Multiple Sink Placement with Latency and Reliability Guarantee in Lossy Wireless Sensor Networks

Yu-Shu Chen*, Shih-Ying Chang†, Tzu-Wen Chang*, and Ming-Jer Tsai*
*Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan, ROC
†Information and Communications Research Laboratories, Industrial Technology Research Institute, Hsinchu, Taiwan, ROC
lazyeagles@gmail.com, itriA00550@itri.org.tw, g9962515@cs.nthu.edu.tw, mjtsai@cs.nthu.edu.tw

***Abstract*—In this paper, we study a multiple sink placement problem with latency and reliability guarantee in a lossy wireless sensor network and its applications in an advanced metering infrastructure (AMI) network. We show the problem is NP-hard and propose an algorithm for the problem. The proposed algorithm jointly considers the routing protocol RPL [1], the MAC protocol TSCH [2], and the scheduling approach DeTAS [3] to meet the latency and reliability requirements in a real AMI network. We conduct simulations on the real data for the establishment of an AMI network. Simulation results show the proposed algorithm has a good performance in the number of selected concentrators and the balanced number of smart meters connected to the selected concentrators.**

I. Introduction

In a smart grid, smart meters send the metering data to concentrators (i.e., data aggregation points) via the advanced metering infrastructure (AMI) network, and concentrators send the metering data collected from smart meters to the utility provider data center via the backbone network [4]. Due to the high cost of concentrators, multiple concentrator placement is a critical problem for the establishment of an AMI network. Today, the wireless mesh network technology can be applied to the AMI network by equipping smart meters and concentrators with low power and low cost transceivers [5]. However, due to the low power of transceivers, lossy links exist in the AMI network. Therefore, the multiple concentrator placement problem in an AMI network is equivalent to the multiple sink placement problem in a lossy wireless sensor network (LWSN). Moreover, according to the SG Network System Requirements Specification [6], each smart meter should send metering data to a concentrator under requirements of 98% success every three hours and 99.5% success over two days. In this paper, we undertake the design of an algorithm for the multiple sink placement problem in an LWSN such that the latency and reliability requirements are satisfied.

Many methods of multiple sink placement have been proposed in WSNs [7]–[11], where a cluster is first constructed for each sink, some clusters are then selected to contain all sensors, and the selected clusters are finally pruned to partition all sensors. And, many routing protocols are proposed to send data from sensors to sinks in the pruned clusters such that energy efficiency [12], bandwidth efficiency [13], or Quality-of-Service (QoS) requirement [14] is achieved. Note that if the satisfaction of the latency requirement is not confirmed in the cluster construction, the routing protocol is unable to meet the latency requirement. Thus, when constructing a cluster, it is required to confirm the satisfaction of the latency requirement for each sensor to be included in the cluster by jointly considering the routing protocol, the MAC protocol, and the scheduling approach, as proposed in [8].

In LWSNs, since lossy links are unreliable, many routing protocols are proposed to send data from sensors to sinks to meet the latency and reliability requirements [15]–[17]. However, to the best of our knowledge, no algorithm confirms the satisfaction of both the latency and reliability requirements for each sensor to be included in a cluster in cluster construction. Obviously, to ensure that these routing protocols really meet the latency and reliability requirements in LWSNs, we need to fill the gap by jointly considering the routing protocol, the MAC protocol, and the scheduling approach to confirm the satisfaction of both the latency and reliability requirements for each sensor to be included in a cluster in cluster construction. Our main contributions are described as follows:

- We study the scenario of deploying concentrators in the establishment of an AMI network, model the scenario as the Minimum Sink Placement with Latency and Reliability Guarantee (MSPLRG) problem, and show the MSPLRG problem is NP-hard.
- We propose an algorithm for the MSPLRG problem by jointly considering the routing protocol RPL [1], the MAC protocol TSCH [2], and the scheduling approach DeTAS [3]. The proposed algorithm is the first one in the literature able to meet the latency and reliability requirements in a real AMI network.
- We conduct extensive simulations to evaluate the performance of the proposed algorithm for the establishment of an AMI network using the real locations of premises (i.e., smart meters) and poles (i.e., concentrators) from the government open data [18], where the communication module and the radio propagation are based on the specification of CC2420 RF transceiver [19] and the practical propagation model [20], respectively.

