8-1-2014

Distributed and Asynchronous Data Collection in Cognitive Radio Networks with Fairness Consideration

Zhipeng Cai
Georgia State University

Shouling Ji

Jing He
Kennesaw State University, jhe4@kennesaw.edu

Lin Wei

Recommended Citation

Cai, Z., Ji, S., He, J., Wei, L., & Bourgeois, A. G. (2013). Distributed and asynchronous data collection in cognitive radio networks with fairness consideration. http://dx.doi.org/10.1109/TPDS.2013.75

Distributed and Asynchronous Data Collection in Cognitive Radio Networks with Fairness Consideration

Zhipeng Cai, *Member, IEEE*, Shouling Ji, *Student Member, IEEE*, Jing (Selena) He, Lin Wei, and Anu G. Bourgeois, *Senior Member, IEEE*

Abstract—As a promising communication paradigm, Cognitive Radio Networks (CRNs) have paved a road for Secondary Users (SUs) to opportunistically exploit unused licensed spectrum without causing unacceptable interference to Primary Users (PUs). In this paper, we study the distributed data collection problem for asynchronous CRNs, which has not been addressed before. We study the Proper Carrier-sensing Range (PCR) for SUs. By working with this PCR, an SU can successfully conduct data transmission without disturbing the activities of PUs and other SUs. Subsequently, based on the PCR, we propose an Asynchronous Distributed Data Collection (ADDC) algorithm with fairness consideration for CRNs. ADDC collects a snapshot of data to the base station in a distributed manner without the time synchronization requirement. The algorithm is scalable and more practical compared with centralized and synchronized algorithms. Through comprehensive theoretical analysis, we show that ADDC is order-optimal in terms of delay and capacity, as long as an SU has a positive probability to access the spectrum. Furthermore, we extend ADDC to deal with the continuous data collection issue, and analyze the delay and capacity performances of ADDC for continuous data collection, which are also proven to be order-optimal. Finally, extensive simulation results indicate that ADDC can effectively accomplish a data collection task and significantly reduce data collection delay.

Index Terms—Cognitive radio networks, data collection, snapshot data collection, continuous data collection, distributed algorithm, asynchronous wireless network, delay, capacity

✦

1 INTRODUCTION

WIRELESS spectrum is one of the most precious resources for wireless networks. With the fast growth of wireless networks, communications over the free (unlicensed) spectrum have become more and more crowded. On the contrary, according to the report from the Federal Communications Commission (FCC) [2], the utilization of the spectrum assigned to licensed users varies from 15 percent to 85 percent temporally and geographically, which is very inefficient. This necessitates a new communication paradigm, named Cognitive Radio Networks (CRNs), enabling a user equipped with a cognitive radio to sense and learn the communication environment, and further exploit the instantaneous assigned (licensed)

spectrum opportunistically without causing unacceptable interference to the licensed users [3].

Under the CRN communication paradigm, a *secondary network*, which consists of Secondary Users (SUs) (unlicensed users) equipped with cognitive radios, is coexisted with a *primary network*, which consists of Primary Users (PUs) (licensed users). The SUs sense and exploit spectrum opportunistically and share the same space, time, and spectrum with PUs. For an SU, when it has some data for transmission, it begins to sense the communication environment. If there is a spectrum opportunity, i.e., the data transmission of the SU will not cause unacceptable interference to any PU, meanwhile, the receiver of this data transmission is out of the interference range of any primary transmitter, then the SU can initiate a data transmission. During the data transmission of an SU, if a PU comes back to transmit/receive data, the SU has to immediately hand-off the spectrum being occupied to guarantee that its data transmission does not interfere with the communications of PUs.

Numerous efforts have been spent for different issues in CRNs, including spectrum sensing [6], [7], [8], [9], spectrum access, scheduling, and management [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], capacity/throughput/delay scaling laws [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], network connectivity [40], [41], routing protocols [42], [43], [44], multicast communication [45], [46], and etc. Data collection is a common and important operation in wireless networks, as well as in CRNs, which can be used to gather data from an entire network. Many

---

• Z. Cai and A.G. Bourgeois are with the Department of Computer Science, Georgia State University, Atlanta, GA 30303 USA. E-mail: zcai@gsu.edu; abourgeois@cs.gsu.edu.
• S. Ji is with the Department of Computer Science, Georgia State University, Atlanta, GA 30303 USA, and also with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0765 USA. E-mail: sji@gatech.edu.
• J. He is with the Department of Computer Science, Kennesaw State University, Kennesaw, GA 30144-5591 USA. E-mail: jhe4@kennesaw.edu.
• L. Wei is with the Department of Agricultural and Biosystems, South DAKOTA State University. E-mail: lin.wei@sdstate.edu.

practical applications in Biomass monitoring, agriculture management and production, public safety, disaster rescue, ecology and mobile health care involve the data collection demand to grasp global information. Interestingly, few efforts have been devoted to data collection for CRNs, especially practical distributed data collection in asynchronous CRNs.

Intuitively, CRNs are distributed asynchronous systems and thus prefer distributed algorithms. First, CRNs tend to be large-scale distributed wireless systems, for which it is difficult and expensive (sometimes, even impossible) to obtain real-time network information for network management and configuration. Second, the deploy environment of CRNs may be ever-changing, which may induce significant degradation of centralized and synchronized algorithms. Third, in CRNs, some existing SUs might leave the network and some new SUs might join the network at any time. In this case, centralized and synchronized algorithms cannot adapt to these network changes in real time. Finally and most importantly, SUs should not cause any unacceptable interference to PUs, which makes the status of a CRN change more frequently and unpredictably. Hence, centralized and synchronized algorithms are not preferable for CRNs. Therefore, we investigate the distributed data collection issue in asynchronous CRNs in this paper.

Without overall network information and time synchronization, it is very complicated to study distributed data collection algorithms in asynchronous CRNs. There are three main challenges. First, data transmissions of SUs should not cause any unacceptable interference to PUs. Therefore, how to guarantee that a secondary network does not interrupt the primary network in a distributed manner is a challenge. Second, centralized algorithms can make an overall optimized decision, however, it is difficult for distributed algorithms to guarantee an overall optimized solution by using only local network information. Hence, how to design an effective distributed data collection algorithm for CRNs is another challenge. Third, in an asynchronous CRN, many data collisions, interference, and retransmissions occur due to lack of time synchronization, followed by capacity degradation and unfairness among data flows. Therefore, how to overcome the problems induced by lack of time synchronization and meanwhile taking fairness into consideration is also a challenge.

To address the aforementioned challenges, we propose a distributed data collection algorithm for CRNs without the time synchronization requirement. First, we study the *Proper Carrier-sensing Range* (PCR) for an SU. Working with the PCR and the Re-Start (RS) mode[1], an SU can successfully conduct data transmissions as long as there are no active PUs or SUs within its PCR, and will not cause unacceptable interference to any activity of PUs. Subsequently, we propose a Connected Dominating Set (CDS)-based data collection algorithm, named *Asynchronous Distributed Data Collection* (ADDC), for CRNs. Through theoretical analysis, we show that ADDC is order-optimal when an SU has a positive probability to access the spec-

trum. Thirdly, we extend ADDC to deal with the continuous data collection problem, and analyze the delay and capacity performances of ADDC for continuous data collection, which are also proven to be order-optimal. We also conduct extensive simulations to validate the performance of ADDC. Specifically, the contributions of this paper are summarized as follows.

