Data Aggregation Scheduling in Duty-Cycled Multihop Wireless Networks Subject to Physical Interference

Lixin Wang†, Hanyu Liang‡ and Peng-Jun Wan*

***Abstract*—Minimum-Latency Aggregation Scheduling (MLAS) has been well studied when all the networking nodes are always active. However, it is well-known that the nodes often switch between the active state and the sleep state to save energy. A node in duty-cycled scenarios with active/sleep cycles may require transmitting multiple times to send the message to all of its neighbors due to their different active times. MLAS in multihop wireless networks with Duty-Cycled scenarios (MLASDC) has also been well-studied under graph-based interference models such as the protocol interference model. To the best of our knowledge, no approximation algorithms have been proposed for MLASDC subject to physical interference. This is the first paper to develop efficient approximation algorithms for MLASDC subject to physical interference. The data aggregation schedule produced by our algorithm proposed in this paper achieves an approximation ratio at most a constant time of the length of a scheduling period if the maximum degree Δ of the network is bounded.**

***Index Terms*—Data aggregation schedule, duty-cycled scenarios, multihop wireless networks, physical interference model.**

I. Introduction

Data aggregation is a primitive and essential communication task in which a distinguished sink node collects the data aggregated from all the packets at the nodes other than the sink node according to some aggregation functions such as logical and/or, maximum, or minimum. Data aggregation is a vital technique and widely used in various applications of multihop wireless ad hoc and sensor networks (WAHSNs). An aggregation schedule not only specifies a spanning in-arborescence of the network topology for routing, but also is a link schedule of all the links in such a spanning in-arborescence [12]. Assume that all the communications proceed in synchronous time-slots and each node can transmit at most one packet of a fixed size in each time-slot. The problem of computing a data aggregation schedule with minimum latency in multihop WAHSNs is referred to as **M**inimum-**L**atency **A**ggregation **S**chedule (**MLAS**) which has been well-studied when the nodes are always active [1][2][6][12][13].

However, it is well-known that the nodes often switch between the active state and the sleep state to save energy. The duty-cycled scenarios have been emerging as a prevalent energy-saving method in multihop WAHSNs and actually implemented in many important applications such as the GreenOrbs WSN projects [4]. A node in duty-cycled scenarios with active/sleep cycles may require transmitting multiple times to send the packet to all of its neighbors due to their different active time. Therefore, all the previous known scheduling algorithms for MLAS without duty-cycled scenarios are no longer suitable for duty-cycled multihop WAHSNs.

With duty-cycled scenarios, we assume that the networking nodes determine the active/sleep time without coordination in advance. In this paper, we adopt the following popular and realistic duty cycle model (as most research papers in this area do): the whole scheduling time is divided into multiple scheduling periods of the same length. A scheduling period T is further divided into fixed $|T|$ time slots, i.e., $T = \{0, 1, \cdots, |T| - 1\}$. *Furthermore, every node $v \in V$ randomly and independently chooses exactly one time slot in T to be active, and wakes up at this time slot in every scheduling period to receive the message*. *If a node v needs to send a message as required, it can wake up at any time slot to transmit the message as long as the receiver node is awake and there is no collision for this transmission*. MLAS in multihop Duty-Cycled WAHSNs (referred to as MLASDC) has been well-studied under the graph-based interference models such as the protocol interference model [3][7][8][9][10].

The *physical interference model* offers a more realistic representation of wireless communication. Under the physical interference model, a transmission is successful if and only if the *SINR* (*Signal-to-Interference-plus-Noise-Ratio*) at the

†TSYS School of Computer Science, Columbus State University, GA 31907. Emails: wang_lixin@columbusstate.edu.
‡School of Computer Science and Engineering, University of New South Wales, New South Wales, Australia. E-mail: hanyuhlv@gmail.com.
*Department of Computer Science, Illinois Institute of Technology, Chicago, Illinois, USA. E-mail: wan@cs.iit.edu.

