Spectrum-Aware Data Replication in Intermittently Connected Cognitive Radio Networks

Jing Zhao and Guohong Cao
Department of Computer Science and Engineering
The Pennsylvania State University
E-mail: {juz139,gcao}@cse.psu.edu

***Abstract*—The opening of under-utilized spectrum creates an opportunity for unlicensed users to achieve substantial performance improvement through cognitive radio techniques. In cognitive radio ad-hoc networks, with node mobility and low node density, the network topology is highly dynamic and end-to-end connection is hard to maintain. We propose data replication techniques to address these problems and improve data access performance in such intermittently connected cognitive radio network. Although data replication has been extensively studied in traditional disruption tolerant networks, existing techniques cannot be directly applied here since they do not consider the effects of primary user appearance on data replication. In this paper, we formulate spectrum-aware data replication as an optimization problem which tries to maximize the average data retrieval probability, subject to storage and time constraints. Since the problem is hard to solve based on mixed integer programming, we further design a distributed replication scheme based on the metric of replication benefit. Extensive simulations based on synthetic and realistic traces show that our scheme outperforms existing schemes in terms of data retrieval probability in various scenarios.**

I. Introduction

The past few years have witnessed the proliferation of wireless devices (e.g., cell phones, tablets, and laptops), accompanied by the explosion of wireless applications such as location-based services, mobile healthcare, remote education, home entertainment systems, etc. Most of these devices are unlicensed and have to operate in the public ISM bands which are becoming increasingly congested. Meanwhile, some other licensed spectrum bands (e.g., TV bands) are extremely under-utilized. To address this problem, FCC approved unlicensed use of licensed spectrum through cognitive radio techniques [1], which enable dynamic configuration of the operating spectrum.

Depending on the network architecture, cognitive radio networks can be either infrastructure-based or ad-hoc based, where cognitive radio ad-hoc networks [2], [3], [4], [5] do not require any infrastructure support. They allow unlicensed users to detect available licensed channels and then establish connections by themselves. In some cases, users are only intermittently connected when they move into the communication range of each other (called *contact*). Such intermittently connected cognitive radio network can be viewed as a special case of Disruption Tolerant Network (DTN) [6], which has applications in battlefield, disaster recovery, environmental monitoring, habitat monitoring, transportation, 3G offloading, etc. Due to mobility and limited range of the wireless communication, the contact duration is usually short. Thus, it is hard to transmit large amount of data such as video, especially considering that most mobile devices use unlicensed ISM bands for peer-to-peer communication. With cognitive radio techniques, the licensed spectrum can be opportunistically exploited to increase the data transmission capacity among these mobile devices. However, data access will be more complex, since we not only need to consider the probability of nodes reaching the destination, but also consider the data transmission capacity which is affected by the primary user appearance.

We propose data replication techniques to improve the performance of data access, in terms of data access delay and data availability, in intermittently connected cognitive radio networks. Here, the key problem is where to replicate the data to minimize the data access delay and increase the data availability. Although data replication has been extensively studied in traditional disruption tolerant networks [7], [8], [9], [10], existing techniques cannot be directly applied since they do not consider the effects of primary user appearance on data replication. In intermittently connected cognitive radio networks, unlicensed users at different regions are generally affected by the primary users at that area during the data transmission time. Such spatial and temporal varying spectrum availability affects the data access delay and the data retrieval probability, and hence affects the data replication strategy.

For example, suppose node A frequently contacts other nodes compared to node B. If spectrum availability is not considered, replicating data at A is better since the replicated data (at A) can be easily accessed by other nodes and hence the data access delay will be shorter. However, the information of spectrum availability may change the decision on where to replicate the data. Suppose the contacts between A and others often occur within the activity regions of the primary users. Then, less amount of data can be transmitted upon contact, and the replicated data at A has less chances to be retrieved by other nodes. In contrast, suppose the contacts between B and others often occur outside the activity regions of the primary users. Then, the replicated data at B has better chances to be retrieved by others during the contact. Thus, we should jointly consider node contact pattern and primary user appearance in determining where to replicate data, which brings more

This work was supported in part by the US National Science Foundation (NSF) under grant number CNS-1320278 and by Network Science CTA under grant W911NF-09-2-0053.

978-14799-3360-0/14/$31.00 ©2014 IEEE

challenges in designing appropriate data replication strategies.

We propose a spectrum-aware data replication scheme to address the aforementioned challenge. Due to high node mobility and limited channel capacity, the amount of data that can be transmitted during a contact is limited. As a result, we not only have to determine where to replicate, but also how much to replicate at a node. The decision will be based on the node contact frequency, the primary user appearance, and the node mobility pattern. The contributions of the paper are three-fold:

- To the best of our knowledge, this is the first paper to study spectrum-aware data replication in intermittently connected cognitive radio networks. We formulate the spectrum-aware data replication problem to determine the optimal replication location, which maximizes the average data retrieval probability, subject to storage and time constraints.
- We calculate the data transmission capacity and the data retrieval probability by using discrete-time Markov chains to model node mobility and primary user appearance.
- We further propose a distributed packet-level replication scheme, whose effectiveness is validated through extensive simulations.

The remainder of the paper is organized as follows. Section II reviews related work. In Section III, we provide an overview of our work. Section IV formulates and analyzes the spectrum-aware data replication problem. Section V presents the proposed spectrum-aware replication scheme in detail. We show evaluation results in Section VI, and conclude the paper in Section VII.

II. Related Work

Most existing solutions on data access in cognitive radio networks assume the existence of an end-to-end path between the data source and data requesters. They focus on designing efficient routing protocols to minimize the routing delay or maximize the throughput. For example, an on-demand protocol [11] has been proposed to minimize the end-to-end delay through joint route selection and spectrum assignment. In [12], Pefkianakis et al. proposed a routing protocol to select the route with the highest available spectrum. However, neither of them considers node mobility. In [13], Chowdhurya et al. proposed a geographic forwarding based spectrum aware routing protocol for cognitive radio ad-hoc networks that can adapt to dynamic spectrum availability and node mobility. However, their protocol is based on AODV [14] which has to establish a route to the destination, and hence not suitable for intermittently connected networks.

Data replication has been widely used to improve the performance of data access in mobile ad-hoc networks [15], [16], [17] and DTNs [7], [8], [9], [10]. In [15], the problem of finding the optimal replication location is formalized as a special case of the NP-hard *connected facility location problem* [18], and then solved by using a greedy algorithm which is within a factor of 6 of the optimal solution. If there are multiple data items, multiple nodes may share and coordinate their replicated data [16], [17]. Data replication is also studied in cognitive radio networks to meet the delay constraints [19]. However, these works do not consider mobility. For DTNs, distributed data replication schemes have been proposed in [7], [8] by assuming a complete data item can be transmitted when two nodes contact each other. In reality, the contact duration is usually short (due to mobility and limited range of peer-to-peer wireless communication), so a complete data item may not be transmitted upon contact. Zhuo et al. [10] addressed this problem through erasure coding and packet-level replication. However, in intermittently connected cognitive radio networks, the data transmission capacity is also affected by primary user appearance, which has not been considered in these works.