- We derive and obtain the sufficient condition for the PCR of SUs. Working with this PCR, it can be guaranteed that the activities of PUs will not be interfered, and on the other hand, SUs can also conduct data transmissions without interference when they have spectrum access opportunities.
- We propose the ADDC algorithm for CRNs. ADDC first constructs a CDS-based *data collection tree*. With this data collection tree, it schedules the data collection process in a distributed manner without time synchronization. With each SU working with the PCR on the RS mode and elegantly determining the back-off timer, all the data packets can be effectively collected by the base station with fairness consideration.
- Comprehensive theoretical analysis of ADDC demonstrates that ADDC is order-optimal as centralized data algorithms for traditional wireless networks. We show that the induced delay of ADDC is upper bounded by $O\left(\frac{(2\beta_{\kappa}+24\beta_{\kappa+1}-1)n\cdot\tau}{p_o}\right)$, where β_{κ} and $\beta_{\kappa+1}$ are two constants related to the value of the PCR, n is the number of SUs in a secondary network, τ is the time duration of a time slot, and $p_o > 0$ is the probability that an SU has a spectrum access opportunity during a time slot. It follows that the achievable data collection capacity, which is defined as the average data receiving rate at the base station [48], of ADDC is $\Omega\left(\frac{p_o}{2\beta_{\kappa}+24\beta_{\kappa+1}-1}\cdot W\right)$, where W is the bandwidth of the spectrum. Since a single-radio base station can receive at most one data packet in a time slot, the data collection capacity of ADDC is order-optimal.
- We extend ADDC to deal with the continuous data collection issue. Even continuous data collection introduces more traffic load, more wireless interference, and the data accumulation problem, ADDC can still successfully achieve order-optimal continuous data collection capacity of $\Omega\left(\frac{np_o}{(n-1)(2\beta_{\kappa}+24\beta_{\kappa+1}-1)}\cdot W\right)$.
- Extensive simulations are conducted to examine the performance of ADDC. The simulation results show that ADDC can effectively gather all the data packets to the base station and significantly reduce data collection delay. Particularly, compared with Coolest [42], the improvements on the induced delay of ADDC varies from 171 percent to 314 percent with respect to different system parameters for snapshot data collection, and the improvement on the delay of ADDC is 301.24 percent for continuous data collection.

The rest of the paper is organized as follows. In Section 2, we describe the network model and interference model. The derived PCR and ADDC are presented in Section 3, followed by theoretical analysis of the delay and capacity performances of ADDC. In Section 4, we extend ADDC to deal with the continuous data collection issue, and analyze

---

1. With the RS mode, a receiver will switch to receive the stronger signal as long as the Signal-to-Interference Ratio (SIR) threshold for the stronger signal can be satisfied [47].
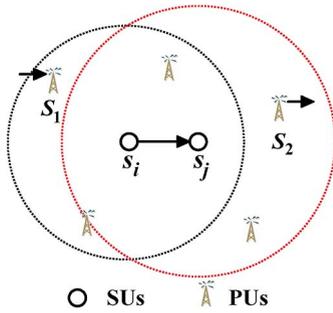
Fig. 1. Communication between s_i and s_j.

the delay and capacity performances of ADDC for continuous data collection. We conclude this paper and point out some possible future research directions in Section 5. Furthermore, in Section **Related Work of the Supplementary File** which is available in the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TPDS.2013.75, we summarize the related works and remark the differences between this work and the existing literatures. We examine ADDC via simulations in Section **Simulation and Analysis of the Supplementary File** available online.

2 SYSTEM MODEL

In this paper, we consider a secondary network consisting of n SUs coexisted with a primary network consisting of N PUs deployed in an area with size $A = c_0 n$, where c_0 is a predetermined constant value. Both networks share the same time, space, and spectrum.

2.1 Primary Network

The primary network consists of N *independent and identically distributed* (i.i.d.) PUs (*licensed users*), denoted by $S_1, S_2, \ldots, S_N$. Define $V_p = \{S_1, S_2, \ldots, S_N\}$. To guarantee the secondary network has some spectrum opportunity, we assume the primary network satisfies the *locally finite* property, i.e., $N/A < \infty$. This is reasonable since otherwise the secondary network may never have an opportunity to access the spectrum. The maximum transmission radius of PUs is R and all the PUs have a fixed power P_p.

The network time is slotted and the duration of a time slot is τ. We use a generalized probabilistic model to describe the data transmission activities of PUs. During a particular time slot, each PU transmits data (performing as a transmitter) with probability p_t. Generally speaking, given a specific probabilistic distribution, such as the Poisson distribution and the Uniform distribution, of the activities of the primary network, p_t can be determined accordingly. Furthermore, at the very beginning of each time slot, if a PU is not scheduled to transmit a data packet, it will keep silent in that time slot.

2.2 Secondary Network

The secondary network consists of $n > N$ single-radio SUs denoted by $s_1, s_2, \ldots, s_n$ and one single-radio *base station* (sink) denoted by s_b. All the SUs and the base station are also i.i.d.. The maximum transmission radius of SUs is r. Therefore, the secondary network can be modeled as a graph $G_s = (V_s, E_s)$, where $V_s = \{s_b, s_1, s_2, \ldots, s_n\}$, and E_s is

the link set consisting of all the possible links formed by SUs in V_s. We assume G_s is connected. All the SUs have the same working power denoted by P_s. Note that, for $\forall s_i, s_j \in V_s$, even if there is a link between s_i and s_j as shown in Fig. 1, it does not imply that s_i can successfully transmit data to s_j. Besides the possible wireless interference from other SUs as in traditional wireless networks, two more reasons may cause the failure: 1) the data transmission from s_i may interfere with some receiving activity of a PU, e.g., S_1 in Fig. 1; 2) the data receiving activity at s_j may be interfered by some transmission activity of PUs, e.g. S_2 in Fig. 1. Therefore, to carry out a data transmission between an SU pair, having spectrum opportunity is necessary.

We formally define a *data collection task* as follows. At a particular time slot t, every SU in the secondary network produces a data packet of size B, which is smaller than the size of a primary data packet. The set of all the n data packets produced by SUs at time t is called a *snapshot*. The task of gathering all the n data packets of a snapshot to the base station without any data aggregation is called a *data collection task*. The *data collection delay* is the time consumption to finish a data collection task. The *data collection capacity* is defined as the average data receiving rate at the base station during a data collection process.

2.3 Interference Model

The available spectrum bandwidth for both the primary network and the secondary network is assumed to be $W = B/\tau$. Therefore, the upper bound of data collection capacity is W since the base station can receive at most one data packet during a time slot. In this paper, we take the *physical interference model* into account. At time t, suppose $\mathcal{S}_p^t$ (respectively, $\mathcal{S}_s^t$) is the set of all the PUs (respectively, SUs) trying to transmit data to some other PUs (respectively, SUs). For $\forall S_i \in \mathcal{S}_p^t$, assume its intended receiver is S_j. Then, S_j can successfully receive data from S_i at time t only if the Signal-to-Interference Ratio (SIR) at S_j associated with S_i satisfies

$$\frac{P_p \cdot D(S_i, S_j)^{-\alpha}}{\sum\limits_{S_k \in \mathcal{S}_p^t, S_k \neq S_i} P_p D(S_k, S_j)^{-\alpha} + \sum\limits_{s_k \in \mathcal{S}_s^t} P_s D(s_k, S_j)^{-\alpha}} \geq \eta_p,$$

where $D(\cdot, \cdot)$ is the Euclidean distance between two nodes, $\alpha > 2$ is the path loss exponent, and η_p is the threshold SIR value for the primary network. Similarly, for $\forall s_i \in \mathcal{S}_s^t$, assume its intended receiver is s_j. s_j can successfully receive data from s_i at time t only if the SIR at s_j associated with s_i satisfies

$$\frac{P_s \cdot D(s_i, s_j)^{-\alpha}}{\sum\limits_{S_k \in \mathcal{S}_p^t} P_p D(S_k, s_j)^{-\alpha} + \sum\limits_{s_k \in \mathcal{S}_s^t, s_k \neq s_i} P_s D(s_k, s_j)^{-\alpha}} \geq \eta_s,$$

where η_s is the threshold SIR value for the secondary network.

