Time Efficient Data Aggregation Scheduling in Cognitive Radio Networks

Mingyuan Yan†, Shouling Ji†‡, and Zhipeng Cai†
†Department of Computer Science, Georgia State University, Atlanta, Georgia 30303, USA
Email: {myan2, zcai}@cs.gsu.edu
‡School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30308, USA
Email: sji@gatech.edu

Abstract—Cognitive Radio Networks (CRNs) have been proposed as a promising technology to alleviate the spectrum shortage and under-utilization problem. A large amount of research effort has been put on CRNs. However, little attention has been paid to the data aggregation problem, which is one of the most important communication protocols in wireless networks due to its efficiency on both energy conservation and latency reduction. In this paper, we investigate the minimum latency data aggregation scheduling problem in CRNs. Considering the hardness of this problem, two practical distributed algorithms under the Unit Disk Graph interference model and the Physical Interference Model are proposed. Extensive simulation results indicate that the proposed algorithms have good latency performance in various network scenarios.

I. INTRODUCTION

CRNs have been considered as a promising solution to alleviate the spectrum shortage and under-utilization problem [1][5]. In Cognitive Radio Networks (CRNs), unlicensed users (Secondary Users, SUs) coexist with Primary Users (PUs). As long as communications among PUs are ensured, SUs are able to access and exploit the unoccupied licensed spectrums through a cognitive radio (that is capable of sensing/accessing available spectrums and switching between spectrums) in an opportunistic manner. Communication protocols such as unicast, broadcast, and multicast have been widely investigated [2]-[4]. However, limited attention has been paid to data aggregation in CRNs.

Data aggregation has been considered as an effective strategy for saving energy and reducing medium access contention in wireless networks. It has been widely investigated in conventional wireless networks [6]-[9]. Due to the scarcity of spectrum opportunity, data aggregation is also considered as having a broad prospective in CRNs. Nevertheless, none of the existing works can be intuitively applied to CRNs for many reasons. First, the communication opportunities in CRNs are not symmetric. Whereas, most of the existing works assume symmetric communication links. Second, PUs in conventional wireless networks only have to consider interference from other PUs. However, an SU has to compete for spectrum resources with both PUs and other SUs in a CRN. Therefore, investigation on distributed data aggregation in CRNs is necessary and meaningful.

In this paper, we concentrate on the investigation of data aggregation in CRNs under a general and practical network model. Due to the fact that time sensitivity plays an vital role in many applications, our work mainly focuses on finding a distributed scheduling plan for data aggregation with the objective of minimizing the total transmission delay. Particularly, we consider a CRN consisting of SUs and PUs with multiple available licensed spectrums. PUs can access the licensed spectrums freely. The SUs, however, have to conduct a spectrum sensing process before accessing the spectrums. During a data aggregation process, if a spectrum is available (has not been used by PUs), an SU has to decide “to transmit or to wait” considering both interference and transmission delay.

The contributions of this paper can be summarized as follows:

• We formalize the Minimum Latency Data Aggregation Scheduling (MLDAS) problem in CRNs.
• Subsequently, the MLDAS problem under the Unit Disk Graph (UDG) interference model is studied. A Connected Dominating Set (CDS)-based hierarchy is introduced as the logical routing structure during the aggregation process. Based on the hierarchy, a distributed scheduling algorithm is presented.
• Considering wireless signal’s fading property, the Physical Interference Model (PhIM) has also been discussed in this paper.

The rest of this paper is organized as follows: the network model and problem formalization are presented in Section II. In Section III and Section IV, the MLDAS problem under the UDG and PhIM interference models are investigated, respectively. Finally, the performance of the proposed algorithms are evaluated in Section V, followed by the conclusion which is presented in Section VI.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. Network Model

In this paper, we consider a dense CRN. Assuming the time is slotted with slot length $T$, where $T$ is the time needed for an SU to sense an available spectrum and finish a transmission successfully, but only long enough for a PU to conduct one data transmission.