978-1-7281-5212-7/19/$31.00 ©2019 IEEE
DOI 10.1109/MSN48538.2019.00041

intended receiver exceeds a threshold so that the transmitted signal can be decoded with an acceptable bit error probability. Since the SINR ratio depends on which transmissions are being scheduled concurrently in each time slot, it is impossible to build a conflict graph under the SINR model. This makes the analysis of scheduling algorithms under the SINR model much more challenging than in graph-based interference models. We assume all the nodes transmit at a fixed uniform power P. The path loss is then determined by a positive reference loss parameter η, and the path-loss exponent κ, which is a constant greater than 2 but less than 6 typically. Specifically, when a node u transmits a signal at power P, the power of this signal captured by another node v is $\eta P \|uv\|^{-\kappa}$, where $\|uv\|$ denotes the scaled Euclidean distance between u and v so that the beaconing radius is normalized to one. Let ξ denote the background noise, σ the threshold of the SINR in order to correctly interpret the wanted signal.

To the best of our knowledge, this is the first paper that develops efficient approximation algorithms for MLASDC under the SINR model. The data aggregation schedule produced by the scheduling algorithm proposed in this paper achieves approximation ratio at most a constant multiple of $|T|$ if the maximum degree Δ of the network is bounded.

The remaining of this paper is organized as follows. In Section II, we give a literature review for known data aggregation scheduling algorithms. In Section III, we present sufficient conditions for a set of transmissions to succeed when they occur simultaneously under the SINR model. A spanning tree rooted at the sink node s is constructed in Section IV. A short data aggregation schedule for MLASDC is developed in V. Finally, we conclude our paper and discuss some future research directions in Section VI.

II. Related Works

MLAS in multihop WAHSNs has been extensively studied in the literature and most of the known algorithms implicitly assumed that all the networking nodes are always active [1][2][6][12][13]. For a multihop wireless network with all planar nodes having uniform transmission radii equal to one, its communication topology is a unit-disk graph (UDG). Let s denote the sink node of the data aggregation, n the number of nodes, and R the graph radius of the communication topology with respect to s. Both R and $\log n$ are two lower bounds for any optimal aggregation schedule for MLAS even when ρ is arbitrary. For $\rho = 1$, [1] and [6] proposed respectively two data aggregation schedules of total latency at most $(\Delta - 1)R$ and $23R + \Delta - 18$, where Δ is the maximum degree. [2] first studied the single-level aggregation and proposed an energy-efficient protocol for aggregator selection (EPAS) protocol. Then, this paper generalized it to an aggregation hierarchy and extend EPAS to hierarchical EPAS. The optimal number of aggregators with generalized compression and power–consumption models was derived, and distributed algorithms for aggregator selection were proposed in [2]. For $\rho = 1$, Wan et al. [12] developed three approximation algorithms which produced aggregations schedules of total latency at most $15R + \Delta - 4$, $2R + \Delta + O(\log R)$ and $\left(1 + O\left(\log R / \sqrt[3]{R}\right)\right) R + \Delta$, respectively. All of these known algorithms for MLAS were developed under the protocol interference model with the assumptions that all the nodes are always active.

Under the SINR model with the assumptions that all the nodes are always active, the best known scheduling algorithms for MLAS was proposed by Wan et al. in [13]. This paper developed an approximation algorithm to produce a short data aggregation schedule for MLAS on multihop WAHSNs.

