Ad Hoc Networks 9 (2011) 228–248

Contents lists available at ScienceDirect

Ad Hoc Networks

journal homepage: www.elsevier.com/locate/adhoc

Routing in cognitive radio networks: Challenges and solutions

Matteo Cesana [a,*], Francesca Cuomo [b], Eylem Ekici [c]

[a] ANT Lab, Dipartimento di Elettronica e Informazione, Politecnco di Milano, Piazza L. da Vinci 32, 20133 Milan, Italy
[b] INFOCOM Dpt., University of Rome "La Sapienza", Via Eudossiana 18, 00184 Rome, Italy
[c] Department of Electrical and Computer Engineering, The Ohio State University, 205 Dreese Laboratory, 2015 Neil Avenue, Columbus, OH 43210, United States

ARTICLE INFO

Article history:
Received 11 June 2010
Accepted 30 June 2010
Available online 6 July 2010

Keywords:
Cognitive radio networks
Routing protocols
Routing metrics
Cross-layering

ABSTRACT

Cognitive radio networks (CRNs) are composed of cognitive, spectrum-agile devices capable of changing their configurations on the fly based on the spectral environment. This capability opens up the possibility of designing flexible and dynamic spectrum access strategies with the purpose of opportunistically reusing portions of the spectrum temporarily vacated by licensed primary users. On the other hand, the flexibility in the spectrum access phase comes with an increased complexity in the design of communication protocols at different layers. This work focuses on the problem of designing effective routing solutions for multi-hop CRNs, which is a focal issue to fully unleash the potentials of the cognitive networking paradigm. We provide an extensive overview of the research in the field of routing for CRNs, clearly differentiating two main categories: approaches based on a full spectrum knowledge, and approaches that consider only local spectrum knowledge obtained via distributed procedures and protocols. In each category we describe and comment on proposed design methodologies, routing metrics and practical implementation issues. Finally, possible future research directions are also proposed.

1. Introduction

Current wireless networks are regulated by governmental agencies mainly according to a fixed spectrum assignment policy. Licenses are granted the rights for the use of various, often relatively small, frequency bands on a long term basis over vast geographical regions. In recent years, the huge success of wireless applications has caused an exponential increase in requests to regulatory authorities for spectrum allocation. In parallel, the use of wireless technologies operating in unlicensed bands, especially in the ISM band, has been prolific with a wide range of applications developed in different fields (e.g, WLANs, mesh networks, personal area networks, body area networks, sensor networks, etc.), which caused overcrowding in this band. On the other hand, the usage of licensed spectrum is quite uneven and depends heavily on the specific wireless technologies, their market penetration, and the commercial success of the operators to which the frequencies have been assigned. Recent studies by the Federal Communications Commission (FCC) highlight 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% and 85% [1].

To address this situation, the notion of dynamic spectrum access (DSA) has been proposed. With DSA, unlicensed users may use licensed spectrum bands opportunistically in a dynamic and non-interfering manner. From a technical perspective, this is possible thanks to the recent advancements in the field of software-defined radios (SDRs). SDRs 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 limited delay [2,3]. Resulting so-called Cognitive Radio (CR) transceivers have the capability of completely changing their transmitter parameters (operating spectrum, modulation, transmission power, and

* Corresponding author.
E-mail addresses: cesana@elet.polimi.it (M. Cesana), francesca.cuomo@uniroma1.it (F. Cuomo), ekici@ece.osu.edu (E. Ekici).

1570-8705/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.adhoc.2010.06.009

communication technology) based on interactions with the surrounding spectral environment. They can sense a wide spectrum range, dynamically identify currently unused spectrum blocks for data communications, and intelligently access the unoccupied spectrum called Spectrum Opportunities (SOP) [4].

Devices with cognitive capabilities can be networked to create Cognitive Radio Networks (CRNs), which are recently gaining momentum as viable architectural solutions to address the limited spectrum availability and the inefficiency in the spectrum usage [5]. The most general scenario of CRNs distinguishes two types of users sharing a common spectrum portion with different rules: Primary (or licensed) Users (PUs) have priority in spectrum utilization within the band they have licensed, and Secondary Users (SUs) must access the spectrum in a non-intrusive manner. Primary Users use traditional wireless communication systems with static spectrum allocation. Secondary Users are equipped with CRs and exploit Spectrum Opportunities (SOPs) to sustain their communication activities without interfering with PU transmissions.

Most of the research on CRNs to date has focused on single-hop scenarios, tackling PHYsical (PHY) layer and/or Medium Access Control (MAC) layer issues, including the definition of effective spectrum sensing, spectrum decision and spectrum sharing techniques [6,7]. Only very recently, the research community has started realizing the potentials of multi-hop CRNs which can open up new and unexplored service possibilities enabling a wide range of pervasive communication applications. Indeed, the cognitive paradigm can be applied to different scenarios of multi-hop wireless networks including Cognitive Wireless Mesh Networks featuring a semi-static network infrastructure [8], and Cognitive radio Ad Hoc Networks (CRAHNs) characterized by a completely self-configuring architecture, composed of CR users which communicate with each other in a peer to peer fashion through ad hoc connections [9]. To fully unleash the potentials of such networking paradigms, new challenges must be addressed and solved. In particular, effective routing solutions must be integrated into the work already carried out on the lower layers (PHY/MAC), while accounting for the unique properties of the cognitive environment.

In the remainder of the paper, we focus on the issues related to the design and maintenance of routes in multi-hop CRNs. The purpose of this work is twofold: First, we aim at dissecting the most common approaches to routing in CRNs, clearly highlighting their design rationale, and their strengths/drawbacks. Then, by leveraging the literature in the field, we comment on possible future research directions.

2. Routing challenges in multi-hop CRNs

The reference network model reported in Fig. 1 features secondary devices which share different spectrum bands (or SOPs) with primary users. Several spectrum bands ($1,\ldots,M$) may exist with different capacities C_1, C_2, C_M, and the SUs may have different views of the available spectrum bands due to inherent locality of the spectrum sensing process. Typically the PUs are assumed motionless while the SUs may vary their position before and during a transmission.

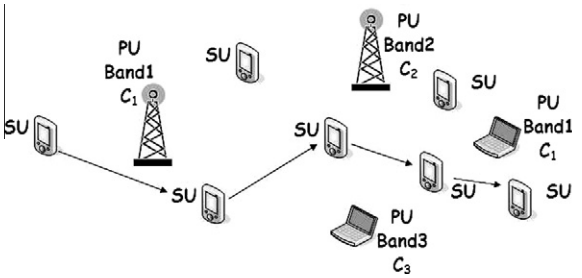


Fig. 1. Information routing in multi-hop CRNs.

In this scenario, *the* ***problem of routing*** *in multi-hop CRNs targets the creation and the maintenance of wireless multi-hop paths among SUs by deciding both the relay nodes and the spectrum to be used on each link of the path.*

Such problem exhibits similarities with routing in multi-channel, multi-hop ad hoc networks and mesh networks, but with the additional challenge of having to deal with the simultaneous transmissions of the PUs which dynamically change the SOPs availability.

In a nutshell, the main challenges for routing information throughout multi-hop CRNs include:

- *Challenge 1*: the **spectrum-awareness;** designing efficient routing solutions for multi-hop CRNs requires a tight coupling between the routing module(s) and the spectrum management functionalities such that the routing module(s) can be continuously aware of the surrounding physical environment to take more accurate decisions. Within this field, three scenarios may be possible:
  - the information on the spectrum occupancy is provided to the routing engine by external entities (e.g., SUs may have access to a data base of white spaces of TV towers [10]);
  - the information on spectrum occupancy is to be gathered locally by each SU through local and distributed sensing mechanisms;
  - a mixture of the previous two.

  In any case, any routing solution designed for multi-hop CRNs must be highly coupled to the entire *cognitive cycle* of spectrum management [4].
- *Challenge 2:* the **set up of "quality" routes** in dynamic variable environment; the very same concept of "route quality" is to be re-defined under CRN scenario. Indeed, the actual topology of multi-hop CRNs is highly influenced by PUs' behavior, and classical ways of measuring/assessing the quality of end-to-end routes (nominal bandwidth, throughput, delay, energy efficiency and fairness) should be coupled with novel measures on path stability, spectrum availability/PU presence. As an example, if the PU activity is moderate-to-low, the topology of the secondary users' network is almost static, and classical routing metrics adopted for wireless mesh networks could be leveraged.

On the other hand, if PUs become active very frequently, routing techniques for disconnected networks could be favorable [11].

- *Challenge 3:* the **route maintenance/reparation**; the sudden appearance of a PU in a given location may render a given channel unusable in a given area, thus resulting in unpredictable route failures, which may require frequent path rerouting either in terms of nodes or used channels. In this scenario, effective signalling procedures are required to restore "broken" paths with minimal effect on the perceived quality.

In the following sections of this paper, several routing solutions are commented keeping an eye on the aforementioned three main challenges. In Fig. 2, we broadly categorize the proposed solutions into two main classes depending on the assumptions taken on the issue of spectrum-awareness (*Challenge 1*):

- *full spectrum knowledge*;
- *local spectrum knowledge*.

In the former case, a spectrum occupancy map is available to the network nodes, or to a central control entity, which could be represented by the centrally-maintained spectrum data bases recently promoted by the FCC to indicate over time and space the channel availabilities [10] in the spectrum below 900 MHz and around 3 GHz. The considered architectural model is a static cognitive multi-hop network where the spectrum availability between any given node pair is known.

The routing approaches building on this assumption leverage theoretical tools to design efficient routes, differentiating on the basis of which kind of theoretical tool is used to steer the route design. A first class encompasses all solutions based on a graph abstraction of the cognitive radio network. The second sub-class instead employs mathematical programming tools to model and design flows along the cognitive multi-hop network. Although these approaches are often based on a centralized computation of the routing paths, their relevance is in the fact that they provide upper bounds and benchmarks for the routing performance.

On the other hand, routing schemes based on local spectrum knowledge include all those solutions where information on spectrum availability is locally "constructed" at each SU through distributed protocols. Thus, the routing module is tightly coupled to the spectrum management functionalities. Indeed, besides the computation of the routing paths, the routing module should be able to acquire network state information, such as currently available frequencies for communication and other locally available data, and exchange them with the other network nodes. While the network state in traditional ad hoc networks is primarily a function of node mobility and traffic carried in the network, network state in multi-hop CRNs is also influenced by primary user activity. How this activity is and which are the suitable models to represent it are key components for the routing design.

A further classification of the proposals in the local spectrum knowledged family can be based on the specific measure of the route "quality" used to set up "quality routes" (*Challenge 2*). Four classes can be broadly recognized: form left to right in Fig. 2, we have routing solutions aiming at controlling the interference the multi-hop CRNs create, delay-based and throughput-based routing schemes where the routing module targets the minimization of the end-to-end delay and the maximization of the achievable throughput, respectively; and finally, those solutions where the quality of the paths is strictly coupled to its availability over time and to its stability (*Challenge 3*).

In summary, as we move from left to right in the *local spectrum knowledge* sub-classes, the routing solutions feature increasing spectrum awareness of the dynamic scenario created by the intermittent PUs which can affects seriously the service offered by the multi-hop CRNs. In these cases, the channel properties such as the holding

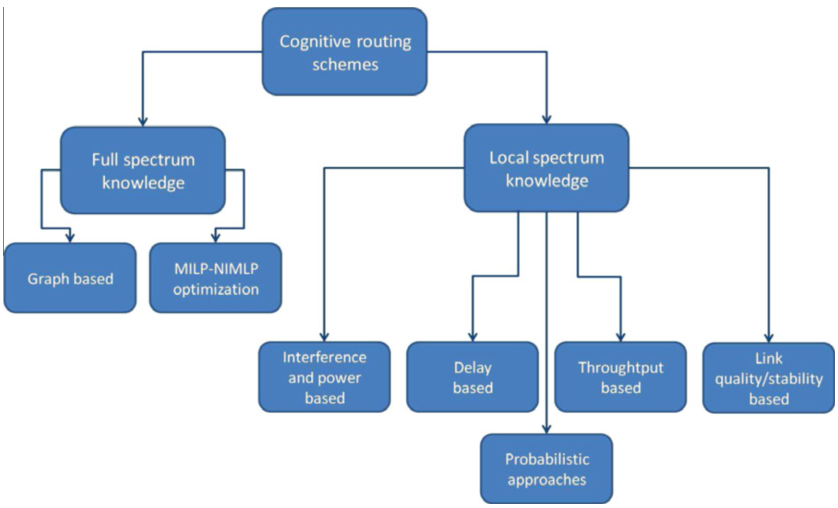


Fig. 2. Classification of cognitive routing schemes.

time, the available capacity and more generally the statistics of channel conditions, are considered in the proposed routing solution.

