Distributed Data Collection and Its Capacity in Asynchronous Wireless Sensor Networks

Shouling Ji
Department of Computer Science
Georgia State University
Atlanta, Georgia 30303, USA
Email: sji@cs.gsu.edu

Zhipeng Cai
Department of Computer Science
Georgia State University
Atlanta, Georgia 30303, USA
Email: zcai@cs.gsu.edu

***Abstract*—Most of the existing works studying the *data collection capacity* issue have an ideal assumption that *the network time is slotted and the entire network is strictly synchronized* explicitly or implicitly. Such an assumption is mainly for *centralized synchronous* WSNs. However, WSNs are more likely to be distributed asynchronous systems. Thus, in this paper, we investigate the achievable data collection capacity of realistic *distributed asynchronous* WSNs. To the best of our knowledge, this is the first work to study the data collection capacity issue for distributed asynchronous WSNs. Our main contributions are threefold. First, to avoid data transmission collisions/interference, we derive an $\mathcal{R}_0$-*Proper Carrier-sensing Range* ($\mathcal{R}_0$-PCR) under the *generalized physical interference model* for the nodes in a data collection WSN, where $\mathcal{R}_0$ is the *satisfied threshold data receiving rate*. Taking $\mathcal{R}_0$-PCR as its *carrier-sensing range*, any node can initiate a data transmission with a guaranteed data receiving rate. Second, based on the obtained $\mathcal{R}_0$-PCR, we propose a *Distributed Data Collection* (DDC) algorithm with fairness consideration for asynchronous WSNs. Theoretical analysis of DDC surprisingly shows that its asymptotic achievable network capacity is $\mathbb{C} = \Omega(\frac{1}{26\beta_{\kappa+1}} \cdot W)$, where $\beta_{\kappa+1}$ is a constant value depends on $\mathcal{R}_0$ and W is the bandwidth of a wireless communication channel, which is order optimal and independent of network size. Thus, DDC is scalable. Finally, we conduct extensive simulations to validate the performance of DDC. Simulation results demonstrate that DDC can achieve comparable data collection capacity as that of the most recently published centralized and synchronized data collection algorithm.**

I. Introduction

One of the most important functions provided by Wireless Sensor Networks (WSNs) is directly gathering data from the physical world. Generally, data gathering can be categorized as *data collection* [4][6][7], which gathers all the data from a network without any data aggregation or merging, and *data aggregation* [1][2], which obtains some aggregation values, *e.g.* MAX, MIN, SUM, and *etc*. To evaluate network performance, *network capacity*, which reflects the data transmission/collection/broadcast rate, is usually adopted, *e.g. multicast capacity* [11][13][12], *unicast capacity* [22][23][34], *broadcast capacity* [35], and *data collection capacity* [4][6][7]. For *data collection capacity*, it is defined as the average data receiving rate at the sink, *i.e.* data collection capacity reflects how fast data been collected by the sink. Without confusion, we use data collection capacity and network capacity interchangeably throughout the following of this paper.

Following the seminal work [3] by Gupta and Kumar, many works emerged to study the network capacity issue under various network scenarios, *e.g.* multicast, unicast, broadcast, and data collection/aggregation. However, to our knowledge, most of the existing works studied the network capacity issue under an ideal assumption that *the network time is slotted, and the entire network is strictly synchronized* explicitly or implicitly, *i.e.* they are mainly for *centralized synchronous* wireless networks. Under the above ideal assumption, many centralized algorithms with nice network capacity bounds are designed and analyzed for a communication mode (*e.g.* multicast, unicast, broadcast, and data collection/aggregation). In the sense of providing theoretical frameworks/bounds for the design of communication protocols, these works are still sound. However, in practical, wireless networks are more likely to be distributed systems, *e.g.* CSMA networks, WiFi networks, wireless ad hoc networks, WSNs, and *etc*. Furthermore, for wireless networks, especially WSNs which consist of vulnerable sensor nodes, it is difficult and not realistic to achieve ideal strict time synchronization due to the unstable deploying environments, clock drift, and other technical limits. Therefore, to comprehensively and profoundly understand the performance of practical WSNs, it is important to investigate the achievable network capacity of *distributed asynchronous* WSNs. Particularly, we study the achievable data collection capacity for distributed asynchronous WSNs in this paper.

When we study the data collection capacity issue for distributed asynchronous WSNs, many new challenges are introduced compared with that in centralized synchronous WSNs. We summarize the main challenges as follows.

- **C1:** unlike that in centralized synchronous WSNs, where we can acquire the overall information of a network and further make an optimized decision for data transmissions, in distributed asynchronous WSNs, we can only schedule data transmissions according to local information. It is followed that the overall optimized scheduling cannot be guaranteed. Therefore, how to design an effective distributed algorithm for data collection is a challenge.
- **C2:** since we cannot maintain a uniform time clock for all the nodes in distributed asynchronous WSNs, every sen-

sor node will conduct data transmissions based on its own time clock and local information. Intuitively, this kind of communication mode will lead to many data collisions and retransmissions, followed by capacity degradation, unfairness among data flows, and *etc*. Thus, how to avoid the disadvantages introduced by time asynchronous is an important concern when designing distributed data collection algorithms.

- **C3:** following challenges **C1** and **C2**, the third challenge is how to theoretically analyze the achievable network capacity bounds for a data collection algorithm in distributed asynchronous WSNs. Since the data collection algorithm works in a distributed manner, it is difficult, sometimes even impossible, to know the exactly time a data transmission happened, as well as the time duration of a data transmission. Hence, to obtain the achievable data collection capacity not only requires the analysis technique, but also depends on the carefully designed data transmission mechanism.

To address these challenges, we propose a scalable and order-optimal distributed algorithm, named *Distributed Data Collection* (DDC), with fairness consideration and capacity analysis under the *generalized physical interference model*. To the best of our knowledge, this is the first attempt to provide detailed protocol design and rigorous capacity analysis for data collection in distributed asynchronous WSNs. DDC works in a CSMA-like manner, except for the RTS/CTS communication mode and the necessity to reply an ACK packet after receiving a data packet. In DDC, when a node has some data packets for transmission, it sets up a backoff timer, and senses the wireless channel with a predefined *Carrier-sensing Range* (CR). If the channel is free when the backoff timer expires, this node conducts a data transmission. Under this transmission manner, DDC gathers all the of data of the network to the sink (*i.e.* base station). Particularly, we summarized the main contributions of this paper as follows.

- The carrier-sensing range is an important parameter in DDS, which has a significant impact on the performance of data collection operations. To avoid data transmission collisions/interference, especially the collisions/interference caused by *hidden-node problems*, we derive the $\mathcal{R}_0$-*Proper Carrier-sensing Range* ($\mathcal{R}_0$-PCR) under the *generalized physical interference model* for the nodes in a data collection WSN, where $\mathcal{R}_0$ is the *satisfied threshold data receiving rate*. By working with $\mathcal{R}_0$-PCR as its CR, any node can initiate a data transmission with guaranteed data receiving rate at least $\mathcal{R}_0$, as long as there is no ongoing transmissions within its CR.
- Based on the obtained $\mathcal{R}_0$-PCR, we propose a scalable and order-optimal *Distributed Data Collection* (DDC) algorithm with fairness consideration for asynchronous WSNs. DDC works in a CSMA-like manner, and effectively gathers all the data to the sink. Theoretical analysis of DDC surprisingly shows that its asymptotic achievable network capacity is $\mathbb{C} = \Omega(\frac{1}{26\beta_{\kappa+1}} \cdot W)$, where $\beta_{\kappa+1}$ is a constant value depends on $\mathcal{R}_0$, and W is the bandwidth of a wireless communication channel. Since the upper bound capacity of data collection is $O(W)$ [6][7], which implies the achievable data collection capacity of DDC is order-optimal. Furthermore, since $\mathbb{C}$ is independent with the network size, DDC is scalable.
- We also conduct extensive simulations to validate the performance of DDC in distributed asynchronous WSNs. Simulation results indicate that DDC can achieve comparable data collection capacity as that of the latest centralized and synchronized data collection algorithm.

