Supplementary File of Continuous Data Collection Capacity of Dual-Radio Multi-Channel Wireless Sensor Networks

Shouling Ji, Student Member, IEEE, Zhipeng Cai, Yingshu Li†, Senior Member, IEEE, and Xiaohua Jia, Senior Member, IEEE

Abstract—The performance of data collection in Wireless Sensor Networks (WSNs) can be measured by network capacity. However, few existing works dedicatedly consider the Continuous Data Collection (CDC) capacity for WSNs under the protocol interference model. In this paper, we propose a multi-path scheduling algorithm for SDC in single-radio multi-channel WSNs and derive its network capacity which is a tighter lower bound compared with the previously best result [10]. We also propose a novel CDC method for dual-radio multi-channel WSNs. It significantly speeds up the data collection process, and achieves a capacity of \( \frac{12M}{[3.63n^2+c_3+c_4]/H} \) when \( \Delta_n \leq 12 \) or \( M_{\Delta_n} [3.63n^2+c_3+c_4]/H \) when \( \Delta_n > 12 \), where \( n \) is the number of the sensors, \( M \) is a constant value and usually \( M \ll n \), \( \Delta_n \) is the maximum number of the leaf nodes having a same parent in the data collection tree, \( W \) is the channel bandwidth, \( H \) is the number of available orthogonal channels, \( \rho \) is the ratio of the interference radius over the transmission radius, \( c_3 = \frac{8\pi}{\sqrt{3}} + \pi + 2 \), and \( c_4 = \frac{8\pi}{\sqrt{3}} + 2\pi + 6 \). Extensive simulation results indicate that the proposed algorithms improve network capacity significantly compared with existing works.

Index Terms—Wireless sensor networks, snapshot data collection, continuous data collection, capacity analysis.

1 RELATED WORK

1.1 Capacity for Single-Radio Single-Channel Wireless Networks

Following the seminal work [20] by Gupta and Kumar, extensive works emerged to study the network capacity issue. The works in [21]-[24] focus more on the MAC layer to improve the network capacity. In [21], the network capacity with random-access scheduling is investigated. In this work, each link is assigned a channel access probability. Based on which some simple and distributed channel access strategies are proposed. Another similar work is [22], in which the authors studied the capacity of CSMA wireless networks. The authors formulated the models of a series of CSMA protocols and study the capacity of CSMA scheduling versus TDMA scheduling. They also proposed a CSMA scheme which combines a backbone-peripheral routing scheme and a dual carrier-sensing and dual channel scheme. In [23], 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 links in one time-slot. Unlike [23], the authors in [24] considered the power-control problem. A family of approximation algorithms were presented to maximize the capacity of an arbitrary wireless networks.

The works in [25], [26], [27], [28], [29], and [30] study the multicast and/or unicast capacity of wireless networks. The multicast capacity for wireless ad hoc networks under the protocol interference model and the Gaussian channel model are investigated in [25] and [26] respectively. In [25], the authors showed that the network multicast capacity is \( \Theta(\sqrt{n/\log n}) \) when \( k = O(n/\log n) \) and is \( \Theta(W) \) when \( k = \Omega(n/\log n) \), 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 [26], the authors showed that when \( k \leq \theta_1 n/\log n \) and \( n_s \geq \theta_2 n^{1/2+\beta} \), the capacity that each multicast session can achieve is at least \( c_5 n^{\beta} \), where \( k \) is the number of the receivers in one multicast session, \( n \) is the number of the nodes in the network, \( n_s \) is the number of the multicast sessions, \( \theta_1 \), \( \theta_2 \) and \( c_5 \) are constants and \( \beta \) is any positive real number. Another similar work [27] 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. Considering the problem of characterizing the unicast capacity scaling in arbitrary...
wireless networks, the authors proposed a general cooperative communication scheme in [28]. The authors also presented a family of schemes that address the issues between multi-hop and cooperative communication when the path-loss exponent is greater than 3. In [29], 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 \( n^{2-k} \)-dimensional balanced multicast capacity regions under the Gaussian fading channel model. A more general \((n, m, k)\)-casting capacity problem was investigated in [30], where \( n \), \( m \), and \( k \) denote the total number of the nodes in the network, the number of destinations for each communication group, and the actual number of communication-group members that receive information respectively. In [30], the upper and lower bounds for the \((n, m, k)\)-cast capacity were obtained for random wireless networks.