However, none of the known works described above has taken the duty-cycled scenarios into consideration. With the duty-cycled scenarios, MLASDC has been well-studied under graph-based interference models [3][7][8][9][10]. [10] proposed a data aggregation routing and duty cycle scheduling heuristics for energy efficiency and communication latency for MLASDC. [11] developed a mixed integer nonlinear mathematical formulation of duty cycle scheduling with data-aggregation routing for MLASDC with adjustable radii of the nodes. [3] proposed a load-balanced and latency-efficient data aggregation scheduling for MLASDC. [9] proposed a lifetime balanced data aggregation scheme (LBA) for duty-cycle sensor networks. [7] investigated the minimum-lentancy data aggregation scheduling in multi-channel duty-cycled WSNs. The paper proposed an efficient data aggregation scheduling algorithm which exploits the fewest-children-first rule to choose the forwarding nodes to benefit the link scheduling. [8] proposed a distributed delay-efficient scheme to solve the MLASDC problem in duty-cycled WSNs, but all performance analysis of the data aggregation scheduling algorithm proposed in [8] was only based on simulation studies. All these known algorithms for MLASDC were proposed under the protocol interference model.

III. Independent Sets of Nodes and Links Under the Physical Interference Model

In this section, we introduce the independent sets of nodes and links presented in [13], and the lemmas for the sufficient conditions of a set of transmissions to succeed when they occur simultaneously subject to physical interference. These lemmas have been proved in [13] and will be used in this paper. Let

$$R = \left(\frac{\eta P}{\sigma \xi}\right)^{1/\kappa}.$$

Then in the absence of interference, for any pair of distinct nodes u and v, they can directly communicate with each other if and only if $\|uv\| \leq R$. Therefore, R is referred to as the *maximum transmission radius* of the network.

Given a real number $r > 0$. Let R' be the maximum edge length of a Euclidean minimum spanning tree of V. That is, R' is the smallest value of r such that the r-$disk$ graph on V is connected. In order to ensure the connectivity of the network, it is required that $R' \leq R$ for the SINR model. In this paper, we assume $R' \leq (1-\varepsilon)R$ for some very small constant $\varepsilon > 0$.

Let Γ denote the set of all mutual distances of the nodes in V at least R' but less than R, i.e.,

$$\Gamma = \{\|uv\| : u, v \in V \ and \ R' \leq \|uv\| < R\}.$$

Then it is easy to see that

$$1 \leq |\Gamma| \leq n(n-1)/2, \ where \ n = |V|.$$

A set I of nodes in V is said to be *r-independent* under the physical interference model if

(1) the mutual distances of the nodes in I are greater than r, and (2) when all the nodes in I transmit simultaneously, the transmission by each node $u \in I$ can be successfully received by all the nodes within a distance of r from u.

Next we present sufficient conditions for a set of transmissions to succeed when they occur simultaneously. Let

$$\zeta(x) = \sum_{j=1}^{\infty} \frac{1}{j^x}.$$

For each $r \in [R', R)$, define

$$\rho = 1 + \left(\frac{\sigma\left(16\zeta(\kappa-1) + 8\zeta(\kappa) - 6\right)}{1 - (r/R)^{\kappa}}\right)^{1/\kappa}. \quad (1)$$

The following lemma, proved in [13], gives a sufficient condition for a set of nodes to be r-independent.

Lemma 1: Given an $r \in [R', R)$. A set I of nodes is *r-independent* under the physical interference model if their mutual distances are all greater than ρr, where ρ is defined by EQ(1).

A set B of links is said to be *independent* under the physical interference model if

(1) all the links in B are disjoint, and (2) when all the transmitting ends of the links in B transmit simultaneously, the receiving end of each link $a \in B$ can successfully receive the message from the transmitting end of the link a.

The following lemma, also proved in [13], provides a sufficient condition for a set of links to be independent under the physical interference model.

Lemma 2: Given an $r \in [R', R)$. Suppose that B is a set of disjoint links whose lengths are at most r. If all the receiving ends of the links in B have mutual distances greater than ρr, the set B of links is independent under the physical interference model.

