

Supplementary File for “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, 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 data of a snapshot to the base station in a distributed manner without any 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 the continuous data collection issue, and analyze the delay and capacity performance of ADDC in continuous data collection, which are also proven to be order-optimal. Finally, extensive simulation results indicate that ADDC can effectively finish 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 RELATED WORK

Ever since the CRN communication paradigm was proposed, extensive research has been conducted on spectrum sensing [6]-[9], spectrum access, scheduling and management [10]-[20], capacity/throughput/delay scaling laws [25]-[35], network connectivity [40], [41], routing protocols [42]-[44], and multicast communication [45], [46]. In this section, we summarize the representative works for different issues in CRNs as well as the latest data collection algorithms for traditional wireless networks. We also remark the differences between this work and existing works at the end of this section.

1.1 Spectrum Sensing

Since the spectrum is one of the most precious resources in CRNs, to sense and learn the instantaneous spectrum opportunities in the communication environment is crucial in designing an efficient CRN. In [6]-[9], the authors studied the spectrum sensing problem. In [6], the authors investigated the problem of optimal Cooperative Sensing Scheduling (CSS) and parameter design to achieve energy efficiency in CRNs using the framework of Partially Observable Markov Decision Process (POMDP). In [7], a distributed consensus-based cooperative spectrum sensing scheme that copes with both fixed and random bidirectional connections among SUs is proposed. In [8], the authors studied the throughput-efficient sequential channel sensing and probing problem. An optimal use-or-skip decision strategy that maximizes a CRN's average throughput is derived by formulating the sequential sensing and probing process as a rate-of-return problem, which can be solved by optimal stopping theory. In [9], the authors proposed an efficient periodic in-band sensing algorithm that optimizes sensing-frequency and sensing-time by minimizing sensing overhead while meeting the detect ability requirements. In the proposed algorithm, the noise uncertainty and inter-CRN interference which affect detection performance are also considered.

A preliminary version of this work has been accepted to publish at The 32nd International Conference on Distributed Computing Systems (IEEE ICDCS 2012) [1].

Z. Cai is with the Department of Computer Science, Georgia State University, 34 Peachtree Street, Suite 1412, Atlanta, GA 30303. Email: zcai@gsu.edu.

S. Ji is with the Department of Computer Science, Georgia State University, 34 Peachtree Street, Suite 1428, Atlanta, GA 30303 and the School of Electrical and Computer Engineering, Georgia Institute of Technology, 266 Ferst Drive Atlanta, GA 30332-0765. Email: sjj@gatech.edu.

J. (S.) He is with the Department of Computer Science, Kennesaw State University, MD 1101, CL, Bldg. 11, Rm. 3015, Kennesaw, GA 30144-5591. Email: jhe4@kennesaw.edu.

L. Wei is with the Department of Agricultural and Biosystems, South DAKOTA State University Email: lin.wei@sdstate.edu.

A. Bourgeois is with the Department of Computer Science, Georgia State University, 34 Peachtree Street, Suite 1449, Atlanta, GA 30303. Email: abourgeois@cs.gsu.edu.

1.2 Spectrum Access, Scheduling, and Management

An elegant spectrum access/scheduling/management scheme can improve spectrum utilization efficiency and reduce the interference to PUs caused by SUs. This issue attracts much attention [10]–[24]. In [10], the authors investigated the performance limitation on the throughput of CRNs under the PU packet collision constraint. They proposed an optimum spectrum access strategy under generic PU traffic patterns with perfect sensing, and a modified threshold-based spectrum access strategy which achieves close-to-optimal performance without perfect sensing. In a CRN where multiple SUs exploit random access to contend for spectrum usage over available PU channels, the SU queuing delay performance is analyzed in [11]. The authors took a fluid queue approximation approach to study the steady-state delay of SUs for the cases with a single PU channel and multiple PU channels. In [12], the authors studied the optimal transmission scheduling in CRNs under the spectrum leasing model. They proposed a cooperative scheme in which SUs share the time slot with PUs in return for cooperation. In [13], the authors studied the spectrum access problem for mobile CRNs. In [14], practical unicast and convergeicast scheduling schemes for SUs are introduced. In [15], the authors studied the optimal control and scheduling problem for multi-hop CRNs to maximize the throughput of SUs while stabilizing the CRN subject to collision rate constraints required by PUs. In addition, the authors proposed a class of feasible suboptimal algorithms to reduce the algorithm complexity.

The pricing-based spectrum access control in CRNs with random access is studied in [16] under two market models: a monopoly PU market and a multiple PUs market where each PU sells its temporarily unused spectrum to SUs. They proposed a pricing-based spectrum trading mechanism that enables SUs to contend for channel usage in a distributed manner, which naturally mitigates the complexity and time overhead associated with centralized scheduling. In [17], the joint power/channel allocation problem is studied, where the spectrum allocation problem is modelled as a non-cooperative game and a Price-based Iterative Water-Filling (PIWF) algorithm is proposed, which allows users to converge to the Nash Equilibrium. In [18], the authors investigated how to design a distributed algorithm for CRNs with the objective of maximizing data rates for a set of user communication sessions. They studied this problem via a cross-layer optimization approach, with joint consideration of power control, scheduling, and routing. In [19], the authors proposed a centralized polynomial-time approximate algorithm to solve the problem. In [20], the authors formulated price competition in a CRN as a game, taking into account both bandwidth uncertainty and spatial reuse. They analyzed the game in a single slot as well as its repeated case. For each case, the authors proved the existence of a Nash equilibrium and provided

a method to explicitly compute it.