Primary network. In the primary network, we consider $N$ randomly deployed PUs, labeled as $P_1, P_2, \ldots, P_N$, operating on $K$ orthogonal parallel licensed spectrums, indexed by
\{C_1, C_2, \ldots, C_K \}$. The transmission radius and interference radius of a PU are denoted as $R$ and $R_I$, respectively. In each time slot $t$, a PU is either active or inactive. An active PU either sends or receives data (not both) on one of the $K$ spectrums, and an inactive PU keeps silent (neither sending nor receiving data) for $T$. Let $P_i^j = \{ P_i^j, P_i^2, \ldots, P_i^K \}$ denote the set of PU becoming active, and $P_i^k$ represents the probability that $P_i$ becomes active on channel $C_k$. Particularly, an active PU can only stay on one of the $K$ spectrums at a given time slot.

**Secondary network.** The secondary network under consideration consists of $n$ randomly deployed SUs, labeled as $S_1, S_2, \ldots, S_n$, among which an SU called base station (SUSBS) denoted as $S_0$ wants to collect aggregated information from the network. Each SU is equipped with a single, half-duplex cognitive radio, with transmission radius $r$ and interference radius $r_I$, respectively. Due to the radio limitation, an SU can either send or receive data, not both, from all directions at one time. Particularly, in each time slot $t$, an SU opportunistically accesses one of the licensed spectrums unoccupied by active PUs in a sensing-before-transmission manner. For two SUs $S_i$ and $S_j$, let vector $P_{ij} = \{ p_{ij}^1, p_{ij}^2, \ldots, p_{ij}^K \}$ represent the channel accessing probability of the transmission from sender $S_i$ to receiver $S_j$, where $p_{ij}^k$ denotes the probability that $S_i$ and $S_j$ simultaneously get the accessing chance of channel $C_k$ at particular time $t$. $p_{ij}^k$ is initialized to $(1 - P_i^k) * \prod_{p_{ij} \in \{ P_u || S_u \rightarrow P_i || \leq r \}} (1 - P_u^k)$, which indicates the expected probability that the transmission will not interrupt the primary network.

A *logical link* exists between $S_i$ and $S_j$ iff the Euclidean Distance between $S_i$ and $S_j$ denoted by $|| S_i - S_j || \leq r$. Then, the secondary network can be represented by $G = (S, E_i)$, where $S = \{ S_1, S_2, \ldots, S_n \}$ is the union of all the SUs, and $E_i$ is the set of logical links among all the SUs.

### B. Problem Formulation

Let $T_{ij}^{s,k,t}$ denote a transmission in the secondary network from sender $S_i$ to receiver $S_j$ ($1 \leq i, j \leq n$ and $i \neq j$) on spectrum $k$ at time $t$, and $T_{uv}^{p,k,t}$ ($1 \leq u, v \leq N$ and $u \neq v$) represent a transmission in the primary network from sender $P_u$ to receiver $P_v$ on spectrum $k$ at time $t$.

**Definition 2.1:** *Secondary User - Primary User (SU-PU) collision.* Given $T_{ij}^{s,k,t}$ and $T_{uv}^{p,k,t}$, if $|| S_i - P_v || \leq r_I$ or $|| S_j - P_u || \leq r_I$, it is called an SU-PU collision.

Then, a *physical link* on channel $C_k$ from a sender $S_i$ to a receiver $S_j$ at $t$ exists if the following two conditions are satisfied: a logical link exists between $S_i$ and $S_j$, and a channel $C_k$ is available to both $S_i$ and $S_j$. The unsymmetrical and opportunistic accessing property of physical links make the data aggregation problem in CRNs even more challenging.

**Definition 2.2:** *Secondary User - Secondary User (SU-SU) collision.* Let $S_i$, $S_j$, $S_x$, and $S_y$ denote four different SUs in the secondary network. Given two transmissions $T_{ij}^{s,k,t}$ and $T_{xy}^{s,k,t}$, if $|| S_i - S_y || \leq r_I$ or $|| S_j - S_x || \leq r_I$, then an SU-SU collision occurs.

Based on the defined network model and definitions, the **Minimum Latency Data Aggregation Scheduling (MLDAS)** problem in a CRN can be formalized as follows:

Given a secondary network denoted by $G = (S = \{ S_1, S_2, \ldots, S_n \}, E_i)$, an MLDAS in a CRN can be defined as a set of scheduled links $S = \{ S_1, S_2, \ldots, S_n \}$, where $S_t$ ($1 \leq t \leq L$) contains collision-free transmissions among SUs scheduled at $t$. Furthermore, to be an MLDAS, the following constraints are required:

**1) 1:** $\forall t (1 \leq t \leq L)$, neither SU-PU collision nor SU-SU collision is caused by any transmission $T_{ij}^{s,k,t} \in S_t$.