II. PRELIMINARY

We employ the routing protocol RPL [1], the MAC protocol TSCH [2], and the scheduling approach DeTAS [3], as described in Sec. II-A, Sec. II-B, and Sec. II-C, respectively, for LWSNs. RPL (Routing Protocol for Low power and Lossy Networks) is developed by the IETF Routing Over Low power and Lossy Networks Working Group for providing a mechanism that supports multipoint-to-point traffic in the low-power and lossy network. TSCH (IEEE Standard 802.15.4e TimeSlotted Channel Hopping) is created by the 6TiSCH working group for providing a mechanism for medium access control in low power and lossy networks. In addition, we employ DeTAS based on RPL and TSCH since we can evaluate the delivery probability of the data from a sensor to a sink, as described in Sec. II-D, so that we can decide whether to include the sensor in the cluster of the sink. Besides, DeTAS ensures the smallest end-to-end latency, keep the queue utilization as small as possible, and make collision free. Please refer to Fig. 1 for an example of the employed mechanisms.

A. The Routing Protocol

Using RPL, a Directed Acyclic Graph (DAG), consisting of disjoint destination oriented DAGs (DODAGs), is constructed. In a DAG, each sink is the root of a DODAG and each sensor is included in exactly one DODAG and connects to the sink via the shortest path, where the distance is in terms of the hop count or expected transmission count (ETX) [21]. In this paper, the ETX is used since the employment of ETX yields the high-throughput routing paths [21]. The ETX of a link is the inverse of the delivery probability of the link, and the ETX of a path is the sum of the ETXs of the links on the path.

B. The MAC Protocol

TSCH is used to avoid packet collision. Using TSCH, the operating frequency is divided into channels, where sensors assigned in different channels do not interfere with each other. Besides, time is sliced into time slots so that TSCH is suitable for the deterministic and periodical traffic pattern. In the network, all sensors are synchronized and can transmit, receive, or sleep in one time slot. A slotframe, consisting of consecutive time slots, continuously repeats over time, where the size of a slotframe depends on the applications [3] [22].

C. The Scheduling Approach

DeTAS is used to schedule the transmission, receipt, or sleep of sensors in each time slot. Using DeTAS, the sensors in the routing path from a sensor to the sink in a DODAG are allocated consecutive unoccupied time slots in the slotframe.

D. The Data Delivery Probability

Suppose that sensor s is in the DODAG of sink d and the shortest path from s to d is $\langle v_0 = s, v_1, ..., v_{h-1}, v_h = d\rangle$. The event that a packet is delivered from s to d during t' slotframes can be divided into the sub-events that the packet is delivered from s to v_i but fails to be delivered to v_{i+1} in the 1st slotframe and from v_i to d during the remaining $t'-1$ slotframes for $i = 0, ..., h$. Let $p^{(k)}_{x,y}$ denote the delivery probability of a packet from x to y during k slotframes. Then, $p^{(t')}_{s,d}$, the delivery probability of a packet from s to d during t' slotframes, can be described in the following recurrence:

$$p^{(t')}_{s,d} = \begin{cases} 1 & \text{if } s = d, \\ \sum_{i=0}^{h} p^{(1)}_{s,v_i}(1 - p_{v_i,v_{i+1}})p^{(t'-1)}_{v_i,d} & \text{if } s \neq d, \end{cases} \quad (1)$$

where $p_{v_i,v_{i+1}} = 0$ as $i = h$, $p^{(1)}_{s,v_i}(1 - p_{v_i,v_{i+1}})$ is equal to the probability that a packet is delivered from s to v_i but fails to be delivered to v_{i+1} in the 1st slotframe, and $p^{(t'-1)}_{v_i,d}$ is equal to the probability that a packet is delivered from v_i to d during the remaining $t'-1$ slotframes. Clearly, $p^{(1)}_{s,v_i}$ can be evaluated by the following equation:

$$p^{(1)}_{s,v_i} = \begin{cases} 1 & \text{if } i = 0, \\ \prod_{j=0}^{i-1} p_{v_j,v_{j+1}} & \text{otherwise.} \end{cases} \quad (2)$$