In [21][22], the authors studied the cooperation stimulation problem. They proposed a reputation-based spectrum access framework where the studied problem is modeled as an indirect reciprocity game. By formulating a SU's decision making as a Markov decision process, the authors obtained the optimal action rule, according to which the SUs will cooperate with the PUs to improve the spectrum utilization efficiency. In [23][24], the authors studied how to dynamically control SUs to access the spectrum and analyzed the induced interference caused during the access process. By modeling the SUs' communication behavior as an ON-OFF process, they showed that SUs' communication behavior is a renewal process.

1.3 Capacity/Throughput/Delay Scaling Laws

The capacity, throughput, and delay scaling issues for CRNs are studied in [25]–[35] and references therein. By introducing preservation regions around primary receivers and avoidance regions around primary base stations, the authors proposed two modified multi-hop routing protocols for SUs in [25]. Based on percolation theory, they showed that when the secondary network is denser than the primary network and the primary network throughput is subject to a fractional loss, both networks can simultaneously achieve the same throughput scaling law as a stand-alone network. Similarly, by percolation theory, the authors in [26] designed an optimal-throughput strategy for secondary networks depending on which type of strategy is adopted in primary networks. They showed that there exists a threshold of the density of SUs according to the density of PUs such that the secondary network can achieve the multicast capacity of the same order as it is stand-alone. The authors in [27] studied the capacity and delay-throughput scaling laws for CRNs. Under certain assumptions, they showed that both primary and secondary networks can achieve the same capacity and delay-throughput scaling laws. In [28], the authors investigated transmission capacity with outage probability constraints for both the primary and the secondary networks. The authors in [29] studied the achievable transmission capacity of a secondary network when it provides cooperative relaying to the primary network under the underlay spectrum sharing model. In [30], the authors explored the Single hop Transport Throughput (STT) of CRNs with outage constraints imposed on both the primary and the secondary networks. They first derived the limitation of STT, single hop transport capacity, together with a practical upper bound for it. Then, they studied STT with SUs randomly deployed. The authors in [31] demonstrated in the case that the transmission range of the secondary network is smaller than that of the primary network in order, a hybrid protocol model suffices to guide the secondary network to achieve the same throughput and delay scaling as a stand-alone network, as long as the primary network

operates in a generalized round-robin TDMA manner. In [32], the authors studied the stable throughput tradeoffs in CRNs with cooperative relaying. They first analyzed the non-cooperation case where all nodes transmit their own packets to their respective destinations. Subsequently, they analyzed a CRN where the SU cooperatively relays some of the primary's packets. In [33], the authors studied the distributions and limitations of information dissemination latency and speed in a CRN where PUs are static and SUs are mobile. The authors in [35] established an Information propagation Speed (IPS) model in CRNs, and obtained the maximum network IPS that maximizes IPS across a network topology over an infinite plane.

1.4 Connectivity

Following the work of [39], which studies the connectivity of ad hoc networks via percolation theory, the authors in [40], [41] and references therein studied the connectivity issue of CRNs. They exploited theories and techniques from continuum percolation and ergodicity to derive the scaling law of the minimum multihop delay with respect to the source-destination distance in CRNs.

1.5 Routing Protocols

In [42]-[44], the authors considered the routing protocols for CRNs. To better characterize the unique features of CRNs, the authors in [42] proposed some new routing metrics, which include accumulated spectrum temperature, highest spectrum temperature, and mixed spectrum temperature to account for the time-varying spectrum availability. In [43], the authors investigated the problem of finding the Least-Priced Path (LPP) between a source and a destination in CRNs. They obtained an optimal route selection and payment determination mechanism that minimizes the price tag of the selected route and at the same time guarantees truthful cost reports from SUs. By modelling the vacancy of licensed bands with a series of random variables, introducing corresponding scheduling constraints and flow routing constraints, the authors in [44] studied the joint routing and link scheduling problem of multihop CRNs.

1.6 Multicast Communication

Multicast is an important operation in wired/wireless networks, as well as in CRNs [45]. The authors in [45] proposed an optimization framework for multicast operations for CRNs. In the proposed framework, a multicast is accomplished by carefully tuning the power. Concurrently, SUs opportunistically perform cooperative transmissions using locally idle primary channels in order to mitigate multicast loss and delay effects. In another work [46], the authors studied the video multicast issue in CRNs. With the objectives of optimizing the overall received video quality and achieving proportional fairness among multicast users, they model the CRN

video multicast as an optimization problem considering important design factors such as scalable video coding, video rate control, spectrum sensing, dynamic spectrum access, modulation, scheduling, retransmission, and PU protection.

1.7 Data Collection in Traditional Wireless Networks

The data collection problem is well studied for traditional wireless networks recently. In [4], the authors showed the hardness of the data collection problem and proposed some of path-based data collection algorithms. Later on, the authors in [5][48][49] improved the path-based scheduling idea and designed a series of data collection algorithms with better capacity/latency bounds. In [36][37], the authors studied the data collection problem under the more practical physical interference model. By applying the network partition idea, they designed an order-optimal data collection algorithm with respect to network capacity. In [38], considering the transitional region phenomenon in wireless sensor networks, the authors studied the data collection problem for probabilistic wireless sensor networks. They implemented a data collection method and analyzed its achievable network capacity under the probabilistic network model. Recently, the authors in [52][53] considered the distributed data collection problem. They demonstrated that order-optimal data collection capacity can also be achieved by distributed algorithms. In [54], a Bernoulli sampling based data collection method was proposed, and this method dramatically reduced the amount of data transmitted in-network since the sampling technique was adopted. In [55], a data collection algorithm on continuous curves, which are more appropriate to express the continuous property of the monitored physical world, is considered. As indicated in [3], the data collection scheduling algorithms for traditional wireless networks cannot be applied in CRNs. The main reasons are that (i) the spectrum opportunities are dynamic in CRNs. However, the spectrum in traditional wireless networks are usually static; and (ii) secondary activities should not cause any unacceptable interference to the primary activities in CRNs. On the other hand, in the works for traditional wireless networks, all the nodes are equal with respect to exploit spectrum opportunities.