The rest of this paper is organized as follows. In Section II, we summarize the related work. In Section III, the considering network model is discussed. In Section IV, the proper carrier-sensing range is derived. According to the obtained proper carrier-sensing range, a distributed asynchronous data collection algorithm is proposed in Section V, followed by the theoretical analysis. In Section VI, we validate the performance and scalability of the proposed data collection by simulations. Finally, this paper is concluded in Section VII.

II. Related Work

In this section, we summarize the existing works based on the communication mode.

A. Data Collection Capacity

Data collection capacity is studied in [4], [5], [6], [7], [8], and [9] for centralized synchronous WSNs. In [4], the authors proposed a load-balanced data gathering algorithm. Considering data compression, the authors showed that the network capacity can be improved. In [5], the authors considered the collision-free delay-efficient data gathering problem and proposed a family of path scheduling algorithms. The authors in [6] and [7] extended the work in [5], and derived tighter upper and lower data collection bounds. In [8], the authors studied data collection capacity of centralized synchronous WSNs based on a grid partition method. They also obtained the achievable data collection capacity under the protocol interference model. The worst-case data collection capacity is investigated in [9]. For data aggregation, the authors in [10] investigated its capacity for centralized synchronous WSNs in the extended network case.

B. Multi/Uni/Broad-cast Capacity

In [11], [12], [13], [14], [15], [16], and [17], the multicast capacity for centralized synchronous wireless networks is studied. The multicast capacity for wireless ad hoc networks under the protocol interference model and Gaussian channel model is investigated in [11] and [12] respectively. In [11], the authors showed that the network multicast capacity is $\Theta(\sqrt{\frac{n}{\log n}} \cdot \frac{W}{k})$ when $k = O(\frac{n}{\log n})$ and is $\Theta(W)$ when $k = \Omega(\frac{n}{\log n})$, where W is the bandwidth of a wireless channel, n is the number of the nodes in a network, and k is the number of the nodes involved in one multicast session. In [12], the authors showed that when $k \leq \theta_1 \frac{n}{(\log n)^{2\alpha+6}}$ and $n_s \geq \theta_2 n^{1/2+\beta}$, the capacity that each multicast session

can achieve is at least $c_8\frac{\sqrt{n}}{n_s\sqrt{k}}$, where k is the number of the receivers in one multicast session, n is the number of the nodes in a network, n_s is the number of the multicast sessions, θ_1, θ_2 and c_8 are constants and β is any positive real number. Another similar work [13] studies the upper and lower bounds of multicast capacity for hybrid wireless networks consisting of ordinary wireless nodes and multiple base stations connected by a high-bandwidth wired network. A more general (n, m, k)-casting capacity problem (including multicast and unicast) is investigated in [14] and [15]. In [14], the upper and lower bounds of (n, m, k)-casting is obtained for ad hoc networks and in [15], a scheduling partition method for large-scale wireless networks is proposed. The multicast capacity of multi-channel wireless networks is studied in [16], where the authors represented the upper bound capacity of per multicast as a function of several system parameters. In [17], the authors investigated the optimal multicast capacity and delay tradeoffs in mobile ad hoc networks from a global perspective.

The achievable capacity of multiple unicasts in centralized synchronous wireless networks is investigated in [18]-[32]. In [18], the impact of the number of channels, the number of interfaces, and the interface switching delay on the capacity of centralized synchronous wireless networks are investigated. In [19], the authors considered the scheduling problem where all the communication requests are single-hop and all the nodes transmit at a fixed power level. They proposed an algorithm to maximize the number of concurrent transmitting links in one time-slot. Unlike [19], the authors in [20] and [21] considered the power-control problem. A family of approximation algorithms were presented to maximize the capacity of arbitrary wireless networks. Considering the problem of characterizing the unicast capacity scaling in arbitrary wireless networks, the authors proposed a general cooperative communication scheme in [22]. In [23], the authors studied the balanced unicast and multicast capacity of a wireless network consisting of n randomly placed nodes, and obtained the characterization of the scaling of the n^2-dimensional balanced unicast and $n2^n$-dimensional balanced multicast capacity regions under the Gaussian fading channel model. A general framework to characterize the capacity of wireless ad hoc networks with arbitrary mobility patterns is studied in [24]. By relaxing the "homogeneous mixing" assumption in most existing works, the capacity of a heterogeneous network is analyzed. Another work [25] studies the relationship between the capacity and the delay of mobile wireless ad hoc networks, where the authors studied how much delay must be tolerated under a certain mobile pattern to achieve an improvement of the network capacity. Similarly as the work in [25] which considers that Lévy mobility and human mobility share several common features, the authors in [26] focused on the delay-capacity tradeoffs for mobile wireless networks with Lévy walks and Lévy flights. In [27] and [28], the authors studied the connectivity and capacity problem of multi-channel wireless networks. They considered a multi-channel wireless network with some constraints on channel switching, proposed some routing and channel assignment strategies for multiple unicast communications and derived the per-flow capacity. In [29], the authors first proposed a multi-channel network architecture, called MC-MDA, where each node is equipped with multiple directional antennas, and then obtained the capacity of multiple unicast communications. Similarly as in [29], the authors in [30] studied the local sufficient rate constraints that can be constructed at each node to ensure a feasible flow allocation for multi-radio multi-channel wireless networks. In [31], the throughput capacity of 3D regular ad hoc networks and 3D heterogeneous ad hoc networks is derived for the first time under the generalized physical interference model. The capacity scaling of multi-hop cellular networks is investigated in [32], and the authors further extended their method to the capacity of heterogeneous multi-hop cellular networks. Unlike the above mentioned works, the authors in [33] and [34] examined multi-unicast capacity for random access networks. In [33], the scheduling of a set of communication links is investigated. Similar as that in [33], the authors in [34] studied multi-unicast capacity for CSMA wireless networks. They formulated the models of a series of CSMA protocols and investigated the multi-unicast capacity of CSMA scheduling versus TDMA scheduling. In [35], the broadcast capacity for ad hoc networks is derived with the fixed data rate channel and the Gaussian channel, respectively.

III. Network Model

In this paper, we consider a connected WSN consisting of one sink node serving as the *base station* denoted by s_0, and n sensor nodes denoted by $s_1, s_2, \cdots, s_n$ respectively, deployed in an area with size $A = c_1 n$, where c_1 is a constant. Furthermore, we assume all the nodes are *independent and identically distributed* (*i.i.d.*). Each node is equipped with one radio, and working with a fixed power P. All the data transmissions are conducted over a common wireless channel with bandwidth W *bits/second*. The size of a data packet is B bits, and thus the transmission duration of a data packet is $\tau = B/W$ seconds. The maximum transmission radius of a node is set to be r (r is associated with the lowest data transmission rate determined by the following defined *generalized physical interference model*). Hence the network can be modeled by a graph $G = (V, E)$, where $V = \{s_i | i = 0, 1, 2, \cdots, n\}$ and E includes all the possible links formed by any pair of nodes in V. A node s_i ($i \in [1, n]$) is said to be *active* at time t *iff* s_i is transmitting a data packet to some other node at time t. Thus, we use $\mathcal{S}^t = \{s_k | s_k \textit{ is active at time } t\}$ to denote the set of active nodes at time t.