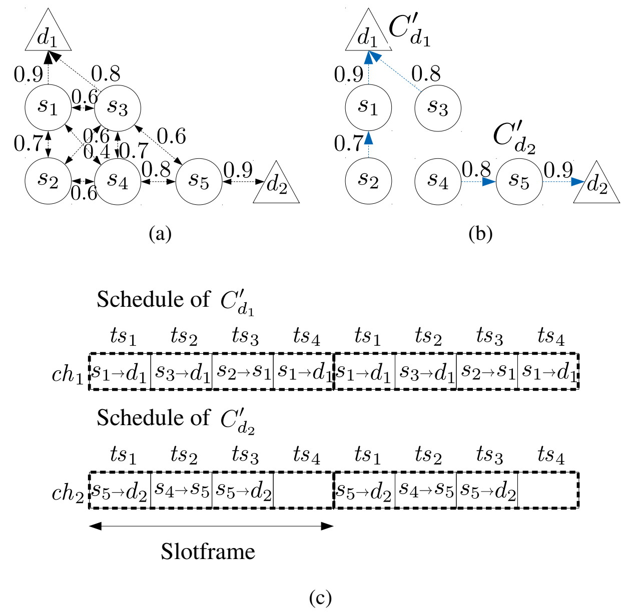

Fig. 1. Example of RPL, TSCH, and DeTAS. (a) A sample LWSN, where the number besides a link denotes the delivery probability of the link. (b) One feasible DAG, consisting of two DODAGs, using RPL. Note that each sensor is included in exactly one DODAG. Besides, the ETXs of links (s_2, s_1) and (s_1, d_1) are equal to $1/0.7$ and $1/0.9$, respectively. Thus, path $\langle s_2, s_1, d_1\rangle$ has the ETX (i.e., length) of $1/0.7 + 1/0.9$ and is the shortest path between s_2 and d_1. Note that each sensor in a DODAG connects to the sink via the shortest path. (c) One feasible DeTAS schedule based on TSCH, where the slotframe with size of 4 time slots repeats twice during 8 time slots. Note that s_2 and s_1 in the routing path $\langle s_2, s_1, d_1\rangle$ are allocated consecutive time slots ts_3 and ts_4 in the slotframe to transmit data to s_1 and d_1, respectively.

III. THE PROBLEM

We first demonstrate the scenario of the establishment of an AMI network in Sec. III-A. Then, we define the problem based on the scenario and study its hardness in Sec. III-B.

A. The Scenario and Network Model

Consider an AMI network in a smart grid. In the AMI network, smart meters at consumers' premises collect the data of power consumption and periodically transfer the data to data aggregation points termed as concentrators. Concentrators transfer the data received from smart meters to the remote data center via a backbone network. The communication technology employed in the network is Zigbee. In addition, we employ RPL [1], TSCH [2], and DeTAS [3] as the routing protocol, the MAC protocol, and the scheduling approach, respectively. According to the SG network system requirements specification [6], each smart meter should send metering data to a concentrator to meet latency and reliability requirements.

The AMI network is realized by an LWSN denoted by $G = (S \cup D, E)$, where S is the set of sensors (i.e., smart meters), D is the set of sinks (i.e., concentrators), E is the set of directed links (i, j) with $i \in S$ and $j \in S \cup D$, and each link (i, j) is associated with the delivery probability $p_{i,j}$ of a packet from i to j in one time slot.

B. The Problem and Hardness

The studied problem and the MUSC problem (which is an NP-hard problem used in the proof of Theorem 1) are described in Definitions 1 and 2, respectively. Besides, the hardness of the problem is shown in Theorem 1.

Definition 1. Suppose that we are given an LWSN $G = (S \cup D, E)$ with RPL, TSCH, and DeTAS employed, a required latency of t time slots, and a required reliability δ. Then, the Minimum Sink Placement with Latency and Reliability Guarantee (MSPLRG) problem asks for a minimum-size set of sinks $D' \subseteq D$ and a slotframe size K such that there exists a feasible DAG (consisting of DODAGs rooted at sinks in D') such that the delivery probability of a packet from each sensor to its sink during t time slots is not less than δ.

Definition 2. Suppose that we are given a universe set U and a collection C of subsets of U. Then, the Minimum Unweighted Set Cover (MUSC) problem [23] asks for a minimum-size set $C' \subseteq C$ such that $\bigcup_{U' \in C'} U' = U$.

Theorem 1. The MSPLRG problem is NP-hard.