In [44], the authors investigated the network capacity scaling in mobile wireless ad hoc networks under the protocol interference model with infrastructure support. In [45], the authors studied the network capacity of hybrid wireless networks with directional antenna and delay constraints. Unlike previous works, the authors in [46] studied the capacity of multi-unicast for wireless networks from the algorithmic aspects, and they designed provably good algorithms for arbitrary instances. The broadcast capacity of wireless networks under the protocol interference model is investigated in [47], where the authors derived the upper and lower bounds of the broadcast capacity in arbitrary connected networks. When the authors in [48] studied the data gathering capacity of wireless networks under the protocol interference model, they concerned the per source node throughput in a network where a subset of nodes send data to some designated destinations while other nodes serve as relays. To gather data from WSNs, a multi-query processing technology is proposed in [49]. In that work, the authors considered how to obtain data efficiently with data aggregation and query scheduling. Under different communication organizations, the authors in [50] derived the many-to-one capacity bound under the protocol interference model. Another work studied the many-to-one capacity issue for WSNs is [51], where the authors considered to use data compression to improve the data gathering efficiency. They also studied the relation between a data compression scheme and the data gathering quality. In [52], the authors studied the scaling laws of WSNs based on an antenna sharing idea. In that work, the authors derived the many-to-one capacity bounds under different power constraints. In [53], the authors studied the multi-unicast capacity of MANETs under the physical interference model, called motioncast. They considered the network capacity of MANETs in two particular situations, which are the LSRM (local-based speed-restricted) model and the GSRM (global-based speed-restricted) model. The multi-unicast capacity of wireless networks is studied in [54] via percolation theory. By applying percolation theory, the authors obtained a tighter capacity bound for arbitrary wireless networks.

The SDC capacity of WSNs is studied in [6], [10], [11], [5], [31], [7], [8] and [32]. In [6], the authors considered the collision-free delay-efficient data gathering problem. Furthermore, they proposed a family of path scheduling algorithms to collect all the data to the sink and obtained the network capacity through theoretical analysis. The authors of [10], [11] extended the work of [6]. They derived tighter upper and lower bounds of the capacity of data collection for arbitrary WSNs. [5] is a work studying how to distribute the data collection task to the entire network to achieve load balancing. In this work, all the sensors transmit the same number of data packets during the data collection process. In [31] and [7], [8], the authors investigated the capacity of data collection for WSNs under protocol interference model and physical interference model, respectively. They proposed a grid partition method which divides the network into small grids to collect data and then derived the network capacity. The worst-case capacity of data collection of a WSN is studied in [32] under the physical and protocol interference models.

The capacity and energy efficiency of wireless ad hoc networks with multi-packet reception under the physical interference model is investigated in [33]. With the multi-packet reception scheme, a tight bound of the network capacity is obtained. Furthermore, the authors showed that a tradeoff can be made between increasing the transport capacity and decreasing the energy efficiency. In [34], a scheduling partition method for large-scale wireless networks is proposed. This method decomposes a large network into many small zones, and then localized scheduling algorithms which can achieve the order optimal capacity as a global scheduling strategy are executed in each zone independently. A general framework to characterize the capacity of wireless ad hoc networks with arbitrary mobility patterns is studied in [35]. By relaxing the “homogeneous mixing” assumption in most existing works, the capacity of a heterogeneous network is analyzed. Another work [36] 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.

### 1.2 Capacity for Multi-Channel Wireless Networks

Since wireless nodes can be equipped with multiple radios, and each radio can work over multiple orthogonal channels, multi-radio multi-channel wireless networks attract many research interests recently [40][41][42][43]. In [40], the authors studied the data aggregation issue in multi-channel WSNs under the protocol interference model. Particularly, they designed a constant factor approximation scheme for data aggregation in multi-channel WSNs modeled by Unit Disk Graphs (UDGs).
Unlike [40], we study the data collection capacity issue for WSNs. In [41], [42], and [43] the authors investigated the joint channel assignment and routing problem for multi-radio wireless mesh networks, software-defined radio networks, and multi-channel ad hoc wireless networks, respectively. They focused on the channel assignment and routing issues, while in data collection, especially continuous data collection, we focus on how to solve the data accumulation problem at the sensors near the sink to improve the achievable network capacity.