3 DISTRIBUTED DATA COLLECTION

In this section, we first construct a CDS-based data collection tree as the routing structure for data collection. Subsequently, we derive the Proper Carrier-sensing Range (PCR) for SUs. By working with the PCR, an Asynchronous
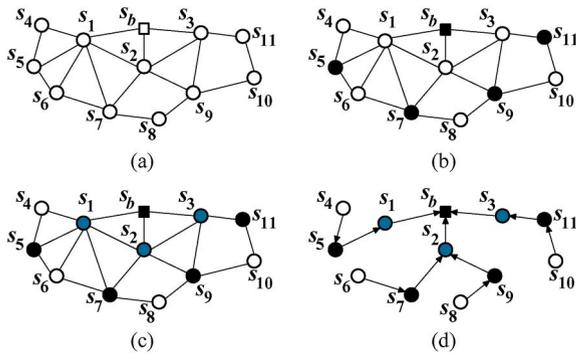
Fig. 2. Construction of a CDS-based data collection tree. s_b is the base station. The black nodes in (b) are *dominators*, the blue nodes in (c) are *connectors*, and the white nodes in (c) are *dominatees*.

Distributed Data Collection (ADDC) algorithm is proposed. ADDC can guarantee that the activities of PUs will not be interfered. Meanwhile, the SUs that transmit data simultaneously are also guaranteed to be interference-free. Finally, we theoretically analyze the performance of ADDC. It shows that ADDC can achieve order-optimal data collection capacity as centralized data collection algorithms [36], [38], [48], [49] for traditional wireless networks.

3.1 Data Collection Tree

In ADDC, we take a CDS-based data collection tree as the routing infrastructure. For graph G_s, a *Dominating Set* (DS) of G_s is a subset U of V_s such that $\forall s_i \in V_s$, either $s_i \in U$ or s_i is adjacent to some node in U. If the induced subgraph of G_s on U is connected, then U is called a *Connected Dominating Set* (CDS). Since CDS is a good candidate serving as a connected *virtual backbone* in a wireless network, it has attracted a lot of attention.

We use the method proposed in [50] to construct a CDS-based data collection tree. Taking the secondary network shown in Fig. 2a as an example, the construction process can be done in three steps as follows. First, make a *Breadth First Search* (BFS) starting from the base station s_b, and identify a *Maximal Independent Set* (MIS) $\mathcal{D}$ of G_s. The nodes in the MIS are called *dominators* (evidently, the base station is also a dominator). As shown in Fig. 2b, the set of all the black nodes is the dominator set $\mathcal{D}$ of the secondary network in Fig. 2a. Second, find a set $\mathcal{C}$ consisting of *connectors* to connect the dominators in $\mathcal{D}$ to form a CDS. For example, in Fig. 2c, the set of blue nodes is the connector set. Finally, the nodes in $V_s \setminus (\mathcal{C} \cup \mathcal{D})$ are identified as *dominatees*. By choosing a dominator node for each dominatee in its 1-hop neighborhood as its parent node, we can construct a CDS-based data collection tree as shown in Fig. 2d.

For the constructed CDS-based data collection tree, it has several interesting properties as shown in Lemma 1 [50].

Lemma 1. *For $\forall s_i \in \mathcal{D}$, s_i is adjacent to at most 12 connectors, among which one is s_i's parent node if $s_i \neq s_b$.*

3.2 Proper Carrier-Sensing Range

In our proposed distributed data collection algorithm, all the SUs have to carrier-sense the spectrum to obtain a spectrum opportunity. Therefore, we need to study how to

properly set the carrier-sensing range with the objectives: 1) the secondary network does not cause unacceptable interference to the activities of the primary network; 2) all the SUs transmitting data simultaneously are interference-free; and meanwhile 3) the carrier-sensing range is as small as possible, which implies SUs can obtain more spectrum opportunities. We have the following definitions.

Definition 3.1 Concurrent Set. $\mathbb{C} = \{u : u \in V_p \cup V_s\}$ *is a concurrent set if all the nodes in $\mathbb{C}$ can carry out data transmissions simultaneously and successfully during a time slot.*

The definition of a concurrent set implies that all the SUs in $\mathbb{C}$ have spectrum opportunities, and the activities of all the SUs and PUs in $\mathbb{C}$ are interference-free.

Definition 3.2 $\mathcal{R}$-Set. *For $\mathcal{R} > 0$, $\mathbb{C} = \{u : u \in V_p \cup V_s\}$ is an $\mathcal{R}$-set if $\forall u, v \in \mathbb{C}$, $u \neq v$, then $D(u, v) \geq \mathcal{R}$.*

Definition 3.3 Proper Carrier-sensing Range (PCR). *$\mathcal{R}$ is a PCR if any $\mathcal{R}$-set is a concurrent set.*

Now, we derive the conditions on PCR for SUs. To this end, we have two constraints. The first one is to guarantee the SUs will not cause unacceptable interference to the primary network, which is stated in Lemma 2. The second one is all the SUs that transmit data simultaneously are interference-free, which can be satisfied as shown in Lemma 3.

Lemma 2. *Let $\mathbb{C} = \{u : u \in V_p \cup V_s\}$ be any $\mathcal{R}$-set. Assume all the nodes in $\mathbb{C}$ initiate data transmissions simultaneously. Then, to guarantee SUs will not cause unacceptable interference to PUs, it is sufficient to have the PCR*

$$\mathcal{R} \geq \left(1 + \sqrt[\alpha]{\frac{c_2 \eta_p}{c_1}}\right) \cdot R, \tag{1}$$

where $c_1 = \frac{P_p}{\max\{P_p, P_s\}}$ and $c_2 = 6 + 6\left(\frac{\sqrt{3}}{2}\right)^{-\alpha}\left(\frac{1}{\alpha-2} - 1\right)$.

Proof. Please refer to the Supplementary File available online. □

Lemma 3. *Let $\mathbb{C} = \{u : u \in V_p \cup V_s\}$ be any $\mathcal{R}$-set. Assume all the nodes in $\mathbb{C}$ initiate data transmissions simultaneously. Then, to guarantee every SU in $\mathbb{C} \cap V_s$ can conduct data transmission successfully (i.e., every SU in $\mathbb{C} \cap V_s$ has a spectrum opportunity and all the SUs transmit data without interference), it is sufficient that*

$$\mathcal{R} \geq \left(1 + \sqrt[\alpha]{\frac{c_2 \eta_s}{c_3}}\right) \cdot r, \tag{2}$$

where $c_3 = \frac{P_s}{\max\{P_p, P_s\}}$.

Proof. By the similar technique as in Lemma 2, this lemma can be proven. □

Let

$$\kappa = \max\left\{\left(1 + \sqrt[\alpha]{\frac{c_2 \eta_p}{c_1}}\right) \cdot \frac{R}{r}, 1 + \sqrt[\alpha]{\frac{c_2 \eta_s}{c_3}}\right\}. \tag{3}$$

From Lemma 2 and Lemma 3, if we set the PCR $\mathcal{R} = \kappa \cdot r$, then any $\mathcal{R}$-set is a concurrent set. Consequently, we set $\mathcal{R} = \kappa \cdot r$ in the following of this paper. We further show
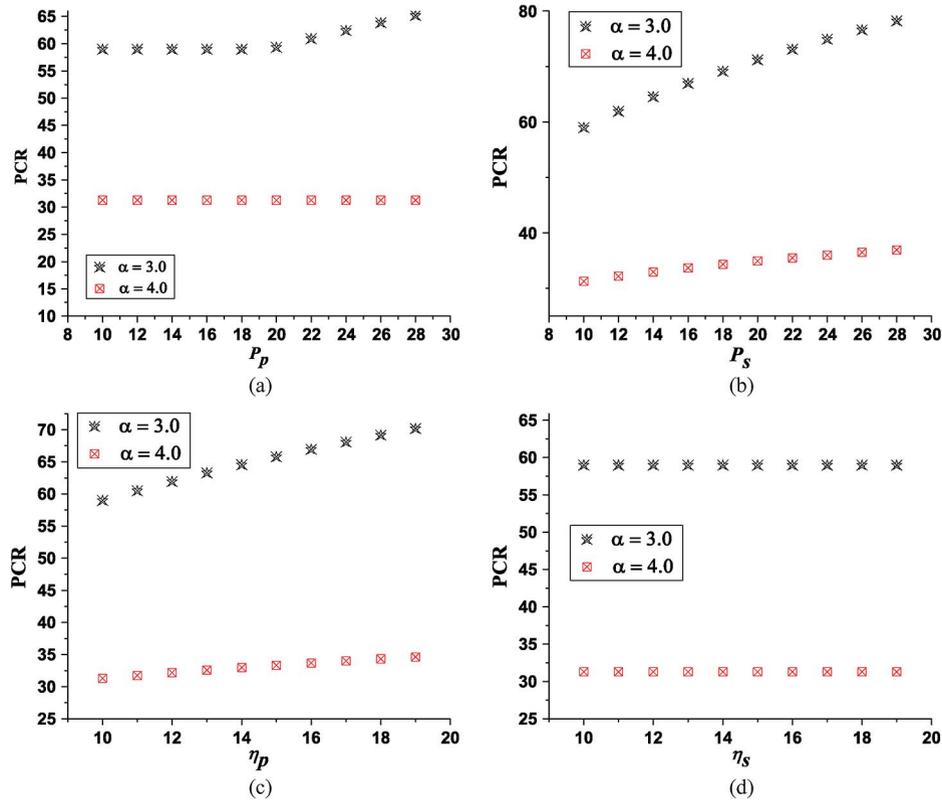
Fig. 3. PCR value. The default settings are $\alpha = 4$, $P_p = 10$, $R = 12$, $\eta_p = 10$ dB, $P_s = 10$, $r = 10$, and $\eta_s = 10$ dB. (a) PCR vs. P_p. (b) PCR vs. P_s. (c) PCR vs. η_p. (d) PCR vs. η_s.