Proof. We show that the MUSC problem can be reduced to the MSPLRG problem in polynomial time. For each instance $MUSC(U, C)$ of the MUSC problem, we construct an instance $MSPLRG(t, \delta, S, D, E)$ of the MSPLRG problem as follows. First, set t to $|S|$ and δ to 1. Second, set S such that a sensor s is in S if and only if an element u in U, where u is called s's corresponding element in U. Third, set D such that a sink d is in D if and only if an element c (i.e., a subset of U) is in C, where c is called d's corresponding element in C. Fourth, set E such that a directed edge from a sensor s to a sink d, (s, d), is in E if and only if s's corresponding element u in U is in d's corresponding element c in C. Fifth, set $p_{s,d}$ to 1 for each directed edge (s, d) in E. Clearly, this instance can be constructed in polynomial time. In addition,

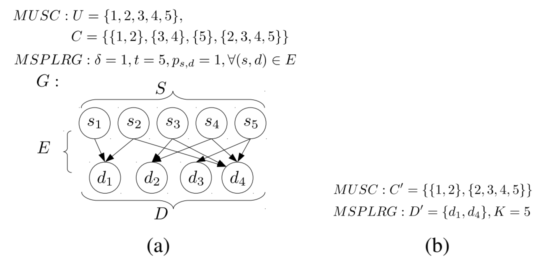

Fig. 2. Example of the reduction from the MUSC problem to the MSPLRG problem. (a) The construction of $MSPLRG(t, \delta, S, D, E)$. (b) The correspondence between a solution C' of $MUSC(U, C)$ and a solution (D', K) of $MSPLRG(t, \delta, S, D, E)$.

let C' be a subset of C and D' be a subset of D such that an element d is in D' if and only if d's corresponding element c is in C'. Obviously, $|C'| = |D'|$ and C' is a feasible solution in $MUSC(U, C)$ if and only if (D', K), where $K = |S|$, is a feasible solution in $MSPLRG(t, \delta, S, D, E)$. Please refer to Fig. 2 for an example of the reduction. This completes the proof. □

IV. The Proposed Algorithm

The proposed algorithm (termed MSPLRG-G) consists of three phases. In the phase of cluster construction, a cluster (which is a directed routing tree rooted at the sink) is constructed for each sink, as described in Section IV-A. In the phase of sink selection, some clusters are selected to contain all sensors, as described in Section IV-B. In the phase of DODAG establishment, the DODAGs of the selected sinks are established, as described in Section IV-C.

A. The Cluster Construction

Note that any sensor s in the constructed DODAG (C'_d) of the selected sink d must have the following three properties: 1) The delivery probability of the data from s to d within the required latency is not smaller than the required reliability (termed reliability property), 2) each sensor in the path from s to d is allocated consecutive unoccupied time slots in the slotframe (termed schedule property), and 3) each sensor in the path from s to d is in C'_d (termed path property). Note that the ETX of a path equals the total expected times of transmitting the data via the path. And, the smaller the EXT of a path, the higher the throughput of the data sent via the path. Thus, in the construction of the cluster of d (C_d), it is good to give high priority to the inclusion of the sensor with the smallest distance from d (i.e., the smallest sum of the ETXs of the links on the path to d). In addition, since sensors in C'_d are chosen among the sensors in C_d, any sensor in C_d also must possess the reliability, schedule, and path properties. Thus, to include a sensor s in C_d, the required number of unoccupied time slots in the slotframe is equal to the number of hops between s and d. Hence, to construct C_d with a large size, we also need to give high priority to the inclusion of the sensor with the smallest number of hops from d. To sum up, in the construction of C_d, we repeatedly include the sensor s with the smallest distance from d, or, in case of parity, the smallest number of hops from d, or, in case of parity, the smallest ID, where s can be included only if s has the reliability, schedule, and path properties in C_d, until no sensor can be included.

However, as can be seen in Fig. 3, infeasible DODAG establishment occurs since the DODAGs in a DAG must be disjoint. Observe that infeasible DODAG establishment occurs when a sensor s in the cluster of a sink d (C_d) is unable to connect to d in the constructed DODAG of d (C'_d) due to that the parent of s in C_d has been included in the DODAG of some other sink d' ($C'_{d'}$) while s cannot be included in $C'_{d'}$. Note that this situation does not occur if s can be included in $C'_{d'}$. Clearly, if the set of the clusters containing the parent of s is a subset of the set of the clusters containing s (termed subset property), s can be included in the DODAG containing the parent of s. Thus, in addition to the reliability, schedule, and path properties, a sensor in a cluster must have the subset property. To ensure any sensor in each cluster has the subset property, when a cluster C_d includes a sensor s that has been included in another cluster $C_{d'}$, the cluster needs to include all descendants of s in $C_{d'}$.

Theorem 2. Each sensor in the cluster of a sink has the reliability, schedule, path, and subset properties.

