Minimizing Aggregation Latency under the Physical Interference Model in Wireless Sensor Networks

Baobing Wang and John S. Baras
Department of Electrical and Computer Engineering and the Institute for Systems Research
University of Maryland, College Park
{briankw, baras}@umd.edu

Abstract*—Wireless Sensor Networks (WSNs) have been widely recognized as a promising technology that can enhance various aspects of today's electric power systems, making them a vital component of the smart grid. Efficient aggregation of data collected by sensors is crucial for a successful WSN-based smart grid application. Existing works on the Minimum Latency Aggregation Scheduling (MLAS) problem in WSNs usually adopt the protocol interference model, which is a tremendous simplification of the physical reality faced in wireless networks. In contrast, the more realistic physical interference model has been proved to have the potential to increase the network capacity. In this paper, we propose a distributed algorithm to minimize the data aggregation latency under the physical interference model, which jointly considers routing, power assignment and transmission scheduling. We theoretically prove that our algorithm solves the MLAS problem correctly and the latency is bounded by $3(K+1)^2(\Delta + \log \frac{\sqrt{2}}{K+1}) + 6K^2 + 4K + 2$, where K is a model-specific constant and Δ is the logarithm of the ratio between the lengths of the longest and shortest links in the network. Simulation results demonstrate that our algorithm can significantly reduce the aggregation latency compared to other schemes under the physical interference model. In networks where n nodes are uniformly distributed, our algorithm achieves an average latency between $O(\log^3 n)$ and $O(\log^4 n)$. We also discuss how to improve the energy efficiency through load-balancing techniques.

I. Introduction

The collaborative and low-cost nature of WSNs brings significant advantages over traditional communication technologies used in today's electric power systems, including rapid deployment, self-organization, flexibility, and aggregated intelligence via parallel processing [1]. Wireless sensor motes can be installed on smart grid equipment and collect the critical parameters for remote system monitoring, smart metering, equipment fault diagnostics, etc. [2]–[5].

To support real-time widespread sensing and reduce energy consumption, efficient aggregation of information collected by sensors is crucial for a successful WSN-based smart grid application, which aims to reduce the latency of data aggregation. The Minimum Latency Aggregation Scheduling (MLAS) problem is defined as follows. Given a wireless sensor network that consists of a set of sensor motes and a base station, assuming each mote has a piece of data to be aggregated and transmitted to the base station, the MLAS problem is to design a transmission schedule for all motes such that all data are received successfully and the total number of timeslots for link transmissions is minimized. Each data packet can be completely transmitted in one timeslot.

Extensive research has been done on the MLAS problem, such as [6]–[9], using both single channel and multiple channels. Most existing works usually adopt the *protocol interference model* that is oversimplified and thus does not accurately reflect wireless interferences in reality. In contrast, the *physical interference model* is more realistic and captures the cumulative interference between links more accurately. Under the physical interference model, a transmission is successful if the signal-to-noise-plus-interference ratio (SINR) at the receiver exceeds a certain hardware-specific threshold. It has been proved to have the potential to increase the network capacity and thus reduce the scheduling latency [10]–[12]. However, due to the challenge of handling the cumulative interference effect, only a few previous works in the literature have studied data aggregation under this model [13]–[16].

In this paper, we propose a distributed algorithm for the MLAS problem under the physical model in WSNs, which jointly considers routing tree construction, power assignment and transmission scheduling. Unlike [13] and [14] that adopt the *uniform power assignment* scheme (all nodes transmit with the same transmission power), we consider the *linear power assignment* scheme, which assigns the power level of each sender proportionally to the link's path-loss factor and can further increase the network capacity.

The performance of our algorithm is both analyzed theoretically and studied experimentally. We prove that our algorithm can solve the MLAS problem correctly and achieves a latency bound of $3(K + 1)^2(\Delta + \log \frac{\sqrt{2}}{K+1}) + 6K^2 + 4K + 2$. Here, K is a model-specific constant depending on the SINR threshold and the path-loss exponent, and Δ is the logarithm of the ratio between the lengths of the longest and shortest links in the network. Numerical results demonstrate that our algorithm can significantly reduce the aggregation latency compared to the algorithm proposed in [15] that achieves the same asymptotic bound of $O(\Delta)$, but with a much larger hidden constant. In networks where n sensor motes are uniformly distributed, our algorithm achieves an average latency between $O(\log^3 n)$ and $O(\log^4 n)$ even under high SINR requirements, while [15] achieves an average latency between $O(\log^5 n)$ and $O(\log^6 n)$. In addition, the computation in our algorithm is much simpler as compared to [15], which makes our solution more suitable for resource-constrained sensor motes. Furthermore, we discuss how to improve the energy efficiency through load-

978-1-4577-1703-1/12/$26.00 ©2012 IEEE

balancing techniques. To the best of our knowledge, this is currently the best distributed algorithm for the MLAS problem under the physical interference model in the literature.