1.8 Remarks

Very few of the above mentioned works consider the data collection issue in CRNs, especially distributed data collection in asynchronous CRNs. Furthermore, most of the existing works for CRNs either are centralized algorithms or require time synchronization. However, as pointed out in Section 1, CRNs tend to be distributed systems and prefer distributed and asynchronous algorithms. Motivated by this fact, we propose a distributed data collection algorithm for CRNs without time synchronization requirement. Through theoretical

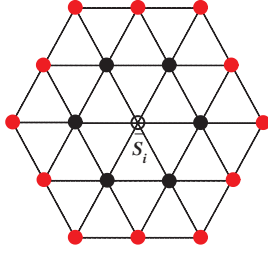


Fig. 1. Hexagon packing with respect to \bar{S}_i . The black (respectively, red) nodes are the nodes at the first (respectively, second) layer.

analysis, we show that the proposed algorithm successfully achieves order-optimal data collection capacity as centralized data collection algorithms for traditional wireless networks.

2 PROOF OF LEMMA 2

Proof of Lemma 2 in the Main File: For $\forall u \in \mathbb{C} \cap V_p$, let u' be its corresponding receiver. Then, to guarantee the data transmission from an arbitrary $S_i \in \mathbb{C}$ to S'_i , we have

$$\frac{P_p \cdot D(S_i, S'_i)^{-\alpha}}{\sum_{S_k \in \mathbb{C}, S_k \neq S_i} P_p D(S_k, S'_i)^{-\alpha} + \sum_{S_k \in \mathbb{C}} P_s D(S_k, S'_i)^{-\alpha}} \geq \eta_p.$$

Since $D(S_i, S'_i)$ is the distance between S_i and S'_i , $D(S_i, S'_i) \leq R$, which implies $D(S_i, S'_i)^{-\alpha} \geq R^{-\alpha}$. Furthermore,

$$\begin{aligned} & \sum_{S_k \in \mathbb{C}, S_k \neq S_i} P_p \cdot D(S_k, S'_i)^{-\alpha} + \sum_{S_k \in \mathbb{C}} P_s \cdot D(S_k, S'_i)^{-\alpha} \quad (1) \\ & \leq \sum_{U \in \mathbb{C}, U \neq S_i} \max\{P_p, P_s\} \cdot D(U, S'_i)^{-\alpha} \quad (2) \\ & = \max\{P_p, P_s\} \cdot \sum_{U \in \mathbb{C}, U \neq S_i} D(U, S'_i)^{-\alpha}. \quad (3) \end{aligned}$$

To obtain the upper bound of $\sum_{U \in \mathbb{C}, U \neq S_i} D(U, S'_i)^{-\alpha}$, for every $v \in \mathbb{C}$, we abstract the communication pair/link (v, v') to a node denoted by \bar{v} . In addition, for any \mathcal{R} -set, the densest packing is the *hexagon packing* as shown in Fig. 1. Then, all the nodes in \mathbb{C} can be layered with respect to \bar{S}_i . By mathematical induction, it can be proven that the number of nodes in the l -th layer is at most $6l$ and the distance from \bar{S}_i to any node at the l -th layer is at least $\frac{\sqrt{3}}{2}l \cdot F$ ($l \geq 2$), where $F = \mathcal{R} - R$ (since for $\forall v, w \in \mathbb{C}$, $D(v, w) \geq \mathcal{R}$, therefore, $D(v, w') \geq \mathcal{R} - R$).

Hence, we have

$$\sum_{U \in \mathbb{C}, U \neq S_i} D(U, S'_i)^{-\alpha} \quad (4)$$

$$\leq 6F^{-\alpha} + \sum_{l \geq 2} 6l \left(\frac{\sqrt{3}}{2}lF\right)^{-\alpha} \quad (5)$$

$$= 6F^{-\alpha} + 6\left(\frac{\sqrt{3}}{2}F\right)^{-\alpha} \cdot \sum_{l \geq 2} l^{-\alpha+1} \quad (6)$$

$$\leq 6F^{-\alpha} + 6\left(\frac{\sqrt{3}}{2}F\right)^{-\alpha} \cdot (\zeta(\alpha-1) - 1) \quad (7)$$

$$\leq 6F^{-\alpha} + 6\left(\frac{\sqrt{3}}{2}F\right)^{-\alpha} \cdot \left(\frac{1}{\alpha-2} - 1\right) \quad (8)$$

$$= \left(6 + 6\left(\frac{\sqrt{3}}{2}\right)^{-\alpha} \left(\frac{1}{\alpha-2} - 1\right)\right) \cdot F^{-\alpha}. \quad (9)$$