Proof. Clearly, a sensor s has the reliability, schedule, path, and subset properties when it was just included in the cluster of a sink d (C_d) at time t. Suppose that another sensor s' is included in the cluster of sink d' ($C_{d'}$) at time $t' > t$. We need to show s keeps the reliability, schedule, path, and subset properties in C_d at time t'. Note that if $d' \neq d$, s needs to be included in $C_{d'}$ at time t' if s is the descendant of s' in C_d. Thus, s must keep the reliability, schedule, path, and subset properties in C_d at time t' as $d' \neq d$. Consider $d' = d$. Clearly, any sensor will not removed from the schedule of d once it has been scheduled. Thus, s keeps the schedule property in C_d at time t'. In addition, at time t', s keeps the reliability, path, and subset properties in C_d since each sensor in the path between s and d is still in C_d. This completes the proof. □

B. The Sink Selection

We select sinks (and clusters) using the greedy strategy for the MUSC problem [24], [25]. Precisely, we repeatedly select the cluster that contains the greatest number of sensors not included in any selected cluster until each sensor is included in at least one selected cluster. Recall that we need to determine the slotframe size (K) using TSCH. Thus, we enumerate all possible slotframe sizes to construct and select clusters, and set K to the slotframe size resulting in the smallest number of selected sinks.

Theorem 3. The number of selected sinks is within the factor of $H_{|S|}$ ($= O(\log |S|)$) of the optimal number of selected sinks, where S denotes the number of sensors.

Proof. Let OPT be the optimal number of selected sinks and K^* be the slotframe size resulting in OPT. Let N_{K^*} be the number of sinks selected by the greedy strategy for the MUSC problem as the slotframe size is K^*, and let N be the number of sinks selected by MSPLRG-G. Clearly, $N \leq N_{K^*}$ since we select sinks using the greedy strategy for the MUSC problem by enumerating all possible slotframe sizes. Since $N_{K^*} \leq H_{|S|} OPT$ by [24], [25], we have $N \leq H_{|S|} OPT$. □

C. The DODAG Establishment

We establish a DODAG for each selected sink based on the clusters of selected sinks. Recall that using RPL, the DODAGs must be disjoint. Thus, a sensor contained in two or more clusters of selected sinks can only be included in exactly one DODAG. In addition, we also need to ensure each sensor in a DODAG has the reliability, schedule, and path properties. Clearly, if the DODAG of a sink d (C'_d) is a subtree of the cluster of d (C_d), a sensor s in C'_d has the schedule property, and s has the reliability property if s has the path property. Thus, we only need to ensure for each selected sink d, C'_d is a subtree of C_d and each sensor has the path property in C'_d. Note that each sensor has the subset property in C_d. Thus, a sensor s in C_d can always be included in C'_d that has included the parent of s in C_d to have the path property. This implies each sensor has the reliability, schedule, and path properties in DODAGs if for each selected sink d, C'_d (which initially contains sink d only) repeatedly includes the sensor in C_d which is not included in any DODAG and whose parent in C_d is in C'_d. On the other hand, to balance the sizes of the established DODAGs, the DODAG with the minimum size first includes the best-fit sensor. To identify the best-fit sensor, we first formulate an integer program 3a-3f, where $x_{s,d}$ indicates if sensor s is in C'_d, y is the minimum size of the DODAGs, S is the set of sensors, D' is the set of selected sinks, constraints in 3b ensure that each sensor s is in exactly one DODAG, constraints in 3c ensure that the minimum size of the DODAGs is y, and constraints in 3d ensure that if a sensor s is in C'_d, the parent s' of s in C_d is in C'_d. Then, we obtain a fraction $x^*_{s,d}$ to indicate the worth of s to be in C'_d for each sensor s and selected sink d by solving the linear program, which is obtained from the integer program 3a-3f by relaxing $x_{s,d}$ from an integer in $\{0, 1\}$ to a fraction in $[0, 1]$.

$$\text{maximize} \quad y \tag{3a}$$

$$\text{subject to} \quad \sum_{d \in D'} x_{s,d} = 1, \ \forall s \in S, \tag{3b}$$

$$\sum_{s \in S} x_{s,d} \geq y, \ \forall d \in D', \tag{3c}$$

$$x_{s',d} \geq x_{s,d}, \ \forall s \in S, d \in D', \tag{3d}$$

$$y \geq 0, \tag{3e}$$

$$x_{s,d} \in \{0, 1\}. \tag{3f}$$

Thus, if C'_d has the minimum size, it first includes the sensor s with the greatest value of $x^*_{s,d}$, or, in case of parity, the minimum distance from d, or, in case of parity, the minimum number of hops from d, or, in case of parity, the minimum ID. If two or more DODAGs have the same size, the DODAG that has a sensor with the greatest value of $x^*_{s,d}$,