One last family of protocols operating with local spectrum knowledged leverages probabilistic routing approaches where SUs opportunistically transmit over any available spectrum band during the short period of the latter existence. Such approaches are feasible and useful in those cases where the surrounding PUs have short idle periods and, as a consequence, the availability of the corresponding SOPs is limited in time [11].

3. Routing schemes based on full spectrum knowledge

As already mentioned in the previous section, the FCC 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 data bases indicating over time and space the channel availabilities [10]. Before sending or receiving data, cognitive opportunistic devices will be required to access these databases to determine available channels.

Under this scenario, the central availability of up-to-date information on spectrum occupancy completely decouples the spectrum assessment modules (sensing, sharing) form the routing decisions/policies which can be locally optimized. This section comments on those routing approaches which start off from the assumption of full knowledge on the spectrum occupancy, further proposing analytical tools to optimize/steer the routing decisions.

3.1. Graph-based routing approaches

Route design in classical wired/wireless networks has been tackled widely resorting to graph-theoretic tools. Graph theory provides extremely effective methodologies to model the multi-hop behavior of telecommunication networks, as well as powerful and flexible algorithms to compute multi-hop routes. The general approach to designing routes in wireless multi-hop networks consists of two phases: *graph abstraction* and *route calculation*. *Graph abstraction* phase refers to the generation of a logical graph representing the physical network topology. The outcome of this phase is the graph structure $G = (N, V, f(V))$, where N is the number of nodes, V is the number of edges, and $f(V)$ the function which allows to assign a weight to each edge of the graph[1]. *Route calculation* generally deals with defining/designing a path in the graph connecting source–destination pairs. Classical approaches to route calculation widely used in wired/wireless network scenarios often resort to mathematical programming tools to model and design flows along multi-hop networks.

3.1.1. Routing through layered-graphs

The very same two-phase approach to route design has been leveraged also for multi-hop CRNs. The authors of [12,13] propose a comprehensive framework to address channel assignment and routing jointly in semi-static multi-hop CRNs. In these works, the PU dynamics are assumed to be low enough such that the channel assignment and the routing among SUs can be statically designed. The authors further focus on the case where cognitive devices are equipped with a single half-duplex cognitive radio transceiver, which can be tuned to M available spectrum bands or channels. The proposed framework is based on the creation of a *layered graph* which features a number of layers equal to the number of available channels. Each secondary user device is represented in the layered graph with a node, A, and M additional subnodes, $A_1, A_2, \ldots, A_M$, one for each available channel.

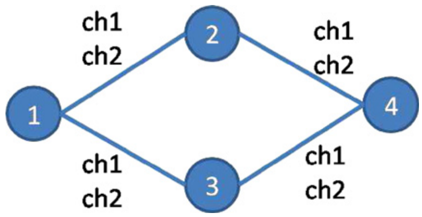
The edges of the layered graph can be of three types: *access, horizontal, and vertical*. Access edges connect each node with all the corresponding subnodes. Horizontal edges between pairs of subnodes belonging to the same logical layer are added to the graph if the two corresponding secondary devices can be tuned to the corresponding channel. Vertical edges connect subnodes of different layers of a single secondary device, and represent the capability for a secondary device to switch from one channel to another to forward incoming traffic. As an example, Fig. 3a reports a simple fournode network topology where all four devices in the network can be tuned to channels $ch1$ and $ch2$. The corresponding layered graph architecture is shown in Fig. 3b. The edges laying on the two horizontal planes representing the two available channels ($ch1$, $ch2$) are horizontal edges, dashed vertical edges are vertical edges, and small dashed ones represent access edges.

As for the edge weights (the function $f(V)$), the weight of horizontal edges should endorse the specific quality of the wireless link, like bandwidth, link availability, link load, etc., whereas the vertical edges could be weighted accounting for different quality parameters including: the cost for switching between channels, or the improvement in the signal to noise ratio when obtained by switching between the two given channels.

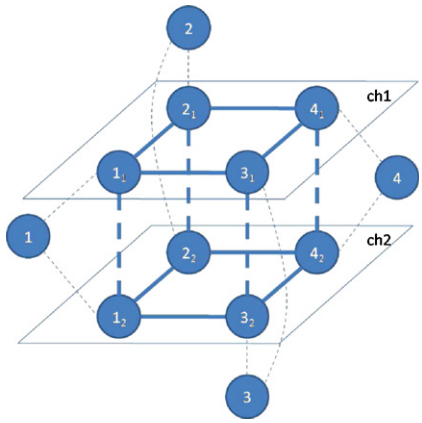
The proposed layered graph is a rather general framework which can be combined with different routing metrics. Further modifications to the layered graph can be introduced to account for specific requirements, such as the need to route outgoing traffic over different channels than the incoming one, or to account for costs associated with specific nodes.

Once the graph is created and the metrics are assigned to each edge, the joint channel assignment/routing problem in the original network topology can be solved by finding multi-hop paths between source–destination couples in the corresponding layered graph. In [13], the authors focus on the case where the metrics for the horizontal links are proportional to traffic load and interference. Here, a centralized heuristic algorithm is proposed based on the calculation of shortest paths in the layered graph. The proposed path-centric route calculation algorithm works iteratively by routing one source–destination flow at a time. Once a flow is routed, a new layered graph is calculated from the previous one by eliminating all unused incoming horizontal/vertical edges and re-calculating the weights assigned to the remaining edges to account for the routed traffic load.

[1] Weights are assigned to reflect the specific quality metrics to be assigned to a wireless link.



(a) Network Topology.



(b) Layered Graph.

Fig. 3. Layered-graph creation.

The proposed layered graph framework is indeed useful to jointly model channel assignment and routing in semi-static multi-hop CRNs, where the topology variability dynamics is low. On the down side, the proposed path-centric routing approach is fundamentally centralized requiring network-wide signalling support to generate the layered-graph. Moreover, the proposed iterative algorithm is suboptimal being based on a greedy approach. Finally, resorting to iterative path computation over graph abstractions may not scale well as the network dimensions increase.

3.1.2. Routing through colored-graphs

A similar approach based on graph structures is proposed in [14], where a *colored graph* is used to represent the network topology. The colored graph $G_c = (N_c, V_c)$, where N_c is the vertex set (one vertex for each network device), and V_c is the edge set. Two vertices in the colored graph may be connected by a number of edges up to M, where M is the number of channels (colors) available for transmission on the specific link. Referring back to Fig. 3a, Fig. 4 corresponds to the colored graph abstracting the physical network topology. The route calculation algorithm follows the same rationale as the one proposed in [13], leveraging a centralized iterative approach. The shortest path is calculated for one source–destination pair on the colored graph resorting to metrics capturing the inter-link interference (the number of adjacent edges on the path using the same color). Once a flow has been routed, the colored graph is updated by re-setting the edge weights, then iterating for all the remaining traffic flows. This approach obviously shares the very same drawbacks as the previously commented one. Namely, the proposed solution approach is centralized and heuristic, meaning that it may lead to suboptimal routing instances.

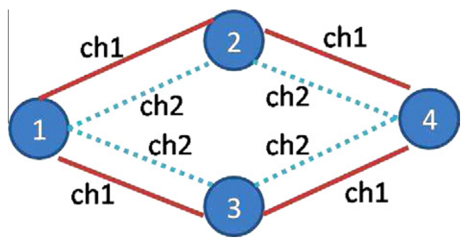


Fig. 4. Colored-graph creation.

3.1.3. Routing and spectrum selection through conflict-graphs

Route and spectrum selection in networks with single transceiver half duplex cognitive radios are addressed also in [15]. Different from the aforementioned pieces of work, the proposed solution decouples routing and channel (spectrum) assignment. In [15], given the network topology, all available routes between source–destination pairs are enumerated, and for each route all available channel assignment patterns are considered. The "best" combination of routing/channel assignment is derived by running a centralized algorithm on a "conflict graph". Each wireless link in the network maps to a vertex in the conflict graph. An edge is defined between two vertices if the corresponding wireless links cannot be active at the same time. The conflict graph is used to derive a conflict-free channel assignment by resorting to a heuristic algorithm to calculate the maximum independent set (or maximum clique). As in the two previous cases, the proposed approach is centralized and assumes full knowledge of the network topology (available spectrum bands, neighboring nodes, etc.). Moreover, the problem of defining the most efficient conflict-free scheduling can be reduced to a problem of calculating the maximum independent set on a properly defined "conflict graph", which is known to be NP-Hard.

3.2. Optimization approaches to routing design

As network topology and spectrum occupation are known *a priori*, optimization models and algorithms can be used to optimally design routes in multi-hop CRNs.

In [16,17], Hou et al. focus on the problem of designing efficient spectrum sharing techniques for multi-hop CRNs.

To this extent, they introduce a Mixed Integer Non-Linear Programming (MINLP) formulation whose objective is to maximize the spectrum reuse factor throughout the network, or equivalently, to minimize the overall bandwidth usage throughout the network. The proposed formulation captures all major aspects of multi-hop wireless networking, i.e., link capacity, interference, and routing.

- *Link capacity:* the formulation forces the total traffic flow not to exceed the capacity of the wireless link it is traveling through. Shannon's law is used to define the link capacity given the nominal bandwidth and the signal to interference ratio. Namely, the capacity of link (i,j) operating on the sub-band m is given by:

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

being W^m the bandwidth of sub-band m, g_{ij} the propagation gain of link (i,j), Q the power spectral density in transmission, and η the Gaussian ambient noise density.

- *Interference* is captured leveraging the concept of *interference range*, R_T, defined as

$$R_T = (Q/Q_T)^{1/\eta},$$

where Q_T is the threshold power spectral density guaranteeing correct reception. Two secondary devices falling within the interference range of one another do interfere, and cannot use the same sub-band for transmission.

- *Routing:* flow balance constraints at each node are used to capture traffic routing in the MINLP formulation; for each source–destination traffic flow, for every node in the network other than source and destination, the flow balance constraint forces the incoming flow to a node to be equal to the outgoing flow; source and destination are respectively flow creation and flow sink points.

As a byproduct, the MINLP formulation ensures the existence of a multi-hop path between any source–destination pair. The use of flow balance constraints to design routes implicitly allows the creation of split routing paths for each source–destination flow; that is, the traffic flow of a source–destination pair may be routed along multiple multi-hop paths. This has the obvious advantage of robustness, but, on the other hand, it is much harder to be implemented in practical packet-switched networks.

As for the solution approach, the authors start off by solving a linear relaxation of the MINLP formulation. Namely, the binary variables which bind each user to transmit over a given sub-band are relaxed to linear values. The resulting formulation is linear (Linear programming, LP), thus it can be easily and effectively solved in polynomial time. The result obtained solving the LP relaxed version of the original problem provides a lower bound on the overall bandwidth usage throughout the network.

To complete the characterization of the MINLP solution, the authors further propose a centralized heuristic based on the concept of "sequential fixing". In a nutshell, the algorithm works iteratively and features two operation phases:

1. set up and solve the relaxed LP version of the original problem as done to obtain the lower bound;
2. sort the assignment variables in descending order;
3. set to 1 (fix) the largest variable in the list, and set to 0 the remaining variables referring to the same user;
4. solve the new LP formulation of the problem with the variables fixed at step 3.

The four steps above are iterated until all the assignment variables are fixed. The authors further propose a technique to speed-up the iterative heuristic, by fixing at each step group of variables.

To summarize, the strengths of works in [16,17] are that the proposed framework is effective in capturing many aspects of networking over multi-hop networks and that the proposed solutions approaches are proved to provide nearly optimal solutions to the joint scheduling/routing problem for multi-hop CRNs. On the down side the proposed scheduling/routing algorithm has to run at a central entity which has perfect knowledge of the network topology (presence, position and traffic pattern of the primary users, presence and position of the secondary users). Moreover, traffic splitting is allowed throughout the secondary network. As expressed above, the assumption of having split traffic between secondary users may be unfeasible in practical secondary networks. Finally, the interference is modeled through the concept of interference range, which automatically excludes effects related to interference accumulation from multiple transmitters far away and the definition of link capacity is based on the assumption that the surrounding interference is Gaussian.