The rest of this paper is organized as follows. We discuss related work in Section II, and formally describe our system model and problem definition in Section III. In Section IV, we present the distributed joint tree construction, power assignment and scheduling algorithm. The correctness and performance of our algorithm are theoretically analyzed and experimentally studied in Section V. Finally, Section VI concludes this paper and future work is discussed.

II. Related Work

A. WSN-based Smart Grid

The harsh and complex electric-power-system environments pose great challenges in the reliability of WSNs in smart grid applications. An extensive survey [17] discussed the current technologies and possible future directions for the smart grid. Gungor et al. [1] performed a comprehensive experimental study and field tests on IEEE 802.15.4-compliant WSNs in real-world power delivery and distribution systems, which provided valuable insights and design guidance for WSN-based smart grid applications. Majumder et al. [5] proposed a Find Reliable Link scheme to ensure reliable communications on error-prone wireless channels, which enables the system to make a quick recovery from node failures or link failures.

Araújo et al. [4] presented a hybrid MAC and control architecture for scalable control over large complex systems with packet drops, where a small number of control loops with high demand of attention are scheduled in a contention-free scheme and well-regulated loops are scheduled in a lossy asynchronous contention-access scheme. iHEM, an in-home energy management system developed in [3], can maintain demand-supply balance and reduce electricity expenses efficiently in the presence of local energy generation capability, prioritized appliances, and real-time pricing. The potential applications and challenges of employing wireless multimedia sensor and actor networks for the smart grid are discussed in [2], which can provide much more information than scalar sensor measurements.

B. Link scheduling and Data Aggregation

Link scheduling under the physical interference model has been studied extensively in recent years. Goussevskaia et al. [18] proved the NP-completeness of this problem and provided an $O(\log n)$ approximation algorithm in [19] for a special case, in which a set of single-hop transmission requests is given and the uniform power scheme is exploited.

Moscibroda et al. [11] demonstrated that the network capacity can be greatly increased using the *non-linear power assignment* scheme (the power level of each node is assigned according to non-liner functions) and presented a novel scheduling algorithm to schedule a strongly connected set of links for multi-hop communications in time $O(\log^4 n)$. Later, they applied link scheduling to the data gathering problem in WSNs, achieving an $O(\log^2 n)$ latency [12]. They also studied topology control under the physical interference model and obtained a theoretical upper bound on the scheduling complexity for arbitrary topologies [10].

The MLAS problem under the physical interference model is first studied by Li et al. in [13], which introduced a scheduling algorithm with a latency bound of $O(R+\Lambda)$ exploiting the uniform power scheme, where R is the network radius and Λ is the maximum node degree in the communication graph. The algorithm in [14] achieves the same latency bound, but with an improved hidden constant by dividing the MLAS problem into the Maximum Weighted Independent Set of Links sub-problem and the Minimum Latency Link Scheduling sub-problem. Lam et al. [16] made the milestone contribution to prove the NP-completeness of the MLAS problem, and provided an algorithm yielding a latency bounded by $O(R+\Lambda)$ in the *dual power assignment* scheme, in which each node is assigned one of two power levels.

Li et al. [15] presented work closest to ours. Using the linear power scheme, they proposed a distributed algorithm with a latency bound of $O(\Delta)$ in networks of arbitrary topology, and a centralized algorithm with a latency bounded by $O(\log^3 n)$ that is the best result in the literature. Our algorithm achieves the same latency bound, but with a much smaller hidden constant, simpler computations and much better performance in practice.

III. System Model

We consider a sensor network consisting of n nodes $V = \{v_1, \ldots, v_{n-1}, v^*\}$ located arbitrarily in the plane, where v^* is the base station that is responsible to collect data from all sensors. Each sensor generates exactly one packet for each measurement and will aggregate its own data with that received from other sensors before transmitting it. The Euclidean distance between two nodes v_i and v_j is denoted by $d(v_i, v_j)$ and the maximal distance between any two nodes in the network is denoted by D. For simplicity and without loss of generality, we assume that the minimal distance between any two nodes is 1 and we define $\Delta = \log d_{max}$ as the link length diversity.