III. Overview

We consider an intermittently connected cognitive radio network consisting of mobile unlicensed users (nodes) whose communications may be affected by the primary users. Each data item is generated by the data source, and requested by other nodes with some query rate. To help data requesters retrieve data within a time constraint, each node provides some storage space for replicating the data items. Due to node mobility and the appearance of primary users, there is no persistent network connection. As a result, a node can only forward data to another node if they are within the communication range and have available communication channels.

The data transmission capacity (i.e., the amount of data that can be transmitted upon contact) depends on the amount of available spectrum. The data transmission capacity will be reduced due to primary user appearance, and a large data item (e.g., large video) may not be transmitted completely upon contact, especially when the contact duration is short. If the data is simply fragmented and only a part of it is transmitted during each contact, the well-known coupon collector problem [20] will appear, where the node may keep looking for the last fragment of the data which is hard to find. To mitigate this problem, we adopt the erasure coding technique [21] to encode data into a large set of coded packets, and then any sufficiently large subset of coded packets can be used to reconstruct the data. Thus, data replication is performed at the coded packet level.

Our goal is to decide which data items and how many packets should be replicated by each node, in order to maximize the average data retrieval probability. We can formulate it as a spectrum-aware data replication problem, which is hard to solve based on mixed integer programming. In our heuristic based approach, each node greedily replicates the packet that brings the maximum replication benefit until the storage space is fully utilized. The replication benefit depends on the data retrieval probability, which is affected by both node movement and primary user appearance. In the next two sections, we first calculate the data retrieval probability by using discrete-time Markov chains to model node movement and primary user appearance, and then describe our distributed packet-level replication scheme in detail.

IV. Problem Formulation and Analysis

In this section, we formally define the spectrum-aware data replication problem and propose a solution to efficiently calculate the data retrieval probability. Finally, we discuss a centralized solution based on mixed integer programming.

A. System Models

Following existing works [22], [23], the movement of a node v is modeled by a discrete-time Markov chain X_v^t, whose states are represented by the locations (the entire area can be divided into a set of grids and each grid defines a location). The set of locations is denoted by $\mathcal{L} = \{1, \ldots, L\}$. We use $p_v^{i,j}$ to denote the probability of node v to transition from location i to location j.

Each location is affected by a number of primary users on each channel. Let $Y_{l,c}^t$ denote the availability of channel c at location l at time t. That is, $Y_{l,c}^t$ is 1 if channel c is available for unlicensed users at location l at time t (not accessed by primary users); otherwise, $Y_{l,c}^t$ is 0. Based on existing works [24], [25], [26], we assume $Y_{l,c}^t$ follows a discrete-time Markov chain with two states 0, 1, and use $q_{l,c}^{i,j}$ to denote the probability of $Y_{l,c}^t$ to transition from state i to state j. Note that an unlicensed channel can also be modeled here. If $q_{l,c}^{0,1} = q_{l,c}^{1,1} = 1$ and $q_{l,c}^{0,0} = q_{l,c}^{1,0} = 0$, channel c can be viewed as an unlicensed channel which is never accessed by primary users.

1) The Network Model: We consider a network of N unlicensed users (nodes) moving around L locations. There are C channels, which may be sometimes accessed by primary users. The set of nodes is denoted by $\mathcal{N} = \{1, \ldots, N\}$, and the set of channels is denoted by $\mathcal{C} = \{1, \ldots, C\}$. If two nodes are at location l at the same time t, the total amount of data that can be transmitted between them at time t is $\beta(\sum_{c \in \mathcal{C}} Y_{l,c}^t)$. Here β denotes the channel bandwidth, i.e., the total amount of data that can be transmitted per channel. We assume all channels have equal bandwidth, and a node can use multiple channels for transmission at the same time. In practice, such flexibility in using multiple spectrum bands can be achieved by k-agile software-defined radios [27].

2) Data Query: There are M data items of the same size G in the network. The set of data items is denoted by $\mathcal{M} = \{1, \ldots, M\}$. If the rate of erasure coding process is $\frac{1}{R}$, the data item of size G is coded into GR/g uniformly sized packets of size g. Then any $S = (1+\epsilon)G/g$ coded packets can be used to reconstruct the original data item, where ϵ is a parameter used in the erasure coding algorithm.

Each data item is generated by some node in the network and requested by others. The query rate for data item d by node v is denoted as $\lambda_{v,d}$. Suppose node v generates a query for data item d at time 0. The query succeeds if and only if node v can reconstruct data item d from the received coded packets within time constraint T.

To simplify presentation, we assume uniform data size and uniform time constraint in this paper. Note that our approach can be easily extended to cases where data items have different sizes and time constraints as follows. Suppose data item d of size G_d can be reconstructed by any S_d coded packets, and its time constraint is T_d. To calculate the data retrieval probability of data item d, the proposed methods can still be applied after replacing G (S, T) with G_d (S_d, T_d).

B. Problem Definition

Definition 1: **Spectrum-Aware Data Replication Problem** We use a $N \times M$ matrix $\mathbf{z}$ to represent the data replication solution, where each element $z_{v,d}$ denotes the number of coded packets of data item d that are replicated by node v. Let $A_{v,d}(\mathbf{z})$ denote the total number of coded packets of data d that node v can retrieve from others within the time constraint T, with data replication solution $\mathbf{z}$. Let ρ_v denote the maximum number of packets that can be replicated by node v. The problem is to maximize the average data retrieval probability, subject to time and storage constraints.

$$\textbf{maximize} \sum_{v \in \mathcal{N}} \sum_{d \in \mathcal{M}} \lambda_{v,d} P(A_{v,d}(\mathbf{z}) \geq S - z_{v,d}) \quad (1)$$

subject to

$$z_{v,d} \in \{0, \ldots, S\}, \forall v \in \mathcal{N}, \forall d \in \mathcal{M} \quad (2)$$

$$\sum_{d \in \mathcal{M}} z_{v,d} \leq \rho_v, \forall v \in \mathcal{N} \quad (3)$$

The objective function (1) is the average data retrieval probability which is to be maximized. Constraint (2) ensures that each node replicates at most S coded packets for each data item since a data item can be reconstructed by any S packets of it. Constraint (3) ensures that the total number of packets replicated by each node does not exceed the storage space of that node.

C. Calculating the Data Retrieval Probability

In this subsection, we calculate the data retrieval probability $P(A_{v,d}(\mathbf{z}) \geq S - z_{v,d})$. We illustrate the calculation steps in Figure 1 (with the notations summarized in Table I) and describe the details as follows.