the PCR values under different situations in Fig. 3 where we can see that the PCR value is bigger when $\alpha = 3.0$ than that of $\alpha = 4.0$. This is because a larger path loss exponent implies a transmitter brings less interference to other ongoing data transmissions. Additionally, $\mathcal{R}$ is a non-decreasing function with respect to P_p, P_s, η_p, and η_s. This can also be seen from Lemma 2 and Lemma 3.

3.3 Data Collection Algorithm

Based on the constructed data collection tree and the PCR $\mathcal{R}$, we propose an *Asynchronous Distributed Data Collection* (ADDC) algorithm (Algorithm 1). In Algorithm 1, τ_c is the time duration of the *contention window* and $\tau_c < \tau$. Since we assume that the size of a secondary data packet is smaller than the size of a primary data packet, we further assume that a secondary data packet can be transmitted during time $\tau - \tau_c$ when the spectrum opportunity is available. Moreover, assume that no two SUs located within each other's PCR have their backoff timers expired at the exactly same time.[2]

The basic idea of ADDC is as follows. On the view of macroscopic, for each SU, e.g. s_i, with data for transmission, after setting the backoff timer, s_i will sense the local communication environment within its PCR, which is determined according to the system parameters, e.g., path

---

Algorithm 1: The ADDC Algorithm

input : CDS-based data collection tree, $\mathcal{R}$
output: a distributed data collection plan

1. s_i $(1 \leq i \leq n)$ sets its carrier sensing range to $\mathcal{R} = \kappa \cdot r$;
2. **while** *s_i has some data for transmission* **do**
3. s_i randomly and uniformly chooses a value $t_i \in (0, \tau_c]$ as its backoff timer;
4.   **while** *the backoff timer does not expire* **do**
5. s_i keeps sensing the spectrum with $\mathcal{R}$;
6.   **if** *during a time slot, there is a PU or SU transmitting some data within its PCR, i.e. s_i does not have a spectrum opportunity or might cause interference to other active SUs* **then**
7. s_i freezes its backoff timer (stopping the countdown process) until the spectrum becomes free;
8.   **if** *the spectrum is free* **then**
9. s_i's backoff timer continues the countdown process;
10.   **if** *the backoff timer expires* **then**
11. s_i transmits one data packet to its parent node when a spectrum opportunity appears;
12. s_i waits for time $\tau_c - t_i$ before it starts another transmission activity;
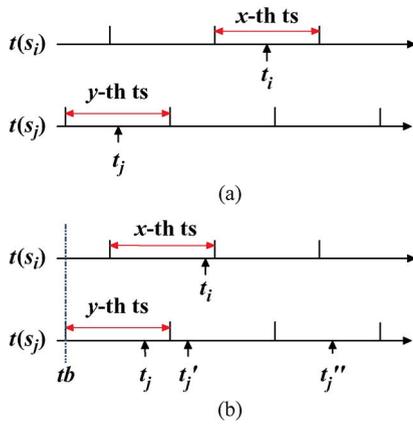
---

2. Collisions due to simultaneous countdown-to-zero can be tackled by an exponential backoff mechanism in which the transmission probability of each SU is adjusted in a dynamic way based on the network business [47].

Fig. 4. Transmission order of s_i and s_j, where $t(s_i)$ and $t(s_j)$ are the time scales of s_i and s_j respectively, $ts = time\ slot$, and $tb = time\ baseline$.

loss exponent, SIR threshold, working power, etc. On one hand, if s_i senses that some node within its PCR is transmitting some data, i.e. the spectrum is occupied by some node, s_i stops the countdown process (freezes the backoff timer). This is mainly because if there is some transmission activity within the PCR of s_i, then the spectrum is not available for s_i and thus the spectrum-not-available time could not be counted in s_i's timer countdown process. As the results of this freeze, 1) s_i will not perform any harmful impacts on the ongoing data transmission; and 2) it can also be guaranteed that no two SUs located within each other's PCR have their backoff timers expired at the exactly same time. On the other hand, if the sensed spectrum is free, s_i continues the timer countdown process. When the backoff timer expires, s_i transmits the data packet when a spectrum opportunity appears and then waits for extra time of duration $\tau_c - t_i$, where t_i is the selected backoff time of s_i. The transmission of s_i can be guaranteed when the spectrum opportunity appears mainly because of the assumptions that 1) no two SUs located within each other's PCR have their backoff timers expired at the exactly same time, and thus s_i will not be interfered by other SUs within its PCR; and 2) each SU decides whether to transmit data at the very beginning of each time slot which implies an SU will not start a data transmission in a time slot after the beginning time point. Furthermore, we assume that $\tau_c < \tau$ and a secondary data packet can be transmitted during time $\tau - \tau_c$. Consequently, s_i will not be interrupted by PUs during its transmission. On the view of microscopic, each SU will compete the spectrum opportunities with other SUs located within its PCR area. This can be controlled by the backoff timer setting together with fairness control (which is shown in Theorem 1). The detailed competition process is demonstrated by the example cases shown in Fig. 4 and analyzed in the proof of Theorem 1.

In Algorithm 1, each SU sets its carrier-sensing range to $\mathcal{R}$ in line 1. In lines 2-12, the data transmission process of one data packet is described. In line 3, the backoff timer of s_i is set. Subsequently, s_i senses the spectrum (line 5). If the spectrum is busy, s_i stops the countdown process (lines 6-7). Otherwise, the countdown process of s_i's backoff timer continues (lines 8-9). If the backoff timer expires, s_i carries

out a data transmission when a spectrum opportunity appears (line 11). Since we assume SUs within each other's PCR will not have their timer expired at the same time, line 11 can be accomplished without interference. Taking fairness into consideration, i.e., to avoid s_i always occupying the spectrum, we still let s_i wait for time $\tau_c - t_i$ after it transmits the current data packet (line 12). This can also be proven in Theorem 1 (Section 3.4).

3.4 Performance Analysis of ADDC

In this subsection, we analyze the delay and capacity performances of ADDC. In our analysis, since τ_c is very small compared with the waiting time for a spectrum opportunity and the data transmission time, we ignore the delay induced by the backoff time. Actually, considering the backoff time only introduces a constant factor to the delay and capacity of ADDC.

First, we introduce a geometric property of disk packing as follows.

Lemma 4: [50]. *Assume that $\mathfrak{D}$ is a disk of radius r_d and $\mathfrak{M}$ is a set of points with mutual distance of at least 1. Then,*

$$|\mathfrak{D} \cap \mathfrak{M}| \le \frac{2\pi r_d^2}{\sqrt{3}} + \pi r_d + 1, \qquad (4)$$

where $|\cdot|$ is the cardinality of a set.