**2) 2:** $\forall t_1, t_2 (1 \leq t_1, t_2 \leq L$, $t_1 \neq t_2)$, given two transmissions $T_{ij}^{s,k_1,t_1}$ and $T_{xy}^{s,k_2,t_2}$, $i \neq x$.

**3) 3:** $\forall t (1 \leq t \leq L)$, if $\bigcup_{t=1}^{L} \{ S_i | T_{ij}^{s,k,t} \in S_t \} = \{ S_1, S_2, \ldots, S_n \} - S_b$.

**4) 4:** All data has been aggregated to $S_b$ at $L$.

**5) 5:** $\arg \min_{S_b} (S_1, S_2, \ldots, S_n) \| S_b \| L$.

Constraint 1 shows MLDAS should be SU-PU collision-free and SU-SU collision free. Constraint 2 indicates the property of data aggregation, that is, each SU sends its aggregation result only once. The data integrity property is ensured by constraint 3, where all SUs have to transmit its data to an assigned receiver. Finally, the aggregation result of all the data is received by the SU-BS according to constraint 4. Constraint 5 denotes that the objective of the MLDAS $S$ is to minimize the total transmission latency.

The MLDAS problem in CRNs is NP-hard. This is based on the fact that the minimum latency data aggregation scheduling problem is NP-hard in WSNs [7][10], which can be considered as a special case of MLDAS in CRNs when the number of PUs is fixed to be zero, and the number of channels is $K = 1$.

### III. DISTRIBUTED DATA AGGREGATION SCHEDULING UNDER THE UDG MODEL

**A. UDG Interference Model**

The Unit Disk Graph (UDG) interference model has been widely used in the existing literatures. Under this model, the interference range and transmission range of wireless devices are denoted by equally likely disks. That is, $R = R_I$ and $r = r_I$.

**B. DA Hierarchy Construction**

Given a graph $G = \{ V, E \}$, where $V$ is the set of vertices and $E$ is the set of edges. A Dominating Set (DS) is defined as a subset $V_D \subseteq V$, such that $\forall v \in V$, either $v$ is in $V_D$ or one of $v$’s neighbors is in $V_D$. Let $N(v)$ represent the set of $v$’s neighbors in $G (\forall v \in V \rightarrow (v, N(v)) \in E)$. That is, if $V_D$ is a DS, then, $\forall v \in V \rightarrow ((v \in V_D) \vee (\exists u (u \in N(v) \wedge u \in V_D)))$.

The *Connected Dominating Set* (CDS)-based aggregation tree has been used in time sensitive data aggregation of both conventional wireless networks [7] and CRNs [11]. Constrained by the characteristics of the tree structure, every vertex
only has one parent except the root. It shows its advantage in traditional wireless networks. However, both the unsymmetrical link availability and opportunistic link accessibility make the intuitively employment of a CDS-based data aggregation tree in CRNs impractical. In our proposed hierarchy, instead of relying on only one parent for transmission, an aggregation set will be considered.

Given a secondary network denoted as \( G = (V = \{S_1, S_2, \ldots, S_n\}, E) \), among which \( S_i \) is the user who wants to get aggregated information from the network, the CDS-based data aggregation hierarchy whose construction is based on the distributed CDS construction algorithm presented in [13], can be concluded into four steps.

Step 1: a DS set \( V_D \) colored in black is found based on node’s ranking (level and node ID) on Breadth First Searching Tree in a distributed way [13], and the other nodes in \( V - V_D \) are colored in gray.

Step 2: let \( r_T \) represent the neighbor of \( S_b \) who has the largest number of black neighbors. A dominating tree \( T^* \) is generated rooted at \( r_T \) by adding black nodes and gray nodes alternatively. Then, the internal SUs of \( T^* \) denoted as \( V_D \cup V_C \) form a CDS [13].