We assume that time is divided into synchronized slots of equal length. Each sensor mote is assigned one timeslot and a power level in that timeslot for its transmission. The *power assignment* P determines the power level P_v of each sensor node $v \in V\backslash\{v^*\}$. An *aggregation schedule* $S = \{S_1, \ldots, S_L\}$ determines which nodes will transmit in each timeslot, where $S_t \subseteq V\backslash\{v^*\}$ $(t = 1, \ldots, L)$ contains all the nodes scheduled to transmit in timeslot t and L is the total time span for the schedule. We adopt the *physical interference model* for wireless communications, in which the received power on the medium is assumed to decay with distance at an exponential rate with path-loss exponent $\alpha > 2$. Whether a packet is received successfully at the receiver depends on the received signal strength, the background noise level, and the cumulative interference caused by simultaneously transmitting nodes. Formally, a packet from s is successfully received by r if the SINR at r is above a certain threshold β, i.e.,

$$\frac{P_s/d(s,r)^\alpha}{N_0 + \sum_{v\in S_t\backslash\{s\}} P_v/d(v,r)^\alpha} \geq \beta \tag{1}$$

where N_0 is the background noise.

The network topology is organized as a tree structure for routing in data aggregation, whose root node is the base station. A valid aggregation schedule should satisfy all the following conditions:

1) Any sensor node should be scheduled exactly once: $\bigcup_{t=1}^{L} S_t = V\backslash\{v^*\}$ and $S_i \cap S_j = \emptyset$ $(\forall i \neq j)$.
2) A non-leaf node must be scheduled after all its child nodes: $\forall r \in S_i, s \in S_j$ and $r = p(s)$, we must have $i > j$, where $p(s)$ is the parent of node s.
3) All the transmissions must be successful, i.e., the SINR at the receiver of each transmission must be strong enough to decode the packet: $\forall t$, $\forall s \in S_t$ and $r = p(s)$, inequality (1) must be satisfied.

The MLAS problem studied in this paper can be formally defined as follows: given a set V of sensor nodes and the base station, and their locations, construct an aggregation tree T, a power assignment P and an aggregation schedule S satisfying the above three conditions, such that the total number of timeslots L is minimized.

IV. Power Assignment and Aggregation Scheduling

In this section, we first present our distributed joint routing tree construction, power assignment and link scheduling algorithm to solve the MLAS problem under the physical interference model. Then we discuss how to improve its energy efficiency through load-balancing techniques.

A. Distributed Algorithm

Our distributed algorithm applies a cluster-based aggregation mechanism. At the beginning, sensor motes in each small area form a cluster with short transmission links. After the cluster head in each cluster has aggregated data from all the members, these cluster heads form larger clusters with longer transmission links. This process repeats until the whole network is covered by a single cluster.

The complete algorithm is shown in Fig. 1, which proceeds in *phases*, each phase corresponding to an iteration of the outermost loop. The purpose of each phase is to gradually reduce the number of active nodes in $\mathcal{A}$. In Phase k ($k = 1, 2, \ldots$), we apply the grid partition to divide the whole network into square cells of side length $l = \frac{2^k}{\sqrt{2}}$, which means that the maximum distance between nodes in the same cell is 2^k. Then we mark all the cells using $(K+1)^2$ colors, guaranteeing that cells of the same color are separated by at least Kl, as shown in Fig. 2. Nodes in the same cell form a cluster and one node is selected as the cluster head.

Each phase consists of at most $(K+1)^2$ *rounds*. The purpose of each round is to schedule the nodes in the cells of the same color sequentially. In Round i, one link (if it exists) is selected from each cell of Color i, and they are scheduled in a new timeslot. The transmission power of the sender of the link $\overrightarrow{(v, h_j)}$ is set to be $\mu d(v, h_j)^\alpha$, where h_j is the cluster head of that cell. The intuition is that when the cells of the same color are sufficiently far away from each other, all these link transmissions will be successful due to the limited cumulative interference at the receivers (proved in Theorem 1). This process repeats until all the links in the cells of Color i have been scheduled. For each scheduled link, its sender is removed from $\mathcal{A}$.