IV. A Spanning Tree for Routing

Given a real number $r > 0$. Let G_r be a connected r-disk graph on V. In this section, we construct a spanning tree T_{span} rooted at the node $s \in V$ that will be used for the routing of data aggregation. Given any node $v \in V$, let $A(v) \in T$ represent its active time slot. For every edge (u, v) of G_r, we define its latency as follows:

$$Latency(u,v) = \begin{cases} A(v) + 1, & if \ u = s; \\ A(v) - A(u), & if \ u \neq s \ and \\ & A(v) - A(u) > 0; \\ A(v) - A(u) + |T|, & otherwise. \end{cases} \quad (2)$$

We consider the latency of each edge as the edge weight and apply the Dijkstra's algorithm on G_r to compute the shortest path tree T_{SPT} rooted at s. Note that total latency of the shortest paths in T_{SPT} from the node s to the other nodes is the minimum latency the other nodes can receive the message from s without collision. Let M denote the maximum latency of all these shortest paths in T_{SPT} from the node s.The shortest path tree T_{SPT} cannot be directly used for routing of data aggregation because of the possible collision among the transmissions.

Next, we construct the spanning tree T_{span} rooted at s that will be used for routing of data aggregation. We divide the nodes of V into different layers $L_0, L_1, L_2, \cdots, L_M$ based on the latency of the shortest paths from s to all other nodes. That is,

- $L_0 = \{s\}$, and
- for each $0 < i \leq M$, L_i is the subset of the nodes in V such that the latency of the shortest path in T_{SPT} from s to these nodes is equal to i.

Note that every node $v \in V$ can only receive a message at its active time-slot $A(v)$. Based on EQ(2), the latency of the shortest path from node s to node v is of the form

$$k(v)\,|T| + A(v) + 1,$$

where $k(v)$ is a non-negative integer depending on v. It is easy to see that for any integer $i \geq 1$, the nodes in the layer

L_i are active at the time-slot $j \equiv i - 1 \bmod |T|$.

For each time slot $0 \le j \le |T| - 1$, let U_j denote the set of all nodes in V that are active in the time-slot j. Let

$$U_j = \bigcup_{k=0}^{K} L_{k|T|+j+1},$$

where K is some non-negative integer satisfying $K|T| + j + 1 \le M$.

Clearly, all possible collisions only occur at the nodes that are active at the current time-slot. For each time slot $j \in T$, we first construct a maximal independent set (MIS) I_j for the nodes in U_j to be used in the construction of T_{span}. The MIS I_j is constructed layer by layer in the top-down manner as follows:

- Initially, $I_j = \varnothing$.
- Sort all the nodes in the layer L_{j+1} in the increasing order of their IDs if it is not empty, say $L_{j+1} = \langle v_1, v_2, \cdots, v_{|L_{j+1}|} \rangle$, and add the first node v_1 to I_j.
- Then add node v_2 to I_j if $I_j \cup \{v_2\}$ is still an independent set. Otherwise, skip v_2 and then consider v_3.
- Repeat the same process for v_3 until the last node in L_{j+1} is either added to I_j or skipped.
- Then repeat the process for all the nodes in $L_{k|T|+j+1}$ with $k = 1, 2, \cdots, K$ until the last node in $L_{K|T|+j+1}$ is processed.

Clearly, $I_j \subset U_j$ is an independent subset. Note that for every layer L_i with $0 < i \le M$, all nodes in L_i are active at the time-slot $j \equiv i - 1 \bmod |T|$. Let $I_{ji} = I_j \cap L_i$. Based on the construction of the MIS I_j of U_j, for each layer $0 < i \le M$ with $j \equiv i - 1 \bmod |T|$, every node $v \in L_i \setminus I_{ji}$ has at least one neighbor in $I_{j1} \cup I_{j2} \cup \cdots \cup I_{ji}$.

Now, we construct T_{span} by specifying the parent of each $v \in V$ other than the root s. T_{span} is also constructed layer by layer in the top-down manner. In each layer L_i with $0 < i \le M$ and $j \equiv i - 1 \bmod |T|$, for any $v \in L_i$,

- if $v \in I_{ji} = I_j \cap L_i$, its parent in T_{span} is the same as its parent in T_{SPT}.
- if $v \in L_i \setminus I_{ji}$, its parent node in T_{span} is chosen to be one of its neighbors in $I_{j1} \cup I_{j2} \cup \cdots \cup I_{ji}$. If v has more than one such neighbors, the one in the smallest layer in T_{SPT} will be chosen to be v's parent in T_{span}.