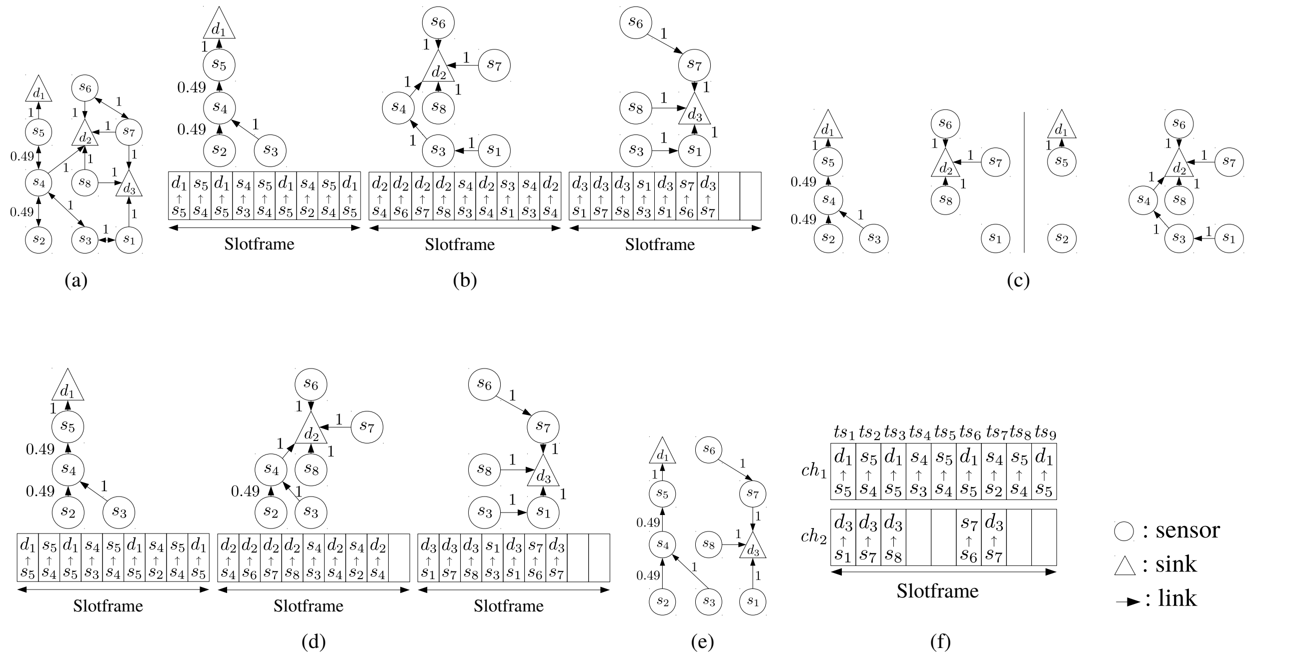

Fig. 3. Example of cluster construction, sink selection, and DODAG establishment, where the required latency is 63 time slots, the required reliability is 0.9, and the slotframe size is 9 time slots. (a) A sample LWSN with 8 sensors $s_1, s_2, \cdots, s_8$ and 3 sinks d_1, d_2, d_3, where the number besides a link denotes the delivery probability of the link. (b) One cluster construction and schedule. Note that s_2 has the reliability property since $p^{(\frac{63}{9})}_{s_2,d_1} = 0.96 \geq 0.9$ by Eq. 1 and has the schedule (or path) property since s_2, s_4, and s_5 are in the path from s_2 to d_1 and s_2, s_4, and s_5 are in the schedule (or cluster) of d_1. Note that each sensor in the clusters of d_1, d_2, and d_3 has the reliability, schedule, and path properties. Also note that the distance between s_1 and d_1 ($= \frac{1}{1} + \frac{1}{1} + \frac{1}{0.49} + \frac{1}{1} = 5.04$) is smaller than that between s_2 and d_1 ($= \frac{1}{0.49} + \frac{1}{0.49} + \frac{1}{1} = 5.08$). Similarly, the distance between s_1 and d_2 is smaller than that between s_2 and d_2. Thus, in constructing clusters of d_1 and d_2, s_1 is considered to be included before s_2. Hence, the cluster of d_2 includes s_1. In contrast, the cluster of d_1 includes s_2 rather than s_1 since s_1 does not have the schedule property (due to that it requires 4 time slots to include s_1 while only 3 time slots remain) at that time. (c) Two infeasible DODAG establishment as clusters of d_1 and d_2 are selected to contains all sensors, where either the DODAG of d_1 contains s_4 (in the left-hand DODAG establishment) or the DODAG of d_2 contains s_4 (in the right-hand DODAG establishment), but not both (resulting in the disconnection between s_1 and d_2 or between s_2 and d_1) due to that the DODAGs of d_1 and d_2 must be disjoint using RPL. (d) Another cluster construction and schedule. Suppose that clusters of d_1, d_2, and d_3 are constructed in order. Note that when constructing the cluster of d_2, s_3 has been included in the cluster of d_1 while s_1 is not. Thus, the cluster of d_2 includes s_2 rather than s_1 since it violates the subset property to include s_1. (e) The feasible DODAG establishment where clusters of d_1 and d_3 are selected to contain all sensors. (f) The schedules of d_1 and d_3. Note that all sensors in a DODAG use the same channel to transmit the data since they are scheduled to transmit the data in different time slots. Also note that sensors in the DODAGs of d_1 and d_3 use channels ch_1 and ch_2, respectively, to avoid interference of each other.