Inequality (7) comes from the fact that $\sum_{l \geq 1} l^{-\alpha+1}$ is a particular case of the *Riemann zeta function* $\zeta(\cdot)$ with parameter $\alpha - 1$. Inequality (8) is based on $\zeta(x) \leq \frac{1}{x-1}$. Let $c_1 = \frac{P_p}{\max\{P_p, P_s\}}$ and $c_2 = 6 + 6\left(\frac{\sqrt{3}}{2}\right)^{-\alpha} \left(\frac{1}{\alpha-2} - 1\right)$. Then, we have

$$\frac{P_p \cdot D(S_i, S'_i)^{-\alpha}}{\sum_{S_k \in \mathbb{C}, S_k \neq S_i} P_p \cdot D(S_k, S'_i)^{-\alpha} + \sum_{S_k \in \mathbb{C}} P_s \cdot D(S_k, S'_i)^{-\alpha}} \quad (10)$$

$$\geq \frac{c_1 R^{-\alpha}}{c_2 F^{-\alpha}}. \quad (11)$$

Therefore, to guarantee the data transmission from S_i to S'_i , it is sufficient to have

$$\frac{c_1 R^{-\alpha}}{c_2 F^{-\alpha}} \geq \eta_p \Leftrightarrow F \geq \sqrt[\alpha]{\frac{c_2 \eta_p}{c_1}} \cdot R \quad (12)$$

$$\Leftrightarrow \mathcal{R} \geq \left(1 + \sqrt[\alpha]{\frac{c_2 \eta_p}{c_1}}\right) \cdot R. \quad (13)$$

Note that S_i is arbitrarily chosen in $\mathbb{C} \cap V_p$, this lemma holds. \square

3 PROOFS OF LEMMA 5, 6, 7, AND 8

3.1 Proof of Lemma 5

Proof of Lemma 5 in the Main File: Let \mathcal{D}_κ be a disk of radius $\kappa \cdot r$ and $\mathcal{D}_{\kappa+1}$ be a disk of radius $(\kappa + 1) \cdot r$. Then, according to Lemma 4, we have $|\mathcal{D}_\kappa \cap \mathcal{D}| \leq \beta_\kappa$ and $|\mathcal{D}_{\kappa+1} \cap \mathcal{D}| \leq \beta_{\kappa+1}$. Now, we consider the number of connectors. It is possible for some connectors that are only connected with some dominators located at disk $\mathcal{D}_{\kappa+1}$ however not in disk \mathcal{D}_κ . Since each dominator connects with at most 12 connectors according to Lemma 1, we have $|\mathcal{D}_\kappa \cap \mathcal{C}| \leq 12\beta_{\kappa+1}$. In summary, this lemma holds. \square

3.2 Proof of Lemma 6

Proof of Lemma 6 in the Main File: Based on Lemma 5 and the construction process of the data collection tree, it is straightforward that the number of SUs within the PCR of an SU is upper bounded by $\Delta\beta_\kappa + 12\beta_{\kappa+1}$ (when

count the number, we should exclude the parent node of a dominator SU which is a connector and include the dominator SU itself). Subsequently, we analyze the upper bound of Δ . Let X be a random variable denoting the number of SUs within the communication range of an SU. Since the secondary network is *i.i.d.*, X is a binomial random variable with parameters $(n, p = \frac{\pi r^2}{A})$. Then, applying the Chernoff bound, we have

$$\Pr(X > a) \leq \min_{\xi > 0} \frac{\mathbf{E}[e^{\xi X}]}{e^{\xi a}} \quad (14)$$

$$= \min_{\xi > 0} \frac{[1 + (e^{\xi} - 1)p]^n}{e^{\xi a}} \quad (15)$$

$$\leq \min_{\xi > 0} \exp[(e^{\xi} - 1)np - \xi a]. \quad (16)$$

Let $\xi = 2$ and $a = \log n + \frac{\pi r^2(e^2 - 1)}{2c_0}$, we have $\Pr(X > \log n + \frac{\pi r^2(e^2 - 1)}{2c_0}) \leq \exp(-2 \log n) \leq \frac{1}{n^2}$. Considering that $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$, it follows $\Pr(X \leq \log n + \frac{\pi r^2(e^2 - 1)}{2c_0}) = 1$ according to the Borel-Cantelli Lemma. \square

3.3 Proof of Lemma 7

Proof of Lemma 7 in the Main File: Since the primary network is *i.i.d.*, the expected number of PUs within the PCR of an SU is $\frac{\pi(\kappa r)^2 N}{c_0 n}$. In addition, a PU conducts a data transmission with probability p_t during a time slot. It follows that the expected probability of a spectrum opportunity appearance for an SU is $p_o = (1 - p_t)^{\pi(\kappa r)^2 N / c_0 n}$. Considering the locally finite property of the primary network, we have $p_o > 0$, which implies every SU has a positive probability to access the spectrum and the expected waiting time is τ/p_o . \square

3.4 Proof of Lemma 8

Proof of Lemma 8 in the Main File: After all the packets at $V_s \setminus (\mathcal{D} \cup \mathcal{C})$ have been collected to $\mathcal{D} \cup \mathcal{C}$, only the SUs in $\mathcal{D} \cup \mathcal{C}$ may have data packets for transmission. From Lemma 5, there are at most $\beta_{\kappa} + 12\beta_{\kappa+1} - 1$ dominator and connector SUs within the PCR of an SU in $\mathcal{D} \cup \mathcal{C}$. From the proof of Theorem 1, this lemma can be proven. \square

4 SIMULATION AND ANALYSIS

In this section, we validate the performance of the proposed ADDC via simulations. In all the simulations, we consider a secondary network consisting of n SUs and a base station coexisted with a primary network consisting of N PUs. Both of the networks are *i.i.d.* in a square area with size A . All the SUs and PUs share the same time, space, and spectrum and SUs cannot cause any unacceptable interference to PUs. The network time is assumed to be slotted and the time duration of a time slot is 1 millisecond (ms). During a time slot, each PU initiates a data transmission with probability p_t or keeps silent with probability $1 - p_t$. In addition, the propagation

time of a data packet (no matter produced by a PU or SU) is less than 1 ms. For other system parameters, *e.g.* $A, \alpha, N, P_p, R, \eta_p, p_t, n, P_s, r, \eta_s$, and *etc* (which have the same meanings as defined before), we will specify them in each group of simulations (see Fig. 2).