Clearly, T_{span} is a spanning tree over V. It will be used for routing of data aggregation. Next we introduce the first-fit distance-d coloring of a finite planar set of nodes in V in the lexicographic order that was presented in [13].

Let U be a finite subset of V. In the lexicographic order of U, all nodes in U are sorted from the left to the right with ties broken by the ordering from the bottom to the top. Suppose that the sequence $< u_1, u_2, \cdots, u_k >$ is the lexicographic order of the nodes in U. *The first-fit distance-d coloring of U* in this ordering uses colors represented by positive integers and runs as follows:

- Assign the color 1 to u_1 in the sequence;
- For $i = 2$ up to k, assign to u_i the smallest possible color not used by any node u_j with $j < i$ and $\|u_i u_j\| \le d$.

The following lemma, proved in [13], gives an upper bound on the number of colors used by the first-fit distance-ρr coloring of any independent set of G_r in the lexicographic ordering.

Lemma 3: Let U be any independent set of G_r. Then, the first-fit distance-ρr coloring of U in the lexicographic ordering uses at most β_ρ colors, where

$$\beta_\rho = \left\lfloor \frac{\pi \rho^2}{\sqrt{3}} + \left(\frac{\pi}{2} + 1 \right) \rho \right\rfloor + 1. \quad (3)$$

Note that β_ρ is actually an upper bound of the total number of points contained in a half-disk of radius ρ with mutual distances at least one (see [13]).

V. Data aggregation Schedule with Duty-cycled Scenarios

Let s be the sink node of the data aggregation. Given a fixed $r \in [R', R)$, compute the value of ρ defined in EQ(1). We first construct the spanning tree T_{span} rooted at s on the graph G_r. The routing of data aggregation is the inward spanning s-aborescence oriented from T_{span}. The transmissions are scheduled from the children nodes to their parents in T_{span} layer by layer *based on the receiving nodes* because a node can only receive the aggregated message when it is active. Any node in a layer cannot be scheduled to transmit until it has received all the messages from its children. At each layer L_i with $0 < i \le M$, the nodes not in I_j will have the priority to receive the aggregated message because some nodes in I_{ji} may be the parents of the nodes at the same layer. Thus, the nodes in layers $L_{M-1},\ L_{M-2}, \cdots,\ L_0$ will receive the aggregated message sequentially in the decreasing order of the layer index. Note that only the children nodes of the nodes in U_j in T_{span} will be scheduled to transmit during the time-slot j.

Now we are ready to describe the first-fit algorithm to compute a short data aggregation schedule for MLASDC under the SINR model as follows:

The transmissions are scheduled in the bottom-up manner and the nodes in layer L_M transmit first. At layer L_M, the nodes not in I_M transmit first, and then the nodes in I_M transmit. The data aggregation completes when the sink node

s receives the aggregated message from all other nodes. For each $i = M$ down to 1, perform the following two operations at layer L_i:

(1) If $L_i \setminus I_{ji} \neq \varnothing$, let $P_i = Parent(L_i \setminus I_{ji})$. Then all the nodes in P_i are independent based on the construction of T_{span}. We partition P_i into ϕ subsets $P_i', P_i'', \cdots, P_i^{(\phi)}$ such that for each $1 \leq t \leq \phi$, every node in $P_i^{(t)}$ has at most one child. For each $1 \leq t \leq \phi$, compute a first-fit distance-ρr coloring of $P_i^{(t)}$ in the lexicographic order. Assume $P_i^{(t)}$ is partitioned into m r-independent sets $I_1', I_2', \cdots, I_m'$ under the SINR model for some $m > 0$. For each of the r-independent sets I_c' with $1 \leq c \leq m$, the child node of a node in I_c' is scheduled to transmit at the j-th time-slot if such a child node exists. Since each node in T_{span} except for s has at least one parent node, we have $\phi \leq \Delta - 1$, where Δ is the maximum degree of G_r.