or, in case of parity, the minimum distance from the sink, or, in case of parity, the minimum number of hops from the sink, or, in case of parity, the minimum ID, first includes the sensor. After the DODAGs of the selected sinks include all sensors, each DODAG is assigned a channel using the graph coloring strategy in [26] such that any two neighboring DODAGs are assigned two distinct channels, where two DODAGs are neighbors if they respectively include two sensors that are neighbors.

Theorem 4. MSPLRG-G constructs a DAG consisting of disjoint DODAGs of selected sinks.

Proof. Observe that a sensor s is included in the DODAG of a selected sink d (C'_d) only if s is in the cluster of d (C_d) and the parent of s in C_d is in C'_d. Thus, C'_d is a subtree of C_d for each selected sink d. This implies that each sensor in any DODAG has the reliability, schedule, and path properties. We only need to show each sensor is included in exactly one DODAG. Each sensor is included in at most one DODAG since a sensor is included in a DODAG only if it is not included in any DODAG. It suffices to show each sensor s can always be included in at least one DODAG. Suppose that s is unable to be included in any DODAG. Clearly, s is in the cluster of at least one selected sink, say d. Since s has the subset property in the cluster of d (C_d) and s is unable to be included in the DODAG of d (C'_d), the parent of s in C_d is unable to be included in C'_d. This implies the child of d that is an ancestor of s in C_d is unable to be included in C'_d by repeating the similar argument, a contradiction. □

V. Simulations

A. Simulation Setup

Network instances: We conduct simulations on the real data of the locations of premises (as the locations of smart meters) from the Google Maps and the locations of poles (as the locations of concentrators) from the government open data [18] which are compiled and available in [27]. We assume all smart meters and concentrators are equipped with CC2420 RF transceivers [19]. To evaluate the delivery probability of a link, we first evaluate $SNR_{R,\mathrm{dB}}(d)$ by

TABLE I
THE SETTINGS OF PARAMETERS

$P_{T,\mathrm{dBm}}$ (output power)	$0\,\mathrm{dBm}$ [19]
$S_{R,\mathrm{dBm}}$ (receiver sensitivity)	$-95\,\mathrm{dBm}$ [19]
$G_{T,\mathrm{dB}}$ (transmitter antenna gain)	$3\,\mathrm{dB}$ [19]
$G_{R,\mathrm{dB}}$ (receiver antenna gain)	$3\,\mathrm{dB}$ [19]
PL_0	21.3 [20]
n_0	3.6 [20]
σ	7.4 [20]
k (Boltzmann's constant)	$1.38 \times 10^{23}\,\mathrm{J\,K^{-1}}$ [20]
T_0	$293\,\mathrm{K}$ [20]
F (noise figure)	$5\,\mathrm{dB}$ [20]