From Lemma 4, we can derive the following lemma, which states the number of dominators and connectors within the PCR of an SU in Algorithm 1.

Lemma 5. *In Algorithm 1, the number of dominators and connectors within the PCR of an SU is upper bounded by $\beta_\kappa + 12\beta_{\kappa+1}$, where β_x is a function on x with $\beta_x = \frac{2\pi x^2}{\sqrt{3}} + \pi x + 1$.*

Proof. Please refer to the Supplementary File available online. □

Based on Lemma 5 and applying the Chernoff bound, we can obtain the upper bound of the number of SUs within the PCR of an SU as shown in Lemma 6.

Lemma 6. *In Algorithm 1, the number of SUs within the PCR of an SU is upper bounded by $\Delta\beta_\kappa + 12\beta_{\kappa+1}$, where Δ is the maximum degree of the CDS-based data collection tree and $\Delta \le \log n + \frac{\pi r^2(e^2-1)}{2c_0}$ with probability 1.*

Proof. Please refer to the Supplementary File available online. □

An SU obtains a spectrum opportunity only when no ongoing data transmission is initiated by any PU within its PCR during a time slot. Consequently, we show the expected waiting time and probability of a spectrum opportunity appearance for an SU in Lemma 7.

Lemma 7. *To obtain a spectrum opportunity, the expected waiting time for an SU in Algorithm 1 is τ/p_o, where $p_o = (1 - p_t)^{\pi(\kappa r)^2 N/c_o n}$ is the expected probability that an SU has a spectrum opportunity during a time slot.*

Proof. Please refer to the Supplementary File available online. □

Based on Lemma 7, we assume the waiting time for an SU is τ/p_o before a spectrum opportunity appears. Therefore, the results we obtained in the following are from the view of expectation. From Lemma 6 and Lemma 7, we have Theorem 1, which states the upper bound of the waiting time for an SU to transmit one data packet.

Theorem 1. *In Algorithm 1, any SU having data for transmission can transmit at least one data packet to its parent node within time $(2\Delta\beta_\kappa + 24\beta_{\kappa+1} - 1)\tau/p_o$.*

Proof. For convenience, we first assume the secondary network is a stand-alone network.[3] For an arbitrary SU s_i, assume s_j is another arbitrary SU within s_i's PCR. We try to prove that "$\mathfrak{P}$: if s_i and s_j are competing for the spectrum, then, before s_i obtains the spectrum to transmit one data packet, s_j can transmit at most two data packets".

Assume both s_i and s_j have multiple data packets for transmission. During the x-th time slot in terms of s_i's time scale, s_i's backoff timer is set to $t_i \in (0, \tau_c]$. During the y-th time slot in terms of s_j's time scale, s_j's backoff timer is set to $t_j \in (0, \tau_c]$. Evidently, if the x-th time slot has no overlap with the y-th time slot as shown in Fig. 4a, then s_i and s_j are not competing for the spectrum at this moment. Therefore, we only have to prove $\mathfrak{P}$ when the x-th time slot of s_i has some overlap with the y-th time slot of s_j.

Without loss of generality, we assume t_j is before t_i as shown in Fig. 4b[4], $t_i < (\tau_c - t_j) + \tau_c = 2\tau_c - t_j$. Take the start point of the y-th time slot as the *time baseline* (starting time). Then, according to Algorithm 1, s_j transmits a data packet at time t_j and sets its next backoff timer as $\varepsilon_1 \in (0, \tau_c]$ after waiting for time $\tau_c - t_i$, i. e. s_j will transmit the second data packet at time $t_j' = \tau_c + \varepsilon_1$. Since $t_i < 2\tau_c - t_j$, it is still possible that $t_j' < t_i$ as shown in Fig. 4b. Similarly, after transmitting the second data packet, s_j waits for time $\tau_c - \varepsilon_1$ and sets its backoff timer as ε_2, i.e. s_j will transmit the third data packet at time $t_j'' = 2\tau_c + \varepsilon_2$. Now, since $t_j'' = 2\tau_c + \varepsilon_2 > 2\tau_c > t_i$ as shown in Fig. 4b, s_i will transmit one data packet before s_j transmits the third data packet. Then, $\mathfrak{P}$ is true.

Because s_i and s_j in $\mathfrak{P}$ are arbitrarily chosen, and at most $\Delta\beta_\kappa + 12\beta_{\kappa+1} - 1$ SUs are within the PCR of an SU according to Lemma 6, we conclude that every SU with data for transmission can transmit at least one packet to its parent node within $2(\Delta\beta_\kappa + 12\beta_{\kappa+1} - 1) + 1 = 2\Delta\beta_\kappa + 24\beta_{\kappa+1} - 1$ time slots. Now, we remove the assumption that the secondary network is a stand-alone network. Based on Lemma 7, the expected time consumption for each SU to transmit one data packet is upper bounded by $(2\Delta\beta_\kappa + 24\beta_{\kappa+1} - 1)\tau/p_o$. □

From Theorem 1, we can see that Algorithm 1 also takes fairness into consideration when it collects data. Therefore, the waiting time for a node with data transmission is upper bounded. It follows that skewed data flows can be avoided

in ADDC. Furthermore, the following corollary can be obtained straightforwardly.

Corollary 1. *The expected time to collect all the data packets from $V_s \setminus (\mathcal{D} \cup \mathcal{C})$ to $\mathcal{D} \cup \mathcal{C}$ is upper bounded by $(2\Delta\beta_\kappa + 24\beta_{\kappa+1} - 1)\tau/p_o$.*

By the similar technique in the proof of Theorem 1, we can obtain the following lemma, which states the waiting time for an SU in $\mathcal{D} \cup \mathcal{C}$ to successfully transmit a data packet.

Lemma 8. *After all the packets at $V_s \setminus (\mathcal{D} \cup \mathcal{C})$ have been collected to $\mathcal{D} \cup \mathcal{C}$, the expected time of an SU in $\mathcal{D} \cup \mathcal{C}$ to transmit a data packet to its parent node is upper bounded by $(2\beta_\kappa + 24\beta_{\kappa+1} - 1)\tau/p_o$.*

Proof. Please refer to the Supplementary File available online. □

Then, based on Corollary 1 and Lemma 8, the following theorem can be proven, which shows the delay and capacity performances of ADDC.

Theorem 2. *The expected delay induced by ADDC is upper bounded by $O((2\beta_\kappa + 24\beta_{\kappa+1} - 1)n\tau/p_o)$. This implies the achievable data collection capacity of ADDC is $\Omega\left(\frac{p_o}{2\beta_\kappa + 24\beta_{\kappa+1} - 1} \cdot W\right)$, which is order-optimal.*

Proof. From Corollary 1, the expected time to collect data packets from $V_s \setminus (\mathcal{D} \cup \mathcal{C})$ to $\mathcal{D} \cup \mathcal{C}$ is upper bounded by $(2\Delta\beta_\kappa + 24\beta_{\kappa+1} - 1)\tau/p_o$. Meanwhile, the base station also collects at least Δ_b data packets from its children, where Δ_b is the degree of s_b in the data collection tree. After that, the base station can receive at least one data packet every $(2\beta_\kappa + 24\beta_{\kappa+1} - 1)\tau/p_o$ time according to Lemma 8. In summary, the time consumption of ADDC is upper bounded by $(2\Delta\beta_\kappa + 24\beta_{\kappa+1} - 1)\tau/p_o + (n - \Delta_b)(2\beta_\kappa + 24\beta_{\kappa+1} - 1)\tau/p_o = O((2\beta_\kappa + 24\beta_{\kappa+1} - 1)n\tau/p_o)$. Therefore, the achievable data capacity of ADDC is lower bounded by $\frac{n \cdot B}{O((2\beta_\kappa + 24\beta_{\kappa+1} - 1)n\tau/p_o)} = \Omega\left(\frac{p_o}{2\beta_\kappa + 24\beta_{\kappa+1} - 1} \cdot W\right)$. Since the upper bound of data collection capacity is W, the achievable capacity of ADDC is order-optimal. □

From Theorem 2, the proposed ADDC successfully achieves order-optimal data collection capacity even working in a distributed and asynchronous manner. Compared with the existing order-optimal centralized algorithms, ADDC is scalable and more practical for unstable and frequently changed CRNs.