Step 3: external nodes of \( T^* \) color themselves white by checking whether their childrenList is empty. Subsequently, a message is exchanged between \( S_b \) and \( r_T \), so that \( S_b \) becomes the root of \( T^* \), and \( r_T \) becomes a child of \( S_b \). After that, each node updates its level information according to the modified \( T^* \), where \( S_b\.level \) is initialized to 0.

Step 4: each SU updates its neighbors’ color and level information according to its neighborList. Subsequently, the aggregation set and childrenList of an SU is updated according to the following policies:

- A white SU sets its childrenList to be empty and adds all the black SUs with no greater level in its neighborhood to its \( A \), where \( A \) is the aggregation Set which contains a set of potential parents for an SU.

- A black SU adds all the white SU in its neighborList with no smaller level and gray SUs with greater level to its childrenList and updates its \( A \) with all gray SUs having smaller level in its neighborList.

- A gray SU gets its new childrenList by adding all black SUs with greater level on its neighborList and generates its \( A \) by including all black SUs with smaller level in its neighborList.

Let \( \theta \) denote the maximum number of points in a disk of radius \( R \) whose mutual distance is greater than one. Then, we have \( \theta \leq \frac{\pi}{\sqrt{3}}R^2 + \pi R \) +1 [14]. Based on the selection of \( A \), let \( S_i \) and \( S_j \) denote different SUs in the secondary network. The following lemma holds.

\begin{lemma}
\begin{enumerate}
\item \( \forall S_i \in V_D - \{S_b\}, \exists S_j ((S_j \in V_C) \land (S_j \in S_i\.childrenList)) \), then \( 1 \leq |A(S_i)| \leq 20 \), otherwise, \( 1 \leq |A(S_i)| \leq 19 \), where \( |A(S_i)| \) denotes the cardinality of \( A(S_i) \). Intuitively, \( |A(S_i)| = 0 \).
\item \( \forall S_i \in V_C, 1 \leq |A(S_i)| \leq 4 \). Particularly, \( |A(S_i)| = 1 \) if \( S_i \in S_b\.childrenList \).
\item \( \forall S_i \in V_E \) where \( V_E = V - \{V_D \cup V_C\} \) is set of leaves in the tree, \( 1 \leq |A(S)| \leq 5 \).
\item Let \( L_{max} = \{S_i\.level \mid \forall S_j \in S_i\.childrenList \}, \) then \( L_{max} \leq 2D \).\( D \) is the diameter of the network [14].
\end{enumerate}
\end{lemma}

The CDS-based DA hierarchy (as shown in Fig. 1(c), where the number and its subscript in the circle represent the ID of an SU and its level, respectively.) is used as the logical routing structure in our data aggregation process.

\section{UDSA Scheduling}

\begin{definition}
SU-PU Collision Set (SPCS). Given a transmission \( T^{(s,k,t)} \) its SPCS is defined as \( \text{SPC}(T^{(s,k,t)}_i) = \{P_u(\{T^{(p,k,t)}_j \land (||S_i - P_u|| \leq r_j)\} \lor (\{T^{(p,k,t)}_u \land (||S_j - P_u|| \leq r_i)\}) \) where the set of active PUs on spectrum \( k \) at time \( t \) that will interfere with or be interfered by \( T^{(s,k,t)}_i \).
\end{definition}

\begin{definition}
SU-SU Collision Set (SSCS). Given a transmission \( T^{(s,k,t)}_j \) its SSCS is defined as \( \text{SSC}(T^{(s,k,t)}_j) = \{S_x (T^{(s,k,t)}_y \land (||S_x - S_y|| \leq r_j) \land (||S_j - S_x|| \leq r_i)) \) contains the set of SUs which will conflict with transmission \( T^{(s,k,t)}_j \) if they are scheduled on spectrum \( k \) at time \( t \). Particularly, according to the data aggregation hierarchy constructed in Section III-B, given a transmission \( T^{(s,k,t)}_i \), where \( S_j \in A(S_i) \), \( \text{SSC}(T^{(s,k,t)}_j) = \{S_x (T^{(s,k,t)}_y \land (|S_x - S_j\.neighborList|) \land (S_x \in N(S_j\.childrenList))) \}
\end{definition}