(2) If $I_{ji} = I_j \cap L_i \neq \varnothing$, since I_{ji} is independent, we partition the set I_{ji} into ϕ subsets $I_{ji}', I_{ji}'', \cdots, I_{ji}^{(\phi)}$ such that for each $1 \leq t \leq \phi$, the parent nodes of the nodes in $I_{ji}^{(t)}$ are pairwise distinct. Then for each $1 \leq t \leq \phi$, compute a first-fit distance-ρr coloring of $I_{ji}^{(t)}$ in the lexicographic order. Assume $I_{ji}^{(t)}$ is partitioned into x r-independent sets $I_1'', I_2'', \cdots, I_x''$ under the SINR model for some $x > 0$. For each of the r-independent sets I_c'' with $1 \leq c \leq x$, the nodes in I_c'' are scheduled to transmit at the j-th time-slot. Similarly, we have $\phi \leq \Delta - 1$.

Finally, for each $r \in \Gamma$, we compute a data aggregation schedule using the algorithm above. We choose the one with minimum latency among all these data aggregation schedules as Γ is a finite set with $|\Gamma| \leq n(n-1)/2$.

The next theorem asserts the correctness of the above algorithm and establishes an upper bound on the latency of the data aggregation schedule produced by this algorithm.

Theorem 4: The scheduling algorithm described above is correct and outputs a collision-free data aggregation schedule. The total latency of the data aggregation schedule output by this algorithm is at most $2\beta_\rho |T| (\Delta - 1)M$. The scheduling algorithm achieves approximation ratio at most $2\beta_\rho |T| (\Delta - 1)$.

Proof: First we prove the correctness of the algorithm. It is sufficient to prove that for each layer i from M down to 1, all the nodes in the layer L_i have successfully completed their transmissions.

(1) At the bottom-most layer L_M, let $j_M \equiv M - 1 \mod |T|$. Note that all the node in L_M are active in the time-slot j_M. Based on the above scheduling algorithm for data aggregation, we consider the following two steps for the nodes in the bottom layer L_M:

(a) If $L_M \setminus I_{j_M,M} \neq \varnothing$, let $P_M = Parent(L_M \setminus I_{j_M,M})$. Then all the parent nodes in P_M are independent based on the construction of the tree T_{span}. We partition the set P_M into ϕ subsets $P_M', P_M'', \cdots, P_M^{(\phi)}$ such that for each $1 \leq t \leq \phi$, every parent node in $P_M^{(t)}$ has at most one child. For each $1 \leq t \leq \phi$, compute a first-fit distance-ρr coloring of $P_M^{(t)}$ in the lexicographic order. Assume $P_M^{(t)}$ is partitioned into m r-independent sets $I_1', I_2', \cdots, I_m'$ under the physical interference model for some integer $m > 0$. For each of the r-independent sets I_c' with $1 \leq c \leq m$, the child node of every node in I_c' is scheduled to transmit at the j-th time-slot of a scheduling period if such a child node exists. By Lemma 2, all the scheduled transmissions are successful. By Lemma 3, $m \leq \beta_\rho$.

(b) If $I_{j_M,M} \neq \varnothing$, we partition the set $I_{j_M,M}$ into ϕ subsets $I_{j_M,M}', I_{j_M,M}'', \cdots, I_{j_M,M}^{(\phi)}$ such that for each $1 \leq t \leq \phi$, the parent nodes of the nodes in $I_{j_M,M}^{(t)}$ are pairwise distinct. Then for each $1 \leq t \leq \phi$, compute a first-fit distance-ρr coloring of $I_{j_M,M}^{(t)}$ in the lexicographic order. Assume $I_{j_M,M}^{(t)}$ is partitioned into x r-independent sets $I_1'', I_2'', \cdots, I_x''$ under the physical interference model for some integer $x > 0$. For each of the r-independent sets I_c'' with $1 \leq c \leq x$, the nodes in I_c'' are scheduled to transmit at the j-th time-slot. Similarly, all the scheduled transmissions are successful and $x \leq \beta_\rho$.