Mathematical programming is leveraged also in [18], where a Mixed Integer Linear Programming (MILP) formulation is derived for the problem of achieving throughput optimal routing and scheduling for secondary transmissions. The objective function aims at maximizing the achievable rate of source–destination pairs, under the very same interference, capacity and routing constraints as defined above. The authors directly use the formulation to design route/channel assignment patterns for small-to-medium size network scenarios by resorting to commercial solvers.

4. Routing schemes based on local spectrum knowledge

This section overviews those routing solutions where the retrieval of information on spectrum occupancy is performed in a distributed way, and, similarly to classical ad hoc networks, distributed approaches are introduced to make local radio resource management decisions on partial information about the network state. In multi-hop CRNs, such functionality is crucial since the local spectrum conditions acquired via radio sensing can be highly variable in time and space. The presented solutions are categorized according to the specific metric used to assess route quality.

4.1. Interference and power based solutions

Routing solutions of this kind mainly leverage routing metrics based on consumed power to perform transmission and/or the perceived/generated interference along a multi-hop path through secondary users.

4.1.1. Minimum power routing

As an example, the main objective of the work of Pyo and Hasegawa [19] is to discover minimum weight paths in cognitive wireless ad hoc networks. A detailed system-level picture is presented where the communication system is partitioned into operating system and communication system. The operating system is responsible for selecting the wireless communication interface to be used at a given time. Different interfaces are used to access various Wireless Systems (WS) such as cellular (e.g., CDMA, TDMA, FDMA) or WLAN (i.e., IEEE 802.11 b/g). Each of the interfaces is associated with a different communication range, as well. The use of a Common Control Channel (CCC) plays a central role in the work. A dedicated interface, referred to as Common Link Control Radio (CLCR) is used for communication between CR terminals to sustain cognitive radio network related functions. The two main functions using CLCR interface are the neighbor discovery and path discovery and establishment. To discover a large neighborhood, CLCR uses a high transmission power to reach out all the potential neighbors. Nodes share with each other their connectivity over different radio interfaces when they exchange messages through the CLCR. The signaling to establish paths between two end points also happens over the CLCR.

The weight of a link is defined as a function of the transmission power of the different WSs an SU may use to communicate with a neighbor node. The paper assumes a free space propagation model for the transmission power of $WS[i]$ which increases with the distance as follows:

$$P_{TX_{WS[i]}} = P_{RX_{WS[i]}} \cdot \left(\frac{4\pi d}{\lambda_{WS[i]}}\right)^2, \tag{1}$$

where $i = 1, \ldots, W$ is the indices of W WSs available at a terminal, $P_{TX_{WS[i]}}$ is the transmission power of $WS[i]$, $P_{RX_{WS[i]}}$ is the received signal power at a receiver, $\lambda_{WS[i]}$ is the wavelength of $WS[i]$, and d is the distance between the transmitter and the receiver.

A routing weight based on the required power to reach a specific destination is associated with different WSs. The proposed routing protocol locally finds the path to minimize the routing weight between a source and a destination. The route discovery procedure is very similar to link state routing algorithms where this newly introduced weight is used. The model does not take into account the primary users, their behavior, or the interference caused by/to other CR nodes. However, such information is implicitly incorporated into routing decisions during neighbor discovery stage. This work introduces a very nicely outlined system model based on multiple interfaces. The performance of the proposed system is highly dependent on the neighbor discovery procedure and its refresh rates as there are no other maintenance or recovery procedures defined in the routing protocol to react to PU activity. Furthermore, the power-level based cost metric is not sufficient to address challenges of multi-hop cognitive radio networks.

4.1.2. Bandwidth footprint minimization

The distributed algorithm presented in [20] addresses the scheduling, power control, and routing problems simultaneously. The routing module is based on the notion of the Bandwidth Footprint Product (BFP). The "footprint" refers the interference area of a node for a given transmission power. Since each node in the network uses a number of bands for transmission and each band has a certain footprint corresponding to its transmission power, the objective is to minimize network-wide BFP, which is the sum of BFPs for all nodes in the network.

The proposed approach increases session rates with an iterative procedure. A Conservative Iterative Procedure (CIP) and an Aggressive Iterative Procedure (AIP) have been proposed to decide on the route selection, link scheduling, and the power allocation. CIP increases the rate of a session with the smallest scaling factor so as not to affect other sessions. On the other hand, AIP increases the rate of a session by allowing a limited decrease in other sessions' rates. Both CIP and AIP are composed of modules to determine the target decisions.

Authors base their routing module on an Incremental Link Cost (ILC) for pushing more data rate onto a link defined as the incremental BFP per additional data rate, which only requires local information and can be computed in a distributed manner. ILC is considered zero if a frequency band already has excess capacity. Routing module in CIP finds the session l with the minimum scaling factor $K(l)$, for which the rate can be increased without affecting other flows. It further distributes the available capacity to flows starting with the session with the smallest scaling factor. On the other hand, a rate scaling $K(l)$ is done under AIP at the expense of other sessions, making sure that the scaling of the affected sessions does not fall below $K(l)$. Session rate scaling of AIP aims at redistribution of resources to improve the overall rate. Both procedures then utilize a so-called Minimalist Scheduling procedure to assign frequency bands to sessions along the decided paths. For CIP, this assignment is performed only if nodes have reached their maximum transmission power limits on a given band. In such cases, a new frequency band is allocated to a session, and the information is propagated to upstream and downstream hops along the path of the session so that adjustments can be made to resource allocation throughout the path. AIP's minimalist scheduling algorithms is similar to that of CIP except for the case where the capacity is reached for a link. AIP opens a new channel to use only of the reducing the rate of a session is not possible. These decisions are then further refined in the Power Control/Scheduling module in both procedures. The main approach here is first to assign available capacity on a channel. If this fails, then transmission power is increased to increase the rate of the session. Finally, if this fails, then alternative channels are considered to migrate the session to achieve the target increase in session rate. The differences in implementation of power control/scheduling between the CIP and AIP is the flexibility of AIP to reduce allocations of existing flows.

The operation of the proposed algorithm is based on the iterative selection of sessions to scale as shown in Fig. 5. First, CIP is used to scale the rate of sessions in the net-

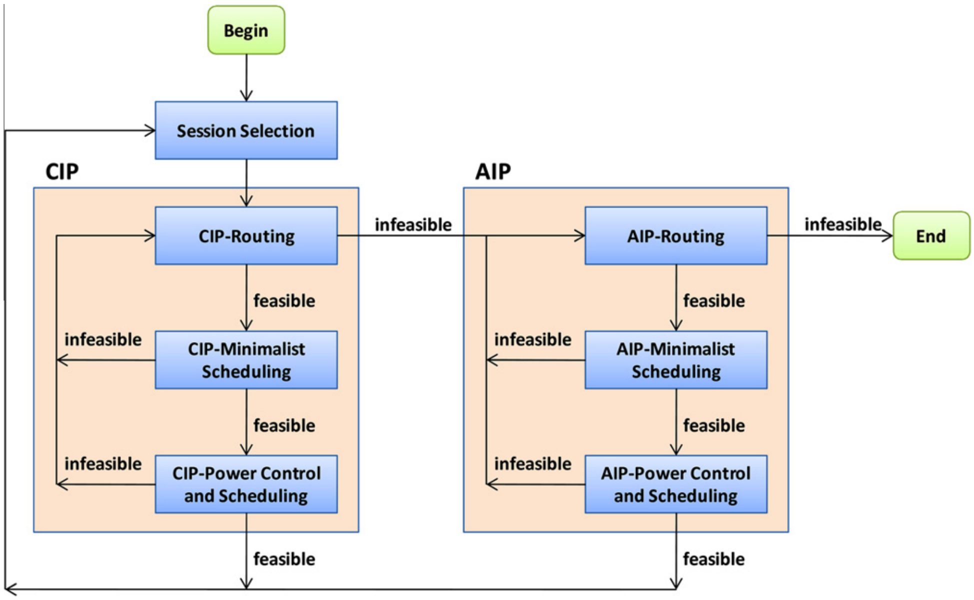


Fig. 5. CIP and AIP session management defined in [20].

work. After each successful iteration, a new session is selected and processed. When CIP can no longer find a session to increase the rate for, AIP takes over and reallocates assigned resources to improve the overall rate. The authors show that the results emerging from their iterative procedure are close to an upper bound of the MINLP formulation of the problem at hand. Since the procedure can be run in a distributed manner, the proposed algorithm carries desirable properties in implementation. However, as many other algorithms that focus on periods of constant PU activity, this algorithm also requires that spectral availability does not change throughout the algorithm operation. Moreover, the scheduling decisions are based on the abstraction of the wireless channel as a fixed capacity resource, which is clearly at odds with reality. The prior work on scheduling and power allocation problem reveals that it is a non-trivial task to make instantaneous decisions for resource allocation in wireless networks. In this work, however, power allocation and scheduling are presented as small steps in the entire optimization work. Consequently, the actual implementation complexity of the proposed algorithm is expected to be higher. Coupled with the dynamic resource availability in multi-hop CRNs, the proposed algorithm is more suitable for offline performance predictions than distributed resource allocation.

4.1.3. Controlled interference routing

Interference constraints are at the basis of the work in [21] where the authors analyze the tradeoff between single-hop and multi-hop transmission for SUs constrained by the interference level that PUs can tolerate. Authors analyze the potentialities of a multi-hop relaying by deriving the geometric conditions under which a SU is admitted into a spectrum occupied by a PU. On the basis of these geometric results authors propose two routing methods termed Nearest-Neighbor Routing (NNR) and Farthest-Neighbor Routing (FNR). In the NNR scheme a transmitter attempts to find the nearest-neighbor inside a sector of a radius D_{max} depending on the considered QoS parameters and the positioning parameters of the SUs and PUs. As opposite to NNR, the FNR scheme searches for the farthest-neighbor within the range D_{max}. Performance results show that FNR achieves a better end to end channel utilization and reliability while NNR has a better energy efficiency. Another result of the paper is the computation of the performance gain of a multi-hop CRN with relaying over a multi-hop CRN without relaying when parameters like the channel utilization, the energy efficiency and the delay are considered. Although the proposed routing schemes are mainly based on a static geometric view of the network, without considering any dynamics in the spectrum occupation, and the considered QoS parameters are relevant to the transmission quality at the physical level (SINR and channel outage probability) this paper identifies basic principles for selecting a multi-hop routing in a CRN.

4.2. Delay based solutions

The quality of routing solutions can also be measured in terms of delays to establish and maintain multi-hop routes and to send traffic through the very same routes. Besides "classical" delay components for transmitting information in wireless networks, novel components related to spectrum mobility (channel switching, link switching) should be accounted for in multi-hop CRNs. Delay-aware routing metrics are proposed in [22–25], which consider different delay components including:

1. the *Switching Delay* that occurs when a node in a path switches from one frequency band to another;
2. the *Medium Access Delay* based on the MAC access schemes used in a given frequency band;

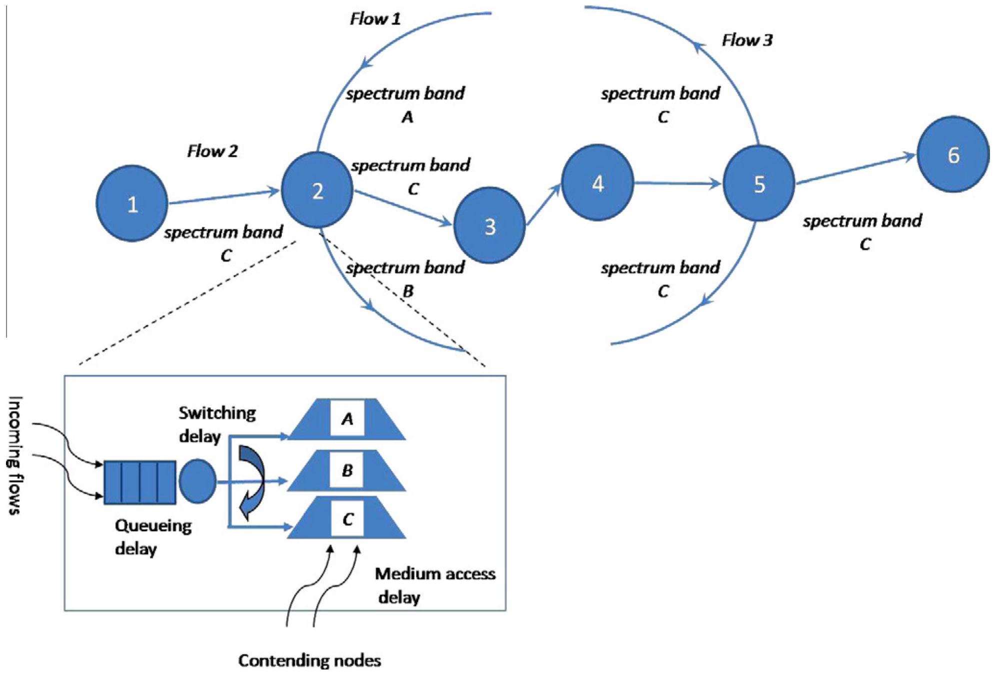


Fig. 6. Delay components in a CR node.