To capture the wireless interference in wireless networks, the *protocol interference model* and *physical interference model* are frequently used. For simplicity, these two models abstract a data transmission as a binary function, with values *successful* or *failed*. Instead of modeling a data transmission process as a binary function, the *Generalized Physical Interference model* (GPI) is more accurate to characterize a practical data transmission. Suppose node s_i is transmitting

a data packet to node s_j at time t, *i.e.* $s_i \in \mathcal{S}^t$, and $\mathcal{R}^t_{i,j}$ is the data receiving rate of s_j from s_i at time t. Then, under the GPI model, $\mathcal{R}^t_{i,j}$ is determined by

$$\mathcal{R}^t_{i,j} = W \cdot \log(1 + SINR^t_{i,j}) \quad (1)$$

where $SINR^t_{i,j}$ is the *Signal-to-Interference-plus-Noise Ratio* (SINR) value at s_j associated with s_i and is defined as

$$SINR^t_{i,j} = \frac{P \cdot D(s_i, s_j)^{-\alpha}}{N_0 + \sum\limits_{s_k \in \mathcal{S}^t, s_k \neq s_i} P \cdot D(s_k, s_j)^{-\alpha}} \quad (2)$$

where N_0 is the background noise, α is the path loss exponent and usually $\alpha \geq 3$, and $D(\cdot,\cdot)$ is the Euclidian distance between two nodes.

Suppose the time consumption to gather all the n data packets produced at s_i $(1 \leq i \leq n)$ is $\mathcal{T}$, then the achievable data collection capacity $\mathbb{C}$ can be defined as $nB/\mathcal{T}$, *i.e.* the data collection capacity reflects how fast that data can be gathered by the sink.

IV. Carrier-sensing Range

Since we study data collection in distributed asynchronous WSNs, every node s_i $(i \in [1, n])$ in a WSN will sense the activities of other nodes within its *Carrier-sensing Range* (CR) when it has some data packets for transmission. Only when there is no ongoing data transmissions within its CR, s_i can initiate a data transmission. Thus, how to determine the CR for each sensor node, to make all the concurrent transmitters out of the CR of each other to simultaneously conduct data transmissions with a data rate no less than a threshold, is crucial for the performance of a distributed data collection scheme. Intuitively, a small CR implies a high degree of spatial reuse which further implies small SINR values, and followed by low data receiving rates at the receivers. On the other hand, a large CR implies a low degree of spatial reuse, which further implies large SINR values, and followed by high data receiving rates. Therefore, in this section, we study how to set a *Proper Carrier-sensing Range* (PCR) for each node to guarantee a satisfied data receiving rate and meanwhile the highest spatial reuse degree. For clarity, we make some definitions as follows.

Definition 4.1: **$\mathcal{R}_0$-feasible state.** The set of all the active nodes $\mathcal{S}^t$ (defined in Section III) is a $\mathcal{R}_0$-feasible state if all the nodes in $\mathcal{S}^t$ can simultaneously transmit data and the data receiving rate at each of their corresponding receivers is no less than $\mathcal{R}_0$. In a $\mathcal{R}_0$-feasible state $\mathcal{S}^t$, $\forall s_i \in \mathcal{S}^t$, assume s_i is transmitting a data packet to s_j, then $\mathcal{R}^t_{i,j} \geq \mathcal{R}_0$.

Based on Definition 4.1, if the lowest tolerative data transmission rate of a WSN is $\mathcal{R}_0$, then the data collection process can be represented by a series of $\mathcal{R}_0$-feasible states $\mathcal{S}^t$ $(t = \tau, 2\tau, 3\tau, \cdots, m\tau)$, where $m = \lceil \mathcal{T}/\tau \rceil$.

Definition 4.2: **R-set ($\mathcal{S}_R$).** Assume R is the carrier-sensing range of a WSN represented by $G = (V, E)$. An R-set, denoted by $\mathcal{S}_R$, is any maximal subset of V that satisfies $\forall s_i, s_j \in \mathcal{S}_R$ $(s_i \neq s_j)$, $D(s_i, s_j) \geq R$.

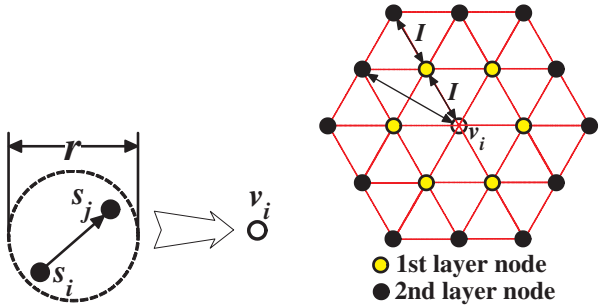
Definition 4.3: **$\mathcal{R}_0$-Proper Carrier-sensing Range ($\mathcal{R}_0$-PCR).** The carrier-sensing range R of a WSN is a $\mathcal{R}_0$-proper carrier-sensing range if for any R-set $\mathcal{S}_R$, it is always a $\mathcal{R}_0$-feasible state.

From Definition 4.3, if R is a $\mathcal{R}_0$-PCR, then s_i can initiate a data transmission with a guaranteed data receiving rate no less than $\mathcal{R}_0$ as long as there is no other active nodes within R of s_i. Then, given a threshold data receiving rate $\mathcal{R}_0$, the $\mathcal{R}_0$-PCR can be determined by the following Theorem 1. In the following analysis, as that in [36], we assume the background noise is very small compared with the transmission power ($N_0 \ll P$). Thus N_0 can be ignored ($N_0 = 0$).

Theorem 1: $\mathcal{R}_0$-PCR $\geq (\sqrt[\alpha]{c_2(2^{\mathcal{R}_0/W} - 1)} + 1) \cdot r$, where c_2 is a constant.

Proof Sketch: Let $R = \mathcal{R}_0$-PCR and $I = R - r$. To make any R-set $\mathcal{S}_R$ always a $\mathcal{R}_0$-feasible state, for $\forall s_i \in \mathcal{S}_R$, assuming its destination node is s_j, then, we have

$$\mathcal{R}_{i,j} \geq \mathcal{R}_0 \Leftrightarrow \frac{D(s_i, s_j)^{-\alpha}}{\sum\limits_{s_k \in \mathcal{S}_R, s_k \neq s_i} D(s_k, s_j)^{-\alpha}} \geq 2^{\mathcal{R}_0/W} - 1 \quad (3)$$



(a) (b)

Fig. 1. (a) Link abstraction and (b) hexagon packing.