In the simulations, we use ADDC to denote our proposed asynchronous and distributed data collection algorithm for CRNs. In ADDC, the contention window τ_w is assumed be 0.5 millisecond. Since there is no existing data collection algorithm for CRNs currently to the best of our knowledge, we compare ADDC with the most recently published routing algorithm for CRNs (necessary modification is required), denoted by Coolest [42]¹. In Coolest, the path with the most balanced and/or the lowest spectrum utilization by PUs is preferred for a data transmission. To finish a data collection task, in Coolest, we assume each SU of the secondary network produces a data packet that will be transmitted to the base station. In the following, each group of simulations is repeated for 10 times and the results are the average values.

4.1 Data Collection Delay vs. Network Size (N and n)

When we change the number of PUs or SUs in a CRN, the changes of data collection delay of ADDC and Coolest are shown in Fig. 2 (a) and (b), respectively. From Fig. 2 (a) and (b), we can see that when the number of PUs N and the number of SUs n increase, the induced data collection delay of both ADDC and Coolest increases. This is because (i) more PUs implies more data transmission activities of the primary network (the default $p_t = 0.3$), which further implies an SU has to wait longer for a spectrum opportunity; and (ii) when n increases, the traffic of a data collection task becomes heavier, *i.e.* the data communication in the secondary network becomes more crowded. It follows that more data collection delay is induced. ADDC has a better performance than that of Coolest. This is because ADDC takes a distributed data collection manner with fairness consideration, which can reduce the data accumulation effect and enable as many SUs as possible to transmit data simultaneously. On the other hand, since Coolest prefers the path with the most balanced and/or the lowest spectrum utilization by PUs, many SUs might choose the same path. This will make the data accumulation effect more serious. From Fig. 2 (a) and (b), we can also see that the increase trend in Fig. 2 (a) is much faster than Fig. 2 (b). This demonstrates that the waiting time for spectrum opportunities is the majority in the overall data collection delay. On average, ADDC induces 266% and 282% less delay compared with Coolest, respectively.

1. It might be not very fair to compare our algorithm with Coolest since Coolest is a routing algorithm which is not dedicated to the distributed data collection problem. We choose Coolest mainly because the lacking of distributed data collection algorithms for CRNs currently and Coolest can also finish the data collection task with little modification.