After Phase k, $\mathcal{A}$ contains only the cluster heads of the cells in Phase k. In Phase $k+1$, the network is covered by

Input: Node set V, the base station v^* and their locations
Output: Tree link set T, node powers P and schedule S

$\mathcal{A} \leftarrow V\backslash\{v^*\}$; $k \leftarrow 1$; $t \leftarrow 1$; $T \leftarrow \emptyset$; $P \leftarrow \emptyset$; $S \leftarrow \emptyset$;
$K \leftarrow \left(1 + 4\beta\left(\frac{\alpha\left(1+2^{\frac{\alpha}{2}}\right)}{\alpha-1} + \frac{\pi}{2(\alpha-2)}\right)\right)^{\frac{1}{\alpha}}$; $\mu \leftarrow N_0\beta K^\alpha$;
while $|\mathcal{A}| > 1$ **do** ▹ Phase k
  Partition the network with cells of side length $\frac{2^k}{\sqrt{2}}$;
  Mark the cells using $(K+1)^2$ colors;
  for $i = 1 \rightarrow (K+1)^2$ **do**
    notEmpty $\leftarrow$ **true**; ▹ Round i
    while *notEmpty* **do**
      $S_t \leftarrow \emptyset$; ▹ Timeslot t
      for all Cell j with Color i **do**
        Assign any node h_j as the head if no head yet;
        Select any non-scheduled cluster member v;
        $S_t \leftarrow S_t \cup \{v\}$; $|\mathcal{A}| \leftarrow |\mathcal{A}|\backslash\{v\}$; $T \leftarrow T \cup \{\overrightarrow{(v, h_j)}\}$;
        Assign transmission power: $P_v \leftarrow \mu d(v, h_j)^\alpha$;
      end for ▹ Timeslot t is filled up
      if $S_t \neq \emptyset$ **then**
        $S \leftarrow S \cup \{S_t\}$; $t \leftarrow t + 1$;
      else
        notEmpty $\leftarrow$ **false**;
      end if
    end while ▹ Round i finished
  end for
  $k \leftarrow k + 1$; ▹ Phase k finished
end while
$w \leftarrow$ the only node in $\mathcal{A}$; $P_w \leftarrow \mu d(w, v^*)^\alpha$;
$S_t \leftarrow \mathcal{A}$; $T \leftarrow T \cup \{\overrightarrow{(w, v^*)}\}$; $S \leftarrow S \cup \{S_t\}$;

Fig. 1. Distributed power assignment and aggregation scheduling

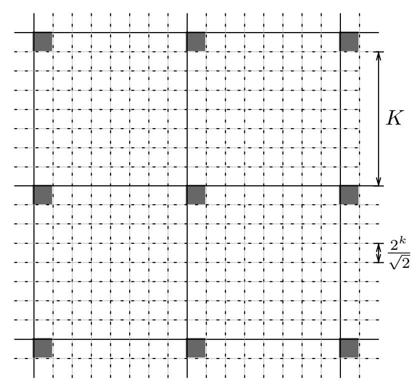


Fig. 2. Grid partition and coloring in Phase k with side length $\frac{2^k}{\sqrt{2}}$

cells of side length $\frac{2^{k+1}}{\sqrt{2}}$ and a new head is selected for data aggregation in each cell. In the last phase (when the outermost loop finishes), only one node will remain, which has collected all the data in the network, and will transmit the aggregated result to the base station in one hop. An example of a small sensor network that can be processed in 4 phases is provided in Fig. 3 to illustrate our algorithm. As shown in [15], the algorithm can be implemented in a fully distributed fashion.

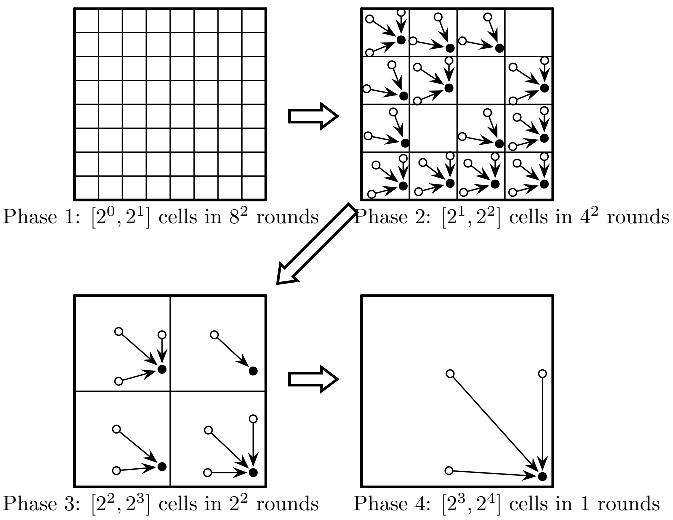


Fig. 3. Example: a small network with $D = \frac{2^4}{\sqrt{2}}$. Nodes in the first phase are not shown here for clarity. Black nodes are cluster heads.

B. Load-Balancing for Energy Efficiency

In [15], the node closest to the base station is selected as the cluster head in each cell. The main drawback of this method is that the lifetime of the cluster head is significantly shortened due to its large energy dissipation compared to the cluster members. To solve this problem, many dynamic cluster-based routing algorithms for energy efficiency have been proposed. In this paper, we apply a load-balancing mechanism based on the residual energy on each node. Since many algorithms are available in the literature, we only briefly describe the main steps here.