3. *Queueing Delay* based on the output transmission capacity of a node on a given frequency band.

Fig. 6 shows an example of these three delay components at a CR node. Node 2 relays flow 1 by receiving data on frequency band *A* and transmitting data on frequency band *B*. It uses the same spectrum band *C* for flow 2. On the other hand, node 5 relays all crossing flows on frequency band *C*. The delay at node 2 is dominated by switching delay, while the medium access delay is dominant in node 5. In addition to these delays, there exists also the queuing delay depending on the output capacity available on a given frequency band and on the number of flows sharing this capacity and on their workload.

4.2.1. Solutions accounting for switching and access delay

The novelty of work in [22,23] is the introduction of a metric for multi-hop CRN which is aware of both the switching delay between frequency bands ($D_{switching}$) and backoff delay (medium access delay) within a given frequency band ($D_{backoff}$). At a relay node *i*, a metric representing the cumulative delay along a candidate route is computed as:

$$D_{route,i} = DP_i + DN_i. \qquad (2)$$

The first term takes into account the switching delay and backoff delay caused by the path and depends on the frequency bands assigned to all nodes along the path. As a consequence, $DP_i = D_{switching,\ i} + D_{backoff,\ i}$. If the path is composed of *H* hops, the switching delay along the path is:

$$D_{switching,i} = \sum_{j=i}^{H} k|Band_j - Band_{j+1}|, \qquad (3)$$

where *k* is a constant with the suggested value of 10 ms/10 MHz. We notice that in some practical cases the switching time may be not a function of how wide the separation in frequency between two channels is (unless this requires a new transceiver to be activated). In this case the switching delay becomes a constant. The backoff delay depends instead on the bandwidth on the current frequency band, the number of consecutive nodes sharing the same frequency, and the packet size. The derivation of the expression $D_{backoff,\ i}$ is reported in [23]. The second term in the Eq. (2) accounts for the switching and backoff delays caused by existing flows at the relay node *i*. For the $D_{switching}$ formulation, the authors assume that the node scheduler serves the active bands in a round robin manner. The frequency band from a node's active bands is denoted as $Band_i$. The number of active bands is assumed to be *M*. The $D_{switching}$ is formulated as:

$$D_{switching} = 2k|Band_M - Band_1|, \qquad (4)$$

and becomes a constant when there is no difference in switching from closer frequencies with respect to far away ones. $D_{backoff}$ is defined as the time from the moment a packet is ready to be transmitted to the moment the packet starts its successful transmission. It is obtained as:

$$D_{backoff}(Num_i) = \frac{1}{(1-p_c)\cdot\left[1-(1-p_c)^{\frac{1}{Num_i-1}}\right]}\cdot W_0, \qquad (5)$$

where Num_i is the number of contending nodes, p_c is the collision probability, and W_0 represents them minimum contention window size of a typical CSMA/CA wireless access.

4.2.2. Solutions accounting for queuing delay

The metric in (2) is generalized in [24,25] where $D_{switching}$ and $D_{backoff}$ are integrated with a queuing delay arising at an intersecting relay node which serves n incoming flows. The expression of this queuing delay (named $D_{queueing}$) is computed in [25]. The generalized cost function then becomes:

$$C_{generalized} = D_{route,i} + D_{queueing}. \quad (6)$$

From the definition of this generalized metric, it is clear that assigning a new active frequency band to a flow results in a larger M and increases the $D_{switching}$ of Eq. (4). On the other hand, letting the flow use existing active frequency band $Band_i$ increases Num_i, making larger $D_{backoff}$ and $D_{queueing}$. The effectiveness of this generalized metric is proven in the performance analysis of the paper in [25], where it is shown that the queueing delay estimation is fairly accurate, and the end-to-end delay provided by the proposed routing protocol outperforms traditional routing solutions.

Another contribution of the work [25] is the proposal of a local coordination of neighbor nodes started by an intersecting node. This node decides whether to accommodate an incoming new flow or to redirect it to its neighbors to relief locally the workload. This local coordination includes the operation of exchanging cost evaluation information with neighborhood and the redirection of the flow to a selected neighbor of the intersecting node. Both routing and spectrum assignment are based on the adoption of an on-demand protocol that is a variation of the Ad-hoc On-demand Distance Vector (AODV).

During the path set-up local state information are piggybacked into the route request packets and delivered to the destination node. It is important to note that this protocol does not rely on a simple list of intermediate nodes for routing: The Route Requests (RREQ), which are sent via broadcasting, contain locally obtained network state and deliver this detailed information to the destination, where they are processed to compute paths. The protocol operation starts with the source node broadcasting a RREQ message. As it is being forwarded, intermediate nodes add their own spectrum opportunities – SOPs, a list of currently available and unavailable channels – to the RREQ messages. Once a RREQ message reaches the destination, it estimates a set of cumulative delays based on possible local frequency bands it can use, following a queuing-based delay estimation method and using the metric of Eq. (6). Once it chooses the best possible frequency band it can use, it sends a Route Reply (RREP) message on the reverse path of the RREQ packet. All nodes along the reverse path process the RREP packet following the procedures of the destination. The similarities with the AODV protocols end at this point. The protocol envisions the possibility of changing the routing decisions as the RREP is forwarded along the reverse path. The rationale behind this lies in the fact that nodes carrying more than one flow may have to switch between two or more frequency bands, which incurs a larger delay. Therefore, when a RREP packet is received by an intersection node, it checks its own neighbors to see if there is a better alternative to carry the flow in question. If any of the neighbors of the node that processes the RREP can provide a better delay, then the flow is routed over this new node and the previous hop is also notified of this change. Such an occurrence has been shown in Fig. 7. Here, the RREP packet traverses the same path as the RREQ packet up to node 3. At this point, node 3 estimates the delay to be large and locates another one of its neighbors, node 3', which can carry the flow. Hence, node 3 notifies its upstream node 4 about this better alternative, upon which node 4 forwards the RREP packet over node 3'. The paths traversed by RREQ and RREP packets are shown in Fig. 7, as well.

4.2.3. Effective transmission time routing

A distributed resource management strategy to support video streaming in multi-hop cognitive radio networks is presented in [26]. Given the characteristics of the traffic flows, the main objective is to minimize the end-to-end delay experienced by each video flow based on its classes. The authors argue that a centralized solution would not be realistic in this case since it would require a network-wide mechanism to distribute the necessary information to drive the resource management algorithm. Therefore, a distributed approach is introduced to make local radio re-

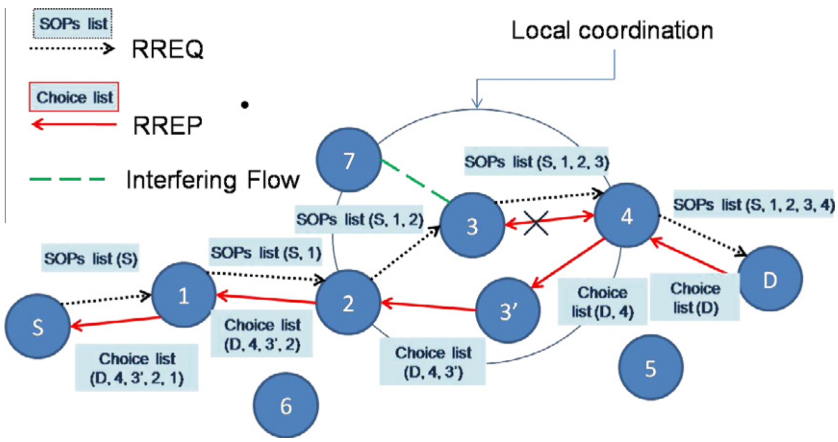


Fig. 7. Example route establishment following [25].

source management decisions on partial information about the network state. The proposed distributed solution accounts for:

- The trade-off between accuracy and cost in gathering information to support radio resource management. Ideally, the larger the information horizon, the better is the visibility of the network condition, and consequently, the more accurate is the radio resource management action taken by each node. On the down side, a large information horizon requires higher signalling traffic. The authors include an accurate model to capture this trade-off in the distributed radio resource strategy.
- A learning approach according to which the SU can dynamically tune their actions on the basis of the observation of their neighbors' behavior. Active Fictitious Play (AFP) techniques are used to evaluate the propensity for a given neighbor to take a specific action (e.g., tuning to a specific channel). Such propensity is then leveraged when deciding on the action.

Under the proposed routing scheme, SUs interact with each other to adjust their transmission parameters to minimize the end-to-end delay such that K different delay sensitivity classes can be supported. The work analyzes the tradeoff between sensing accuracy and signalling overhead. To establish routes, the authors propose a metric called *Effective Transmission Time* (ETT) that reflects the delay experienced by priority k packets departing from node n over link e:

$$ETT_{nk}(e,f) = \frac{L_k}{T_n(e,f)\cdot(1-p_n(e,f))}, \quad (7)$$

where f is a frequency band that can be used to establish link e. $T_n(e,f)$ and $p_n(e,f)$ represent the transmission rate and the packet error rate of node n using frequency band f over link e, respectively. $T_n(e,f)$ and $p_n(e,f)$ are estimated at MAC/PHY layer. L_k is the average packet length in class k. The value of *ETT* depends on the action taken by node n to relay a delay sensitive packet. An action A_n is defined as the selection of a link and a frequency band, i.e., $A_n = (e,f—e \in \mathbf{E}_n, f \in \mathbf{F}_n)$, where $\mathbf{E}_n$ is the set of all links adjacent to n and $\mathbf{F}_n$ is the set of all frequency bands that can be used by n. The idea is to optimize the end-to-end delay by learning the "environment" (channel conditions and source characteristic) and the actions of competing nodes within x hops of a given node n. For the kth class, the metric $J(k,(I)_n(x))$ denotes the benefit (reward) of local information $(I)_n(x)$ gathered from the neighbor nodes at distance x. $J(k,(I)_n(x))$ is computed as the difference between the optimal expected delay (denoted as $K_n(k,\cdot)$) computed with the information at distance $x-1$ and the one computed with the information learned at distance x:

$$J(k,(I)_n(x)) = K_n(k,x-1) - K_n(k,x). \quad (8)$$

The value of $K_n(k,x)$ decreases as x increases since by having more information from a larger neighborhood, it is possible to better optimize the value of the end-to-end delay. Consequently, $J(k,(I)_n(x))$ is always nonnegative. The tradeoff analysis of having increasing values of x results in the definition of a suitable information horizon which determines the best value of x to be used for a given application. The reward of information is zero beyond the information horizon. Furthermore, the cost for the information exchange in the horizon space is integrated into the considered metric. With these properties, the defined metric can capture the physical and MAC layer behaviors by selecting a suitable action A on the basis of its effects on Eq. (7). The work also provides a relatively detailed structure that defines cross-layer interactions between different modules as shown in Fig. 8.

It is also worth noting that the information horizon concept is also presented a means to capture mobility of nodes. Mobility of nodes has been overlooked in many routing and resource management solutions for multi-

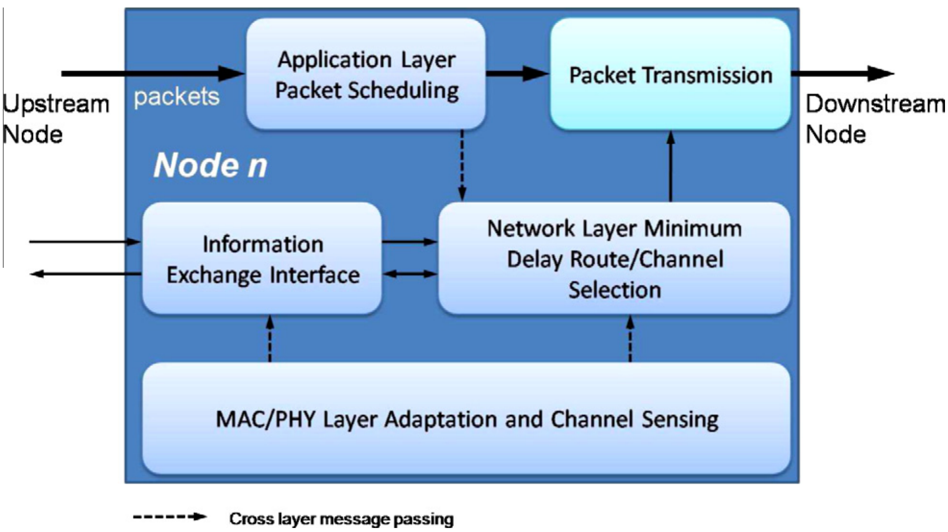


Fig. 8. Cross-layering of [26].

hop CRNs. It is argued that, for higher mobility scenarios, the information exchange happens more frequently to capture changes from the selected target neighborhood. However, simulation results from the same reference suggest that the adjustment of the information horizon depends on more than the mobility, including the properties of the information stream. Therefore, the effectiveness of information horizon concept to capture effects of mobility is inconclusive, subject to further study of the method.