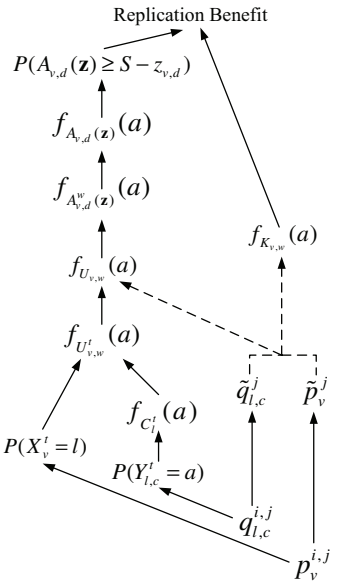


Fig. 1. Steps for calculating the data retrieval probability and evaluating the replication benefit (in Section V-B)

TABLE I
TABLE OF NOTATIONS

Symbols	Meaning
$A_{v,d}(\mathbf{z})$	total number of packets of data item d that node v can retrieve from others within the time constraint, given data replication solution $\mathbf{z}$
$A^w_{v,d}(\mathbf{z})$	total number of packets of data item d that node v can retrieve from node w within the time constraint, given data replication solution $\mathbf{z}$
$U_{v,w}$	total amount of data that can be transmitted from node w to node v within the time constraint
$U^t_{v,w}$	total amount of data that can be transmitted from node w to node v at time t
C^t_l	number of available channels that can be used at location l at time t
X^t_v	location of node v at time t
$Y^t_{l,c}$	availability of channel c at location l at time t
$p^{i,j}_v$	probability of node v to transition from location i to j
$q^{i,j}_{l,c}$	probability of $Y^t_{l,c}$ to transition from state i to j
$\tilde{p}^j_v$	stationary probability of node v to be at location j
$\tilde{q}^j_{l,c}$	stationary probability of channel c to be at state j at location l
$K_{v,w}$	maximum number of packets that can be transmitted from node w to node v within the time constraint

1) Calculation of $P(A_{v,d}(\mathbf{z}) \geq S - z_{v,d})$: In Definition 1, there remains a problem of how to derive the closed form expression of $P(A_{v,d}(\mathbf{z}) \geq S - z_{v,d})$. Note that

$$P(A_{v,d}(\mathbf{z}) \geq S - z_{v,d}) = \sum_{a=S-z_{v,d}}^{\infty} f_{A_{v,d}(\mathbf{z})}(a) \tag{4}$$

where $f_{A_{v,d}(\mathbf{z})}(a)$ is the probability mass function of $A_{v,d}(\mathbf{z})$. Now the problem becomes how to calculate $f_{A_{v,d}(\mathbf{z})}(a)$.

Let $A^w_{v,d}(\mathbf{z})$ be the total number of coded packets of data d that node v can retrieve from node w within the time constraint T, given data replication solution $\mathbf{z}$. Since $A_{v,d}(\mathbf{z}) = \sum_{w \neq v} A^w_{v,d}(\mathbf{z})$, we know from [20] that

$$f_{A_{v,d}(\mathbf{z})}(a) = \bigotimes_{w \neq v} f_{A^w_{v,d}(\mathbf{z})}(a) \tag{5}$$

where $f_{A^w_{v,d}(\mathbf{z})}(a)$ is the probability mass function of $A^w_{v,d}(\mathbf{z})$. Equation (5) indicates that $f_{A_{v,d}(\mathbf{z})}(a)$ is a discrete convolution of $f_{A^w_{v,d}(\mathbf{z})}(a)$ (for $\forall w \neq v$).

Let $U_{v,w}$ be the total amount of data that can be transmitted from node w to node v within the time constraint. Then we have

$$f_{A^w_{v,d}(\mathbf{z})}(a) = \begin{cases} \int_{ga}^{g(a+1)} f_{U_{v,w}}(u)du & 0 \leq a \leq z_{w,d} \\ \int_{gz_{w,d}}^{\infty} f_{U_{v,w}}(u)du & a = z_{w,d} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where $f_{U_{v,w}}(a)$ is the probability mass function of $U_{v,w}$. The meaning of equation (6) is as follows. If $0 \leq a \leq z_{w,d}$, in order for node v to receive a coded packets of data item d from node w, the total amount of data that can be transmitted ($U_{v,w}$) has to be within the range of $[ga, g(a+1))$ (g is the packet size). If node v receives $z_{w,d}$ coded packets from node w, $U_{v,w}$ has to be at least $gz_{w,d}$. It is impossible for the number of packets received from node w to exceed $z_{w,d}$ (which is the number of packets replicated by node w).

Let $U^t_{v,w}$ be the total amount of data that can be transmitted from node w to node v at time t. Since $U_{v,w} = \sum_{t=0}^{T} U^t_{v,w}$, we have

$$f_{U_{v,w}}(a) = \bigotimes_{t \in \{0,\ldots,T\}} f_{U^t_{v,w}}(a) \tag{7}$$

where $f_{U^t_{v,w}}(a)$ is the probability mass function of $U^t_{v,w}$.

Let C^t_l be the number of available channels that can be used at location l at time t. Then we have

$$f_{U^t_{v,w}}(\beta a) = \sum_{l \in \mathcal{L}} f_{C^t_l}(a) P(X^t_v = l) P(X^t_w = l) \tag{8}$$

where $f_{C^t_l}(a)$ is the probability mass function of C^t_l. $P(X^t_v = l)$ can be calculated as follows. Suppose each node has equal probability to be at all locations at time 0. That is, $P(X^0_v = l) = 1/L$ for $\forall v \in \mathcal{N}, \forall l \in \mathcal{L}$.

Since X^t_v follows a discrete-time Markov chain, $P(X^t_v = l)$ can be calculated from $P(X^0_v)$ using the following recurrence:

$$P(X^t_v = j) = \sum_{i \in \mathcal{L}} p^{i,j}_v P(X^{t-1}_v = i), \forall t \in \{1, \ldots, T\} \tag{9}$$

Since $C^t_l = \sum_{c \in \mathcal{C}} Y^t_{l,c}$, we have

$$f_{C^t_l}(a) = \bigotimes_{c \in \mathcal{C}} f_{Y^t_{l,c}}(a) \tag{10}$$

where $f_{Y^t_{l,c}}(a)$ is the probability mass function of $Y^t_{l,c}$. Suppose all channels are available at each location at time 0. That is, $P(Y^0_{l,c} = 0) = 0$ and $P(Y^0_{l,c} = 1) = 1$ for $\forall l \in \mathcal{L}, \forall c \in \mathcal{C}$.

Since $Y^t_{l,c}$ follows a discrete-time Markov chain, $P(Y^t_{l,c} = a)$ (a.k.a., $f_{Y^t_{l,c}}(a)$) can be calculated from $P(Y^0_{l,c})$ using the following recurrence:

$$P(Y^t_{l,c} = j) = \sum_{i \in \{0,1\}} q^{i,j}_{l,c} P(Y^{t-1}_{l,c} = i), \forall t \in \{1, \ldots, T\} \tag{11}$$

To summarize, the closed form expression of $P(A_{v,d}(\mathbf{z}) \geq S - z_{v,d})$ can be derived from a series of substitutions using Equations (4)-(11). However, the calculation of $P(A_{v,d}(\mathbf{z}) \geq S - z_{v,d})$ is still too complicated. We reduce its computational complexity through the following approximation techniques.

