore, adding the costs on arcs of each subpath, we have

$$\sum_{i=e_0}^{|\mathcal{E}|-1} C[P_\Upsilon(e), P_\Upsilon(e+1)] = \sum_{i=e_0}^{|\mathcal{E}|-1} C[P_{\Upsilon'}(e), P_{\Upsilon'}(e+1)]$$

and is minimum. □

Corollary V.2 Proof: Suppose two PU activity processes Υ and Υ' are equivalent from epoch e_0 to epoch $|\mathcal{E}|$, but not necessarily equivalent for epochs $e < e_0$. Furthermore, let the optimum routes selected under the two PU activity processes at epoch e_0 be equal, i.e., $(P_\Upsilon(e_0) = P_{\Upsilon'}(e_0))$. Let us proceed by contradiction and assume that the future route selection depends on selections made at epochs $e < e_0$. This means that routes selected after epoch e_0 in the two optimum sequences $\{P_\Upsilon(e)\}, e \in \mathcal{E}$ and $\{P_{\Upsilon'}(e)\}, e \in \mathcal{E}$ can be different, in particular, the sums of transition costs of $\{P_\Upsilon(e)\}_{e\in\{e_0..|\mathcal{E}|\}}$ and $\{P_{\Upsilon'}(e)\}_{e\in\{e_0..|\mathcal{E}|\}}$ can be different. This contradicts Theorem V.1, thus future route selections cannot depend on the past. □

REFERENCES

[1] FCC, Washington, DC, "Spectrum policy task force report," FCC 02-155, 2002.

[2] J. Mitola, "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, KTH, Stockholm, Sweden, 2000.

[3] I. Akyildiz, M. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 40–48, Apr. 2008.

[4] I. Akyildiz, W. Lee, M. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Comput. Netw.*, vol. 50, pp. 2127–2159, 2006.

[5] W. Y. Lee and I. F. Akyildiz, "Optimal spectrum sensing framework for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3845–3857, Oct. 2008.

[6] G. Ghurumuruhan and Y. G. Li, "Cooperative spectrum sensing in cognitive radio: Part I: Two user networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2204–2213, Jun. 2007.

[7] Z. Ji and K. Liu, "Dynamic spectrum sharing: A game theoretical overview," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 88–94, May 2007.

[8] H. Kim and K. G. Shin, "Efficient discovery of spectrum opportunities with MAC-layer sensing in cognitive radio networks," *IEEE Trans. Mobile Comput.*, vol. 7, no. 5, pp. 533–545, May 2008.

[9] H. Su and X. Zhang, "Cross-layer based opportunistic MAC protocols for QoS provisionings over cognitive radio wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 118–129, Jan. 2008.

[10] M. Cesana, F. Cuomo, and E. Ekici, "Routing in cognitive radio networks: Challenges and solutions," *Ad Hoc Netw.*, vol. 9, no. 3, pp. 228–248, 2011.

[11] Y. Shi and Y. Hou, "A distributed optimization algorithm for multi-hop cognitive radio networks," in *Proc. IEEE INFOCOM*, 2008, pp. 1292–1300.

[12] M. Xie, W. Zhang, and K.-K. Wong, "A geometric approach to improve spectrum efficiency for cognitive relay networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 268–281, Jan. 2010.

[13] G. Cheng, W. Liu, Y. Li, and W. Cheng, "Spectrum aware on-demand routing in cognitive radio networks," in *Proc. IEEE DySPAN*, 2007, pp. 571–574.