The distributed scheduling algorithm UDSA is presented in Algorithm 1. In UDSA, each node \( S_i \) first checks if its childrenList is empty, that is, whether all of its children have finished data transmission (Line 3). If \( S_i\.childrenList = 0 \),
Algorithm 1: UDSA Scheduling

begin
input : $G = (S = \{S_1, S_2, ..., S_n\}, E_i), S_0 \in S$
output: schedule $S = \{S_1, S_2, ..., S_n\}$
spec = \emptyset, sprob = 1, flag = 0, $t = 1$
for each $S_i : S_i \in \{S_1, ... , S_n\} - \{S_0\}$ do
if $S_i . childrenList = \emptyset$ then
for $i : 1 \rightarrow K$ do
$S_j = ParentSelection()$
if $C_i . available$ then
if $SPC(T_{ij}^{(s,spec,t)}) = \emptyset$ then
$sprob = p^k_{ij}$
flag = 1;
else
Update $\mathcal{P}$ (Lines 8).
Send FINISH to $A$;
$S_f = S_j \cup T_{ij}^{(s,spec,t)}$;
$S_i . sleep();$
t + +;
else
Wait for FINISH from $S_i . childrenList$;
if Received FINISH from $S_c, S_e \in S_i . childrenList$ then
$S_i . childrenList = S_i . childrenList - S_c$; 
go to 3;
else
if no FINISH from $S_c, S_e \in S_i . childrenList$ for $T$ then
[go to 23;]
end
end
end
end

$S_i$ is ready to send. Otherwise, $S_i$ has to wait and collect “FINISH” messages from $S_i . childrenList$ (Lines 23). When $S_i$ is ready, a spectrum sensing and selection process will be activated (Lines 4-20). During the spectrum sensing process, $S_i$ tries to find transmission opportunity with all the nodes in $A(S_i)$ on all the $K$ spectrums. Then, a $ParentSelection()$ process will get the “best” parent denoted as $S_j$ and best communication spectrum $C_{spec}$ dynamically from $A(S_i)$ (Lines 6-7). The policy for best parent selection is defined as: if more than one node in $A(S_i)$ has the opportunity to receive data, the one who has the fewest children and worst transmission opportunity, that is, shortest $childrenList$ and lowest $p^k_{ij}$, will be chosen. The purpose for choosing the channel which is the most difficult to get is to keep other spectrums which are easier to get for other SUs. This strategy may also reduce potential SU-SU collisions. Subsequently, a further “conflicting free” verification step will be carried out by checking $SPC(T_{ij}^{(s,k,t)})$ and $SSC(T_{ij}^{(s,k,t)})$ (Line 8). If both collision sets are empty, $S_j$ is considered as the candidate parent and channel $C_{spec}$ is the candidate channel. (Lines 8-11). After passing all the condition checking, transmission $T_{ij}^{(s,spec,t)}$ can be processed without conflicting with PUs or SUs (Line 17). Subsequently, $S_i$ broadcasts a “FINISH” message to all the nodes in $A(S_i)$ to inform them no longer have to wait for it, and then $S_i$ goes to sleep since it already has finished its transmission (Line 18-20). Otherwise, $S_i$ has to go back and restart from the spectrum sensing due to the dynamic spectrum opportunity in the CRN (Lines 13-15).

In UDSA, once there is a contention between two transmissions, priority will be given to a transmission according to the following order: sender colored in white, gray then black; larger level; lower channel accessing probability; longer $childrenList$; or larger ID.

IV. MLDAS UNDER THE PHIM

Since wireless signal can be easily affected by the environment and fade during data transmission, the Physical Interference Model (PhIM) has been considered as another important model in CRNs.

A. Physical Interference Model (PhIM)

In the PhIM, the quality of signal received by a receiver is measured by the Signal to Interference and Noise Ratio (SINR) - the ratio of the expected signal strength over the total unwanted (including interference from other users and noise from the environment) signal strength. For simplicity, we ignore the influence of the environment and consider the Signal to Interference Ratio (SIR) instead. Under the PhIM with SIR, a receiver can receive data successfully only if the SIR ratio of the transmission is greater than a threshold. Particularly, according to the network model specified in this paper, the SIR constraints for $P_u$ of transmission $T_{ij}^{(p,k,t)}$ and $S_i$ of transmission $T_{ij}^{(s,k,t)}$ are shown by InEq. 1 and InEq. 2, respectively.