2) Approximate Calculations: We simplify the calculation of $P(A_{v,d}(\mathbf{z}) \geq S - z_{v,d})$ through approximate calculation of $U_{v,w}$. Specifically, $U_{v,w}$ is approximately calculated based on the stationary probabilities related to node movement and primary user appearance.

In general, $\lim_{n\to\infty} P(X^{t+n}_v = j | X^t_v = i)$ ($\lim_{n\to\infty} P(Y^{t+n}_{l,c} = j | Y^t_{l,c} = i)$) exists and is independent of i. Define

$$\tilde{p}^j_v = \lim_{n\to\infty} P(X^{t+n}_v = j), \forall j \in \mathcal{L} \tag{12}$$

$$\tilde{q}^j_{l,c} = \lim_{n\to\infty} P(Y^{t+n}_{l,c} = j), \forall j \in \{0, 1\} \tag{13}$$

where $\tilde{p}^j_v$ ($\tilde{q}^j_{l,c}$) can be solved by

$$\begin{cases} \tilde{p}^j_v = \sum_{i \in \mathcal{L}} p^{i,j}_v \tilde{p}^i_v, \forall j \in \mathcal{L} \\ \sum_{j \in \mathcal{L}} \tilde{p}^j_v = 1 \end{cases} \tag{14}$$

$$\begin{cases} \tilde{q}^j_{l,c} = \sum_{i\in\{0,1\}} q^{i,j}_{l,c}\tilde{q}^i_{l,c}, \forall j \in \{0,1\} \\ \sum_{j\in\{0,1\}} \tilde{q}^j_{l,c} = 1 \end{cases} \quad (15)$$

If t is large enough, $P(X^t_v = j)$ ($P(Y^t_{l,c} = j)$) will be very close to the stationary probability $\tilde{p}^j_v$ ($\tilde{q}^j_{l,c}$). In disruption tolerant networks, the time constraint is usually loose (i.e., t is usually large), so we can use $\tilde{p}^j_v$ ($\tilde{q}^j_{l,c}$) to approximate $P(X^t_v = j)$ ($P(Y^t_{l,c} = j)$). Then the expected amount of data that can be transmitted from node w to node v at time t (i.e., $E(U^t_{v,w})$) is equal to $\beta\sum_{l\in\mathcal{L}}\sum_{c\in\mathcal{C}}\tilde{p}^l_v\tilde{p}^l_w\tilde{q}^1_{l,c}$.

Let $\mathcal{U}$ denote $E(U^t_{v,w})$, and $\mathcal{P}$ denote $\sum_{l\in\mathcal{L}}\tilde{p}^l_v\tilde{p}^l_w$. We assume that nodes v and w contact each other with probability $\mathcal{P}$, and that upon contact at time t, the total amount of data that can be transmitted from node w to node v is equal to $\mathcal{U}$. Under this model, it is impossible for the amount of data that can transmitted from node w to node v within the time constraint to exceed $\mathcal{U}T$. Thus, $f_{U_{v,w}}(a) = 0$ if $a > \mathcal{U}T$. Now suppose $a \le \eta T$. When $a \in ((n-1)\mathcal{U}, n\mathcal{U}]$ ($n \ge 0$), the total number of contacts between nodes v and w is n, and the probability for making this number of contacts is $\binom{T}{n}\mathcal{P}^n(1-\mathcal{P})^{T-n}$. To summarize, the probability mass function $f_{U_{v,w}}(a)$ can be approximated by the following function:

$$f_{U_{v,w}}(a) = \begin{cases} \frac{1}{\mathcal{U}}\binom{T}{\lceil\frac{a}{\mathcal{U}}\rceil}\mathcal{P}^{\lceil\frac{a}{\mathcal{U}}\rceil}(1-\mathcal{P})^{T-\lceil\frac{a}{\mathcal{U}}\rceil} & a \le \mathcal{U}T \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

Furthermore, let $K_{v,w}$ be the maximum number of packets that can be transmitted from node w to node v within the time constraint T. $K_{v,w}$ can be calculated as:

$$K_{v,w} = \begin{cases} \lfloor\frac{U_{v,w}}{g}\rfloor & \lfloor\frac{U_{v,w}}{g}\rfloor < S \\ S & \text{otherwise} \end{cases} \quad (17)$$

By substituting Equation (17) into Equation (16), $P(K_{v,w} = a)$ is equal to $\frac{\gamma_{v,w}(a)}{\sum_{k=0}^{S}\gamma_{v,w}(k)}$, where $\gamma_{v,w}(a)$ is defined as follows:

(i) $\mathcal{U}T \ge gS$ (we can obtain S packets within the time constraint T):

$$\gamma_{v,w}(a) = \begin{cases} \binom{T}{\lceil\frac{ga}{\mathcal{U}}\rceil}\mathcal{P}^{\lceil\frac{ga}{\mathcal{U}}\rceil}(1-\mathcal{P})^{T-\lceil\frac{ga}{\mathcal{U}}\rceil} & a < S \\ \sum_{k=\lceil gS/\mathcal{U}\rceil}^{T}\binom{T}{k}\mathcal{P}^k(1-\mathcal{P})^{T-k} & a = S \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

(ii) $\mathcal{U}T < gS$ (we cannot obtain S packets within the time constraint T):

$$\gamma_{v,w}(a) = \begin{cases} \binom{T}{\lceil\frac{ga}{\mathcal{U}}\rceil}\mathcal{P}^{\lceil\frac{ga}{\mathcal{U}}\rceil}(1-\mathcal{P})^{T-\lceil\frac{ga}{\mathcal{U}}\rceil} & a \le \lfloor\mathcal{U}T/g\rfloor \\ 0 & \text{otherwise} \end{cases} \quad (19)$$

D. Mixed Integer Programming Formulation

The spectrum-aware data replication problem has the following mixed integer programming formulation.

Let $\mathcal{X}$ represent the set of all possible patterns of node movement. Each element $X \in \mathcal{X}$ is denoted by an $N \times T$ vector $(X^t_v)_{N\times T}$, where X^t_v denotes the location at which node v is located at time t as aforementioned in Section IV-A.

Let $\mathcal{Y}$ represent the set of all possible patterns of primary user appearance. Each element $Y \in \mathcal{Y}$ is denoted by an $L \times C \times T$ vector $(Y^t_{l,c})_{L\times C\times T}$, where $Y^t_{l,c}$ denotes the availability of channel c at location l at time t as aforementioned in Section IV-A.