Now, we derive the lower bound of $\frac{D(s_i,s_j)^{-\alpha}}{\sum_{s_k \in \mathcal{S}_R, s_k \neq s_i} D(s_k,s_j)^{-\alpha}}$. Evidently, $D(s_i, s_j)^{-\alpha} \geq r^{-\alpha}$ since r is the maximum transmission range of a node (defined in Section III). Furthermore, if we abstract a data transmission link as a node as shown in Fig. 1(a), then, for the nodes in $\mathcal{S}_R$, the densest packing of nodes is the hexagon packing [36] with edge length I as shown in Fig. 1(b). Subsequently, the nodes in $\mathcal{S}_R$ can be layered with respect to v_i (abstracted by the transmission link from s_i to s_j), with the l-th layer having at most $6l$ nodes. Furthermore, the distance between v_i and any node at the l-th layer is no less than $\frac{\sqrt{3}}{2} lI$. Then, we have

$$\sum_{s_k \in \mathcal{S}_R, s_k \neq s_i} D(s_k, s_j)^{-\alpha} \quad (4)$$

$$\leq 6 \cdot I^{-\alpha} + \sum_{l \geq 2} 6l \cdot (\frac{\sqrt{3}}{2} lI)^{-\alpha} \quad (5)$$

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

In Formula 6, $\sum\limits_{l \geq 2} l^{-\alpha+1} = \zeta(\alpha - 1) - 1$, where $\zeta(\cdot)$ is the *Riemann zeta function*. Considering that $\alpha \geq 3$, then

$\zeta(\alpha-1) \le \zeta(2) = \frac{\pi^2}{6}$. It follows that $\sum_{l\ge 2} l^{-\alpha+1} \le \frac{\pi^2}{6} - 1$. Thus, we have

$$\sum_{s_k \in \mathcal{S}_R, s_k \neq s_i} D(s_k, s_j)^{-\alpha} \tag{7}$$

$$\le (6 + (\pi^2 - 6)(\frac{\sqrt{3}}{2})^{-\alpha}) \cdot I^{-\alpha} = c_2 \cdot I^{-\alpha}, \tag{8}$$

where $c_2 = (6 + (\pi^2 - 6)(\frac{\sqrt{3}}{2})^{-\alpha})$. It follows that

$$\frac{D(s_i, s_j)^{-\alpha}}{\sum_{s_k \in \mathcal{S}_R, s_k \neq s_i} D(s_k, s_j)^{-\alpha}} \ge \frac{r^{-\alpha}}{c_2 \cdot I^{-\alpha}}. \tag{9}$$

Therefore, to make Formula 3 valid, it is sufficient to have

$$\frac{r^{-\alpha}}{c_2 \cdot I^{-\alpha}} \ge 2^{\mathcal{R}_0/W} - 1 \tag{10}$$

$$\Leftrightarrow I^{-\alpha} \le \frac{r^{-\alpha}}{c_2(2^{\mathcal{R}_0/W} - 1)} \tag{11}$$

$$\Leftrightarrow I \ge (\frac{1}{c_2(2^{\mathcal{R}_0/W} - 1)})^{-1/\alpha} \cdot r \tag{12}$$

$$\Leftrightarrow I \ge \sqrt[\alpha]{c_2(2^{\mathcal{R}_0/W} - 1)} \cdot r. \tag{13}$$

Therefore,

$$\mathcal{R}_0\text{-PCR} = R = I + r \tag{14}$$

$$\ge \sqrt[\alpha]{c_2(2^{\mathcal{R}_0/W} - 1)} \cdot r + r \tag{15}$$

$$= (\sqrt[\alpha]{c_2(2^{\mathcal{R}_0/W} - 1)} + 1) \cdot r. \tag{16}$$

It follows this lemma holds. □

From Theorem 1, we know that given a threshold data receiving rate $\mathcal{R}_0$, we can determine an $\mathcal{R}_0$-PCR, which is at least a constant times r. Since a small CR implies a high degree of spatial reuse, we set $\mathcal{R}_0$-PCR $= (\sqrt[\alpha]{c_2(2^{\mathcal{R}_0/W} - 1)} + 1) \cdot r$ in this paper. Furthermore, Fig. 2 depicts the relation between $\mathcal{R}_0$ and $\mathcal{R}_0$-PCR, where the X-axis represents the threshold data receiving rate $\mathcal{R}_0$, and the Y-axis represents the corresponding $\mathcal{R}_0$-PCR. From Fig. 2, with the increase of $\mathcal{R}_0$, the associated $\mathcal{R}_0$-PCR increases accordingly for every α value. This is because a high data receiving rate requires that CR should be sufficient large to avoid interferences, which also implies a low degree of spatial reuse. Additionally, a large α also implies a small $\mathcal{R}_0$-PCR. This is because the interference impact will decrease quickly with the increase of α, which can also be derived from Equation 2.

V. Distributed Data Collection and Capacity

According to the obtained $\mathcal{R}_0$-PCR in Section IV, if we set the CR of a WSN as $\mathcal{R}_0$-PCR, then all the nodes in an R-set ($R = \mathcal{R}_0$-PCR) can simultaneously transmit data at a guaranteed threshold data receiving rate without interference by letting each node working on the Re-Start (RS) mode [36]. Thus, in this section, we propose a CSMA-like data collection algorithm for distributed asynchronous WSNs, which has an order-optimal capacity.

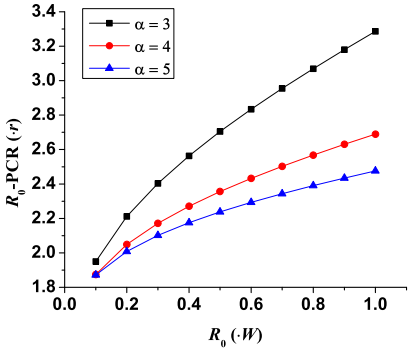
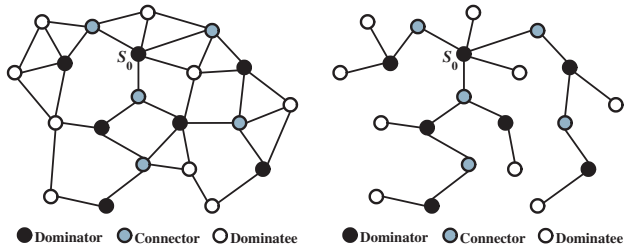


Fig. 2. $\mathcal{R}_0$ *vs.* $\mathcal{R}_0$-PCR.

A. Distributed Data Collection

Before giving the distributed data collection algorithm, for a WSN represented by $G = (V, E)$, we construct a *Connected Dominating Set* (CDS)-based data collection tree, denoted by T, according to the following steps [1]: (i) Construct a *Breadth-First-Search* (BFS) tree on G beginning at the sink s_0, and obtain a *Maximal Independent Set* (MIS) $\mathcal{D}$ according to the search sequence. Note that $\mathcal{D}$ is a *Dominating Set* (DS) of G. For the nodes in $\mathcal{D}$, we call them *dominators*. Taking the network shown in Fig. 3(a) as an example, the black nodes are nominators. (ii) To form a CDS, we choose nodes, named *connectors*, as less as possible in V to connect all the nodes in $\mathcal{D}$. The set of all the connectors is denoted by $\mathcal{C}$. As shown in Fig. 3(a), the blue nodes are the connectors chosen to connect all the dominators. Then, each dominator, except for s_0, has a connector as its parent node in T. On the other hand, each connector has a dominator as its parent node in T. (iii) For any other node in $V \backslash (\mathcal{D} \cup \mathcal{C})$, called a *dominatee*, randomly choose a dominator within its communication range as its parent node. Then, the CDS-based data collection tree T rooted at s_0 is formed. For the network shown in Fig. 3(a), the constructed CDS-based data collection tree is shown in Fig. 3(b).



(a) (b)

Fig. 3. CDS-based data collection tree.