4   EXTENSION: PERFORMANCE OF ADDC FOR CONTINUOUS DATA COLLECTION

For completeness, in this section, we study the performance of ADDC for continuous data collection in asynchronous CRNs. A *continuous data collection task* is a task to collect the data of $Q(Q \geq 1)$ continuous snapshots to the base station without any data aggregation. Similarly, the *continuous data collection delay* is the time consumption to finish a continuous data collection task. The *continuous data collection capacity* is defined as the average data receiving rate at the base station during a continuous data collection

---

3. Here, a stand-alone secondary network means the primary network does not exist or is not working, which implies all the SUs always have spectrum opportunities.

4. For the case that t_i is before t_j, it is clear that s_i will transmit one data packet before s_j according to Algorithm 1. Then, $\mathfrak{P}$ is true.

process. To distinguish from continuous data collection, we refer to the previous discussed data collection as *snapshot data collection* in this subsection.

Compared with snapshot data collection, continuous data collection introduces more communication traffic, therefore more wireless interference and longer transmission time. In addition, due to the fact that the base station can receive at most one data packet during a time slot (under the single-radio single-channel assumption), data packets are easier to accumulate at the nodes close to the base station.

Since ADDC is a distributed and asynchronous algorithm, each SU in ADDC makes actions only based on the status of the local wireless communication environment. Therefore, ADDC still works for continuous data collection tasks because compared with snapshot data collection, continuous data collection just introduces more traffic loads without changing the wireless communication environment. In the following of this subsection, we analyze the delay and capacity performances of ADDC for continuous data collection.

Suppose a continuous data collection task consists of $Q \geq 1$ continuous snapshots. Then, the delay and capacity of ADDC to finish this continuous data collection task are shown in Theorem 3.

Theorem 3. *To collect $Q \geq 1$ continuous snapshots, the expected delay induced by ADDC is upper bounded by $\mathcal{T}_1 + \mathcal{T}_2$, where $\mathcal{T}_1 = (2\Delta\beta_\kappa + 24\beta_{\kappa+1} - 1) \cdot Q\tau/p_o$ and $\mathcal{T}_2 = (Q(n-1) - \Delta_b + 1)(2\beta_\kappa + 24\beta_{\kappa+1} - 1) \cdot \tau/p_o$. It implies that the achievable continuous data collection capacity of ADDC is lower bounded by $\Omega\left(\frac{np_o}{(n-1)(2\beta_\kappa + 24\beta_{\kappa+1} - 1)} \cdot W\right)$, which is order-optimal.*

Proof. In a continuous data collection task consisting of Q snapshots, each node in $V_s \setminus (\mathcal{D} \cup \mathcal{C})$ has Q data packets for transmission. Therefore, based on Theorem 1 and Corollary 1, it is straightforward that the time consumption to collect all the data packets at $V_s \setminus (\mathcal{D} \cup \mathcal{C})$ to $\mathcal{D} \cup \mathcal{C}$ is upper bounded by $\mathcal{T}_1 = Q \cdot (2\Delta\beta_\kappa + 24\beta_{\kappa+1} - 1) \cdot \tau/p_o$.

Now, we analyze how many data packets can be collected by the base station within time $\mathcal{T}_1$. Based on the same proof of Theorem 2, the base station can collect at least Δ_b data packets during the first $(2\Delta\beta_\kappa + 24\beta_{\kappa+1} - 1) \cdot \tau/p_o$ time, where Δ_b is the degree of the base station in the data collection tree. Subsequently, at least one data packet can be collected by the base station every $(2\beta_\kappa + 24\beta_{\kappa+1} - 1) \cdot \tau/p_o$ time according to Theorem 1 and the assumption that the secondary network is connected (which implies the data collection tree is connected). Therefore, at least $\Delta_b + Q - 1$ data packets can be collected by the base station within time $\mathcal{T}_1$. In other words, after time $\mathcal{T}_1$, there are no data packets at nodes in $V_s \setminus (\mathcal{D} \cup \mathcal{C})$, and there are at most $nQ - \Delta_b - Q + 1 = Q(n-1) - \Delta_b + 1$ data packets at nodes in $(\mathcal{D} \cup \mathcal{C}) \setminus \{s_b\}$ not been collected by the base station. According to Lemma 8, the base station can receive at least one data packet every $(2\beta_\kappa + 24\beta_{\kappa+1} - 1) \cdot \tau/p_o$ time after time $\mathcal{T}_1$. Therefore, after time $\mathcal{T}_1$, the time consumption to collect all the data packets at $(\mathcal{D} \cup \mathcal{C}) \setminus \{s_b\}$ to the base station s_b is upper bounded by

$\mathcal{T}_2 = (Q(n-1) - \Delta_b + 1) \cdot (2\beta_\kappa + 24\beta_{\kappa+1} - 1) \cdot \tau/p_o$. In summary, the delay induced by ADDC to finish a continuous datacollection task consisting of Q snapshots is upper bounded by $\mathcal{T}_1 + \mathcal{T}_2$.

Consequently, the achievable continuous data collection capacity of ADDC is

$$\frac{nQ \cdot B}{\mathcal{T}_1 + \mathcal{T}_2} = \frac{nQ \cdot p_o}{O(2\beta_\kappa Q \cdot \log n) + 2\beta_\kappa Q \cdot (n-1)} \cdot W \quad (5)$$

$$= \Omega\left(\frac{np_o}{(n-1)(2\beta_\kappa + 24\beta_{\kappa+1} - 1)} \cdot W\right). \quad (6)$$

Considering that the upper bound of capacity for the data collection problem is $O(W)$ [48], [49], ADDC has already achieved the order-optimal continuous data collection capacity. □

From Theorem 3, we can see that even continuous data collection introduces more traffic, more wireless interference, and the data accumulation problem, ADDC can still achieve order-optimal continuous data collection capacity, which is independent of the number of snapshots in a continuous data collection task. Therefore, ADDC is scalable not only with respect to the number of SUs in a secondary network, but also with respect to the number of snapshots in a continuous data collection task.

5 CONCLUSION AND FUTURE WORK

CRNs introduce a novel promising communication paradigm, where SUs can opportunistically access unused licensed spectrum without harming the communications among PUs. In this paper, we study the distributed data collection problem for asynchronous CRNs. First, we study how to set a Proper Carrier-sensing Range (PCR) for SUs. Subsequently, based on the derived PCR, we propose an Asynchronous Distributed Data Collection (ADDC) algorithm for CRNs with fairness consideration. Through theoretical analysis, we show that ADDC successfully achieves order-optimal delay and capacity, which implies ADDC is scalable and practical. Thirdly, we extend ADDC to deal with the continuous data collection issue, and analyze the delay and capacity performances of ADDC for continuous data collection, which are proven to be order-optimal. Finally, the simulation results demonstrate that ADDC can effectively accomplish a data collection task and significantly reduce delay.