After the cluster head $h \in S_{t_1}$ has served for a certain time, which is an application-specific defined parameter, it initiates the procedure to select a new cluster head in the cell. The node $h^* \in S_{t_2}$ with the most residual energy will become the new cluster head. Each cluster member v (including h) sets $\mu d(v, h^*)^\alpha$ as its new transmission power. Node h^* sets its transmission power to $\mu d(v, p(h))^\alpha$, where $p(h)$ is the parent of h in the aggregation tree. Then h^* and h transmit in timeslot t_1 and t_2, respectively. Other nodes in the network are not affected and the resulting scheduling is still valid (by Theorem 1). Autonomous fault recovery from node failures can be achieved in a similar way.

V. Performance Analysis

In this section, we first prove the correctness of the proposed scheduling algorithm and analyze the bound of data aggregation latency. Then we briefly compare the performance of our method with another important algorithm in the literature.

A. Correctness

Theorem 1 (Correctness): The algorithm in Fig. 1 can construct a valid data aggregation tree and correctly schedule all the transmissions under the physical model with:

$$K = \left(1 + 4\beta\left(\frac{\alpha\left(1 + 2^{\frac{\alpha}{2}}\right)}{\alpha - 1} + \frac{\pi}{2(\alpha - 2)}\right)\right)^{\frac{1}{\alpha}} \quad \text{and } \mu = N_0 \beta K^\alpha \quad (2)$$

Proof: From [15], we know that the resulting aggregation tree is valid, satisfying Condition (1) and (2) described in Section III. We now prove that all transmissions are successful under the physical interference model (Condition (3)).

Consider any link $\overrightarrow{(s,r)}$, where the sender s is scheduled in timeslot t (i.e., $s \in S_t$). For $\forall v \in S_t \setminus \{s\}$, r and v are located in cells of the same color and their distance $d(v,r) \geq \sqrt{i^2 + j^2}Kl$, where $-\infty < i, j < \infty$ (i and j cannot both be zero) and l is the side length of a cell in the current grid partition. Obviously, $d(s,r) \leq \sqrt{2}l$ because they are located in the same cell. Since at most one node is transmitting in each cell in timeslot t, we get the cumulative interference I^+ at the receiver r from all the other nodes transmitting in timeslot t as follows:

$$\begin{aligned}
I^+ &= \sum_{v \in S_t \setminus \{s\}} \frac{P_v}{d(v,r)^\alpha} \leq \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \frac{N_0 \beta K^\alpha \left(\sqrt{2}l\right)^\alpha}{\left(\sqrt{i^2 + j^2}Kl\right)^\alpha} \\
&= N_0 \beta 2^{\frac{\alpha}{2}} \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \left(\sqrt{i^2 + j^2}\right)^{-\alpha} \\
&= 4N_0 \beta 2^{\frac{\alpha}{2}} \left(1 + \sum_{i=2}^{\infty} i^{-\alpha} + 2^{-\frac{\alpha}{2}} + 2\sum_{i=2}^{\infty} \left(\sqrt{1 + i^2}\right)^{-\alpha} \right. \\
&\quad \left. + \sum_{i=2}^{\infty} \sum_{j=2}^{\infty} \left(\sqrt{i^2 + j^2}\right)^{-\alpha}\right) \\
&\leq 4N_0 \beta 2^{\frac{\alpha}{2}} \left(\frac{\alpha\left(1 + 2^{-\frac{\alpha}{2}}\right)}{\alpha - 1} + \frac{\pi 2^{-\frac{\alpha}{2}}}{2(\alpha - 2)}\right) \\
&= 4N_0 \beta \left(\frac{\alpha\left(1 + 2^{\frac{\alpha}{2}}\right)}{\alpha - 1} + \frac{\pi}{2(\alpha - 2)}\right)
\end{aligned}$$

Substituting the above inequality in inequality (1), we get the SINR at the receiver r as follows:

$$\begin{aligned}
SINR &= \frac{P_s / d(s,r)^\alpha}{N_0 + I^+} \\
&\geq \frac{N_0 \beta K^\alpha}{N_0 + 4N_0 \beta \left(\frac{\alpha\left(1+2^{\frac{\alpha}{2}}\right)}{\alpha-1} + \frac{\pi}{2(\alpha-2)}\right)} \\
&= \frac{\beta K^\alpha}{1 + 4\beta \left(\frac{\alpha\left(1+2^{\frac{\alpha}{2}}\right)}{\alpha-1} + \frac{\pi}{2(\alpha-2)}\right)} = \beta
\end{aligned}$$

Therefore, each link transmission is successful under the physical interference model. ■

B. Aggregation Latency

Lemma 1: If the minimum distance between any two nodes is 1, there can be at most 6 nodes within a cell in the grid partition in the first phase.