Let $R^{X,Y}_{v,d}(\mathbf{z})$ denote whether node v can retrieve enough coded packets (at least S packets) to reconstruct data item d within the time constraint, given node movement pattern X, primary user appearance pattern Y, and replication solution $\mathbf{z}$. $R^{X,Y}_{v,d}(\mathbf{z})$ can be calculated as follows:

$$R^{X,Y}_{v,d}(\mathbf{z}) = \begin{cases} 1 & \sum_{w\in\mathcal{N}}\min(z_{w,d}, \frac{1}{g}U^{X,Y}_{v,w}) \ge S \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

where $U^{X,Y}_{v,w}$ is the amount of data that can be transmitted from node w to node v within the time constraint, given node movement pattern X and primary user appearance pattern Y. Note that the number of packets of data item d that can be transmitted from node w to node v should be bounded by $z_{w,d}$ (the number of packets of data item d that are replicated by node w). $U^{X,Y}_{v,w}$ is equal to $\beta\sum_{t=0}^{T}\sum_{l\in\mathcal{L}}\sum_{c\in\mathcal{C}}Y^t_{l,c}I_{X^t_v,X^t_w}$, where $Y^t_{l,c}$ denotes the availability of channel c at location l at time t as aforementioned in Section IV-A, and $I_{X^t_v,X^t_w}$ is an indicator function. If $X^t_v = X^t_w$, $I_{X^t_v,X^t_w} = 1$; otherwise, $I_{X^t_v,X^t_w} = 0$.

Let P_X denote the probability that the node movement follows pattern X, and Q_Y denote the probability that the primary user appearance follows pattern Y. Here we assume the availability of different channels at each location is independent, and the channel availability at different locations is independent. Then, P_X is equal to $\prod_{v\in\mathcal{N}}\prod_{t=1}^{T}p^{X^{t-1}_v,X^t_v}_{Y^{t-1}_{l,c},Y^t_{l,c}}$, and Q_Y is equal to $\prod_{l\in\mathcal{L}}\prod_{c\in\mathcal{C}}\prod_{t=1}^{T}q^{Y^{t-1}_{l,c},Y^t_{l,c}}_{l,c}$.

The spectrum-aware data replication problem can be redefined as follows:

$$\textbf{maximize} \sum_{v\in\mathcal{N}}\sum_{d\in\mathcal{M}}\sum_{X\in\mathcal{X}}\sum_{Y\in\mathcal{Y}}\lambda_{v,d}P_XQ_YR^{X,Y}_{v,d}(\mathbf{z}) \quad (21)$$

subject to

$$z_{v,d} \in \{0,\ldots,S\}, \forall v \in \mathcal{N}, \forall d \in \mathcal{M} \quad (22)$$

$$\sum_{d\in\mathcal{M}} z_{v,d} \le \rho_v, \forall v \in \mathcal{N} \quad (23)$$

$$\forall X \in \mathcal{X}, \forall Y \in \mathcal{Y}, \forall v \in \mathcal{N}, \forall d \in \mathcal{M}:$$

$$R^{X,Y}_{v,d}(\mathbf{z}) \in \{0,1\} \quad (24)$$

$$\sum_{w\in\mathcal{N}}\min(z_{w,d}, \frac{1}{g}U^{X,Y}_{v,w}) \ge SR^{X,Y}_{v,d}(\mathbf{z}) \quad (25)$$

Constraints (24) and (25) ensure that $R^{X,Y}_{v,d}(\mathbf{z}) = 0$ if $\sum_{w\in\mathcal{N}}\min(z_{w,d}, \frac{1}{g}U^{X,Y}_{v,w}) \ge S$ is unsatisfied. Otherwise, $R^{X,Y}_{v,d}$ must be equal to 1 in order to maximize the objective function. However, the min function makes Constraint (25) nonlinear, so we introduce auxiliary variables $h^{X,Y}_{v,w,d}(\mathbf{z})$ and replace Constraint (25) with the following constraints.

$$\sum_{w\in\mathcal{N}} h^{X,Y}_{v,w,d}(\mathbf{z}) \ge SR^{X,Y}_{v,d}(\mathbf{z}) \quad (26)$$

$$h^{X,Y}_{v,w,d}(\mathbf{z}) \le z_{w,d}, \forall w \in \mathcal{N} \quad (27)$$

$$h^{X,Y}_{v,w,d}(\mathbf{z}) \le \frac{1}{g}U^{X,Y}_{v,w}, \forall w \in \mathcal{N} \quad (28)$$

Now the problem becomes a mixed integer programming problem, which is NP-hard in general. This problem is more complicated due to its exponential number of variables (constraints). For example, since $\mathcal{X}$ and $\mathcal{Y}$ have $L^{N\times T}$ elements and $2^{L\times C\times T}$ elements respectively, the number of variables $h^{X,Y}_{v,w,d}(\mathbf{z})$ is $N^2ML^{N\times T}2^{L\times C\times T}$. Even for a small sized problem with $N = 10$, $M = 10$, $L = 5$, $C = 10$, $T = 50$, the number is 1.15×10^{1105}, which is too big to be loaded into general computer memory by any optimization software (e.g., CPLEX). To address this challenge, we propose the following distributed replication scheme based on some heuristics.

V. Spectrum-Aware Replication Scheme

In this section, we present our distributed spectrum-aware replication scheme.

A. Main Idea

Our data replication scheme is a distributed algorithm that runs locally at each node. Specifically, each node greedily replicates the packet that brings the maximum replication benefit until the storage is fully utilized. Here the key problem is how to evaluate the replication benefit accurately.

A straight-forward solution is based on the increased data retrieval probability if the packet is replicated at the node. However, the average data retrieval probability can only be calculated using the knowledge of all nodes' replication strategies and mobility patterns, which is impossible in a distributed environment.

In our scheme, each node only uses a reasonable amount of information to evaluate the replication benefit, which is based on the number of useful packets contributed by the replication.

B. Replication Benefit

We introduce the concept of contribution and contribution gain, and then give the definition of replication benefit. The contribution (contribution gain) represents the capability of a node to contribute all its replicated packets (the newly replicated packet) to another node.

Definition 2: **Contribution**

The contribution provided by node v to node w in terms of data item d, denoted by $B^w_{v,d}(z_{v,d})$, is the expected number of coded packets of d that v can transmit to w within the time constraint.

$$\begin{aligned} B^w_{v,d}(z_{v,d}) &= E(A^v_{w,d}(z_{v,d})) \\ &= \sum_{a=1}^{z_{v,d}-1} aP(K_{v,w} = a) + \sum_{a=z_{v,d}}^{S} z_{v,d}P(K_{v,w} = a) \end{aligned} \tag{29}$$

where $P(K_{v,w} = a)$ is derived in Section IV-C2.

Definition 3: **Contribution Gain**

The contribution gain provided by node v to w in terms of data item d, denoted by $\Delta B^w_{v,d}(z_{v,d})$, is the increment in $B^w_{v,d}$ by replicating an extra coded packet of d at node v.

$$\Delta B^w_{v,d}(z_{v,d}) = B^w_{v,d}(z_{v,d} + 1) - B^w_{v,d}(z_{v,d}) \tag{30}$$

Note that $\Delta B^w_{v,d}(S) = 0$, since S packets are enough to reconstruct the original data item. The contribution gain provided by a node to itself is defined to be 1, since the node can always use the newly replicated packet to reconstruct the original data item.