The future work of this paper will be conducted along the following directions. First, in this work, we implement and analyze ADDC under a specific network distribution and the physical interference model. Although this is reasonable to investigate and resolve the data collection issue for CRNs, we would like to study the performance of ADDC under different network distributions and interference models in the future. Second, the delay and capacity bounds derived in this work are from the view of expectation. This is because the probability that an SU has an opportunity to access the spectrum during a time slot is obtained from the view of expectation. Therefore, another future research direction is to investigate the delay and capacity performance of ADDC in the worst case. Third, in this work, we focus on snapshot and continuous

data collection. For other important communication modes, e.g. multicast, unicast, data aggregation, etc., we will design distributed and asynchronous algorithms in the future work of this paper. Fourth, we assume perfect spectrum sensing in this work which might induce a high cost. In the future, we would like to remove this assumption and study the data collection problem in CRNs where SUs may fail to sense the actual PUs' activities (named, *false alarm* and *missed detection*). Finally, in this paper, it is also assumed that each PU determines whether to initiate a data transmission only at the very beginning of each time slot. Moreover, if an SU senses that there is a primary data transmission (which is started at the very beginning of a time slot) within its PCR, it will freeze its backoff timer and keep silent during that time slot. This implies a PU will not return to transmit data on the licensed spectrum during the time when an SU is transmitting a data packet. Therefore, in the future, we also want to remove this assumption and investigate the data collection issue for CRNs where the network is completely asynchronous and PUs can initiate data transmissions anytime.
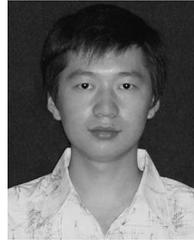
ACKNOWLEDGMENT

REFERENCES

[1] Z. Cai, S. Ji, J.S. He, and A.G. Bourgeois, ''Optimal Distributed Data Collection for Asynchronous Cognitive Radio Networks,'' in *Proc. IEEE ICDCS*, 2012, pp. 245-254.

[2] FCCET Docket No 03-237 Notice of Proposed Rule Making and Order, Washington, DC, USA, Dec. 2003.

[3] I.F. Akyildiz, W.-Y. Lee, M.C. Vuran, and S. Mohanty, ''NeXt Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey,'' *Comput. Netw.*, vol. 50, no. 13, pp. 2127-2159, Sept. 2006.

[4] X. Zhu, B. Tang, and H. Gupta, ''Delay Efficient Data Gathering in Sensor Networks,'' in *Proc. MSN*, 2005, pp. 380-389.

[5] S. Chen, S. Tang, M. Huang, and Y. Wang, ''Capacity of Data Collection in Arbitrary Wireless Sensor Networks,'' in *Proc. IEEE INFOCOM*, 2010, pp. 52-60.

[6] T. Zhang and D.H.K. Tsang, ''Optimal Cooperative Sensing Scheduling for Energy-Efficient Cognitive Radio Networks,'' in *Proc. IEEE INFOCOM*, 2011, pp. 2723-2731.

[7] Z. Li, F.R. Yu, and M. Huang, ''A Coorperative Spectrum Sensing Consensus Scheme in Cognitive Radios,'' in *Proc. IEEE INFOCOM*, 2009, pp. 2546-2550.

[8] T. Shu and M. Krunz, ''Throughput-Efficient Sequential Channel Sensing and Probing in Cognitive Radio Networks under Sensing Errors,'' in *Proc. ACM MOBICOM*, 2009, pp. 37-48.

[9] H. Kim and K.G. Shin, ''In-Band Spectrum Sensing in Cognitive Radio Networks: Energy Detection or Feature Detection,'' in *Proc. ACM MOBICOM*, 2008, pp. 14-25.

[10] S. Huang, X. Liu, and Z. Ding, ''Optimal Transmission Strategies for Dynamic Spectrum Access in Cognitive Radio Networks,'' *IEEE Trans. Mobile Comput.*, vol. 8, no. 12, pp. 1636-1648, Dec. 2009.

[11] S. Wang, J. Zhang, and L. Tong, ''Delay Analysis for Cognitive Radio Networks with Random Access: A Fluid Queue View,'' in *Proc. IEEE INFOCOM*, 2010, pp. 1-9.

[12] K. Khalil, M. Karaca, O. Ercetin, and E. Ekici, ''Optimal Scheduling in Cooperate-to-Join Cognitive Radio Networks,'' in *Proc. IEEE INFOCOM*, 2011, pp. 3002-3010.

[13] A.W. Min, K.-H. Kim, J.P. Singh, and K.G. Shin, ''Opportunistic Spectrum Access for Mobile Cognitive Radios,'' in *Proc. IEEE INFOCOM*, 2011, pp. 2993-3001.

[14] S. Ji, A.S. Uluagac, R. Beyah, and Z. Cai, ''Practical Unicast and Convergecast Scheduling Schemes for Cognitive Radio Networks,'' *J. Comb. Optim.*, vol. 26, no. 1, pp. 161-177, July 2013.

[15] D. Xue and E. Ekici, ''Guaranteed Opportunistic Scheduling in Multi-Hop Cognitive Radio Networks,'' in *Proc. IEEE INFOCOM*, 2011, pp. 2984-2992.

[16] L. Yang, H. Kim, J. Zhang, M. Chiang, and C.W. Tan, ''Pricing-Based Spectrum Access Control in Cognitive Radio Networks with Random Access,'' in *Proc. IEEE INFOCOM*, 2011, pp. 2228-2236.

[17] F. Wang, M. Krunz, and S. Cui, ''Spectrum Sharing in Cognitive Radio Networks,'' in *Proc. IEEE INFOCOM*, 2008, pp. 36-40.

[18] Y. Shi and Y.T. Hou, ''A Distributed Optimization Algorithm for Multi-Hop Cognitive Radio Networks,'' 2008, pp. 1966-1974.

[19] T. Shu and M. Krunz, ''Coordinated Channel Access in Cognitive Radio Networks: A Multi-Level Spectrum Opportunity Perspective,'' in *Proc. IEEE INFOCOM*, 2009, pp. 2976-2980.

[20] G.S. Kasbekar and S. Sarkar, ''Spectrum Pricing Games with Bandwidth Uncertainty and Spatial Reuse in Cognitive Radio Networks,'' in *Proc. ACM MOBIHOC*, 2010, pp. 251-260.

[21] B. Zhang, Y. Chen, and K.J.R. Liu, ''An Indirect Reciprocity Game Theoretic Framework for Dynamic Spectrum Access,'' in *Proc. IEEE ICC*, 2012, pp. 1747-1751.

[22] B. Zhang, Y. Chen, and K.J.R. Liu, ''An Indirect-Reciprocity Reputation Game for Cooperation in Dynamic Spectrum Access Networks,'' *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4328-4341, Dec. 2012.

[23] C. Jiang, Y. Chen, K.J.R. Liu, and Y. Ren, ''Analysis of Interference in Cognitive Radio Networks with Unknown Primary Behavior,'' in *Proc. IEEE ICC*, 2012, pp. 1721-1725.

[24] C. Jiang, Y. Chen, K.J.R. Liu, and Y. Ren, ''Renewal-Theoretical Dynamic Spectrum Access in Cognitive Radio Networks with Unknown Primary Behavior,'' *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 406-416, Mar. 2013.

[25] S.-W. Jeon, N. Devroye, M. Vu, S.-Y. Chung, and V. Tarokh, ''Cognitive Networks Achieve Throughput Scaling of a Homogeneous Network,'' *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5103-5115, Aug. 2008.

[26] C. Wang, X.-Y. Li, S. Tang, and C. Jiang, ''Multicast Capacity Scaling for Cognitive Networks: General Extended Primary Network,'' in *Proc. IEEE MASS*, 2010, pp. 262-271.

[27] C. Yin, L. Gao, and S. Cui, ''Scaling Laws for Overlaid Wireless Networks: A Cognitive Radio Network Versus a Primary Network,'' *IEEE/ACM Trans. Netw.*, vol. 18, no. 4, pp. 1317-1329, Aug. 2010.

[28] T. Jing, X. Chen, Y. Huo, and X. Cheng, ''Achievable Transmission Capacity of Cognitive Mesh Networks with Different Media Access Control,'' in *Proc. IEEE INFOCOM*, 2012, pp. 1764-1772.

[29] X. Chen, T. Jing, Y. Huo, W. Li, X. Cheng, and T. Chen, ''Achievable Transmission Capacity of Cognitive Radio Networks with Cooperative Relaying,'' in *Proc. CROWNCOM*, 2012, pp. 1-6.

[30] C. Li and H. Dai, ''Transport Throughput of Secondary Networks in Spectrum Sharing Systems,'' in *Proc. IEEE INFOCOM*, 2011, pp. 2732-2740.

[31] W. Huang and X. Wang, ''Throughput and Delay Scaling of General Cognitive Networks,'' in *Proc. IEEE INFOCOM*, 2011, pp. 2210-2218.