Therefore, at this point, all nodes in layer L_M have completed their transmissions and all nodes in layer L_{M-1} are ready to transmit.

(2) For each layer $i = M - 1$ down to 1, we perform the following two steps for the nodes in the layer L_i

(a) If $L_i \setminus I_{ji} \neq \varnothing$, let $P_i = Parent(L_i \setminus I_{ji})$. We partition the set P_i into ϕ subsets $P_i', P_i'', \cdots, P_i^{(\phi)}$ such that for each $1 \leq t \leq \phi$, every parent node in $P_i^{(t)}$ has at most one child. For each $1 \leq t \leq \phi$, compute a first-fit distance-ρr coloring of $P_i^{(t)}$ in the lexicographic order. Assume $P_i^{(t)}$ is partitioned into m r-independent sets $I_1', I_2', \cdots, I_m'$ under the physical interference model for some integer $m > 0$. For each of the r-independent sets I_c' with $1 \leq c \leq m$, the child node of every node in I_c' is scheduled to transmit at the j-th time-slot if such a child node exists. By Lemma 2, all the scheduled transmissions are successful. By Lemma 3, $m \leq \beta_\rho$.

(b) If $I_{ji} \neq \varnothing$, we partition the set I_{ji} into ϕ subsets $I_{ji}', I_{ji}'', \cdots, I_{ji}^{(\phi)}$ such that for each $1 \leq t \leq \phi$, the parent nodes of the nodes in $I_{ji}^{(t)}$ are pairwise distinct. Then for each $1 \leq t \leq \phi$, compute a first-fit distance-ρr coloring of $I_{ji}^{(t)}$ in the lexicographic order. Assume $I_{ji}^{(t)}$ is partitioned into x r-independent sets $I_1'', I_2'', \cdots, I_x''$ under the physical interference model for some integer $x > 0$. For each of the r-independent sets I_c'' with $1 \leq c \leq x$, the nodes in I_c'' are scheduled to transmit at the j-th time-slot. Similarly, all the scheduled transmissions are successful and $x \leq \beta_\rho$.

Therefore, at this point, all nodes in layer L_i have completed their transmissions and all nodes in layer L_{i-1} are ready to transmit.

Upon completion of the step (2) for each layer $i = M - 1$ down to 1, all nodes in the layers $L_M, L_{M-1}, \cdots, L_1$ have completed their transmissions. Thus, the sink node s can successfully receive the aggregated message from all other nodes.

Next we establish an upper bound on the latency of the data aggregation schedule output by this algorithm. For every layer i from M down to 1, it takes at most $2(\Delta - 1)\beta_\rho$ scheduling periods for all nodes in the layer L_i to complete their transmissions. Therefore, the total latency of the entire data aggregation is at most $2\beta_\rho |T| (\Delta - 1) M$.

Since M is a trivial lower bound for MLASDC, the approximation ratio of the data aggregation schedule algorithm is at most $2\beta_\rho(\Delta - 1) |T|$.

This completes the proof of the theorem. ■

Therefore, when the length $|T|$ of the scheduling period T is a constant and the maximum degree Δ is bounded, the above data aggregation scheduling algorithm for MLASDC achieves constant approximation ratio.

VI. Conclusion and Future Research Directions

In this paper, we developed an efficient approximation solution for MLASDC subject to physical interference that achieved approximation ratio at most a constant time of $|T|$ when the maximum degree Δ of the network is bounded. This is the first work that developed efficient approximation algorithm for MLASDC in multihop WAHSNs subject to physical interference. Since the SINR ratio depends on which transmissions are being scheduled concurrently in each time slot, it is impossible to build a conflict graph under the physical interference model. This makes the analysis of scheduling algorithms under the SINR model much more challenging than in graph-based interference models such as the protocol interference model. The data aggregation schedule we developed in this paper are built upon a general technique which enables a unified graph theoretic treatment of the communication scheduling subject to the physical interference constraint.