Assume L_{max} is the height of T, *i.e.* the maximum number of hops from s_0 to any sensor, and $L(s_i)$ is the number of hops from node s_i to the root s_0 in T. Evidently, according to the construction process of T, $\forall s_i \in \mathcal{D}$, $L(s_i)$ is an even number, and $\forall s_j \in \mathcal{C}$, $L(s_j)$ is an odd number. Furthermore,

we define $\mathcal{L}_\iota = \{s_i | L(s_i) = \iota\}$ $(0 \leq \iota \leq L_{max})$. Then, the following lemma [1] shows some properties of T.

Lemma 1: [1] (i) s_0 is adjacent to at most 12 connectors in $\mathcal{C}$; *(ii)* $\forall s_i \in \mathcal{D}, s_i \neq s_0$, s_i is adjacent to at most 11 connectors in $\mathcal{L}_{L(s_i)+1}$.

From Lemma 1, we can readily obtain the following result.

Corollary 1: $\forall s_i \in \mathcal{D}$, s_i is adjacent to at most 12 connectors in T.

Proof: Please refer to [37] for details. □

Based on T, we propose a *Distributed Data Collection* (DDC) algorithm for asynchronous WSNs as shown in Algorithm 1. In Algorithm 1, *counter*(s_i) is a counter that denotes the number of data packets transmitted by s_i, τ_w is the *backoff contention window*, and t_i^j $(1 \leq j \leq$ *counter*$(s_i))$ is the backoff time set for the transmission of the j-th data packet at node s_i. As that in [34] and [36] and because of the same reasons, we assume (i) $\tau_w \ll \tau$ such that τ_w is negligible compared with the data transmission time, and (ii) no two transmitters within the CR of each other have their backoff timers expired at the same time instant[1].

According to Algorithm 1, DDC runs in a CSMA-like manner, except for the RTS/CTS working mode and the necessity to reply an ACK packet after receiving a data packet. This is because that by properly setting the CR and working on the RS mode, a transmission with satisfied data receiving rate can be guaranteed as shown in Section IV.

In Algorithm 1 (here, taking the algorithm running process at node s_i as an example), Line 1-5 are basic settings. Line 6 randomly sets the backoff time for each data transmission. In Line 7-8, the backoff time for each transmission is reset to be $(\tau_w - t_i^{j-1}) + t_i^j$, and this is mainly for fairness (any node will not wait too long when it has some data to transmit) as shown in Theorem 2 and Corollary 3 (see Section V-B). Under this setting, a node cannot transmit multiple data packets in a short time period. Actually, each node can transmit up to one data packet during each backoff contention window. In Line 9-14, s_i begins the countdown process and keeps sensing the channel with $\mathcal{R}_0$-PCR. If the wireless channel is busy sensed by s_i, the countdown process at s_i will be frozen. By this way, when a data transmission is ongoing, all the other nodes having data packets within the CR of the transmitter will stop their countdown process, *i.e.* they can share the waiting time. In Line 15-16, s_i transmits the j-th data packet when the backoff timer expires. Since no two transmitters that within the CR of each other have their backoff timers expired on the same time instant, the transmission of the j-th data packet can be conducted successfully.

Algorithm 1: The DDC Algorithm

input : CDS-based data collection tree T, $\mathcal{R}_0$-PCR
output: a distributed asynchronous data collection plan

1 *counter*(s_i)←0;
2 $s_i (i \in [1, n])$ sets its CR as $\mathcal{R}_0$-PCR according to the required threshold data receiving rate $\mathcal{R}_0$;
3 **while** s_i *has some data packets for transmission* **do**
4 *counter*(s_i)← *counter*(s_i) + 1;
5 j ←*counter*(s_i);
6 s_i randomly sets a backoff time t_i^j for the transmission of the j-th packet at window $(0, \tau_w]$;
7 **if** $j > 1$ **then**
8 $t_i^j \leftarrow (\tau_w - t_i^{j-1}) + t_i^j$;
9 **while** t_i^j *is not countdown to* 0 **do**
10 s_i senses the channel with $\mathcal{R}_0$-PCR;
11 **if** s_i *senses that the channel is busy* **then**
12 s_i stops the countdown process (the backoff timer is frozen) until the channel becomes free again;
13 **if** s_i *senses that the channel is free* **then**
14 $t_i^j - -$;
15 **if** $t_i^j == 0$, i.e. *the backoff timer expires* **then**
16 s_i transmits the j-th data packet to its parent node;

B. Capacity Analysis

In this subsection, we analyze the achievable data collection capacity of the DDC algorithm. Since the upper bound capacity of data collection is $O(W)$ [6][7], we investigate the average achievable capacity and the lower bound capacity of DDC. First, we study the upper bound time consumption to collect all data packets at dominatees to the CDS in DDC, *i.e.* the upper bound time consumption to collect data packets at $V \setminus (\mathcal{D} \cup \mathcal{C})$ to $\mathcal{D} \cup \mathcal{C}$.

Let $R = \mathcal{R}_0$-PCR $= (\sqrt[\alpha]{c_2(2^{\mathcal{R}_0/W} - 1)} + 1) \cdot r$, where $\mathcal{R}_0$-PCR is the CR used in DDC. Then, we have the following lemma, which indicates the average/upper bound number of the sensor nodes, denoted by $\mathbb{A}/\mathbb{U}$, within the CR of a node.

Lemma 2: Let the random variable X denote the number of the sensor nodes within the carrier-sensing area of a node. Then,

(i)

$$\mathbb{A} = \mathbf{E}[X] = \frac{\pi R^2}{c_1}. \tag{17}$$

(ii)

$$\Pr[X > \log n + \frac{\pi R^2(e^2-1)}{2c_1}] \tag{18}$$

$$\leq \Pr[X \geq \log n + \frac{\pi R^2(e^2-1)}{2c_1}] \tag{19}$$

$$\leq \frac{1}{n^2}. \tag{20}$$

Thus, it is almost impossible that the carrier-sensing area of a node contains more than $\log n + \frac{\pi R^2(e^2-1)}{2c_1}$ sensor nodes, i.e. it is almost sure that $\mathbb{U} = \log n + \frac{\pi R^2(e^2-1)}{2c_1}$.

[1]Collisions due to simultaneous countdown-to-zero can be tackled by an exponential backoff mechanism in which the transmission probability of each node is adjusted in a dynamic way based on the network busyness [36].

Proof: Please refer to [37] for details. □

Based on Lemma 2, we can derive the upper bound time consumption to collect all the data packets at $V \setminus (\mathcal{D} \cup \mathcal{C})$ to $\mathcal{D} \cup \mathcal{C}$ in DDC.

Theorem 2: Any node s_i with data packets for transmission can transmit at least one data packet to its parent node within time $2\mathbb{U}\tau = (2\log n + \frac{\pi R^2(e^2-1)}{c_1})\tau$.

Proof: According to the DDC algorithm, for any node s_i with data packets for transmission, it will carrier-sense the node activities within its CR. When the backoff timer of s_i expires and meanwhile the channel sensed by s_i is free, s_i can transmit a data packet successfully. Thus, the problem now is how long it will take for s_i until it actually initiates a data transmission in the worst case, *i.e.* the waiting time of s_i in the worst case. For convenience, assume s_j is any other node within the CR of s_i having data packets for transmission, $t_i, t_j \in (0, \tau_w]$ $(t_i \neq t_j)$ are the backoff time for the current data transmissions of s_i and s_j respectively, and $T(U)$, $T(s_i)$ and $T(s_j)$ are the universal time (standard time), the system time maintained at s_i and s_j respectively. Furthermore, if s_j has more than one data packet for transmission, the backoff time for s_j to transmit a subsequent data packet is denoted by t_{j+1}. Evidently, the transmission sequence of s_i and s_j follows one of the following three cases (Note that no two transmitters within the CR of each other have their backoff timers expiring at the same time instant).