The issue of the capacity of multi-channel wireless networks also attracts a lot of attention [37][38][39] [55][16]. In [37] and [38], the authors studied the connectivity and capacity problem of multi-channel wireless networks. They considered a multi-channel wireless network under constraints on channel switching, proposed some routing and channel assignment strategies for multiple unicast communications and derived the per-flow capacity. The multicast capacity of multi-channel wireless networks is studied in [39]. In this work, the authors represented the upper bound capacity of per multicast as a function of the number of the sources, the number of the destinations per multicast, the number of the interfaces per node, and the number of the available channels. Subsequently, an order-optimal scheduling method is proposed under certain circumstances. In [55], 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 under arbitrary and random network models. The impact of the number of the channels, the number of the interfaces and the interface switching delay on the capacity of multi-channel wireless networks is investigated in [16]. In this work, the authors derived the network capacity under different situations for arbitrary and random networks.

### 1.3 Remarks

Unlike the above mentioned works, our work has the following two main characteristics. First, most of the above mentioned works are specifically for single-radio single-channel wireless networks, while our work considers the network capacity for dual-radio multi-channel WSNs. Second, our work is the dedicated one that investigates the network capacity for CDC in detail under the protocol interference model, whereas most of the previous works study the network capacity for multicast or/and unicast and etc. For the works that study the data collection capacity of wireless networks, they focus on the SDC problem which is a special case of CDC. Compared with them, the results proposed in this paper are more universal.

### 2 Proofs of Lemma 4, Corollary 1, Lemma 5, Lemma 6, and Lemma 7

**Proof of Lemma 4:** Suppose that the node sequence on $P$ is $s_1, s_2, \ldots, s_L, s_0$, where $s_1$ is the leaf node (dominatee), and $s_0$ is the sink node. Considering the building process of $T$, each link in $P$ has either a dominator head or a dominator tail. According to the scheduling scheme of a single path, during the first (odd) round, the links in $P_o$ are scheduled, which implies each non-dominator with at least one packet transmits this packet to its parent node. After the first round, the sink, receives one packet and all the other dominators of the links in $P_o$ have two packets to be transmitted. During the second (even) round, the links in $P_e$ are scheduled, which implies that every dominator in $P_e$ transmits one packet to its parent node. As a result, the sensor $s_i (2 \leq i \leq L)$ has exactly one packet to be transmitted and a new odd-even scheduling round begins. In summary, after every two rounds, the sink receives one packet and the length of the data collection path decreases by 1. Since the length of $P$ is $L$ and $s_0$ is the destination of all the packets which does not have to transmit any data, it takes at most $2L − 1$ rounds to collect all the packets on $P$.

**Proof of Corollary 1:** The proof of Corollary 1 is similar to that of Lemma 4. Note that the intersecting point is not a sink node in this case and thus it needs one round to transmit its packet.

**Proof of Lemma 5:** During every odd (respectively, even) round, the scheduled links are links in $P_o$ (respectively, $P_e$). Since the heads (respectively, tails) of links in $P_o$ (respectively, $P_e$) are dominators, we can schedule all the links in $P_o$ (respectively, $P_e$) in one time slot with at
most \( \beta_{p+1} \) channels in polynomial time by Lemma 3 and Lemma 2. Now, we have \( H \) available channels, which means we can finish the scheduling within \( \left\lceil \frac{\beta_{p+1}}{P} \right\rceil \) time slots. Therefore, the lemma holds.

**Proof of Lemma 6:** Since the sibling nodes at every level have been divided into different subsets and different subsets are scheduled in a certain order, there is no radio confliction among the links in \( A \). Furthermore, for any link in \( A \), either the tail or the head of this link is a dominator according to the building process of the routing tree \( T \). Suppose that \( A' \) is a subset of \( A \) and \( e \) is the link in \( A' \) whose tail, denoted by \( (e) \), or head, denoted by \( h(e) \), is the bottommost dominator among all the dominators in \( A' \). Then, we prove the number of the links interfered with \( e \), i.e. \( \delta(R(A')) \), is at most \( 2\beta_{p+2} - 1 \) case by case as follows.