4.3. Throughput-based solutions

Throughput maximization is the main objective of the routing solutions described in this section.

4.3.1. Path spectrum availability routing

The Spectrum Aware Mesh Routing (SAMER) proposal [27] is a routing protocol that accounts for long term and short term spectral availability. SAMER seeks to utilize available spectrum blocks by routing data traffic over paths with higher spectrum availability, without ignoring instantaneous spectral conditions. The protocol first establishes candidate paths using periodically collected global states, and associating paths with Path Spectrum Availability (PSA) metrics. Then, packets are delivered opportunistically along the path with the highest PSA value and that is available at that point in time. SAMER seeks to utilize available spectrum blocks by routing data traffic over paths with higher spectrum availability. Authors of SAMER define a metric for estimating Path Spectrum Availability (PSA). PSA's goal is to capture:

1. Local spectrum availability: Spectrum availability at a node i depending on the number of available spectrum blocks at i, their aggregated bandwidth and the contention from secondary users, and
2. Spectrum blocks quality depending on their bandwidth and loss rate.

The PSA is expressed as the throughput between a pair of nodes (i,j) across a spectrum block b as:

$$Thr_{(i,j),b} = T_{f,b} \cdot B_{w,b} \cdot (1 - p_{loss,b}), \tag{9}$$

where $B_{w,b}$ is the bandwidth and $p_{loss,\,b}$ the loss probability of the spectrum block b. This latter value can be estimated by measuring the loss rate of broadcast packets between pairs of neighboring nodes. In Eq. (9), $T_{f,\,b}$ is the minimum between the fractions of time during which the node i (j) is free to transmit and/or receive packets through a spectrum block b. The aggregate throughput $Thr_{(i,j)}$ between a pair of neighboring nodes is then computed on the basis of the spectrum blocks available at a node i and then smoothed by multiplying it by a value α (assumed to be 0.4) to capture both the current view and the statistical information of spectrum availability. The Smoothed Aggregate Throughput is then updated as:

$$SThr_{(i,j)} := \alpha \cdot SThr_{(i,j)} + (1 - \alpha) \cdot Thr_{(i,j)}. \tag{10}$$

Spectrum availability for a path P is then defined as the minimum Smoothed Aggregate Throughput for $(i,j) \in P$. In calculating the PSA value, the paths are restricted to be H hops or less. When a node relays a packet, it chooses the next hop based on PSA and local spectral availability. The next hop is chosen locally along the path that has the best PSA value and for which the spectrum is available. Since the channel can be accessed by many SUs, SUs contend for a channel over the CCC. All spectral resource reservations are performed over the CCC before an SU transmits a packet. If SU contention is high, this is reflected in the measurements of bandwidth availability that in turn affect the PSA values of paths. Consequently, the proposed scheme accounts for SU as well as PU activity to rank paths.

In the paper, SAMER is found to outperform the popular hop count and Expected Transmission Time metrics. Furthermore, simulation results suggest that SAMER avoids highly congested and unavailable links. However, overheads associated with forwarding mesh establishment and maintenance have not been considered in depth. Furthermore, details of the channel access, deafness due to the separation of signaling and communication channel, and contention resolution among SUs have not been discussed.

4.3.2. Spectrum utility based routing

Achieving high throughput efficiency is the main goal of protocol ROSA [28]. Opportunities to transmit are assigned based on the concept of spectrum utility and routes are explored based on the presence of spectrum opportunities with the objective of maximizing the spectrum utility. The authors introduce a spectrum utility for the generic link (i,j) defined as the maximum differential backlog between node i and node j; in formulas:

$$U_{ij} = c_{ij}(Q_i^{s^*} - Q_j^{s^*}),$$

where c_{ij} is the achievable capacity for link (i,j), $Q_i(s^*)$ is the current backlog of packets at node i for the session (packet flow) s^* and s^* is the session with the highest differential backlog.

The current value of c_{ij} depends on the scheduling policy, the dynamic spectrum allocation policy, and the power allocation scheme. Indeed,

$$c_{ij}(f, P_i(f)) = \sum_{f \in [f_i, f_i + \Delta_{fi}]} wlog2\left[1 + \frac{P_i(f) L_{ij}(f) G}{N_j(f) + I_j(f)}\right],$$

where G is the processing gain, $L_{ij}(f)$ is the transmission loss from i to j, $P_i(f)$ represents the transmission power node i uses over frequency f, $I_j(f)$ is the perceived interference at j, and $N_j(f)$ is the background noise.

The generic node i performs the following actions:

- it periodically searches for the list of potential next-hops for session s $\{n_1, n_2, \ldots, n_N\}$,
- it calculates the capacity (c_{ij}, where $j \in \{n_1, n_2, \ldots, n_N\}$) over the links towards all the potential neighbors; more specifically, given the current spectrum condition, each SU runs a distributed decision algorithm to decide which spectrum mini-bands should be used for the access and which power level to be used throughout the aforementioned spectrum bands,

- it chooses the actual next hop, j^*, that maximizes the spectrum utility, that is, $(s, j^*) = argmax_j(U^s_{ij})$.

The proposed routing protocol is further coupled with a cooperative sensing technique which leverages both physical sensing information on spectrum occupancy and virtual information contained in signalling packets exchanged by secondary nodes. The exchange of additional virtual information is performed through a common control channel and is used by the local spectrum/power allocation algorithm.

4.4. Link quality/stability based solutions

The channel availability in multi-hop CRNs is significantly different than in traditional wireless multi-channel multi-hop networks. Indeed, nodes in multi-hop CRNs potentially have partially overlapping or non-overlapping sets of available channels, and the available channel set at a SU is of time-varying nature and changes in correlated or uncorrelated manner with respect to sets of other nodes. Consequently, network layer solutions in multi-hop CRNs should be able to cope with the necessity of re-routing in case specific portions of the currently active path are "impaired" by the presence of an activating PU. This section overviews proposed routing solutions which shift the focus to designing stable and quality multi-hop routes.

4.4.1. Solutions with enhanced path recovery functionalities

Throughput maximization by combining end-to-end optimization with the flexibility of link based approaches to address spectrum heterogeneity is proposed in SPEAR (*SPEctrum-Aware routing* [29]). The available spectrum is location dependent and the introduction of primary users typically creates islands of different spectrum availability. As an example in [29] it has been show that using random topologies the probability of finding a route between two nodes by forcing nodes of the path through the use of a single channel is significantly lower with respect to the probability of finding a route hopping on different channels. In this framework the proposal of SPEAR goes in the direction of:

- integrating spectrum discovery with route discovery to cope with spectrum heterogeneity;
- having a coordination of the channel assignments of a per-flow basis, by minimizing inter-flow interference;
- exploiting local spectrum heterogeneity to in order to have a *spectrum diversity* and reduce intra-flow interference.

To achieve these goals SPEAR starts the route set-up by broadcasting and AODV-style route discovery which accumulates information about each node's available channels and their quality. At the end of the different paths towards the destination each RREQ contains a list with the node IDs, the nodes' spectrum availability and the links' quality. Furthermore, to account for inter and intra flow interference nodes intersecting different flows store the time schedules of these flows. These parameters are combined at the destination to select the optimal route (by using for instance graph coloring approaches as in [14]). Unlike traditional on-demand route discovery protocols SPEAR discovers different paths. Redundant paths are not suppressed but are sent to the destination for the best path selection. The selected route is then reserved by using RREP messages. Channel usage is the scheduled at each node; a node can also locally change part of the channels assignment, in case of failures or node mobility, by keeping unchanged the local throughput.

A collaboration between route selection and spectrum decision is considered also in the paper [30]. Authors propose the *Spectrum Tree based On Demand Routing Protocol* (STOD-RA) framework constituted by: (i) a route metric based on statistical PUs activities and SUs QoS requirements; (ii) a spectrum-tree structure in each sensed available channel; (iii) the Spectrum Tree based On Demand Routing Algorithm.

As for the routing metric it combines link stability and spectrum availability. The idea is to predict the availability time of a spectrum band from the statistical history of PU activities. The link cost C_i of the link l_i is calculated as:

$$C_i = \left(O_{ca} + O_p + \frac{P_{kt}}{r_i}\right) \cdot \frac{1}{1 - e_{pti}} \cdot \frac{1}{T_{l_i}}, \tag{11}$$

where:

- O_{ca} and O_p are constant for a specific access technology and represent the channel and protocol overhead, respectively;
- P_{kt} the packet size, which is constant for a specific access technology;
- r_i is the link rate (in Mbps);
- e_{pti} is the packet error rate on the link;
- T_{l_i} is time duration during which a spectrum band is available to the link l_i.

The consequence of the use of T_{l_i} in the metric allows the integration of the link stability. The available time of a spectrum band can be predicted from the statistical history of PU activities. The overall cost C of and end-to-end route composed of k links is:

$$C = \sum_{i=1}^{k} C_i + M \cdot D_{switching}, \tag{12}$$

where M is the number of spectrum band switches along the route and $D_{switching}$ is the switching delay between two different bands (see Eq. (4)).

The spectrum-tree is a lookup structure to keep trace of nodes operating in different spectrum band. A spectrum tree exists for a given spectrum band and has only one root node which keeps the information about the tree topology (e.g., routes to other non-root nodes). Nodes belonging to multiple spectrum-trees and having multi-radios are called "overlapping" nodes and they can work in multiple spectrum-trees simultaneously. A root selection procedure assures that there is only one root in each spectrum tree. This root is a node which belongs to the largest number of spectrum trees and within nodes in this set the one which has a spectrum trees with the longest time duration during which a spectrum band is available. The time duration is a parameter affecting the spectrum tree reconfigurability due to PU activity.

The spectrum-tree is then used for both Intra-spectrum routing and Inter-spectrum one. In the fist case a combination of proactive routing mechanism performed along the tree with a reactive mechanism is used. In the second case the overlapping nodes are considered as cluster-head of two (or multiple spectrum trees) and are in charge of routing packets that should cross different spectrum bands.

STOD-RA uses spectrum-trees also for route recovery purposes. Assuming that the spectral dynamics due to primary user access changes slowly, the system avoids further coordination by confining data communication and routing related signaling to the same frequency bands. In case for instance of a temporarily impossibility to use a spectrum band there is the possibility that all the nodes in spectrum-tree handoff to an available spectrum band. In this way a fast and efficient spectrum-adaptive route recovery method is introduced.

The work presented in [31] presents an algorithm for handoff scheduling and routing in multi-hop CRNs. One of the main contributions of this work is the extension of the spectrum handoff to a multi-link case. Following a classical approach, the problem of minimizing latency for spectrum handoff across the network is shown to be NP-hard and a centralized and a distributed heuristic algorithms have been developed. The centralized algorithm is based on the computation of the maximum non-conflict link set. With this approach, the algorithm iteratively assigns new channels to links. To address the starvation problem, an aging based prioritization scheme is utilized. The distributed algorithm uses a link cost metric that is inversely proportional to the link holding time and link quality. Then, the rerouting algorithm tries to minimize the total link cost along a path from source to destination. The distributed algorithm isolates handoff occurrences to a single link along the rerouted path to ensure connectivity. The simulation results show that the performances of the distributed and centralized solutions are very close to each other for the tested scenarios in a grid topology. Both algorithms also provide improvements over the cases where the proposed algorithms are not utilized. Unfortunately, it is not clear how close these algorithms approach the optimal solutions. The implementation details of the distributed algorithm have not been laid out in detail.