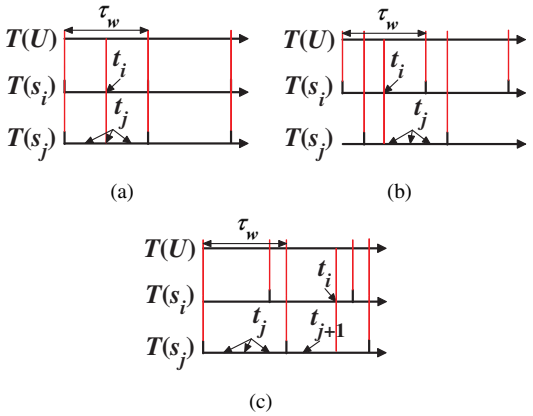


Fig. 4. Transmission sequence of s_i and s_j.

Case 1: s_i and s_j share a synchronized backoff contention window. In this case, as shown in Fig. 4(a), s_i will transmit a data packet before/after s_j transmits a data packet. This is because $t_{j+1} = t_j + (\tau_w - t_j) + t'_{j+1} = \tau_w + t'_{j+1} > t_i$, where $t'_{j+1} \in (0, \tau]$ is the backoff time chosen by s_j for the subsequent data transmission according to the DDC algorithm.

Case 2: s_i and s_j share an asynchronous backoff contention window and $t_i < t_j$. In this case, as shown in Fig. 4(b), s_i will transmit a data packet before s_j according to DDC.

Case 3: s_i and s_j share an asynchronous backoff contention window and $t_i > t_j$. In this case, as shown in Fig. 4(c), when s_i tries to transmit a data packet, it sets a backoff time t_i for the packet and carrier-senses the channel. It turns out that the channel is busy since s_j is transmitting some data. Therefore, we conclude that $0 < t_i - t_j < 2\tau_w$ because the time slots of s_i and s_j have some overlap (otherwise, s_i cannot sense the channel is occupied by s_j when it tries to transmit the data packet). Since $0 < t_i - t_j < 2\tau_w$, it is possible that $t_{j+1} = t_j + (\tau_w - t_j) + t'_{j+1} = \tau_w + t'_{j+1} < t_i$. This implies that s_j may transmit two data packets before s_i transmits one data packet. On the other hand, according to the DDC algorithm, we have $t_{j+2} = t_j + (\tau_w - t_j) + t'_{j+1} + (\tau_w - t'_{j+1}) + t'_{j+2} = 2\tau + t'_{j+2} > t_i$, where t_{j+2} is the time that s_j transmits its third data packet and t'_{j+2} is the backoff time set by s_j for its third data packet transmission. Consequently, s_i will transmit one data packet before s_j transmits the third data packet.

In summary, s_j can transmit at most two data packets before s_i transmits one data packet in the worst case. Considering that there are at most $\mathbb{U}$ sensor nodes within the carrier-sensing area of s_i according to Lemma 2, s_i can transmit at least one data packet to its parent node within time $2\mathbb{U}\tau$ in the worst case in DDC. □

Corollary 2: By DDC, the time consumption of collecting all the data packets at $V \setminus (\mathcal{D} \cup \mathcal{C})$ to $\mathcal{D} \cup \mathcal{C}$ is at most $2\mathbb{U}\tau$.

Proof: Based on the construction process of the data collection tree T, every node in $V \setminus (\mathcal{D} \cup \mathcal{C})$ has a parent node in $\mathcal{D} \cup \mathcal{C}$. Thus, all the data packets at $V \setminus (\mathcal{D} \cup \mathcal{C})$ can be transmitted to nodes in $\mathcal{D} \cup \mathcal{C}$ within time $2\mathbb{U}\tau$ according to Theorem 2. □

After time $2\mathbb{U}\tau$, all the data packets at $V \setminus (\mathcal{D} \cup \mathcal{C})$ will be collected to $\mathcal{D} \cup \mathcal{C}$ according to Corollary 2. Subsequently, we investigate the time consumption to collect all the data packets at $(\mathcal{D} \cup \mathcal{C}) \setminus \{s_0\}$ to the sink s_0.

Lemma 3: [1] Assume that $\mathcal{X}$ is a disk of radius r_d and $\mathcal{M}$ is a set of points with mutual distance of at least 1. Then $|\mathcal{X} \cap \mathcal{M}| \leq \frac{2\pi r_d^2}{\sqrt{3}} + \pi r_d + 1$.

Let $\kappa = \sqrt[\alpha]{c_2(2^{\mathcal{R}_0/W} - 1)} + 1$. It follows that $\mathcal{R}_0$-PCR $= \kappa \cdot r$. Then, we can obtain the following lemma by applying Lemma 3.

Lemma 4: Assume that $\mathcal{X}$ is a disk of radius $\mathcal{R}_0$-PCR, then $|\mathcal{X} \cap (\mathcal{D} \cup \mathcal{C})| \leq 13\beta_{\kappa+1}$, where $\beta_{\kappa+1} = \frac{2\pi(\kappa+1)^2}{\sqrt{3}} + \pi(\kappa+1) + 1$, *i.e.* the number of dominators and connectors within the CR of a node is at most $13\beta_{\kappa+1}$ in DDC.

Proof: Since $\mathcal{X}$ is a disk of radius $\mathcal{R}_0$-PCR, it is possible for some connectors in $\mathcal{X}$ only connect some dominators out of disk $\mathcal{X}$ as shown in Fig. 5. On the other hand, all the dominators adjacent to the connectors in $\mathcal{X} \cap \mathcal{C}$ must locate in a concentric disk of $\mathcal{X}$ with radius $\mathcal{R}_0$-PCR$+r = (\kappa + 1)r$, denoted by $\mathcal{X}'$ as shown in Fig. 5.

Now, if r is normalized to 1, then $\mathcal{X}'$ is a disk of radius $\kappa + 1$ and $\mathcal{D}$ is a set of nodes with mutual distance of at least 1. Then, by Lemma 3, we have $|\mathcal{X}' \cap \mathcal{D}| \leq \beta_{\kappa+1} = \frac{2\pi(\kappa+1)^2}{\sqrt{3}} + \pi(\kappa + 1) + 1$, *i.e.* the number of the dominators within $\mathcal{X}'$ is at most $\beta_{\kappa+1}$. According to Corollary 1, each dominator in T is adjacent to at most 12 connectors, which implies $|\mathcal{X} \cap \mathcal{C}| \leq 12|\mathcal{X}' \cap \mathcal{D}| \leq 12\beta_{\kappa+1}$. Then, $|\mathcal{X} \cap (\mathcal{D} \cup \mathcal{C})| \leq 13\beta_{\kappa+1}$. □

From Lemma 4, we can obtain the following corollary.

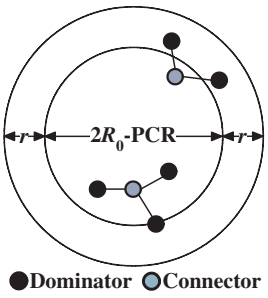


Fig. 5. The number of dominators and connectors within the CR of a node.