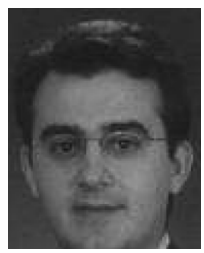
[14] Z. Yang, G. Cheng, W. Liu, W. Yuan, and W. Cheng, "Local coordination based routing and spectrum assignment in multi-hop cognitive radio networks," *Mobile Netw. Appl.*, vol. 13, no. 1–2, pp. 67–81, 2008.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[15] H.-P. Shiang and M. van der Schaar, "Distributed resource management in multihop cognitive radio networks for delay-sensitive transmission," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 941–953, Feb. 2009.
[16] I. Pefkianakis, S. Wong, and S. Lu, "SAMER: Spectrum aware mesh routing in cognitive radio networks," in *Proc. IEEE DySPAN*, 2008, pp. 1–5.
[17] L. Ding, T. Melodia, S. Batalama, J. Matyjas, and M. J. Medley, "Cross-layer routing and dynamic spectrum allocation in cognitive radio ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1969–1979, May 2010.
[18] G.-M. Zhu, I. Akyildiz, and G.-S. Kuo, "STOD-RP: A spectrum-tree based on demand routing protocol for multi-hop cognitive radio networks," in *Proc. IEEE GLOBECOM*, 2008.
[19] A. Abbagnale and F. Cuomo, "Gymkhana: A connectivity-based routing scheme for cognitive radio ad hoc networks," in *Proc. IEEE INFOCOM*, 2010.
[20] K. Chowdhury and M. Felice, "Search: A routing protocol for mobile cognitive radio ad-hoc networks," *Comput. Commun.*, vol. 32, no. 18, pp. 1983–1997, 2009.
[21] M. Ma and D. Tsang, "Joint spectrum sharing and fair routing in cognitive radio networks," in *Proc. IEEE CCNC*, 2008, pp. 978–982.
[22] Y. Shi and Y. Hou, "A distributed optimization algorithm for multi-hop cognitive radio networks," in *Proc. IEEE INFOCOM*, 2008, pp. 1292–1300.
[23] C. Xin, B. Xie, and C. Shen, "A novel layered graph model for topology formation and routing in dynamic spectrum access networks," in *Proc. IEEE DySPAN*, 2005, pp. 308–317.
[24] E. Amaldi, A. Capone, M. Cesana, I. Filippini, and F. Malucelli, "Optimization models and methods for planning wireless mesh networks," *Comput. Netw.*, vol. 52, no. 11, pp. 2159–2171, 2008.
[25] L. Wolsey, *Integer Programming*. New York: Wiley, 1998.
[26] J. Renegar, "Linear programming, complexity theory and elementary functional analysis," *Math. Program.*, vol. 70, no. 3, pp. 279–351, 1995.
[27] S. Even, A. Itai, and A. Shamir, "On the complexity of timetable and multicommodity flow problems," *SIAM J. Comput.*, vol. 5, no. 4, pp. 691–703, 1976.
[28] M. Fisher, "The Lagrangian relaxation method for solving integer programming problems," *INFORMS Manag. Sci.*, vol. 50, no. 12, pp. 1861–1871, 2004.
[29] FCC, Washington, DC, "Unlicensed operation in the TV broadcast bands," FCC 08-260, 2008.
[30] Y. Yuan, P. Bahl, R. Chandra, P. Chou, J. Ferrell, T. Moscibroda, S. Narlanka, and Y. Wu, "KNOWS: Cognitive radio networks over white spaces," in *Proc. IEEE DySPAN*, 2007, pp. 416–427.
[31] C. Perkins, E. Belding-Royer, and S. Das, "RFC3561—Ad hoc on-demand distance vector (AODV) routing," IETF Network Working Group, 2003.

Ilario Filippini received the B.S. and M.S. degrees in telecommunication engineering and Ph.D. degree in information engineering from the Politecnico di Milano, Milan, Italy, in 2003, 2005, and 2009, respectively.

From February to August 2008, he worked as a Visiting Researcher with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus. He is currently an Assistant Professor with the Electronics and Information Department, Politecnico di Milano. His research activities include networking and optimization issues, in particular wireless multihop network planning, optimization, and protocol design.



Eylem Ekici (S'99–M'02–SM'11) received the B.S. and M.S. degrees in computer engineering from Bogaziçi University, Istanbul, Turkey, in 1997 and 1998, respectively, and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, in 2002.

Currently, he is an Associate Professor with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus. His current research interests include cognitive radio networks, wireless sensor networks, vehicular communication systems, and nano communication systems with a focus on modeling, optimization, resource management, and analysis of network architectures and protocols.

Dr. Ekici is an Associate Editor of the IEEE/ACM TRANSACTIONS ON NETWORKING, *Computer Networks*, and *Mobile Computing and Communications Review*.

Matteo Cesana received the M.S. degree in telecommunications engineering and Ph.D. degree in information engineering from the Politecnico di Milano, Milan, Italy, in 2000 and 2004, respectively.

From 2002 to 2003, he worked as a Visiting Researcher with the Computer Science Department, University of California, Los Angeles (UCLA). He is now an Assistant Professor with the Electronics and Information Department, Politecnico di Milano. He is an Associate Editor of *Ad Hoc Networks*. His research activities are in the field of performance evaluation of cellular systems, ad hoc networks protocol design, and evaluation and wireless networks optimization.