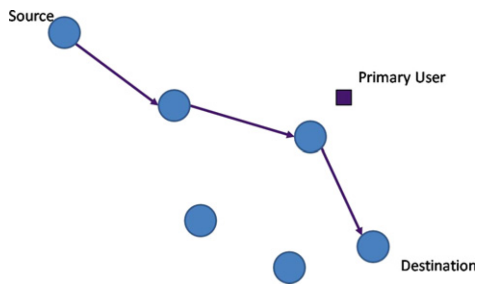
4.4.2. Solutions targeting route stability

The link stability is considered also in the paper of Abbagnale et al. [32] where this parameter is associated, in a innovative way, to the overall path connectivity via a mathematical model based on the Laplacian spectrum of graphs. Paths are measured in terms of their degree of connectivity that in a multi-hop CRN is highly influenced by the PUs behavior. The behavior of a PU is modeled by its average activity factor. The authors introduce a novel metric to weight routes (paths) which is able to capture path stability and availability over time. Indeed, the core idea is to assign weights to routes and paths proportionally to the algebraic connectivity of the Laplacian matrix of the connectivity graph abstracting the secondary network. On the basis of this model authors design a routing scheme, named ***Gymkhana***, which routes the information across paths that avoid network zones that do not guarantee stable and high connectivity. Gymkhana uses a distributed protocol to collect some key parameters related to candidate paths from an origin to a destination. These parameters are then fed into the basic mathematical structure based on Laplacian matrixes which is used to compute efficient routing paths. The main contributions of the work in the framework of the cognitive radio routing are (i) the provision of a simple but effective re-elaboration of the algebraic connectivity in a cognitive context; (ii) the formulation of an utility function which accounts for the path connectivity and the path length that can be effectively used in a cognitive routing protocol. The analysis of significant case studies shows the effectiveness of the proposed approach in achieving the routing goals. Moreover, beside the routing purposes, the provision of a model for measuring the connectivity of a multi-hop CRN can be also used for network planning an dimensioning.

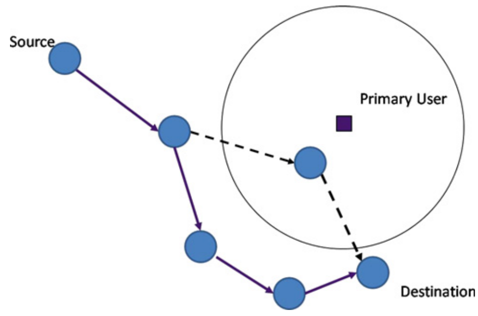
A route stability oriented routing analysis and a protocol are presented in [33], where a novel definition of route stability is introduced based on the concept of ***route maintenance cost***. The maintenance cost represents the effort needed or penalty paid to maintaining end-to-end connectivity in dynamic multi-hop CRNs. The maintenance of a route may involve link switching and channel switching operations as a PUs become active. In the former case, one or more links along the route must be replaced by other ones not interfered with by PUs, whereas in the latter case, the same link can be maintained, but the transmission must be carried over to another spectrum portion. In either case, signalling is required to coordinate with other SUs, which translates to a cost in terms of consumed power, and service interruption time while switching routes. Fig. 9 shows a case where rerouting is needed due a PU becoming active. The two-hop portion of the path (dashed lines) needs to be replaced with the three-hop segment in Fig. 9b. The cost involved in the rerouting phase contributes to the overall maintenance cost.

The authors start off by obtaining optimal minimum maintenance cost paths according to the specified metrics under ideal spectrum sensing conditions and perfect knowledge of the PU activity. Different from other existing proposals, the proposed optimization formulation directly accounts for the dynamics of the network topology. The authors introduce the concepts of network epochs, which are defined as time intervals where the topology of the SU network is stable. From epoch to epoch, the network topology might change due to activation (or de-activation) of licensed primary users. The main focus is on the construction of stable routes, that is, routes between secondary source–destination pairs which can be maintained with the lowest maintenance cost during their lifetime. The maintenance cost includes the cost for switching among different channels on the same wireless link, and the cost for establishing entirely new portions of a path to circumvent a zone blocked by an incoming PU.

The authors propose a MILP formulation for the problem of minimizing the route maintenance cost, under interference, link capacity, and flow balance constraints.



(a) Epoch 1.



(b) Epoch 2.

Fig. 9. Rerouting due to PU activation.

They design a centralized algorithm running in polynomial time to optimally compute minimum maintenance cost routes in multi-hop cognitive networks with perfect information on PU dynamics. Leveraging properties and observations gathered from the optimal routes, assumption on perfect knowledge of the multi-hop CRN state is dropped, and a practical routing metric is proposed. This metric captures the "quality" of a given link l between secondary users as far as route maintenance is concerned. This metric 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. The proposed approach weighs each link by the following metric:

$$w_l = \frac{C_l^{sw} + \alpha C_l^{Rep}}{E[TTS_l]}, \tag{13}$$

where parameter α allows gauging of different cost contributions, and $E[TTS_l]$ represents the average time to switch for the link l. Ideally, the longer the continuous lifetime of a link is, the lower is the incurred maintenance cost. Therefore, the denominator is used to give lower weights to links available for longer time periods. Exact expressions for the components of the metric in Eq. (13) are given in the paper under the assumptions that PU activity can be modeled as a random ergodic ON-OFF process, and the secondary users have knowledge of the first order statistics of the PU activity.

This metric is then used to compute paths at the source side and allows for local modifications to the path as the spectral conditions change. The work presents a unique perspective on path stability in multi-hop CRNs. The insights gained from the analysis are also incorporated into a routing metric to be used by source routing algorithms. However, the proposed practical algorithm does not provide detailed discussion on the dissemination of PU statistics across the multi-hop CRN. Furthermore, as the recovery and update procedures require at times lengthy signaling operations, although costs associated with such switching cases are explicitly considered, the proposed protocol are not well suited for highly dynamic PU activity scenarios. Finally, the analysis and the protocol does not account for the "link capacity" explicitly. We note that the notion of the link capacity is not clearly defined for wireless networks due to scheduling and interference constraints, and therefore cannot be directly incorporated into a MILP formulation.

4.4.3. Routing with mobile SUs

Depending on the specific cognitive scenario, secondary users accessing opportunistically the spectrum may be mobile. Think of handheld devices carried by humans which may want to establish opportunistic link among themselves to support file sharing applications. Thus, a multi-hop CRN needs to be established among mobile users.

SEARCH [34] is a routing protocol that is designed for mobile multi-hop CRNs based on the geographic forwarding principles. The proposed protocol makes routing and channel selection decisions while avoiding regions of PU activity. It also considers a host of nodal mobility cases using predictive Kalman filtering. The main idea behind SEARCH is to discover several paths from source to destination, which are then combined at the destination to form low hop count paths.

The protocol's route setup phase is similar to many ad hoc routing algorithms: The source initiates path search with RREQ packets on every channel that is available at the source. While they are being forwarded towards the destination, the RREQ packets are transmitted only on the original spectrum they were originally transmitted in. The forwarding procedure follows a greedy routing mechanism within a focus region. The focus region is a sector of a circle centered around the line that connects a current node with the destination and of angular range of $2\theta_{max}$. Intermediate nodes forwarding a RREQ packet search for a next hop within their focus regions according to greedy geographic forwarding principles. If a node cannot forward a RREQ packet to another node in its focus region on a particular spectrum band, such nodes mark themselves as *decision points* (DP) and enter PU avoidance phase. The rationale behind this classification is that DPs emerge when an active region lies along the path towards the destination on a given spectrum band.

In the PU avoidance phase, the RREQ packet is routed over nodes lying outside focus regions. Nodes are assumed to know the spectral availability of their neighbors through periodic message exchanges. If a forwarder node finds that none of its neighbors within a focus region is available in a spectrum band, it marks the packet to be in the PU avoidance phase so that a suitable neighbor outside the focus region forwards it. The packet is forwarded in the PU avoidance phase until it reaches a point where the forwarder node can relay the RREQ packet over a neighbor within its focus region. Once the RREQ packets forwarded in different channels reach the destination, the destination node makes a path and channel selection decision to form the end-to-end path. The selection of the end-to-end path is based on the shortest path among all candidate paths discovered over different channels. The candidate path is taken until the first DP is encountered. At this point, all other paths are considered to see if a lower latency path can be reached from this current DP, after accounting for the channel switching delay. If so, the path is augmented with portions of a lower latency path at the DP, where the packets would now be forwarded on a different channel. A similar decision is made whenever a DP is reached while forming the end-to-end path. Once the path is formed, a single RREP is sent back on the reverse route to the source, marking the forwarding path as well as channel switching decisions.

The protocol also envisions local optimization of the path once an initial path is setup with RREP. The SEARCH protocol also accounts for the changing spectral availability and the mobility of the SUs. In case the operational path is affected by a new PU activity, the last forwarded before the affected region becomes a DP, and initiates a new partial route search with RREQ packets. If the resulting path is within a threshold of the old path's latency, then the path is updated with this new patch. Otherwise, a notification sent to the source triggers a new path search. On the other hand, SU mobility is handled by associating forwarders with their locations rather than their IDs. In such a case, if a node moves from its original location more than a predetermined amount, its upstream neighbor replaces it with a new forwarder node within the old scope. If no such new forwarder node is found, then the path is extended towards either the source or destination, whichever is closer. Furthermore, the stability of the path is observed at each hop throughout the session with Kalman filtering and new links are established/maintained with these predictions.

The SEARCH protocol combines several routing techniques effectively to establish and maintain routes in multi-hop CRNs. It explicitly accounts for nodal and spectral dynamics when maintaining paths. The protocol relies on the latency predictions at intermediate nodes to form paths. As in other wireless networks, such predictions are generally very inaccurate and should be used sparingly. Considering the spectral dynamics, latency predictions are bound to be very inaccurate. Furthermore, the protocol requires a detailed set of information about one hop neighbors to be maintained in every node, which incurs a considerable overhead, as well. Nevertheless, this protocol is one of the few examples that accounts for network dynamics of CRNs at such detail and should be considered as a good starting point for further research.

4.5. Probabilistic approaches

In case the exact status of spectrum occupancy is not available or cannot be dynamically reconstructed through distributed schemes, routing solutions may need to be more "myopic" with respect to the spectrum awareness, and routing decisions (and metrics) should be based on probabilistic figure of merit.

In [35] it is defined a routing approach based on a probabilistic estimation of the available capacity of every CR link. A probability-based routing metric is introduced; the metric definition relies on the probability distribution of the PU-to-SU interference at a given SU over a given channel. This distribution accounts for the activity of PUs and their random deployment. This routing metric is used to determine the most probable path to satisfy a given bandwidth demand D in a scenario with N nodes that operate on a maximum of M orthogonal frequency bands of respective bandwidths $W_1, \ldots, W_M$ (in Hz). Authors derive the probability that channel i can support the demand D (expressed in bit/s) as:

$$Pr[C(i) \geqslant D] = Pr\left[P_{I,j}^{(i)} \leqslant \frac{P_{r,j}^{(i)}}{2^{D/W_i} - 1} - N_0\right], \tag{14}$$

where $P_{I,j}^{(i)}$ is the total PU-to-SU interference at SU j over channel i, with $i = 1, \ldots, M$ and $j = 1, \ldots, N$. Authors assume that $P_{I,j}^{(i)}$ follows a lognormal distribution. The probability in (14) can be obtained for every channel of every link by calculating the cumulative distribution function of the lognormal distribution of the PU-to-SU interference. Based on this probability, the routing metric is given a weight of the link between nodes k and j on channel i:

$$l_{k,j}^{(i)} = -\log Pr\left[C_{k,j}^{(i)} \geqslant D + U^{(i)}\right], \tag{15}$$

where $U^{(i)}$ is the system memory that accounts for the cognitive interference in the vicinity of nodes k and j (and is detailed in the paper [35]), while $C_{k,j}^{(i)}$ is the maximum channel capacity given by Shannon's Theorem.

A source-based routing protocol is proposed for the path selection. Link state advertisements are exchanged on a common control channel to acquire the parameters for computing Eqs. (14) and (15). With this phase the source is able to compute the most probable path to the destination. A subsequence phase is dedicated to compute the available capacity over every link in the selected path and augmenting this capacity till the total capacity available on the path is grater than the demand D. During this path augmentation the already accepted flows crossing the link (i,j) are taken into account by using the variable $U_{kj}^{(i)}$. Including this variable in the probability computation naturally pushes the algorithm to use different frequencies on consecutive nodes thus reducing the interference. Through simulations and numerical results, the efficiency of the proposed routing metrics and the algorithm is validated by showing that the most probable path to the destination is selected in all cases [35]. This path yields the best performance in terms of throughput.