Corollary 3: After time $2\mathbb{U}\tau$, every node in $(\mathcal{D}\cup\mathcal{C})\setminus\{s_0\}$ with data packets for transmission can transmit at least one data packet to its parent node within time $26\beta_{\kappa+1}\tau$ in DDC.

Proof: Please refer to [37] for details. □

Based on Lemma 4 and Corollary 3, we can obtain the time consumption to collect all the data packets at $(\mathcal{D}\cup\mathcal{C})\setminus\{s_0\}$ to the sink s_0 as shown in Theorem 3.

Theorem 3: After time $2\mathbb{U}\tau$, it takes at most $26(n-\Delta_0)\cdot\beta_{\kappa+1}\cdot\tau$ time to collect all the data packets at $(\mathcal{D}\cup\mathcal{C})\setminus\{s_0\}$ to the sink s_0 in DDC, where Δ_0 is the degree of s_0 in the data collection tree T.

Proof: As shown in Corollary 2, after time $2\mathbb{U}\tau$, all the nodes in $V\setminus(\mathcal{D}\cup\mathcal{C})$ have no data packets for transmission, and meanwhile, s_0 has received at least Δ_0 data packets according to Theorem 2 since it has Δ_0 children nodes in T. Subsequently, s_0 will receive at least one data packet in every $26\beta_{\kappa+1}\tau$ time according to Corollary 3. Thus, it takes at most $(n-\Delta_0)\cdot 26\beta_{\kappa+1}\tau$ time to collect all the data packets at $(\mathcal{D}\cup\mathcal{C})\setminus\{s_0\}$ to the sink s_0 after time $2\mathbb{U}\tau$. □

Theorem 4: The lower bound of data collection capacity achieved by DDC is $\Omega(\frac{1}{26\beta_{\kappa+1}}\cdot W)$, which is scalable and order-optimal.

Proof: According to Theorem 2 and Theorem 3, to collect all the n data packets to the sink, the time consumption

$$\mathcal{T}\le 2\mathbb{U}\tau+26(n-\Delta_0)\cdot\beta_{\kappa+1}\cdot\tau \tag{21}$$

$$=[(2\log n+\frac{\pi R^2(e^2-1)}{c_1})+26(n-\Delta_0)\cdot\beta_{\kappa+1}]\cdot\tau \tag{22}$$

$$\le[2\log n+\frac{\pi R^2(e^2-1)}{c_1}+26\beta_{\kappa+1}n]\cdot\tau \tag{23}$$

$$=O(26\beta_{\kappa+1}n\cdot\tau). \tag{24}$$

Thus, the achievable data collection capacity of DDC is

$$\mathbb{C}=\frac{nB}{\mathcal{T}}\ge\frac{nB}{O(26\beta_{\kappa+1}n\cdot\tau)}=\Omega(\frac{1}{26\beta_{\kappa+1}}\cdot W). \tag{25}$$

Considering that the upper bound capacity of data collection is $O(W)$ [6][7], and $\beta_{\kappa+1}$ is a constant value depending on $\mathcal{R}_0$, it implies the achievable data collection capacity of the DDC algorithm is order-optimal. Furthermore, since $\mathbb{C}$ is independent of the network size n, DDC is thus scalable. □

VI. Simulation Results

In this section, we conduct simulations to validate the performance of DDC. In all the simulations, we consider the WSNs consisting of one sink node and n sensor nodes which are randomly deployed in a square area with size $A=c_1 n$. Thus the node density is $\frac{1}{c_1}$. The maximum transmission radius of a sensor node is normalized to one and any node can work on the Re-Start (RS) mode with the IPCS technique [36]. During the data collection period, every sensor node produces a data packet whose size is also normalized to one. Furthermore, all the nodes work with the same power $P=1$ and over a common wireless channel with bandwidth normalized to one, which implies the transmission time of a data packet $\tau=1$ in the ideal case. Then, we set the backoff contention window $\tau_w=\frac{1}{10}$ for DDC in all the simulations. For a data transmission, the background noise is negligible compared with the interference brought by concurrent transmissions. Hence, we do not consider the background noise. For other important system parameters, *e.g.* the network size A, the node density $\frac{1}{c_1}$, the number of sensor nodes n, the path loss exponent α, and *etc*., we specify them later in each group of simulations.

The compared algorithm is the *Multi-Path Scheduling* (MPS) algorithm proposed in [7], which is the most recently centralized and synchronized data collection method under the simplified *protocol interference model* for WSNs. In MPS, the interference radius $R_I=\eta\cdot r(\eta\ge 1)$, where η is a constant and r is the communication radius of a node. Thus, in the following simulations, we set $R_I=\mathcal{R}_0$-PCR, which guarantees that MPS can also initiate data transmissions with a satisfied data receiving rate $\mathcal{R}_0$. In the following, each group of simulations is repeated for 100 times and the results are the average values.

A. *DDC Capacity* vs. $\mathcal{R}_0$ *and* α

In this subsection, we consider the WSNs deployed in a square area with size $A=20\times 20$ and the node density is 3. The impacts of $\mathcal{R}_0$ and α on the capacities of DDC and MPS are shown in Table I. According to Table I, with the increase of $\mathcal{R}_0$, the achievable capacities of DDC and MPS increase. Although a large $\mathcal{R}_0$ implies a large $\mathcal{R}_0$-PCR (Fig. 2), which further implies that fewer nodes can conduct transmissions concurrently, nevertheless, a large $\mathcal{R}_0$-PCR also implies short transmission time of a data packet. Furthermore, with the increase of $\mathcal{R}_0$, the decrease of the transmission time of a data packet is faster than the increase of $\mathcal{R}_0$-PCR. It follows that a large $\mathcal{R}_0$ leads to a high capacity finally. From Table I, we can also see that with the increase of α, the achievable capacities of DDC and MPS also increase. This is because, for any transmitter, the interference impact from other concurrent transmitters decreases quickly with the increase of α. Thus, a large α implies a small $\mathcal{R}_0$-PCR, resulting in more nodes being able to initiate transmissions concurrently. Furthermore, from Table I, the capacity difference between DDC and MPS is very small although DDC is a distributed and asynchronous data collection algorithm. This is because we set a proper CR for DDC. By setting the CR of each node as $\mathcal{R}_0$-PCR, as many

TABLE I
DDC CAPACITY *vs.* MPS CAPACITY.

$\mathcal{R}_0$	$\alpha = 3$		$\alpha = 4$	
	DDC	MPS	DDC	MPS
.1	.03173	.03358	.03186	.03287
.2	.0536	.06216	.05644	.06558
.3	.08342	.08814	.08853	.09458
.4	.11512	.11225	.1194	.1223
.5	.119	.13467	.13421	.149
.6	.1326	.15036	.1626	.17238
.7	.16543	.1701	.191	.19987
.8	.19151	.19506	.21778	.22421
.9	.20497	.21352	.23815	.24712
1	.22255	.23321	.25938	.27011

as possible nodes can initiate data transmissions concurrently with a guaranteed data receiving rate at the receivers. Particularly, the average capacity differences between DDC and MPS are 5.25% and 4.99% when $\alpha = 3$ and $\alpha = 4$, respectively, which indicates that DDC even achieves comparable capacity as centralized and synchronized MPS.