$$\frac{\sum_{\forall P_x \in T_{ij}^{(p,k,t)}} P_{p*x} \parallel P_x P_c \parallel^{-\alpha}}{\sum_{\forall S_x \in T_{ij}^{(s,k,t)}} P_{s*x} \parallel S_x S_j \parallel^{-\alpha} + \sum_{\forall P_v \in T_{ij}^{(p,k,t)}} P_{p*y} \parallel P_x S_j \parallel^{-\alpha}} \geq \tau_p.$$  \hspace{1cm} (1)

$$\frac{\sum_{\forall S_x \in T_{ij}^{(s,k,t)}} P_{s*x} \parallel S_x S_j \parallel^{-\alpha}}{\sum_{\forall P_v \in T_{ij}^{(p,k,t)}} P_{p*x} \parallel P_x S_j \parallel^{-\alpha}} \geq \tau_s.$$  \hspace{1cm} (2)

where constant $\alpha > 2$ is the path loss exponent. $P_p$ and $P_s$ are the transmission power of PUs and SUs, respectively. $T_{ij}^{(p,k,t)}$ (respectively, $T_{ij}^{(s,k,t)}$) is the set of transmissions in the primary network (respectively, secondary network) on channel $k$ at $t$. $\tau_p$ and $\tau_s$ are the respective SIR threshold for a PU and an SU.

According to InEq. 1 and InEq. 2, transmission $T_{ij}^{(s,k,t)}$ can be conducted under the PhIM iff the SIR of $S_i$ satisfies Eq. 2. Meanwhile, $T_{ij}^{(s,k,t)}$ will not ruin the SIR requirement of all the ongoing transmissions in both the primary and secondary networks.
B. PDSA scheduling

In CRNs, the transmissions among PUs have absolute priorities and are not allowed to be interrupted by SUs. Similarly, the ongoing transmissions among SUs should not be interfered either. Furthermore, interference from ongoing transmissions should be limited so that the receiver can get the information correctly. Therefore, an SU with data to send needs to perform a carrier-sensing action before transmitting data. A transmission will be carried out only if the increased signal strength from the transmission will not affect the ongoing transmissions (among PUs or SUs), and the noise from ongoing transmissions is limited. On the other hand, the carrier-sensing range is expected to be as small as possible, so that no transmission opportunity will be missed due to over-estimation. Therefore, a proper carrier sensing range which can help with avoiding interference and making the best use of spectrum opportunity is desired. The Proper Carrier sensing Range (PCR) $R_c$ derived from [11], [12], which satisfies the requirements as proved, is used in this paper. Based on [11] and [12], given an $S_i$ with data to send, the concept defined below specifies the collision set of $S_i$, which are senders that will conflict with $S_i$ if they are scheduled on the same spectrum within the same time slot.

**Definition 4.1: SIR Collision Set (SIRCS).** Given a potential transmission $T_{ij}^{s,k,t}$, its SIRCS is defined as

$$SCS(T_{ij}^{s,k,t}) = \{S_u, S_x\} | (T_{ij}^{p,s,k,t}) \land \{S_i P_u \leq R_c\} \land S_i S_x \leq R_c, \}$$

and $SCS(T_{ij}^{s,k,t})$, in PDSA, $SCS(T_{ij}^{s,k,t})$ will be checked. Furthermore, since a spectrum sensing will be conducted first, only interference coming from SUs needs to be verified when a spectrum is available. Thus, we consider $SCS_s(T_{ij}^{s,k,t}) = \{S_x | T_{ij}^{s,k,t} \land S_i S_x \leq R_c\}$ instead of $SCS(T_{ij}^{s,k,t})$. Then we can get the PDSA by replacing line 8 of Algorithm 1 by $SCS_s(T_{ij}^{s,k,t}) = \emptyset$.

V. PERFORMANCE EVALUATION

A. Performance of UDSA

We investigate the performance of UDSA under the single spectrum and multiple spectrums scenarios, separately.