Case 1: \( t(e) \) is a dominator. In this case, assume that \( e' \) is another link in \( A' \) interfered with and \( t(e') \) interferes with it. Since \( t(e) \) is the bottommost dominator, the necessary condition for \( e \) and \( e' \) to be interfering links is that \( t(e') \) locates at the upper half-disk centered at \( t(e) \) with radius \( \rho + 1 \). On the other hand, if \( h(e') \) is a dominator, then the necessary condition for \( e \) and \( e' \) to be interfering links is that \( h(e') \) locates at the upper half-disk centered at \( t(e) \) with radius \( \rho + 2 \). By Lemma 1, the number of the dominators within a half-disk of radius \( \rho + 2 \) is at most \( \beta_{p+2} \). Since every dominator in \( A' \) is associated with at most two links, there are at most \( 2\beta_{p+2} \) links within the half-disk of radius \( \rho + 2 \) centered at \( t(e) \). Therefore, \( \delta(R(A')) \leq 2\beta_{p+2} - 1 \), where minus 1 means \( e \) is also in the half-disk. As a result, \( \delta^*(R(A)) = \max_{A' \subseteq A} \delta(R(A')) \leq 2\beta_{p+2} - 1 \).

Case 2: \( h(e) \) is a dominator. By the similar method as in Case 1, it can be proven that the conclusion also holds in this case.

**Proof of Lemma 7:** By Lemma 6, \( \delta^*(R(A)) \leq 2\beta_{p+2} - 1 \). By Lemma 2, we can use \( 1 + \delta^*(R(A)) \leq 2\beta_{p+2} \) to schedule all the links in \( A \) in one STS simultaneously. Now, we have \( H \) channels, which implies we can schedule all the links in \( A \) in \( \left\lceil \frac{2\beta_{p+2}}{P} \right\rceil \) STSs of \( H \) links in each STS.

**3 Simulations and Results Analysis**

We conducted simulations to verify the performances of the proposed algorithms through implementing them with the C language. For all the simulations, we assume every WSN has one sink, and all the sensor nodes of each WSN are randomly distributed in a square area and the communication radius of each node is normalized to one. Suppose the network MAC layer works with TDMA, i.e., the network time can be slotted. Every node produces one data packet in a snapshot and the size of a packet is normalized to one. Every available channel has the same bandwidth normalized to one. For any two different channels, we suppose they are orthogonal, i.e., the communications initialized over any two channels have no wireless interference. Furthermore, we assume a packet can be transmitted over a channel within a time slot.

The compared algorithms are BFS [10], SLR [37] and CDG [5]. BFS is a SDC algorithm based on a breadth first search tree and the scheduling is carried out path by path [10]. BFS is specifically proposed for single-radio single-channel WSNs. We extend it to the dual-radio multi-channel scenario in our simulations for fairness. SLR is a straight-line routing method for multi-unicast communication in multi-channel wireless networks with channel switching constraints [37]. For data collection, SLR works by setting every sensor having a unicast communication with the sink simultaneously. We also remove the channel switching constraints in SLR for fairness. Furthermore, we also implement the pipelined versions of BFS and SLR (i.e., add the pipeline technique to the data transmission in BFS and SLR, referred to as BFS-P and SLR-P respectively, when evaluate the performance of the proposed pipeline scheduling algorithm. The basic idea of CDG is discussed in Section 4.1 of the main file. The proposed multi-path scheduling algorithm for SDC is referred to as MPS and the proposed pipeline scheduling algorithm for CDC is referred to as PS in the following discussions.

In the remainder of this section, we investigate the achievable capacities of MPS and PS through three groups of simulations respectively. In the simulations, \( H \) is the number of the available channels, \( \rho \) is the interference radius, \( n \) is the number of the sensors in a WSN, AR refers to the square area where a WSN is deployed, and \( N \) is the number of the snapshots in a CDC task.

**3.1 Performance of MPS**

The SDC capacities of MPS, BFS, and SLR in different network scenarios are shown in Fig.1. In Fig.1(a), the capacity of every algorithm increases when the number of the available channels increases. This is because more available channels enable more concurrent transmissions, which accelerates the data collection process resulting in a higher capacity. After the number of the available channels arrives at 4, the capacities of BFS and SLR almost maintain the same level. This is because 4 channels are enough to prevent channel interference. However, radio confliction becomes the main barrier of a higher capacity at this time. MPS achieves a higher capacity compared with BFS and SLR. This is because MPS simultaneously schedules all the paths without radio confliction (except at the sink). Since radio confliction on a single path can be avoided easily, MPS can simultaneously schedule all the links without radio confliction on multiple paths, which implies MPS can make use of channels in a maximum degree. Whereas, BFS just schedules links without radio confliction on one path every time and SLR schedules all the transmission links simultaneously, which leads to serious radio confliction.
On average, MPS achieves 77.49% and 41.95% more capacity than BFS and SLR, respectively.