[32] S. Kompella, G.D. Nguyen, J.E. Wieselthier, and A. Ephremides, ''Stable Throughput Tradeoffs in Cognitive Shared Channels with Cooperative Relaying,'' in *Proc. IEEE INFOCOM*, 2011, pp. 1961-1969.

[33] L. Sun and W. Wang, ''On Distribution and Limits of Information Dissemination Latency and Speed in Mobile Cognitive Radio Networks,'' in *Proc. IEEE INFOCOM*, 2011, pp. 246-250.

[34] C. Li and H. Dai, ''On the Throughput Scaling of Cognitive Radio Ad Hoc Networks,'' in *Proc. IEEE INFOCOM*, 2011, pp. 241-245.

[35] C. Han and Y. Yang, ''Information Propagation Speed Study in Multihop Cognitive Radio Networks,'' in *Proc. IEEE INFOCOM*, 2011, pp. 226-230.

[36] S. Ji, R. Beyah, and Y. Li, ''Continuous Data Collection Capacity of Wireless Sensor Networks under Physical Interference Model,'' in *Proc. IEEE MASS*, 2011, pp. 222-231.

[37] S. Ji, J.S. He, A.S. Uluagac, R. Beyah, and Y. Li, ''Cell-Based Snapshot and Continuous Data Collection in Wireless Sensor Networks,'' *ACM Trans. Sensor Netw.*, vol. 9, no. 4, p. 47, July 2013.

[38] S. Ji, R. Beyah, and Z. Cai, ''Snapshot/Continuous Data Collection Capacity for Large-Scale Probabilistic Wireless Sensor Networks,'' in *Proc. IEEE INFOCOM*, 2012, pp. 1035-1043.

[39] O. Dousse, F. Baccelli, and P. Thiran, ''Impact of Interferences on Connectivity in Ad Hoc Networks,'' in *Proc. IEEE INFOCOM*, 2003, pp. 1724-1733.

[40] W. Ren, Q. Zhao, and A. Swami, ''On the Connectivity and Multihop Delay of Ad Hoc Cognitive Radio Networks,'' in *Proc. IEEE ICC*, 2010, pp. 1-6.

[41] P. Wang, I.F. Akyildiz, and A.M. Al-Dhelaan, ''Percolation Theory Based Connectivity and Latency Analysis of Cognitive Radio Ad Hoc Networks,'' *Wireless Netw.*, vol. 17, no. 3, pp. 659-669, Apr. 2011.

[42] X. Huang, D. Lu, P. Li, and Y. Fang, ''Coolest Path: Spectrum Mobility Aware Routing Metrics in Cognitive Ad Hoc Networks,'' in *Proc. IEEE ICDCS*, 2011, pp. 182-191.

[43] T. Shu and M. Krunz, ''Truthful Least-Priced-Path Routing in Opportunistic Spectrum Access Networks,'' in *Proc. IEEE INFOCOM*, 2010.

[44] M. Pan, C. Zhang, P. Li, and Y. Fang, ''Joint Routing and Link Scheduling for Cognitive Radio Networks under Uncertain Spectrum Supply,'' in *Proc. IEEE INFOCOM*, 2011, pp. 2237-2245.

[45] J. Jin, H. Xu, and B. Li, ''Multicast Scheduling with Cooperation and Network Coding in Cognitive Radio Networks,'' in *Proc. IEEE INFOCOM*, 2010, pp. 1-9.

[46] D. Hu, S. Mao, and J.H. Reed, ''On Video Multicast in Cognitive Radio Networks,'' in *Proc. IEEE INFOCOM*, 2009, pp. 2222-2230.

[47] L. Fu, S.C. Liew, and J. Huang, ''Effective Carrier Sensing in CSMA Networks Under Cumulative Interference,'' in *Proc. IEEE INFOCOM*, 2010, pp. 1-9.

[48] S. Ji, Y. Li, and X. Jia, ''Capacity of Dual-Radio Multi-Channel Wireless Sensor Networks for Continuous Data Collection,'' in *Proc. IEEE INFOCOM*, 2011, pp. 1062-1070.

[49] S. Ji, Z. Cai, Y. Li, and X. Jia, ''Continuous Data Collection Capacity of Dual-Radio Multi-Channel Wireless Sensor Networks,'' *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 10, pp. 1844-1855, Oct. 2012.

[50] P.-J. Wan, S.C.-H. Huang, L. Wang, Z. Wan, and X. Jia, ''Minimum-Latency Aggregation Scheduling in Multihop Wireless Networks,'' in *Proc. ACM MOBIHOC*, 2009, pp. 185-194.

[51] S. Ross, *Introduction to Probability Models,* 9th ed., Amsterdam, The Netherlands: Elsevier, 2007.

[52] S. Ji and Z. Cai, ''Distributed Data Collection and its Capacity in Asynchronous Wireless Sensor Networks,'' in *Proc. IEEE INFOCOM*, 2012, pp. 2113-2121.

[53] S. Ji and Z. Cai, ''Distributed Data Collection in Large-Scale Asynchronous Wireless Sensor Networks Under the Generalized Physical Interference Model,'' *IEEE/ACM Trans. Netw.*, vol. 21, no. 4, pp. 1270-1283, Aug. 2013.

[54] J. Li and S. Cheng, ''(ϵ, δ)-Approximate Aggregation Algorithms in Dynamic Sensor Networks,'' *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 3, pp. 385-396, Mar. 2012.

[55] S. Cheng, J. Li, and Z. Cai, ''$O(\epsilon)$-Approximation to Physical World by Sensor Networks,'' in *Proc. IEEE INFOCOM*, 2013, pp. 3084-3092.

Zhipeng Cai is currently an Assistant Professor at the Department of Computer Science at the Georgia State University, Atlanta, GA, USA. He received the BS degree in computer science from Beijing Institute of Technology, Beijing, China, and the MS and PhD degrees in computing science from University of Alberta, Canada. His research interests include wireless networks and optimization theory. Dr. Cai received a number of awards and honors, including US NSF CAREER Award and NSERC Postdoc Fellowship. He is a member of the IEEE.

Shouling Ji is a PhD Candidate at the Department of Computer Science at Georgia State University, Atlanta, GA, USA. He received the BS and MS degrees in computer science from Heilongjiang University, China, in 2007 and 2010, respectively. His research interests include Wireless Sensor Networks, Cognitive Radio Networks, and Network Evolution Analysis of Social Networks. He is now a Student Member of the ACM, IEEE, and IEEE COMSOC. He is also the Membership Chair of the IEEE Student Branch at Georgia State.

Jing (Selena) He is currently an Assistant Professor at the Department of Computer Science at Kennesaw State University. She received the PhD and MS degrees from the Department of Computer Science at Georgia State University (GSU), USA emphasized in Wireless Networking. Before joining GSU, she received the BS degree in electronic engineering from Wuhan Institute of Technology, Wuhan, Hubei, China. She received another MS degree in computer science concentrated in Artificial Intelligence from Utah State University, Salt Lake City, UT, USA. Her research interests include Wireless Networking, Network Optimization, Social Networks, and Cloud Computing.

Lin Wei is currently an Assistant Professor at the Department of Agricultural and Biosystems at the South Dakota State University, USA. He received the BS degree in agricultural engineering from China Agricultural University, China, and his MS and PhD degrees in biological engineering from Mississippi State University. His research interests include Catalytic fast pyrolysis of agricultural and forest residues for aromatic fuel additives, and Biomass syngas to gasoline through thermochemical pathways.

Anu G. Bourgeois received the BS degree in electrical engineering from Louisiana State University, USA, in 1991 and the MS and PhD degrees in electrical and computer engineering from Louisiana State University, Baton Rouge, LA, in 1997 and 2000, respectively. She is currently an Associate Professor at the Department of Computer Science at Georgia State University, USA. She has been serving on the editorial board of the *International Journal of High Performance Computing and Networking* since 2005. Her research interests include wireless networks, parallel and distributed algorithms, fault tolerance and STEM education. She is a Senior Member of the IEEE.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.