The contribution gain $\Delta B^w_{v,d}(z_{v,d})$ is a non-increasing function, which indicates the contribution gain provided by a node decreases as more packets are replicated at that node. This can be explained as follows. For nodes v and w, the limited contact opportunities restrict the number of packets that can be transmitted between them. Even if node v replicates many packets of data d, some of them may never be transmitted to node w. As a result, replicating more packets is less efficient, so it leads to lower contribution gain.

Now we define the replication benefit based on the contribution gain.

Definition 4: **Replication Benefit**

The replication benefit provided by node v in terms of data d, denoted by $B_{v,d}(z_{v,d})$, is the weighted sum of the contribution gain provided by node v to any node w in terms of each data item d.

$$B_{v,d}(z_{v,d}) = \sum_{w\in\mathcal{N}} \lambda_{w,d}\Delta B^w_{v,d}(z_{v,d}) \tag{31}$$

where $\lambda_{w,d}$ is the query rate of node v to data item d, and can be estimated by $\lambda_{w,d} = n_{w,d}/h$. Here node w counts the number of requests in a period of h time units, and $n_{w,d}$ is the number of requests for data item d during h.

As can be seen, the replication benefit can be calculated with a reasonable amount of information. This is because the replication benefit $B_{v,d}(z_{v,d})$ is based on $K_{v,w}$ (through a series of substitutions using Equations (29)-(31)), whose calculation only depends on the contact pattern between node v and other nodes in the network and the pattern of the primary user appearance (as shown in Section IV-C2). These information can be collected by the node itself, and the query rates of other nodes can be exchanged upon contact.

C. The Distributed Protocol

When two nodes contact, they exchange packets replicated in their storage. With software defined radio, full-duplex communications can be achieved such that the packets can be sent to and received from another node at the same time [28]. Therefore, we only need to focus on downloading packets upon contact.

Generally speaking, each node greedily replicates the packet that brings the maximum replication benefit until the storage is full. Suppose node v has already replicated $z_{v,d}$ packets of data item d. When node v contacts another node, it downloads a packet of the data item d_{max} which has the maximum replication benefit $B_{v,d_{max}}(z_{v,d_{max}})$ from the encountering node. Note that the information required for calculating the replication benefit can be collected by node v beforehand as aforementioned in Section V-B.

If the storage is full, node v decides whether to remove a packet to make room for the newly downloaded packet.

If node v removes a packet of data item d, the accumulated replication benefit of data item d will be decreased by $B_{v,d}(z_{v,d}-1)$. We find out the data item d_{min} which has the minimum $B_{v,d_{min}}(z_{v,d_{min}}-1)$ among all the data items. If $B_{v,d_{min}}(z_{v,d_{min}}-1)$ is less than $B_{v,d_{max}}(z_{v,d_{max}})$, the storage replacement brings benefit, so a packet of data item d_{min} is replaced by that of data item d_{max}. Otherwise, there is no update to the storage.

VI. Performance Evaluations

In this section, we evaluate the performance of our spectrum-aware replication scheme based on synthetic and realistic traces.

A. Schemes for Comparisons

To evaluate the performance of our *spectrum-aware replication scheme (**SPEC**)*, we compare it with three existing replication schemes which do not consider primary user appearance: 1. **DARA**: A contact duration aware replication scheme [10]; 2. **UNI**: A replication scheme where the storage space is evenly allocated among all the data items; 3. **PROP**: A replication scheme where the storage allocation is proportional to the data query rate; i.e., frequently accessed data will be replicated with more storage space. For all these schemes, the data replication is at data packet level and erasure coding is used.

B. Synthetic Trace

1) Simulation Settings: We generate a synthetic trace in which there are 20 mobile nodes and 20 data items in the network. We set 20 locations, and the channel availability at each location is determined by our model for primary user appearance (the transition probabilities among different states are randomly generated). Considering that the node moving speed is relatively slow, we assume it takes 100 time units to transition from one location to another. Each data item is generated by some node which is randomly selected, and can be reconstructed by 20 coded packets. In our simulations, we assume each node has equal storage space and can replicate at most 100 coded packets. Following existing works [10], [16], the data query pattern is based on Zipf-like distribution in which the query rate of the ith most popular data item is proportional to $i^{-\theta}$. Here θ shows how skewed the query pattern is, and is set to 0.8 in default according to studies on real Web traces [29].

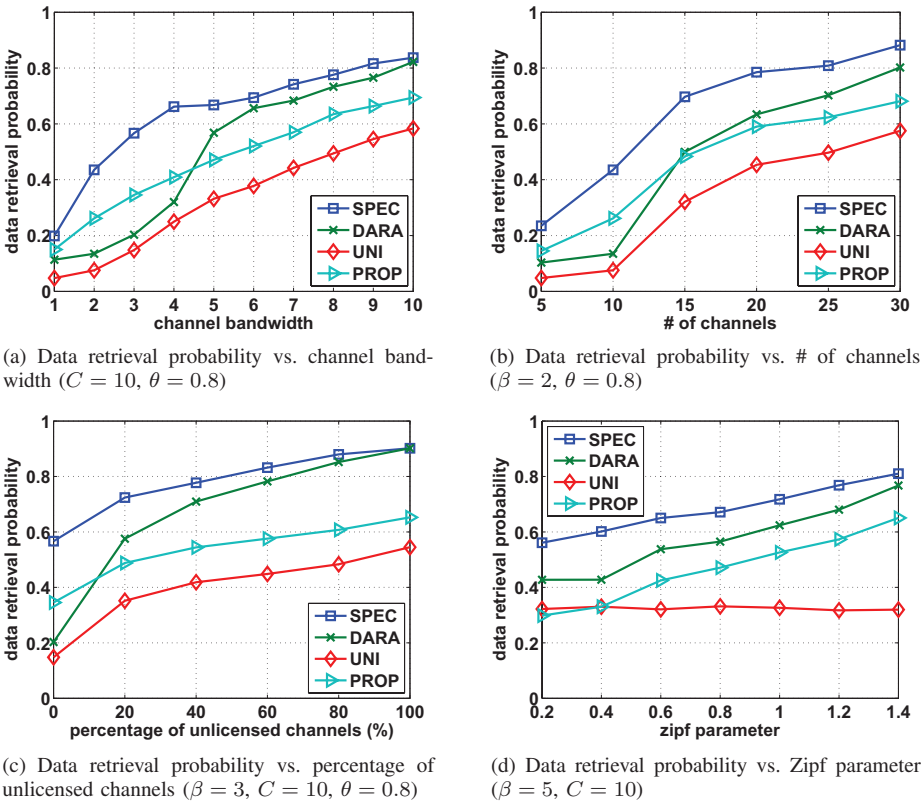
We vary the *channel bandwidth* (β), the *number of channels* (C), and the *Zipf parameter* (θ), to study their effects on the (average) data retrieval probability. The channel bandwidth is the per channel transmission capacity. That is, if there are c available channels, the maximum number of packets that can be transmitted in one time unit is $c\beta$. We also investigate the effect of primary user appearance on the performance. Specifically, we model some channels as unlicensed channels which are never accessed by primary users (following Section IV-A) and study how the percentage of unlicensed channels affects the data retrieval probability. In all simulations, the first half of the trace is used for warmup to collect necessary network information. All the data and queries are generated during the second half of the trace. The presented results are averaged over 100 runs.