However, the fully opportunistic approach makes sense if PUs are highly active, then the availability of SOPs to sustain a full communication session in a single SOP becomes impossible as highlighted in [11]. A possible solution for SUs is to transmit over any available spectrum band during the short SOPs in a fully opportunistic way. In this case every packet of a given flow can be sent on a different channel by exploiting the intrinsic intermittent CRN channels availability. Authors of [11] observe that few researchers have looked at multi-hop CRNs under these assumptions and discuss pros and cons of possible protocols in this direction. The selection of a channel to be opportunistically used can be made by tracing back the history of the channel itself, as sensed by a given node. It is to be noticed that, oppositely with respect to probabilistic approaches as in [35], here a node first looks at the available channels and then selects on the basis of an history. On the contrary probabilistic approaches select a path composed by a set of channels on the basis of the history.

5. Discussion and open research issues

A summary of the protocol solutions for routing in multi-hop CRNs is reported in Table 1. As presented in this table there exist two main categories for routing solutions: (i) proposals focused on static network topologies, with fully available topological information on neighboring SUs and spectrum occupancy (indicated in the table as approaches with a *full spectrum knowledge*); (ii) proposals based on local radio resource management decisions on partial information about the network state (approaches based on *local spectrum knowledge*). In the first case, the problem of designing/modeling CRNs scales down to the classical problem of designing static (wireless) networks, where tools of graph theory and mathematical programming can be leveraged extensively. Even if the implementation of these approaches may result complex and may be scarcely scalable, their importance can be seen in the application to all that scenarios where the SUs have access to data bases storing the spectrum maps, as recently envisaged by the FCC ([10]).

On the other hand there exist several approaches based on local information on spectrum occupancy gathered by each SU through local and distributed sensing mechanisms. In some cases the protocols are able to set up the whole path while in other cases the proposed approaches are based on the selection hop by hop of the next forwarding node. However, a distinguishing characteristic of all routing approaches is that they combine to the routing the selection of the spectrum on each link of the path. This can be done by using different metrics for capturing the characteristics of the available spectrum holes. The most appropriate spectrum bands can be then selected according to both radio environment (interference, power) as well as QoS parameters like throughput, delay, etc.

Also the behaviors of the PUs is a key parameter to be considered for routing data in a multi-hop CRNs. In fact, routes must explicitly provide a measure of protection to the ongoing communication of the PUs while at the SUs side must guarantee stability when the PU behavior varies. This is taken into account in a set of routing solutions where the PUs' statistical behavior and the consequent spectrum fluctuations are considered via suitable models in the routing metrics. Besides this, also the ability to reconfigure the routing paths when a PU becomes active can be a distinguish feature of the routing. As reported in Table 1 only few solutions have this ability. Finally, very few solutions have considered till now the SUs mobility.

We strongly believe that research in the field of modeling/designing CRNs routing still needs major contributions explicitly endorsing network dynamics and variability, which are distinctive features of the multi-hop CRNs. To this extent, open research issues in the field of models and algorithms for route designs in multi-hop CRNs include the following issues:

5.1. True Cross-Layering

Successful operation of a routing solution in CRN highly depends on the exchange of information among multiple layers. A very prominent example of such information exchange is the spectrum sensing information that requires cooperation of PHY and MAC layers to obtain and consumed by routing decisions. As such, almost all routing protocols can be classified as *cross-layer* solutions. However, the flow of information is mainly uni-directional, i.e., information produced by lower layers is consumed by higher layers and no direct feedback is provided back to the lower layers. The effect of routing protocols' actions are only fed back to lower layers through the channel and

Table 1
Summary of protocol solutions for routing in multi-hop cognitive radio networks.

		Protocols	Whole path selection	Next hop selection	Spectrum dynamics awareness	Reconfig. to varying spectrum	Mobility support
Full spectrum knowledge	Graph based	[12–15]	√	–	–	–	–
	MINLP-MILP formulation	[16–18]	√	–	–	–	–
Local spectrum knowledge	Interference and Power Based	[19,20]	√	–	–	–	–
		[21]	–	√	–	–	–
	Delay Based	[22–25]	√	–	–	–	–
		[26]	–	√	√	–	√
	Throughput Based	[27]	√	–	√	–	–
		[28]	–	√	√	–	–
	Link Quality/Stability Based	[29–31,33]	√	–	√	√	–
		[32]	√	–	√	–	–
		[34]	√		√	√	√
	Probabilistic approaches	[35,11]	√	√	√	–	–

resource availability realizations. This is not only a simplistic view of cross-layer interactions, but also a highly vulnerable one, since the two legs of the information flow, i.e., from lower to higher layers and vice versa, occur over significantly different time scales (milliseconds vs. tens of seconds). In the particular example of channel sensing, a true exchange of information between routing and PHY/MAC layers would not only minimize the waste of precious wireless transmission opportunities, but also reduce the scale gap between feedback between layers, thus ensuring higher stability of solutions. We refer to direct interactions between various layers without depending on indirect feedback through channel realizations as ***"true cross-layering***".

In addition to the above-discussed issue of sensing, other components of communication protocols for multi-hop CRNs would also benefit from true cross-layer interactions. Management of SU mobility, spectrum handoff decisions, candidate end-to-end path selection, channel allocation decisions, and incremental allocation/deallocation of resources along paths constitute a partial list of all functions that stand to benefit from true cross-layering. For these and other potential functions, we believe appropriately detailed analytical models are of utmost importance, which constitutes next our major open research issue. With these analytical models, it is possible to estimate the effects of cross-layer interactions and take appropriate steps in exchanging information and controlling the protocol behavior. This is also an important step to propel protocol design from ad hoc to analytically grounded and provable approaches.

5.2. Analytical Models for CRN Environment and Functions

Resource availability in (multi-hop) CRNs is shaped by the behavior of PUs as well as the actions taken by SUs. Existing analytical models aim at (and achieve in a limited sense) describing the behavior of PUs in isolation from SU activity. Existing PU activity models assume simple structures (such as ON/OFF models) necessary to be of theoretical significance to aid in design and evaluation of CRN protocols. Unfortunately, such models' accuracy fall sharply in multi-hop CRNs due to a multitude of reasons. One of the main reasons for accuracy drop is the violation of a simplifying assumption in what can be sensed and what is of importance: SUs are supposed to avoid harmful interference at the *PU receivers*. In a single-hop CRN, it may be reasonable to assume that sensed PU activity is directly correlated to potential interference with PU receivers if SU were to become active. However, this simplifying assumption is clearly incorrect in the case of multi-hop CRNs: Since SUs can be arbitrarily far from PU receivers and since transmission powers are not symmetrical across PUs and SUs, such a direct correlation cannot be assumed. In fact, detailed models that relate sensed PU activity to the potential for interference are much sought-after. Moreover, correlation between sensed PU activity among several SUs can be leveraged to estimate the location of (or channel gain between SUs and) PUs, or to combine the channel sensing effort in the SU network. All these potential benefits are conditioned on models describing the resource availability of multi-hop CRNs, which is not available to date at desired levels of accuracy and therefore still an open research problem. Such comprehensive analytical tools can be integrated into the route design phase. Prediction of future conditions (interference, link quality) would definitely favor the implementation of effective "super-cognitive" solutions that go beyond sensing and reporting current resource availability. To this end, tools from the theory on machine learning and regression could definitely boost the quality of cognitive protocols including routing.

Analytical models for protocol behavior is also an important issue that has been largely overlooked in the current multi-hop CRN literature. While some existing solutions provide provable performance bounds and convergence properties, these are primarily limited to centralized solutions and a few distributed ones such as [26]. On the other hand, changes in resource availability may result in ripple effects throughout the network in resource reallocation. Knowing convergence characteristics of distributed algorithms and designing algorithms to withstand instability that may be caused frequent resource availability fluctuations are only possible through accurate modeling of protocol behaviors[2]. Moreover, realistic interactions between PU and SU protocols needs to be developed, as well. While the commonly accepted CRN principles lead us to believe that PUs will operate as if SUs do not exist (and therefore, SUs can detect PU presence and vacate the channels), such ideal PU behavior cannot be expected in all PU networks. As an example, under the assumption of unchanged PU protocol stack, PU networks that employ CSMA-based channel access are bound to treat SU presence as any other PU presence and back off. A joint model of PU and SU protocol behavior still remains a very important open research problem. A generalization of this consideration leads us to our next major open research problem, namely, interaction of PU and SU Systems.

5.3. Interactions of PU–SU and SU–SU Systems

Our preceding discussion eludes to complex direct and indirect interactions between SU and PU systems, which are very poorly understood. Nevertheless, in a real implementation, it is clear that PU systems will be negatively affected from the presence of SUs, despite every effort to minimize interference with the PU system. Theoretically and from a purely policy perspective, there is no reason why a PU system should *allow* an SU system operate in its interference region: By simply injecting dummy transmissions instead of staying idle, a PU system would ensure that no SU would gain an opportunity to transmit, and consequently, potentially interfere with any real PU transmission.

[2] Lessons learned from the vast wireless networking literature point at a striking disparity between outcomes of "design first, model later" and "model first, design later" approaches. The former would be the case of IEEE 802.11 MAC protocol, which was designed first with an ad hoc approach, but for which a very accurate model (esp. for delay) is still elusive. On the other hand, various scheduling/congestion control algorithms and related protocols have been designed using the latter approach, which achieve analytically provable performance levels automatically.

Such issues are more pronounced in multi-hop CRNs due to their potential size and need to relay the same message over multiple hops through an PU activity area. Methods like voluntary spectrum handoff [36] can be incorporated in to routing decisions to further minimize potential negative effects of SUs on PU systems. However, proper analytical models are not mature enough to minimize such effects arbitrarily. Another approach would be to alleviate some of the barriers between PU and SU systems and ensure that PUs stand to gain in performance with the presence of SUs. Dynamic spectrum leasing is a potentially promising direction that allows SUs to transmit their data in return for cooperation with PUs in relaying their data [37,38]. Dynamic spectrum leasing and other approaches that ensures performance improvement for PUs need further investigation for adaptation to multi-hop CRNs and to the routing problem.

Another aspect of the interactions deals with different constituents of the SU system. When SUs do not necessarily cooperate with each other, the potential negative impact on the PU system increases and the potential gain for the SU systems decreases. *Game Theory* has been proved to be a very versatile tool in similar wireless environments where participants only look out for their own interest. Game theory has already been proposed as a powerful methodology to assess the quality of spectrum access and sharing in CRNs. However, its application to multi-hop CRN routing problems is still in its infancy. We believe that game theoretic tools could be extremely helpful also in the design of routing strategies for CRNs. Different from the spectrum access/sharing situations where users compete for SOPs on a single-hop basis, the problem of routing features a multi-hop competition among contending users (flows). Moreover, the very same interaction patterns among cognitive devices forming a network may be of different nature under different scenarios. As an example, SUs may be cooperative or competitive to set up the network and cognitive capabilities of the CR nodes may be different from each other. Also on the PU side, different behaviors as well as different benefits in hosting SUs on licensed portion of the spectrum may exist. Due to these reasons, game theoretical solutions bear significant potential to solve several issues related to routing in multi-hop CRNs.

5.4. Prototypes and Testbed Implementations

Finally, still much work needs to be carried out in the field of experimentation. Indeed, the integration of prototypes and the testbed implementations with cognitive devices is deemed essential to validate findings and refine models, algorithms, and systems. As demonstrated by the studies carried out in the past for other wireless technologies, we strongly believe that a serious investigation of all technical issues related to CRNs requires validation in real testbeds in addition to simulation and analytical models. To this end, we observe that research initiative have been recently launched to gather detailed measurements on the spectrum usage [39–41]. These measurements are provided to the international cognitive radio and dynamic spectrum access research community and can be used to validate and analyze the performance of the proposed routing solutions. Furthermore, recent papers have demonstrated the cognitive radio over FM bands via the Universal Software Radio Peripheral ([42]) and the feasibility of supporting Wifi connections in TV white spaces [43]. The integration of the cognitive radio routing on these platforms is a next objective.