As for future research directions in this area, the approach we proposed in this paper may be used to develop efficient approximation algorithms for minimum-latency data gathering schedule, data gossiping schedule and other primitive communication tasks in duty-cycled multihop WAHSNs subject to the physical inference constraint.

Acknowledgement: *This work of Dr. Peng-Jun Wan was supported in part by the National Science Foundation of USA under grant CNS-1526638.*

References

[1] X.J. Chen, X.D. Hu, J.M. Zhu: Minimum data aggregation time problem in wireless sensor networks, *Lecture Notes in Computer Science* 3794, pp. 133-142, 2005.

[2] Chen, Y. P., Liestman, A. L., & Liu, J. (2006). A hierarchical energy-efficient framework for data aggregation in wireless sensor networks. *IEEE Transactions on Vehicular Technology*, 55(3), 789-796.

[3] Chen, Z., Yang, G., Chen, L., Xu, J., & Wang, H. (2012, December). A load-balanced data aggregation scheduling for duty-cycled wireless sensor networks. *In Cloud Computing Technology and Science* (CloudCom), 2012 *IEEE 4th International Conference on* (pp. 888-893). IEEE.

[4] The GreenOrb Wireless Sensor Network Projects available at http://www.greenorbs.org/all/projects.htm

[5] K. Han, J. Luo, Y. Liu and A. Vasilakos: Algorithm Design for Data Communications in Duty-Cycled Wireless Sensor Networks: A Survey. *Communications Magazine, IEEE*, 51(7), 2013.

[6] S.C.-H. Huang, P.-J. Wan, C. T. Vu, Y. Li, and F. Yao: Nearly Constant Approximation for Data aggregation Scheduling in Wireless Sensor Networks, *IEEE INFOCOM* 2007.

[7] Jiao, X., Lou, W., Feng, X., Wang, X., Yang, L., & Chen, G. (2018, October). Delay efficient data aggregation scheduling in multi-channel duty-cycled WSNs. In 2018 *IEEE 15th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)* (pp. 326-334). IEEE.

[8] Kang, B., Nguyen, P. K. H., Zalyubovskiy, V., & Choo, H. (2017). A distributed delay-efficient data aggregation scheduling for duty-cycled WSNs. *IEEE Sensors Journal*, 17(11), 3422-3437.

[9] Li, Z., Peng, Y., Qiao, D., & Zhang, W. (2012, March). LBA: Lifetime balanced data aggregation in low duty cycle sensor networks. In 2012 *Proceedings IEEE INFOCOM* (pp. 1844-1852). IEEE.

[10] Lin, Frank Yeong-Sung, and Yean-Fu Wen: Multi-sink data aggregation routing and scheduling with dynamic radii in WSNs. *Communications Letters, IEEE* 10.10 (2006): 692-694.

[11] Wen, Yean-Fu, and Frank Yeong-Sung Lin: Cross-layer duty cycle scheduling with data aggregation routing in wireless sensor networks. *Embedded and Ubiquitous Computing*. Springer Berlin Heidelberg, 894-903 (2006).

[12] P.-J. Wan, C.-H. Huang, L. Wang, Z.-Y. Wan, and X. Jia: Minimum Latency aggregation Scheduling in Multihop Wireless Networks, *ACM MOBIHOC* 2009.

[13] P.-J. Wan, L. Wang and O. Frieder: Fast Group Communication in Multihop Wireless Networks Subject to Physical Interference, *IEEE MASS* 2009.

[14] X. Xu, J. Cao, and P.-J. Wan: Fast Group Communication Scheduling in Duty-Cycled Multihop Wireless Sensor Networks, *WASA* 2012.