2) Simulation Results: Figure 2(a) shows the effect of channel bandwidth on the data retrieval probability. For all schemes, the data retrieval probability increases as the channel bandwidth increases, since more packets can be transmitted upon contact. Among the four schemes, SPEC performs the best, since it considers the effect of primary user appearance on the data replication strategy, which is ignored by the other three schemes. Compared to DARA, UNI and PROP, SPEC improves the data retrieval probability by 75%, 318% and 32% when the channel bandwidth is 1 packet per time unit. When the channel bandwidth reaches 10 packets per time unit, the improvement changes to 2%, 44% and 21%.

When the channel bandwidth is less than 5 packets per time unit, PROP performs the best among the other three schemes. PROP outperforms UNI since PROP allocates more storage space to the data items of high query rate. PROP outperforms DARA due to the following reason. The replication strategy in DARA is based on the data transmission capacity upon each contact. Without considering the primary user appearance, the data transmission capacity cannot be calculated accurately, which affects the performance of DARA. When the channel bandwidth exceeds 5 packets per time unit, DARA outperforms PROP. Increasing the channel bandwidth makes data replication less restricted by the data transmission capacity upon each contact. This reduces the effect of inaccurate calculation of data transmission capacity on the performance of DARA. Meanwhile, DARA considers the node contact pattern which is not considered in PROP, and thus DARA performs better.

Figure 2(b) shows the effect of the number of channels on the data retrieval probability. For all schemes, the data retrieval probability increases as the number of channels increases, since there are generally more available channels to be used for data transmission upon contact. When there are 10 channels in the network, Figure 2(c) shows the effect of unlicensed channels on the data retrieval probability. For all schemes, the data retrieval probability increases as the percentage of unlicensed channels increases, since more packets can be transmitted upon contact by using more available channels. When all channels are unlicensed, DARA and SPEC have the same data retrieval probability of 90% since data replication is only determined by the node contact pattern.

Figure 2(d) shows the effect of Zipf parameter θ on the data retrieval probability. For SPEC, DARA and PROP, the data retrieval probability increases as θ increases. Increasing θ makes the query pattern much skewer, which increases the query rate of popular data items. These three schemes generally replicate more packets of popular data items, so their data retrieval probability increases. For UNI, the performance is similar to that of PROP when the Zipf parameter is 0.2 or 0.4. Small θ indicates similar query rate for all data items, so the replication strategy of UNI is similar to that of PROP. When θ increases, the popular data items are given higher



(a) Data retrieval probability vs. channel bandwidth ($C = 10$, $\theta = 0.8$)

(b) Data retrieval probability vs. # of channels ($\beta = 2$, $\theta = 0.8$)

(c) Data retrieval probability vs. percentage of unlicensed channels ($\beta = 3$, $C = 10$, $\theta = 0.8$)

(d) Data retrieval probability vs. Zipf parameter ($\beta = 5$, $C = 10$)

Fig. 2. Comparison of SPEC and other schemes on the synthetic trace

query rate, but are not treated differently in UNI. Thus, the data retrieval probability of UNI almost stays flat at around 32% with the increase of θ.

C. Realistic Traces

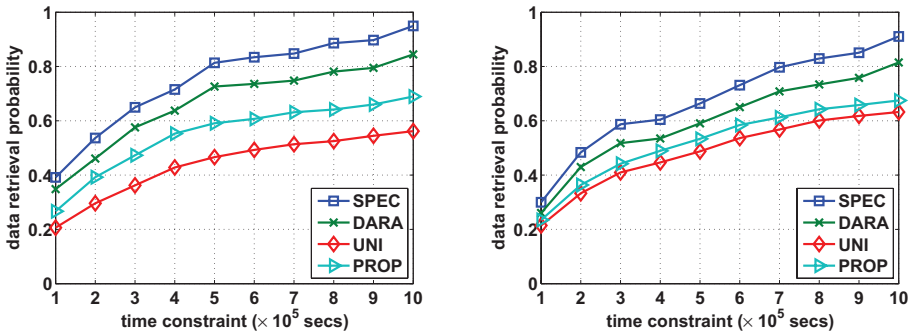
1) Simulation Settings: The performance of our scheme is also evaluated on realistic traces. However, most realistic traces are inappropriate for our simulations. They do not record where each contact happens, and hence it is difficult to model the channel availability upon each contact. We find that in the Dartmouth trace [22] and the UCSD trace [30], each mobile node records the nearby associated wireless access points (APs), which may be used to model the locations. A contact happens if two nodes are at the same location at the same time. The amount of data that can be transmitted upon contact depends on the channel availability at that location, which can be simulated using our model for primary user appearance (the transition probabilities among different states are randomly generated).

The Dartmouth trace was collected by several thousand wireless laptops which were carried by students and faculty at the Dartmouth College campus over five years. In our simulation, we focus on the data collected between September 1, 2002 and December 1, 2002. If two nodes are associated with the APs in the same building, they are assumed to be at the same location. There are 185 locations in total by grouping APs of the same building together. We sort all users in a descending order of trace length, and select the first 50 users for simulation. We set 20 channels and 20 data items. The channel bandwidth is 5 packets per second. That is, if there are c available channels, the data transmission capacity is the combined size of $5c$ packets per second. Each data item is generated by some node which is randomly selected, and can be reconstructed by 20 coded packets. The storage space of each node is the combined size of 100 coded packets. The data query pattern is based on Zipf-like distribution with $\theta = 0.8$.

The UCSD trace was collected by approximately 300 wireless PDAs which were carried by UCSD freshmen for an 11-week period between September 22, 2002 and December 8, 2002. There are 520 APs, and each AP corresponds to one location. Similar to the Dartmouth trace, we sort all users in an descending order of trace length, and select the first 50 users for simulation. The other simulation settings are the same as the Dartmouth trace.

In both Dartmouth trace and UCSD trace, we vary the time constraint to study its effect on the (average) data retrieval probability. In all simulations, the first half of the trace is used for warmup to collect necessary network information. All the data and queries are generated during the second half of the trace. The presented results are averaged over 20 runs.