Proof: We use the Groemer inequality [9]: suppose that C is a compact convex set and U is a set of points with mutual distances at least 1, then

$$|U \cap C| \le \frac{area(C)}{\sqrt{3}/2} + \frac{peri(C)}{2} + 1$$

where $area(C)$ and $peri(C)$ are the area and perimeter of C respectively.

In the grid partition in the first phase, the side length of each cell is $\sqrt{2}$. Then we derive

$$|U \cap C| \le \frac{2}{\sqrt{3}/2} + \frac{4\sqrt{2}}{2} + 1 = 6.1378 < 7$$

Therefore, there can be at most 6 nodes in each cell in the first phase. ■

Lemma 2: In the algorithm in Fig. 1, there can be at most $\log \frac{\sqrt{2}D}{K+1}$ phases that consist of exactly $(K+1)^2$ rounds.

Proof: Since the maximum distance between any two nodes is D, the whole network can be covered by a square with side length D. In Phase i, the side length of each cell is $\frac{2^i}{\sqrt{2}}$. Thus the number of cells in each row (or column) is $\frac{\sqrt{2}D}{2^i}$. If Phase i consists of exactly $(K+1)^2$ rounds, we must have $\frac{\sqrt{2}D}{2^i} \ge K+1$. Thus we derive $i \le \log \frac{\sqrt{2}D}{K+1}$, which means there can be at most $\log \frac{\sqrt{2}D}{K+1}$ such phases. ■

Theorem 2 (Latency Bound): The algorithm in Fig. 1 can achieve an upper-bound on latency of $3(K+1)^2\Delta + 3(K+1)^2 \log \frac{\sqrt{2}}{K+1} + 6K^2 + 4K + 2$ timeslots, where $\Delta = \log D$ and K is a constant given in equation (2).

Proof: The data aggregation latency consists of three parts based on different phases:

i) In Phase 1, we know that there can be at most 5 links transmitting to the head node in each cell from Lemma 1. Since all the cells can be marked using $(K+1)^2$ colors, Phase 1 consists of $(K+1)^2$ rounds, one for each color. In each round, the links selected from cells of the same color can transmit simultaneously from Theorem 1. Therefore, at most $5(K+1)^2$ timeslots are needed to schedule all cells in Phase 1.

ii) From Lemma 2, we know that $(K+1)^2$ rounds are needed for each phase i ($2 \le i \le \log \frac{\sqrt{2}D}{K+1}$). Since each cell in Phase i contains 4 cells from Phase $i-1$, each of which contains at most one head node left, there are at most 4 nodes for each cell in Phase i. Thus there are at most 3 links transmitting to the new head node in each cell, which means $3(K+1)^2\left(\log \frac{\sqrt{2}D}{K+1} - 1\right)$ timeslots are needed for Phase 2 to Phase $\log \frac{\sqrt{2}D}{K+1}$.

iii) In Phase $\log \frac{\sqrt{2}D}{K+1} + 1$, at most $3K^2$ timeslots are needed. In all the following phases, the number of cells in the next phase is just $\frac{1}{4}$ of those in the current phase. In the last phase, there is only one cell left and at most 3 timeslots are needed. Hence, $3\sum_{i=0}^{\log K}\left(2^i\right)^2$ timeslots are totally needed for these phases.

After all data have been transmitted to the root of the aggregation tree, one additional timeslot is required for the root to transmit the result to the base station. Therefore, the overall aggregation latency is as follows:

$$\begin{aligned} L &\le \underbrace{5(K+1)^2}_{(i)} + \underbrace{3(K+1)^2\left(\log \frac{\sqrt{2}D}{K+1} - 1\right)}_{(ii)} + \underbrace{3\sum_{i=0}^{\log K}\left(2^i\right)^2}_{(iii)} + 1 \\ &= 3(K+1)^2 \log \frac{\sqrt{2}D}{K+1} + 2(K+1)^2 + 4K^2 \\ &= 3(K+1)^2 \log D + 3(K+1)^2 \log \frac{\sqrt{2}}{K+1} + 6K^2 + 4K + 2 \\ &= 3(K+1)^2\Delta + 3(K+1)^2 \log \frac{\sqrt{2}}{K+1} + 6K^2 + 4K + 2 \end{aligned}$$