VII. CONCLUSION

Since WSNs in practice tend to be distributed asynchronous systems and most of the existing works study the network capacity issues for centralized synchronized WSNs, we investigate the achievable data collection capacity for distributed asynchronous WSNs in this paper. To avoid data transmission collisions/interferences, we derive an $\mathcal{R}_0$-Proper Carrier-sensing Range ($\mathcal{R}_0$-PCR) under the generalized physical interference model. By taking $\mathcal{R}_0$-PCR as its carrier-sensing range, any node can initiate a data transmission with a guaranteed data receiving rate. Subsequently, based on the obtained $\mathcal{R}_0$-PCR, we propose a scalable Distributed Data Collection (DDC) algorithm with fairness consideration for asynchronous WSNs. Theoretical analysis of DDC shows that its achievable data collection capacity is also order optimal as that of centralized synchronized algorithms. Extensive simulation results demonstrate that DDC has comparable data collection capacity compared with the most recently published centralized and synchronized data collection algorithm, and DDC is scalable in WSNs with different network sizes and node densities.

ACKNOWLEDGEMENT

This work is partly supported by the NSF under grant No. CNS-1152001.

REFERENCES

[1] P.-J. Wan, S. C.-H. Huang, L. Wang, Z. Wan, and X. Jia, Minimum-Latency Aggregation Scheduling in Multihop Wireless Networks, Mobihoc 2009.
[2] C. Liu and G. Cao, Distributed Monitoring and Aggregation in Wireless Sensor Networks, Infocom 2010.
[3] P. Gupta and P. R. Kumar, The Capacity of Wireless Networks, IEEE Transactions on Information Theory, 46(2): 388-404, 2000.
[4] C. Luo, F. Wu, J. Sun, and C. W. Chen, Compressive Data Gathering for Large-Scale Wireless Sensor Networks, Mobicom 2009.
[5] X. Zhu, B. Tang, and H. Gupta, Delay Efficient Data Gathering in Sensor Networks, MSN 2005.
[6] Siyuan Chen, Shaojie Tang, Minsu Huang, and Yu Wang, Capacity of Data Collection in Arbitrary Wireless Sensor Networks, Infocom 2010.
[7] S. Ji, Y. Li, and X. Jia, Capacity of Dual-Radio Multi-Channel Wireless Sensor Networks for Continuous Data Collection, Infocom 2011.
[8] S. Chen, Y. Wang, X.-Y. Li, and Xinghua Shi, Order-Optimal Data Collection in Wireless Sensor Networks: Delay and Capacity, Secon 2009.
[9] T. Moscibroda, The Worst-Case Capacity of Wireless Sensor Networks, IPSN 2007.
[10] C. Wang, C. Jiang, Y. Liu, X.-Y. Li, S. Tang, and H. Ma, Aggregation Capacity of Wireless Sensor Networks: Extended Network Case, Infocom 2011.
[11] X.-Y. Li, S.-J. Tang, and O. Frieder, Multicast Capacity for Large Scale Wireless Ad Hoc Networks, Mobicom 2007.
[12] S. Li, Y. Liu, and X.-Y. Li, Capacity of Large Scale Wireless Networks Under Gaussian Channel Model, Mobicom 2008.
[13] X. Mao, X.-Y. Li, and S. Tang, Multicast Capacity for Hybrid Wireless Networks, Mobihoc 2008.
[14] Z. Wang, H. R. Sadjadpour, and J. J. Garcia-Luna-Aceves, A Unifying Perspective on the Capacity of Wireless Ad Hoc Networks, Infocom 2008.
[15] Y. Xu and W. Wang, Scheduling Partition for Order Optimal Capacity in Large-Scale Wireless Networks, Mobicom 2009.
[16] V. Ramamurthi, S. K. C. Vadrevu, A. Chaudhry, and M. R. Bhatnagar, Multicast Capacity of Multi-Channel Multihop Wireless Networks, WCNC 2009.
[17] Y. Wang, X. Chu, X. Wang, and Y. Cheng, Optimal Multicast Capacity and Delay Tradeoffs in MANETs: A Global Perspective, Infocom 2011.
[18] P. Kyasanur and N. H. Vaidya, Capacity of Multi-Channel Wireless Networks: Impact of Number of Channels and Interfaces, Mobicom 2005.
[19] O. Goussevskaia, R. Wattenhofer, M. M. Halldorsson, and E. Welzl, Capacity of Arbitrary Wireless Networks, Infocom 2009.
[20] M. Andrews and M. Dinitz, Maximizing Capacity in Arbitrary Wireless Networks in the SINR Model: Complexity and Game Theory, Infocom 2009.
[21] E. I. Asgeirsson and P. Mitra, On a Game Theoretic Approach to Capacity Maximization in Wireless Networks, Infocom 2011.
[22] U. Niesen, P. Gupta, and D. Shah, On Capacity Scaling in Arbitrary Wireless Networks, IEEE Transactions on Information Theory, 55(9): 3959-3982, 2009.
[23] -, The Balanced Unicast and Multicast Capacity Regions of Large Wireless Networks, IEEE Transactions on Information Theory, 56(5): 2249-2271, 2010.
[24] M. Garetto, P. Giaccone, and E. Leonardi, On the Capacity of Ad Hoc Wireless Networks Under General Node Mobility, Infocom 2007.
[25] G. Sharma, R. Mazumdar, and N. B. Shroff, Delay and Capacity Trade-Offs in Mobile Ad Hoc Networks: A Global Perspective, IEEE/ACM Transactions on Networking, 15(5): 981-992, 2007.
[26] K. Lee, Y. Kim, S. Chong, I. Rhee, and Y. Yi, Delay-Capacity Tradeoffs for Mobile Networks with Lévy Walks and Lévy Flights, Infocom 2011.
[27] V. Bhandari and N. H. Vaidya, Connectivity and Capacity of Multi-Channel Wireless Networks with Channel Switching Constraints, Infocom 2007.
[28] -, Capacity of Multi-Channel Wireless Networks with Random (c, f) Assignment, Mobihoc 2007.
[29] H.-N. Dai, K.-W. Ng, R. C.-W. Wong, and M.-Y. Wu, On the Capacity of Multi-Channel Wireless Networks Using Directional Antennas, Infocom 2008.
[30] H. Li, Y. Cheng, P.-J. Wan, and J. Cao, Local Sufficient Rate Constraints for Guaranteed Capacity Region in Multi-Radio Multi-Channel Wireless Networks, Infocom 2011.
[31] P. Li, M. Pan, and Y. Fang, The Capacity of Three-Dimensional Wireless Ad Hoc Networks, Infocom 2011.
[32] P. Li, X. Huang, and Y. Fang, Capacity Scaling of Multihop Cellular Networks, Infocom 2011.
[33] D. Chafekar, D. Levin, V. S. A. Kumar, M. V. Marathe, S. Parthasarathy, and A. Srinivasan, Capacity of Asynchronous Random-Access Scheduling in Wireless Networks, Infocom 2008.
[34] C.-K. Chau, M. Chen, and S. C. Liew, Capacity of Large-Scale CSMA wireless Networks, Mobicom 2009.
[35] X.-Y. Li, J. Zhao, Y. Wu, S. Tang, X. Xu, X. Mao, Broadcast Capacity for Wireless Ad Hoc Networks, MASS 2008.
[36] L. Fu, S. C. Liew, and J. Huang, Effective Carrier Sensing in CSMA Networks under Cumulative Interference, Infocom 2010.
[37] S. Ji and Z. Cai, Technical Report: Distributed Data Collection and Its Capacity in Asynchronous Wireless Sensor Networks, http://www.cs.gsu.edu/∼sji/Paper/TechReport-distributed.pdf.