2) Simulation Results: Figure 3 shows the effect of time constraint on the data retrieval probability on the Dartmouth trace and the UCSD trace, respectively. For all schemes, the data retrieval probability increases as the time constraint increases. This is because increasing the time constraint creates more contact opportunities to retrieve the requested data items. Among the four schemes, SPEC performs the best, since it considers the effect of primary user appearance on



(a) Dartmouth trace (b) UCSD trace

Fig. 3. Data retrieval probability vs. time constraint

the data replication strategy, which is ignored by the other three schemes. Compared to DARA, UNI and PROP, SPEC improves the data retrieval probability by 12%, 90%, 47% for the Dartmouth trace (15%, 40%, 28% for UCSD trace) with time constraint 10^5secs. When the time constraint reaches 10^6secs, the improvement changes to 12%, 69%, 38% for Dartmouth trace (12%, 44%, 35% for UCSD trace).

VII. Conclusions

This paper studied the data replication problem in intermittently connected cognitive radio networks. Different from existing replication schemes in traditional DTNs, the proposed spectrum-aware data replication scheme jointly considers node contact pattern and primary user appearance in determining where to replicate the data. We formulated the spectrum-aware data replication problem as an optimization problem which tries to maximize the average data retrieval probability of the network subject to time and storage constraints. We further analyzed how to efficiently calculate the data retrieval probability which is essential for calculating the replication benefit. Also, we designed a distributed packet-level replication scheme, and evaluated its performance on both synthetic and realistic traces. Evaluation results demonstrated that our spectrum-aware data replication scheme outperforms existing schemes in terms of data retrieval probability in various scenarios.

References

[1] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer Networks*, vol. 50, pp. 2127–2159, 2006.

[2] M. Pan, P. Li, Y. Song, Y. Fang, and P. Lin, "Spectrum Clouds: A Session Based Spectrum Trading System for Multi-hop Cognitive Radio Networks," in *IEEE INFOCOM, 2012*.

[3] Z. Liu, X. Wang, W. Luan, and S. Lu, "Transmission Delay in Large Scale Ad Hoc Cognitive Radio Networks," in *ACM MobiHoc, 2012*.

[4] J. Zhao and G. Cao, "Robust Topology Control in Multi-hop Cognitive Radio Networks," in *IEEE INFOCOM, 2012*.

[5] Q. Yan, M. Li, T. Jiang, W. Lou, and T. Hou, "Vulnerability and Protection for Distributed Consensus-based Spectrum Sensing in Cognitive Radio Networks," in *IEEE INFOCOM, 2012*.

[6] K. Fall, "A Delay-Tolerant Network Architecture for Challenged Internets," in *ACM SIGCOMM, 2003*.

[7] J. Reich and A. Chaintreau, "The Age of Impatience: Optimal Replication Schemes for Opportunistic Networks," in *ACM CoNEXT, 2009*.

[8] S. Ioannidis, L. Massoulie, and A. Chaintreau, "Distributed Caching over Heterogeneous Mobile Networks," in *ACM SIGMETRICS, 2010*.

[9] W. Gao, G. Cao, A. Iyengar, and M. Srivatsa, "Supporting Cooperative Caching in Disruption Tolerant Networks," in *IEEE ICDCS, 2011*.

[10] X. Zhuo, Q. Li, W. Gao, G. Cao, and Y. Dai, "Contact Duration Aware Data Replication in Delay Tolerant Networks," in *IEEE ICNP, 2011*.

[11] G. Cheng, W. Liu, Y. Li, and W. Cheng, "Spectrum Aware On-demand Routing in Cognitive Radio Networks," in *IEEE DySPAN, 2007*.

[12] I. Pefkianakis, S. H. Wong, and S. Lu, "Spectrum Aware Routing in Cognitive Radio Mesh Networks," in *IEEE DySPAN, 2008*.

[13] K. R. Chowdhury and M. D. Felice, "SEARCH: A Routing Protocol for Mobile Cognitive Radio Ad-hoc Networks," *Computer Communications*, vol. 32, no. 18, pp. 1983–1997, 2009.

[14] C. E. Perkins and E. M. Royer, "Ad hoc on-demand distance vector routing," in *IEEE Workshop on Mobile Computing Systems and Applications, 1999*.

[15] P. Nuggehalli, V. Srinivasan, and C.-F. Chiasserini, "Energy-Efficient Caching Strategies in Ad Hoc Wireless Networks," in *ACM MobiHoc, 2003*.

[16] J. Zhao, P. Zhang, G. Cao, and C. R. Das, "Cooperative Caching in Wireless P2P Networks: Design, Implementation, and Evaluation," *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 2, pp. 229–241, 2010.

[17] L. Yin and G. Cao, "Supporting Cooperative Caching in Ad Hoc Networks," *IEEE Transactions on Mobile Computing*, vol. 5, no. 1, pp. 77–89, 2006.

[18] C. Swamy and A. Kumar, "Primal-Dual Algorithms for Connected Facility Location Problems," in *International WOrkshop on Approximation Algorithms for Combinatorial Optimization (APPROX), 2002*.

[19] J. Zhao, W. Gao, Y. Wang, and G. Cao, "Delay-Constrained Caching in Cognitive Radio Networks," in *IEEE INFOCOM, 2014*.

[20] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.

[21] J. W. Byers, M. Luby, and M. Mitzenmacher, "A Digital Fountain Approach to Asynchronous Reliable Multicast," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 8, pp. 1528–1540, 2002.

[22] L. Song, D. Kotz, R. Jain, and X. He, "Evaluating location predictors with extensive Wi-Fi mobility data," in *IEEE INFOCOM, 2004*.

[23] Q. Yuan, I. Cardei, and J. Wu, "Predict and Relay: An Efficient Routing in Disruption-Tolerant Networks," in *ACM MobiHoc, 2009*.

[24] R. Urgaonkar and M. J. Neely, "Opportunistic Scheduling with Reliability Guarantees in Cognitive Radio Networks," *IEEE Transactions on Mobile Computing*, vol. 8, no. 6, pp. 766–777, 2009.

[25] A. Laourine, S. Chen, and L. Tong, "Queuing Analysis in Multichannel Cognitive Spectrum Access: A Large Deviation Approach," in *IEEE INFOCOM, 2010*.

[26] T. Zhang and D. H. K. Tsang, "Optimal Cooperative Sensing Scheduling for Energy-Efficient Cognitive Radio Networks," in *IEEE INFOCOM, 2011*.

[27] L. Cao, L. Yang, and H. Zheng, "The Impact of Frequency-Agility on Dynamic Spectrum Sharing," in *IEEE DySPAN, 2010*.

[28] M. Jain, J. I. Choi, T. M. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, Real-time, Full Duplex Wireless," in *ACM MobiCom, 2011*.

[29] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "Web Caching and Zipf-like Distributions: Evidence and Implications," in *IEEE INFOCOM, 1999*.

[30] M. McNett and G. M. Voelker, "Access and Mobility of Wireless PDA Users," *Mobile Computing Communications Review*, vol. 9, no. 2, pp. 40–55, 2005.