1) UDSA vs. CS with Single Spectrum: We evaluate the performance of UDSA by comparing it with CS [10] which is the most state-of-the-art convergecast scheduling algorithm for CRNs. In CS, the area is divided into different cells. Furthermore, SUs in different cells are classified into different levels and concurrent sets. In order to avoid interference, SUs in CS are scheduled from higher level to lower level, and in each level, SUs are scheduled from concurrent set with smaller ID to larger ID. In the simulation, we consider a CRN coexisting with a primary network in a region of 100m × 100m. A primary user is assumed to be active with probability $p_a$ or keep silent with probability $1 - p_a$ at each time slot. Since only a single spectrum is considered in CS, to be fair, we perform the performance evaluation with only one available spectrum as shown in Fig. 2, followed by the multi-spectrums performance examination as shown in Fig. 3. According to the results, UDSA has a better time performance than CS in all the senarios. Particularly, the multi-spectrums shows its good performance on data aggregation in CRNs.

We evaluate the impact of PUs’ activity on UDSA and CS by varying $p_a$ from 0.1 to 0.5, $N = 400$, $N = 80$ and $R = r = 20m$. The results are shown in Fig. 2(a). With the increase of PUs’ activities, the scheduling time needed for both the algorithms goes up in order to avoid disturbing data transmissions in the primary network. However, the amount of time needed for UDSA is 58.8% less compared with CS. That is because the concurrent sets are scheduled in a fixed order in CS. Therefore, the more active the PUs, the longer the time needed for SUs in later concurrent sets to wait for the scheduling of SUs in the former concurrent sets.

We study the performance of UDSA and CS with respect to the interference radius of PUs. In this case, the parameters are $N = 80$, $N = 400$, $r = 20m$, $p_a = 0.3$ and the radius of a PU varies from 12m to 25m. From Fig. 2(b), we can see that the delay of both the algorithms decreases with the
increasing of the interference radius of PUs. The reason is that the increasing interference radius of active PUs causes interference within a larger area. However, even though the reduced concurrency leads to the increasing of latency, UDSA outperforms CS by 57%.

2) Performance of UDSA with Multiple Spectrums: We examine the influence of PUs and SUs on the multi-spectrum UDSA as shown in Fig. 3(a) and Fig. 3(b), respectively. We assume the network is within an area of 100m * 100m, $R = r = 20m$ and $p_a = 0.3$ in both scenarios. In Fig. 3(a), $n = 800$ while $N$ varies from 150 to 300. In Fig. 3(b), $N$ is fixed as 120 while $n$ changes from 1000 to 3000. The results show that the increasing of the number of SUs and PUs results in longer latency. Particularly, SU-PU collisions have heavier negative influence compared with SU-SU collisions.

B. Performance of PDSA

In this subsection, we investigate the performance of PDSA under different network scenarios. For simplicity, the parameters are consistent with Section IV. To be specific, in the following scenarios, we consider the CRN and the primary network in an area of 100 * 100m.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>3.8</th>
<th>3.9</th>
<th>4.0</th>
<th>4.1</th>
<th>4.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>84715</td>
<td>47576</td>
<td>25820</td>
<td>11900</td>
<td>3356</td>
</tr>
<tr>
<td>$p_{active}$</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
</tr>
<tr>
<td>Time</td>
<td>446</td>
<td>2362</td>
<td>26168</td>
<td>340034</td>
<td>5776267</td>
</tr>
</tbody>
</table>

The investigation of how $p_a$ and $\alpha$ influence the latency of PDSA is shown in Table I. In this scenario, we adjust $\alpha$ from 3.8 to 4.2 and $p_a$ from 0.1 to 0.5. The other parameters are $N = 400$, $n = 80$, $P_p = P_s = 10$, and $\tau_p = \tau_s = 8$. The results demonstrate that with the augmentation of $\alpha$, the induced transmission latency is decreased. That is because the interference of a sender to other ongoing transmissions in the CRN decreases with the increase of path loss ratio. On the other hand, the delay gets worse with the growth of PUs’ activities because more time is needed for SUs to get transmission chance.

VI. CONCLUSION

In this paper, we investigate the minimum latency data aggregation problem in CRNs. Two practical distributed algorithms under the Unit Disk Graph interference model and the Physical Interference Model are proposed, respectively. Extensive simulation results indicate that the proposed algorithms have a superior performance in various network scenarios.

ACKNOWLEDGMENT

This work is partly supported by the NSF under grants No. CNS-1152001 and No. CNS-1252292.

REFERENCES