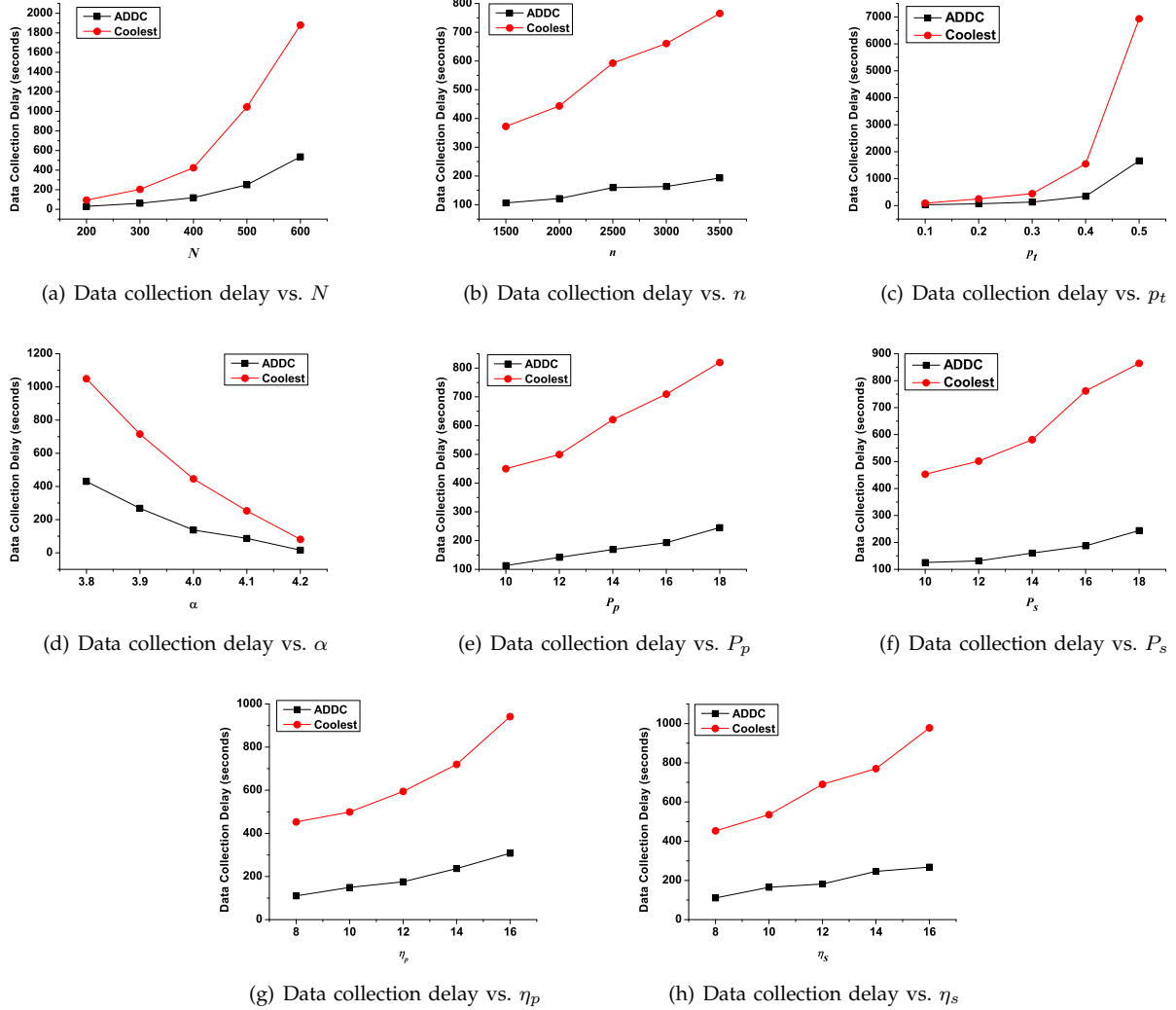


Fig. 2. Data collection delay of ADDC and Coolest. The default settings are $A = 250 \times 250$, $\alpha = 4$, $N = 400$, $P_p = 10$, $R = 10$, $\eta_p = 8\text{dB}$, $p_t = 0.3$, $n = 2000$, $P_s = 10$, $r = 10$, $\eta_s = 8\text{dB}$, and $Q = 1$ (i.e. snapshot data collection).

4.2 Data Collection Delay vs. p_t and α

The impact of p_t (the probability of a PU to initiate a data transmission during a time slot) on data collection delay of ADDC and Coolest is shown in Fig. 2 (c). From Fig. 2 (c), when p_t increases, the induced data collection delay of both ADDC and Coolest increases very fast. This is because the spectrum opportunities for SUs decrease fast with the increase of p_t , i.e. more activities of the primary network. The impact of the path loss exponent α on the data collection delay of ADDC and Coolest is shown in Fig. 2 (d). When α increases, the interference induced by a transmitter to other ongoing transmissions decreases. Therefore, SUs might have more spectrum opportunities and more SUs can conduct data transmissions concurrently without interference. It follows that the delay of both ADDC and Coolest decreases. Again, as shown in Fig. 2 (c) and (d), ADDC has a better performance than Coolest. On average, ADDC takes 314% and 171% less time than Coolest to finish a data collection task in Fig.

2 (c) and (d), respectively.

4.3 Data Collection Delay vs. Transmission Power (P_p and P_s)

The impacts of P_p (power of PUs) and P_s (power of SUs) on the performance of ADDC and Coolest are shown in Fig. 2 (e) and (f), respectively. From Fig. 2 (e) and (f), we can see that when P_p and P_s increase, the induced delay of both ADDC and Coolest increases. This is because a large working power will cause more interference to other transmissions. Therefore, the spectrum opportunities are reduced and less SUs can conduct data transmissions simultaneously. Because of the same reasons stated before, ADDC finishes a data collection task much faster than Coolest. On average, it takes ADDC 260% and 273% less time to finish a data collection task compared with Coolest as shown in Fig. 2 (e) and (f), respectively.

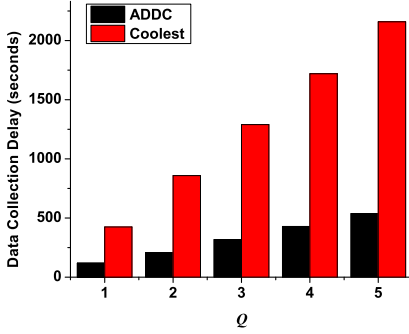


Fig. 3. Continuous data collection delay of ADDC and Coolest. The default settings are $A = 250 \times 250$, $\alpha = 4$, $N = 400$, $P_p = 10$, $R = 10$, $\eta_p = 8\text{dB}$, $p_t = 0.3$, $n = 2000$, $P_s = 10$, $r = 10$, and $\eta_s = 8\text{dB}$.

4.4 Data Collection Delay vs. SIR Threshold Value (η_p and η_s)

When the SIR threshold values η_p and η_s change, the changes of the induced data collection delay of ADDC and Coolest are shown in Fig. 2 (g) and (h), respectively. With the increase of η_p and η_s , the delay of both ADDC and Coolest shows an increasing trend. This is because that a large SIR threshold value implies less wireless interference, *i.e.* less spectrum opportunities for SUs and less SUs can carry out data transmissions simultaneously. Again, ADDC achieves a much better performance than Coolest. On average, ADDC induces 260% and 273% less data collection time compared with Coolest as shown in Fig. 2 (g) and (h), respectively.

4.5 Delay of Continuous Data Collection

The impact of Q (the number of snapshots in a continuous data collection task) on the continuous data collection delay of ADDC and Coolest is shown in Fig. 3. From Fig. 3, we can see that with the increase of Q , the induced delay of both methods increases. This is a straightforward result from the fact that large Q implies more communication traffic. From Fig. 3, we can also see that ADDC has much better performance than Coolest in terms of delay, and the delay increasing trend of Coolest is faster than that of ADDC. Specifically, the average delay increasing of ADDC with respect to Q is 88.33%, while the average delay increasing of Coolest with respect to Q is 102.33%. This is because (i) ADDC is a distributed and asynchronous scheduling algorithm. In ADDC, each SU makes actions only based on the local wireless communication environment within its PCR, and thus as many as possible SUs may be scheduled concurrently. Furthermore, the fairness is considered in ADDC, which can alleviate the data accumulation problem; and (ii) on the other hand, for continuous data collection, the data accumulation problem becomes more serious in Coolest since the SUs in Coolest prefer to choose the most balanced path and/or the path with the

lowest spectrum utilization by PUs. On average, ADDC induces 301.24% less delay compared with Coolest in continuous data collection.

ACKNOWLEDGMENT

This work is partly supported by the National Science Foundation (NSF) under grants No. CNS-1152001 and CNS-1252292.