■

C. Comparison With Cell-AS [15]

A similar algorithm Cell-AS is proposed in [15]. Cell-AS achieves an upper latency bound of $12(\frac{16}{3}X^2 + 12X + 7)\Delta - 32X^2 - 72X - 29$, where $X = (6\beta(1+(\frac{2}{\sqrt{3}})^\alpha \frac{1}{\alpha-2})+1)^{1/\alpha}$. Although both algorithms are bounded by $O(\Delta)$ asymptotically, Cell-AS has a much larger hidden constant. From Table I, we can see that the hidden constant of Cell-AS is at least 15 times of that of our algorithm, which makes our algorithm more practical in real-world deployments. In addition, since grid partitions are applied in our algorithm rather than hexagon partitions as in Cell-AS, the computations become much simpler, which is more suitable for resource-constrained sensor motes.

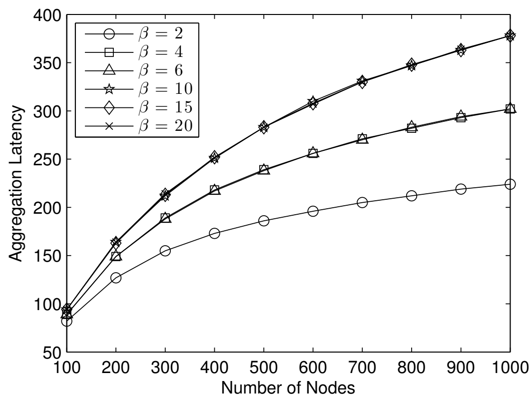
TABLE I
THE COMPARISON OF HIDDEN CONSTANTS UNDER DIFFERENT α AND β

α	β	Our Algorithm	Cell-AS [15]	Ratio
4	2	43.099	712.96	16.543
4	6	65.205	1030.6	15.806
4	10	79.834	1239.2	15.522
4	15	94.113	1441.6	15.318
4	20	105.97	1609	15.184
3	8	154.28	2382	15.439
4	8	73.027	1142.3	15.643
5	8	50.309	804.22	15.986

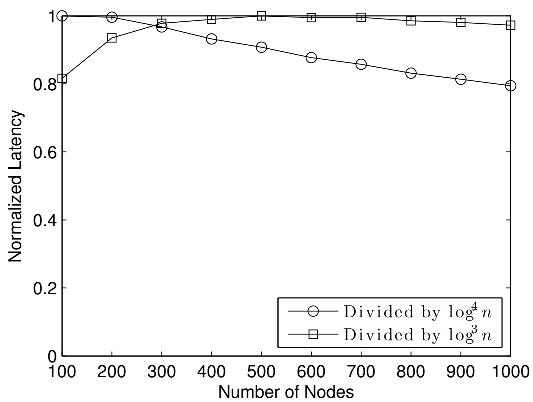
D. Numerical Result

In order to compare the performance of our algorithm with Cell-AS, we apply the same simulation settings as [15]. The simulations are conducted in a $200m \times 200m$ two-dimensional free-space region, in which $n = 100$ to 1000 sensor motes are uniformly distributed. We set $\alpha = 4$ and β varies from 2 to 20. Each simulation result is the average of 100 runs.

Fig. 4a shows the aggregation latency with different β. As expected, the latency increases with larger β corresponding to higher SINR requirement, which will increase the number of



(a) Aggregation Latency $\alpha = 4$



(b) Asymptotic Latency $\alpha = 4, \beta = 20$

Fig. 4. Aggregation latency in uniform distribution

colors needed and thus reduce the transmission concurrency. In addition, compared to [15] (please refer to [15] for their experimental results), our algorithm achieves a much better performance because the concurrency of link scheduling across different cells is significantly increased. Due to the simplicity of grid partition using square cells, we are able to analyze the cumulative interference more accurately, resulting in much less colors required to mark all cells.

Fig. 4b shows the asymptotic performance of our algorithm. The latency when $\alpha = 4$ and $\beta = 20$ (most demanding) is divided by $\log^3 n$ and $\log^4 n$, respectively, and then scaled properly. The curve slightly increases in the case of $\log^3 n$, and decreases in the case of $\log^4 n$. Therefore, our algorithm achieves an average latency between $O(\log^3 n)$ and $O(\log^4 n)$ in uniform distributions, while Cell-AS achieves an average latency between $O(\log^5 n)$ and $O(\log^6 n)$.