References

[1] F.C. Commission, Spectrum policy task force, Technical report, November 2002.
[2] E. Buracchini, The software radio concept, IEEE Communication Magazine 38 (9) (2000) 138–143.
[3] J. Mitola, Cognitive radio, An Integrated Agent Architecture for Software Defined Radio, PhD Dissertation Thesis, KTH, Sweden.
[4] I.F. Akyildiz, W.Y. Lee, M. Vuran, S. Mohanty, A survey on spectrum management in cognitive radio networks, IEEE Communication Magazine 46 (4) (2008) 40–48.
[5] I.F. Akyildiz, W.-Y. Lee, M.C. Vuran, S. Mohanty, Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey, Computer Networks 50 (13) (2006) 2127–2159. doi: 10.1016/j.comnet.2006.05.001.
[6] C. Cormio, K.R. Chowdhury, A survey on MAC protocols for cognitive radio networks, Ad Hoc Networks 7 (7) (2009) 1315–1329, doi:10.1016/j.adhoc.2009.01.002.
[7] S. Haykin, J.H. Reed, G.Y. Li, M. Shafi, Scanning the issue, Proceedings of the IEEE 97 (5) (2009) 784–786, doi:10.1109/JPROC.2009.2015701.
[8] K. Chowdhury, I. Akyildiz, Cognitive wireless mesh networks with dynamic spectrum access, IEEE Journal on Selected Areas in Communications 26 (1) (2008) 168–181, doi:10.1109/JSAC.2008.080115.
[9] I.F. Akyildiz, W.-Y. Lee, K.R. Chowdhury, CRAHNs: cognitive radio ad hoc networks, Ad Hoc Networks 7 (5) (2009) 810–836, doi:10.1016/j.adhoc.2009.01.001.
[10] FCC, FCC 08-260, Unlicensed Operation in the TV Broadcast Bands, November 2008.
[11] H. Khalife, N. Malouch, S. Fdida, Multihop cognitive radio networks: to route or not to route, IEEE Network Magazine 23 (4) (2009) 20–25.
[12] C. Xin, B. Xie, C.-C. Shen, A novel layered graph model for topology formation and routing in dynamic spectrum access networks, in: First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2005, pp. 308–317, doi: 10.1109/DYSPAN.2005.1542647.
[13] C. Xin, L. Ma, C.-C. Shen, A path-centric channel assignment framework for cognitive radio wireless networks, Mobile Networks and Applications 13 (5) (2008) 463–476. doi: 10.1007/s11036-008-0084-y.
[14] X. Zhou, L. Lin, J. Wang, X. Zhang, Cross-layer routing design in cognitive radio networks by colored multigraph model, Wireless Personal Communications 49 (1) (2009) 123–131. doi: 10.1007/s11277-008-9561-7.
[15] Q. Wang, H. Zheng, Route and spectrum selection in dynamic spectrum networks, in: 3rd IEEE Consumer Communications and Networking Conference, CCNC 2006, vol. 1, 2006, pp. 625–629.
[16] Y. Hou, Y. Shi, H. Sherali, Optimal spectrum sharing for multi-hop software defined radio networks, in: 26th IEEE International Conference on Computer Communications, INFOCOM 2007, 2007, pp. 1–9, doi: 10.1109/INFCOM.2007.9.
[17] Y. Hou, Y. Shi, H. Sherali, Spectrum sharing for multi-hop networking with cognitive radios, IEEE Journal on Selected Areas in Communications 26 (1) (2008) 146–155, doi:10.1109/JSAC.2008.080113.
[18] M. Ma, D. Tsang, Joint spectrum sharing and fair routing in cognitive radio networks, in: 5th IEEE Consumer Communications and Networking Conference, CCNC 2008, 2008, pp. 978–982, doi: 10.1109/ccnc08.2007.225.
[19] C.W. Pyo, M. Hasegawa, Minimum weight routing based on a common link control radio for cognitive wireless ad hoc networks, in: IWCMC '07: Proceedings of the 2007 International Conference on Wireless Communications and Mobile Computing, 2007, pp. 399–404, doi: 10.1145/1280940.1281026.
[20] Y. Shi, Y. Hou, A distributed optimization algorithm for multi-hop cognitive radio networks, in: The 27th IEEE Conference on Computer Communications, INFOCOM 2008, 2008, pp. 1292–1300, doi: 10.1109/INFOCOM.2008.186.

[21] M. Xie, W. Zhang, K.-K. Wong, A geometric approach to improve spectrum efficiency for cognitive relay networks, IEEE Transactions on Wireless Communications 9 (1) (2010) 268–281.

[22] H. Ma, L. Zheng, X. Ma, Y. luo, Spectrum aware routing for multi-hop cognitive radio networks with a single transceiver, in: 3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications, CrownCom 2008, 2008, pp. 1–6, doi: 10.1109/CROWNCOM.2008.4562509.

[23] G. Cheng, W. Liu, Y. Li, W. Cheng, Spectrum aware on-demand routing in cognitive radio networks, in: 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2007, 2007, pp. 571–574, doi: 10.1109/DYSPAN.2007.80.

[24] G. Cheng, W. Liu, Y. Li, W. Cheng, Joint on-demand routing and spectrum assignment in cognitive radio networks, in: IEEE International Conference on Communications, ICC '07, 2007, pp. 6499–6503, doi: 10.1109/ICC.2007.1075.

[25] Z. Yang, G. Cheng, W. Liu, W. Yuan, W. Cheng, Local coordination based routing and spectrum assignment in multi-hop cognitive radio networks, Mobile Networks and Applications 13 (1-2) (2008) 67–81. doi: 10.1007/s11036-008-0025-9.

[26] H.-P. Shiang, M. van der Schaar, Distributed resource management in multihop cognitive radio networks for delay-sensitive transmission, IEEE Transactions on Vehicular Technology 58 (2) (2009) 941–953, doi:10.1109/TVT.2008.925308.

[27] I. Pefkianakis, S. Wong, S. Lu, SAMER: spectrum aware mesh routing in cognitive radio networks, in: 3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2008, 2008, pp. 1–5, doi: 10.1109/DYSPAN.2008.90.

[28] L. Ding, T. Melodia, S. Batalama, M.J. Medley, Rosa: distributed joint routing and dynamic spectrum allocation in cognitive radio ad hoc networks, in: MSWiM '09: Proceedings of the 12th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, 2009, pp. 13–20, doi: 10.1145/1641804.1641810.

[29] A. Sampath, L. Yang, L. Cao, H. Zheng, B.Y. Zhao, High throughput spectrum-aware routing for cognitive radio based ad-hoc networks, in: 3th International Conference on Cognitive Radio Oriented Wireless Networks and Communications, CrwownCom 2008, 2008.

[30] G.-M. Zhu, I. Akyildiz, G.-S. Kuo, STOD-RP: a spectrum-tree based on-demand routing protocol for multi-hop cognitive radio networks, in: IEEE Global Telecommunications Conference, GLOBECOM 2008, 2008, pp. 1–5, doi: 10.1109/GLOCOM.2008.ECP.592.

[31] W. Feng, J. Cao, C. Zhang, C. Liu, Joint optimization of spectrum handoff scheduling and routing in multi-hop multi-radio cognitive networks, in: ICDCS '09: Proceedings of the 2009 29th IEEE International Conference on Distributed Computing Systems, 2009, pp. 85–92, doi: 10.1109/ICDCS.2009.64.

[32] A. Abbagnale, F. Cuomo, Gymkhana: a connectivity-based routing scheme for cognitive radio ad hoc networks, in: IEEE Conference on Computer Communications, INFOCOM 2010, pp. 1–5, doi:10.1109/INFCOMW.2010.5466618.

[33] I. Filippini, E. Ekici, M. Cesana, Minimum maintenance cost routing in cognitive radio networks, in: Mobile Adhoc and Sensor Systems, 2009, MASS '09. IEEE 6th International Conference on, 2009, pp. 284 –293, doi: 10.1109/MOBHOC.2009.5336987.

[34] K. Chowdhury, M. Felice, Search: a routing protocol for mobile cognitive radio ad-hoc networks, Computer Communications 32 (18) (2009) 1983–1997, doi:10.1016/j.comcom.2009.06.011. Cognitive Radio and Dynamic Spectrum Sharing Systems <http://www.sciencedirect.com/science/article/B6TYP-4WMM7DS-/2/e2ba46e19d1d0c8189a178202085e8af>.

[35] H. Khalife, S. Ahuja, N. Malouch, M. Krunz, Probabilistic path selection in opportunistic cognitive radio networks, in: IEEE Global Telecommunications Conference, GLOBECOM 2008, 2008, pp. 1–5, doi: 10.1109/GLOCOM.2008.ECP.931.

[36] S.-U. Yoon, E. Ekici, Voluntary spectrum handoff: a novel approach to spectrum management in CRNs, in: IEEE International Conference on Communications, ICC 2010, 2010.

[37] R. Murawski, E. Ekici, Backward-compatible dynamic spectrum leasing for 802.11-based wireless networks, in: IEEE Global Communications Conference, GLOBECOM 2010, accepted for publication.

[38] T. Elkourdi, O. Simeone, An information-theoretic view of spectrum leasing via secondary cooperation, in: IEEE International Conference on Communications, ICC 2010, 2010.

[39] A. Wyglinski, Spectrum Query Utility Interface for Realtime Radio Electromagnetics Web Interface, 2010 <http://www.spectrum.wpi.edu/>.

[40] iNETS, Static Spectrum Occupancy Measurement Data, 2010 <http://download.mobnets.rwth-aachen.de/>.

[41] M. Wellens, P. Mähönen, Lessons learned from an extensive spectrum occupancy measurement campaign and a stochastic duty cycle model, Mobile Networks and Applications 15 (2010) 461–474, doi:10.1007/s11036-009-0199-9.

[42] R. Zhou, O. Mian, X. Li, B. Wang, Z. Wu, A software-defined radio based cognitive radio demonstration over FM band, Wireless Communications and Mobile Computing 10 (2010) 4–15, doi:10.1002/wcm.v10:1.

[43] P. Bahl, R. Chandra, T. Moscibroda, R. Murty, M. Welsh, White space networking with wi-fi like connectivity, SIGCOMM Computer Communication Review 39 (4) (2009) 27–38. doi: 10.1145/1594977.1592573.

Matteo Cesana received his MS degree in Telecommunications Engineering and his Ph.D. degree in Information Engineering from the Politecnico di Milano in July 2000 and in September 2004, respectively. From September 2002 to March 2003 he has been working as a visiting researcher at the Computer Science Department of the University of California in Los Angeles (UCLA). He is now an Assistant Professor of the Electronics and Information Department of the Politecnico di Milano. His research activities are in the field of performance evaluation of cellular systems, ad-hoc networks protocol design and evaluation and wireless networks optimization. He is an Associate Editor of Ad Hoc Networks Journal (Elsevier).

Francesca Cuomo received her "Laurea" degree in Electrical and Electronic Engineering in 1993, magna cum laude, from the University of Rome "La Sapienza", Italy. She earned the Ph.D. degree in Information and Communications Engineering in 1998, also from the University of Rome "La Sapienza". From 1996 to 2005 she has been an Assistant Professor at the INFOCOM Department of this University.

From 2005 she is Associate Professor at the University of Rome "La Sapienza" and teaches courses in Telecommunication Networks. Cuomo has advised numerous master students in computer science, and has been the advisor of 5 Ph.D. students in Networking at Rome University "La Sapienza". Her main research interests focus on: Wireless ad-hoc and Sensor networks, Cognitive Radio Networks, Reconfigurable radio systems, Quality of Service guarantees and real time service support in the Internet and in the radio access, Mobile and Personal Communications, Architectures and protocol for fixed an mobile wireless networks Modeling and Control of broadband networks, Signaling and Intelligent Networks.

She has participated in several National and European projects on wireless network systems, such as the RAMON, VICOM, INSYEME, IST WHYLESS, IST EPERSPECE, IST CRUISE, projects.

She has authored over 60 peer-reviewed papers published in prominent international journals and conferences.

She has been in the editorial board of the Elsevier Computer Networks journal and now is member of the editorial board of the journal Ad-Hoc Networks (Elsevier). She has served on several technical program committees including ACM Wireless Mobile Internet Workshop, ACM MobiHoc, IEEE INFOCOM (from 2008 to 2011), SECON (2009–2010), ICC, GLOBECOM, VTC, Med-Hoc-Net, WONS, WICON, ACM PE-WASUN. She served as reviewer in several international conferences and journals including IEEE Trans. on Wireless Communications, IEEE Journal on Selected Areas on Communications, IEEE Transactions on Mobile Computing, , IEEE Transactions on Networking, ACM Transactions on Sensor Networks.

Eylem Ekici received his B.S. and M.S. degrees in computer engineering from Bogazic ¸i University, Istanbul, Turkey, in 1997 and 1998, respectively, and his Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, in 2002. Currently, he is an Associate Professor with the Department of Electrical and Computer Engineering, The Ohio State University. 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. He is an Associate Editor of IEEE/ACM Transactions on Networking, Computer Networks Journal (Elsevier), and ACM Mobile Computing and Communications Review.