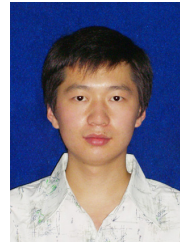
REFERENCES

- [1] Z. Cai, S. Ji, J. (S.) He, and A. G. Bourgeois, Optimal Distributed Data Collection for Asynchronous Cognitive Radio Networks, ICDCS 2012.
- [2] FCC, ET Docket No 03-237 Notice of Proposed Rule Making and Order, December 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*, Elsevier Computer Networks, Vol. 50, pp. 2127-2159, 2006.
- [4] X. Zhu, B. Tang, and H. Gupta, Delay Efficient Data Gathering in Sensor Networks, MSN 2005.
- [5] S. Chen, S. Tang, M. Huang, and Y. Wang, Capacity of Data Collection in Arbitrary Wireless Sensor Networks, IEEE INFOCOM 2010.
- [6] T. Zhang and D. H. K. Tsang, Optimal Cooperative Sensing Scheduling for Energy-Efficient Cognitive Radio Networks, IEEE INFOCOM 2011.
- [7] Z. Li, F. R. Yu, and M. Huang, A Cooperative Spectrum Sensing Consensus Scheme in Cognitive Radios, IEEE INFOCOM 2009.
- [8] T. Shu and M. Krunz, Throughput-Efficient Sequential Channel Sensing and Probing in Cognitive Radio Networks Under Sensing Errors, ACM MOBICOM 2009.
- [9] H. Kim and K. G. Shin, In-Band Spectrum Sensing in Cognitive Radio Networks: Energy Detection or Feature Detection, ACM MOBICOM 2008.
- [10] S. Huang, X. Liu, and Z. Ding, Optimal Transmission Strategies for Dynamic Spectrum Access in Cognitive Radio Networks, IEEE Transactions on Mobile Computing, Vol. 8, No. 12, December 2009.
- [11] S. Wang, J. Zhang, and L. Tong, Delay Analysis for Cognitive Radio Networks with Random Access: A Fluid Queue View, IEEE INFOCOM 2010.
- [12] K. Khalil, M. Karaca, O. Ercetin, and E. Ekici, Optimal Scheduling in Cooperative-to-Join Cognitive Radio Networks, IEEE INFOCOM 2011.
- [13] A. W. Min, K.-H. Kim, J. P. Singh, and K. G. Shin, Opportunistic Spectrum Access for Mobile Cognitive Radios, IEEE INFOCOM 2011.
- [14] S. Ji, A. S. Uluagac, R. Beyah, and Z. Cai, Practical Unicast and Convergecast Scheduling Schemes for Cognitive Radio Networks, Journal of Combinatorial Optimization, 2012.
- [15] D. Xue and E. Ekici, Guaranteed Opportunistic Scheduling in Multi-Hop Cognitive Radio Networks, IEEE INFOCOM 2011.
- [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, IEEE INFOCOM 2011.
- [17] F. Wang, M. Krunz, and S. Cui, Spectrum Sharing in Cognitive Radio Networks, IEEE INFOCOM 2008.
- [18] Y. Shi and Y. T. Hou, A Distributed Optimization Algorithm for Multi-Hop Cognitive Radio Networks, IEEE INFOCOM 2008.
- [19] T. Shu and M. Krunz, Coordinated Channel Access in Cognitive Radio Networks: A Multi-Level Spectrum Opportunity Perspective, IEEE INFOCOM 2009.
- [20] G. S. Kasbekar and S. Sarkar, Spectrum Pricing Games with Bandwidth Uncertainty and Spatial Reuse in Cognitive Radio Networks, ACM MOBIHOC 2010.
- [21] B. Zhang, Y. Chen, and K. J. R. Liu, An Indirect Reciprocity Game Theoretic Framework for Dynamic Spectrum Access, ICC 2012.
- [22] B. Zhang, Y. Chen, and K. J. R. Liu, An Indirect-Reciprocity Reputation Game for Cooperation in Dynamic Spectrum Access Networks, IEEE Transactions on Wireless Communications, Vol. 11, No. 12, 2012.
- [23] C. Jiang, Y. Chen, K. J. R. Liu, and Y. Ren, Analysis of Interference in Cognitive Radio Networks with Unknown Primary Behavior, ICC 2012.