VI. Conclusions and Future Work

In this paper, we proposed a distributed joint routing, power assignment and link scheduling algorithm to minimize the data aggregation latency in WSNs under the physical interference model, which achieves a latency bounded by $O(\Delta)$, with a much smaller hidden constant than another important algorithm which achieves the same bound. Simulation results show that our algorithm can reduce the latency significantly and achieve an average latency between $O(\log^3 n)$ and $O(\log^4 n)$ in networks where nodes are uniformly distributed.

Our future work includes: (*i*) extend our algorithm to deal with multiple base stations and analyze the latency bound, and (*ii*) extend our algorithm to apply the non-linear power assignment that has been proved to be able to further increase the network capacity and thus reduce the aggregation latency.

Acknowledgment

Research partially supported by DARPA and SRC through grant award 013641-001 of the FCRP, by the National Science Foundation (NSF) under grant award CNS-1035655, and by the National Institute of Standards and Technology (NIST) under grant award 70NANB11H148.

References

[1] V. Gungor, B. Lu, and G. Hancke, "Opportunities and challenges of wireless sensor networks in smart grid," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 10, pp. 3557 – 3564, October 2010.

[2] M. Erol-Kantarci and H. T. Mouftah, "Wireless multimedia sensor and actor networks for the next generation power grid," *Elsevier Ad Hoc Networks*, vol. 9, no. 4, pp. 542 – 551, June 2011.

[3] M. Erol-Kantarci and H. Mouftah, "Wireless sensor networks for cost-efficient residential energy management in the smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 314 – 325, June 2011.

[4] J. Araujo, Y. Ariba, P. Park, H. Sandberg, and K. Johansson, "Control over a hybrid MAC wireless network," in *IEEE SmartGridComm*, 2010, pp. 197 – 202.

[5] R. Majumder, G. Bag, and K.-H. Kim, "Power sharing and control in distributed generation with wireless sensor networks," *IEEE Transactions on Smart Grid*, no. 99, pp. 1 – 17, 2012.

[6] B. Wang and X. Jia, "Reducing data aggregation latency by using partially overlapped channels in sensor networks," in *IEEE GLOBECOM*, 2009, pp. 1 – 6.

[7] B. Yu, J. Li, and Y. Li, "Distributed data aggregation scheduling in wireless sensor networks," in *IEEE INFOCOM*, 2009, pp. 2159 – 2167.

[8] K.-W. Fan, S. Liu, and P. Sinha, "Dynamic forwarding over tree-on-DAG for scalable data aggregation in sensor networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 10, pp. 1271–1284, October 2008.

[9] P.-J. Wan, S. C.-H. Huang, L. Wang, Z. Wan, and X. Jia, "Minimum-latency aggregation scheduling in multihop wireless networks," in *ACM MobiHoc*, 2009, pp. 185–194.

[10] T. Moscibroda, R. Wattenhofer, and A. Zollinger, "Topology control meets sinr: the scheduling complexity of arbitrary topologies," in *ACM MobiHoc*, 2006, pp. 310 – 321.

[11] T. Moscibroda and R. Wattenhofer, "The complexity of connectivity in wireless networks," in *IEEE INFOCOM*, 2006, pp. 1 – 13.

[12] T. Moscibroda, "The worst-case capacity of wireless sensor networks," in *ACM/IEEE IPSN*, 2007, pp. 1 – 10.

[13] X.-Y. Li, X. Xu, S. Wang, S. Tang, G. Dai, J. Zhao, and Y. Qi, "Efficient data aggregation in multi-hop wireless sensor networks under physical interference model," in *IEEE MASS*, 2009, pp. 353 – 362.

[14] X. Xu, W. Lou, X. Liu, and S. Tang, "Delay efficient link and aggregation scheduling under physical interference model," in *IEEE MASS*, 2011, pp. 421 – 429.

[15] H. Li, Q. S. Hua, C. Wu, and F. C. M. Lau, "Minimum-latency aggregation scheduling in wireless sensor networks under physical interference model," in *ACM MSWiM*, 2010, pp. 360–367.

[16] N. X. Lam, M. K. An, D. T. Huynh, and T. N. Nguyen, "Minimum latency data aggregation in the physical interference model," in *ACM MSWiM*, 2011, pp. 93–102.

[17] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid – the new and improved power grid: A survey," *IEEE Communications Surveys and Tutorials*, pp. 1 – 37, 2011.

[18] O. Goussevskaia, Y. A. Oswald, and R. Wattenhofer, "Complexity in geometric sinr," in *ACM MobiHoc*, 2007, pp. 100 –109.

[19] O. Goussevskaia, R. Wattenhofer, M. M. Halldórsson, and E. Welzl, "Capacity of arbitrary wireless networks," in *IEEE INFOCOM*, 2009, pp. 1872–1880.