The effect of the interference radius on the capacity is shown in Fig.1(b). With the increase of the interference radius, more transmission interference occurs, which leads to the decrease of the capacities of all the algorithms. Nevertheless, MPS still achieves the largest capacity since it simultaneously schedules multiple paths without radio confliction, which suggests a nice trade-off between BFS and SLR. On average, MPS achieves 67.45% and 37.37% more capacity than BFS and SLR, respectively.

The effect of the number of the sensors on the capacity is shown in Fig.1(c). We can see that the number of the sensors in a network has a little impact on the capacities of MPS and SLR and almost no impact on the capacity of BFS. There are two reasons for this result. First, BFS is a single-path scheduling algorithm. Whatever the number of the sensors is, it schedules only one path every time. Second, the number of the channels is fixed to 2 in all of these three algorithms. This implies that whatever the number of the sensors is, they can simultaneously schedule at most two interfering links without radio confliction. On average, MPS achieves 83.51% and 32.87% more capacity than BFS and SLR, respectively.

### 3.2 Performance of PS

The CDC capacities of PS, CDG, BFS-P, BFS, SLR-P, and SLR in different network scenarios are shown in Fig.2.

![Fig. 1. SDC Capacity (packets/time slot).](image-url)

(a) Capacity vs. \(H (\rho=2, n=4000, AR=30 \times 30)\) (b) Capacity vs. \(\rho (H=3, n=4000, AR=30 \times 30)\) (c) Capacity vs. \(n (\rho=2, H=3, AR=20 \times 20)\)
In this subsection, we investigate the impacts of the number of sensors on the capacity of PS and CDG. As shown in Fig. 3(a), with the increase of $N$ in a CDC task, PS achieves about 87.54% more capacity. This is straightforward from the analysis in Section 4 of the main file. Since PS employs the pipeline technology, the transmissions of continuous snapshots are overlapped, which can significantly reduce the time used to collect all the snapshots data. With more snapshots in a CDC task, the capacity of PS approaches closer and closer to its theoretical asymptotic capacity. For CDG, the number of the snapshots has little impact on its capacity.

Since the performance of CDG is depends on the value of $M$, the capacities of both PS and CDG decrease about 80% with the increase of the value of $M$ as shown in Fig. 3(b). This is because a bigger $M$ implies more packets have to be transmitted for every sensor and longer transmission time for each snapshot is resulted. Nevertheless, considering that the value of $M$ is usually much less than $n$, PS can still achieve a high capacity.

### ACKNOWLEDGMENTS

The authors would like to thank the editor and innuminate reviewers for their time and valuable comments that help to improve the quality of this manuscript.

### REFERENCES


[46] V. S. A. Kumar, M. V. Marathe, S. Parthasarathy, and A. Sriniv


Shouling Ji is a Ph.D. Student 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, Data Management in Wireless Networks, and Cyber-Physical Systems. He is now an ACM student member, an IEEE student member and an IEEE COMSOC student member.

Zhipeng Cai 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. He is currently an adjunct assistant professor in the Department of Computer Science at the Georgia State University, USA. His research interests include wireless networks, bioinformatics and optimization theory. Dr. Cai received a number of awards and honors, including NSERC Postdoc Fellowship, Ph.D Outstanding Research Achievement Award. He also served as a guest editor for Algorithmica and Journal of Combinatorial Optimization.

Yingshu Li received her Ph.D. and M.S. degrees from the Department of Computer Science and Engineering at University of Minnesota-Twin Cities. She received her B.S. degree from the Department of Computer Science and Engineering at Beijing Institute of Technology, China. Dr. Li is currently an Associate Professor in the Department of Computer Science at Georgia State University. Her research interests include Wireless Networking, Cyber-Physical Systems, and Phylogenetic Analyses. Her research has been supported by the National Science Foundation (NSF) of US, the National Science Foundation of China (NSFC), the Electronics and Telecommunications Research Institute (ETRI) of South Korea, and GSU internal grants. Dr. Li is the recipient of an NSF CAREER Award.

Xiaohua Jia received the BSc (1984) and MEng (1987) degrees from the University of Science and Technology of China and received the DSc (1991) degree in information science from the University of Tokyo, Japan. Professor Jia is currently associated with the Department of Computer Science at the City University of Hong Kong. His research interests include distributed systems, computer networks, WDM optical networks, and Internet and mobile computing. He is a senior member of the IEEE.