- [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 Journal of Selected Areas in Communications (JSAC)*, 2012.
- [25] S.-W. Jeon, N. Devroye, M. Vu, S.-Y. Chung, and V. Tarokh, *Cognitive Networks Achieve Throughput Scaling of a Homogeneous Network*, *IEEE Transactions on Information Theory*, Vol. 57, No. 8, pp. 5103-5115, August 2008.
- [26] C. Wang, X.-Y. Li, S. Tang, and C. Jiang, *Multicast Capacity Scaling for Cognitive Networks: General Extended Primary Network*, *IEEE MASS* 2010.
- [27] C. Yin, L. Gao, and S. Cui, *Scaling Laws for Overlaid Wireless Networks: A Cognitive Radio Network versus a Primary Network*, *IEEE/ACM Transactions on Networking*, Vol. 18, No. 4, pp. 1317-1329, August 2010.
- [28] T. Jing, X. Chen, Y. Huo, and X. Cheng, *Achievable Transmission Capacity of Cognitive Mesh Networks With Different Media Access Control*, *IEEE INFOCOM* 2012.
- [29] X. Chen, T. Jing, Y. Huo, W. Li, X. Cheng, and T. Chen, *Achievable Transmission Capacity of Cognitive Radio Networks with Cooperative Relaying*, *Crowncoms* 2012.
- [30] C. Li and H. Dai, *Transport Throughput of Secondary Networks in Spectrum Sharing Systems*, *IEEE INFOCOM* 2011.
- [31] W. Huang and X. Wang, *Throughput and Delay Scaling of General Cognitive Networks*, *IEEE INFOCOM* 2011.
- [32] S. Kompella, G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, *Stable Throughput Tradeoffs in Cognitive Shared Channels with Cooperative Relaying*, *IEEE INFOCOM* 2011.
- [33] L. Sun and W. Wang, *On Distribution and Limits of Information Dissemination Latency and Speed In Mobile Cognitive Radio Networks*, *IEEE INFOCOM* 2011.
- [34] C. Li and H. Dai, *On the Throughput Scaling of Cognitive Radio Ad Hoc Networks*, *IEEE INFOCOM* 2011.
- [35] C. Han and Y. Yang, *Information Propagation Speed Study in Multi-hop Cognitive Radio Networks*, *IEEE INFOCOM* 2011.
- [36] S. Ji, R. Beyah, and Y. Li, *Continuous Data Collection Capacity of Wireless Sensor Networks Under Physical Interference Model*, *IEEE MASS* 2011.
- [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 Transactions on Sensor Networks (TOSN)*, 2012.
- [38] S. Ji, R. Beyah, and Z. Cai, *Snapshot/Continuous Data Collection Capacity for Large-Scale Probabilistic Wireless Sensor Networks*, *IEEE INFOCOM* 2012.
- [39] O. Dousse, F. Baccelli, and P. Thiran, *Impact of Interferences on Connectivity in Ad Hoc Networks*, *IEEE INFOCOM* 2003.
- [40] W. Ren, Q. Zhao, and A. Swami, *On the Connectivity and Multihop Delay of Ad Hoc Cognitive Radio Networks*, *IEEE ICC* 2010.
- [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 Networks*, Vol. 17, pp. 659-669, 2011.
- [42] X. Huang, D. Lu, P. Li, and Y. Fang, *Coolest Path: Spectrum Mobility Aware Routing Metrics in Cognitive Ad Hoc Networks*, *IEEE ICDCS* 2011.
- [43] T. Shu and M. Krunz, *Truthful Least-Priced-Path Routing in Opportunistic Spectrum Access Networks*, *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*, *IEEE INFOCOM* 2011.
- [45] J. Jin, H. Xu, and B. Li, *Multicast Scheduling with Cooperation and Network Coding in Cognitive Radio Networks*, *IEEE INFOCOM* 2010.
- [46] D. Hu, S. Mao, and J. H. Reed, *On Video Multicast in Cognitive Radio Networks*, *IEEE INFOCOM* 2009.
- [47] L. Fu, S. C. Liew, and J. Huang, *Effective Carrier Sensing in CSMA Networks under Cumulative Interference*, *IEEE INFOCOM* 2010.
- [48] S. Ji, Y. Li, and X. Jia, *Capacity of Dual-Radio Multi-Channel Wireless Sensor Networks for Continuous Data Collection*, *IEEE INFOCOM* 2011.
- [49] S. Ji, Z. Cai, Y. Li, and X. Jia, *Continuous Data Collection Capacity of Dual-Radio Multi-Channel Wireless Sensor Networks*, *IEEE Transactions on Parallel and Distributed Systems*, Vol. 23, No. 10, pp. 1844-1855, October 2012.
- [50] P.-J. Wan, S. C.-H. Huang, L. Wang, Z. Wan, and X. Jia, *Minimum-Latency Aggregation Scheduling in Multihop Wireless Networks*, *ACM MOBIHOC* 2009.
- [51] S. Ross, *Introduction to Probability Models* (9th Edition), Elsevier, 2007.
- [52] S. Ji and Z. Cai, *Distributed Data Collection and Its Capacity in Asynchronous Wireless Sensor Networks*, *IEEE INFOCOM* 2012.
- [53] S. Ji and Z. Cai, *Distributed Data Collection in Large-Scale Asynchronous Wireless Sensor Networks, under the Generalized Physical Interference Model*, *IEEE/ACM Transactions on Networking (ToN)*, 2012.
- [54] J. Li and S. Cheng, (ϵ, δ) -Approximate Aggregation Algorithms in Dynamic Sensor Networks, *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, Vol. 23, No. 3, pp. 385-396, 2012.
- [55] S. Cheng, J. Li, and Z. Cai, $O(\epsilon)$ -Approximation to Physical World by Sensor Networks, *IEEE INFOCOM* 2013.



Zhipeng Cai is currently an assistant professor in the Department of Computer Science at the Georgia State University. He received his BS degree in computer science from Beijing Institute of Technology, China, and his 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.



Branch at Georgia State.

Shouling Ji is a Ph.D. Candidate in Department of Computer Science at Georgia State University. He received his B.S. and M.S. of 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 an ACM student member, an IEEE student member and an IEEE COMSOC student member. He is also the Membership Chair of the IEEE Student



State University. Her research interests include Wireless Networking, Network Optimization, Social Networks, and Cloud Computing.

Jing (Selena) He is currently an Assistant Professor in the Department of Computer Science at Kennesaw State University. She received her Ph.D. and M.S. degrees from the Department of Computer Science at Georgia State University (GSU) emphasized in Wireless Networking. Before joining GSU, she received a B.S. degree in Electronic Engineering from Wuhan Institute of Technology, Wuhan, Hubei, China. She received another M.S. degree in computer science concentrated in Artificial Intelligence from Utah



Lin Wei is currently an assistant professor in the Department of Agricultural and Biosystems at the South Dakota State University, USA. He received his 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 bachelor of science degree in electrical engineering from Louisiana State University in 1991 and the masters and PhD degrees in electrical and computer engineering from Louisiana State University in 1997 and 2000, respectively. She is currently an associate professor in the Department of Computer Science at Georgia State University. 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.