$EIRP_{\mathrm{dBm}} + G_{R,\mathrm{dB}} - PL_{\mathrm{dB}}(d) - N_{0,\mathrm{dBm}}$ and $SNR_{REQ,\mathrm{dB}}$ by $S_{R,\mathrm{dBm}} - N_{0,\mathrm{dBm}}$, where $EIRP_{\mathrm{dBm}} = P_{T,\mathrm{dBm}} + G_{T,\mathrm{dB}}$, $PL_{\mathrm{dB}}(d) = PL_0 + n_0 \log_{10}(d)$, in which d is the transmitter-receiver distance, and $N_{0,\mathrm{dBm}} = 30 + \log_{10}(kT_0F)$. Then, we evaluate $F_X(x)$ by $\int_0^\infty [1 - \frac{1}{2}\,\mathrm{erfc}(\frac{x - 10\log_{10} y}{7.4\sqrt{2}})]e^{-y}dy$, where $x = SNR_{REQ} - SNR_R$. Finally, we evaluate the delivery probability of a link by $1 - F_X(x)$. The settings of necessary parameters are described in Table I.

Comparison methods and performance metrics: To evaluate the performance in the number of selected concentrators, we compare MSPLRG-G with MSP-G [7], [8]. In MSP-G, the delivery probability of a link is 1 if the distance between two transceivers over the link is within the transmission range of the transceiver, and 0 otherwise. In our simulations, the transmission range of the transceiver is set to $163\,\mathrm{m}$ since the difference between the output power ($P_{T,\mathrm{dBm}}$) and the path loss ($PL_{\mathrm{dB}}(d)$) is less than the receiver sensitivity ($S_{R,\mathrm{dBm}}$) as the distance between two transceivers over the link is greater than $163\,\mathrm{m}$. In MSP-G, the cluster of each concentrator contains all smart meters within a number of hops depending on the required latency. In addition, to evaluate the performance of the greedy strategy for the selection of concentrators employed in MSPLRG-G, we compare MSPLRG-G with MSPLRG-PSO which uses the Particle Swarm Optimization (PSO) method [9], [10], a randomized algorithm, to select concentrators based on the clusters of concentrators constructed by MSPLRG-G. Besides, we also evaluate the number of channels assigned by the graph coloring strategy in [26] when employing MSP-G, MSPLRG-PSO, and MSPLRG-G.

Moreover, to evaluate the performance in the balanced size of the DODAGs, we compare the DODAG establishment of MSPLRG-G with NAIVE and MIN-FIRST based on the clusters of the selected concentrators constructed by MSPLRG-G. NAIVE chooses one selected concentrators at random and establish the DODAG of the chosen concentrator by removing all smart meters that have been included in other DODAGs. MIN-FIRST is the same as the DODAG establishment of MSPLRG-G except that MIN-FIRST includes a smart meter without evaluating $x^*_{s,d}$ for sensor s and selected sink d.

B. Simulation Results

The number of selected concentrators and assigned channels: Fig. 4 shows the results as the assigned reliability

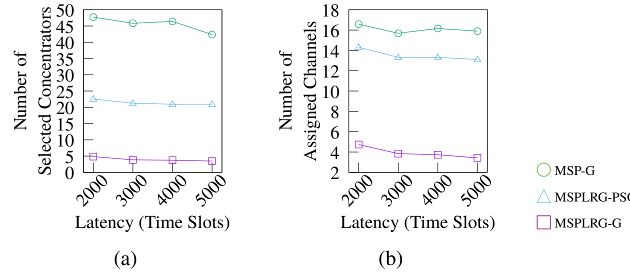

Fig. 4. The impact of the required latency on the numbers of (a) the selected concentrators and (b) the assigned channels for the clusters of the selected concentrators using MSP-G, MSPLRG-PSO, and MSPLRG-G.

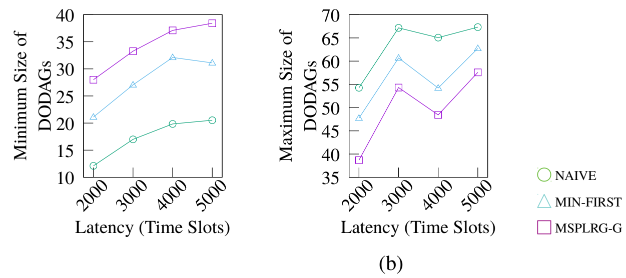

Fig. 5. The impact of the required latency on the (a) minimum and (b) maximum sizes of the DODAG by NAIVE, MIN-FIRST, and MSPLRG-G.

is 0.99 and the required latency is 2000, 3000, 4000, or 5000 times slots. First, since some smart meters in the clusters of the selected concentrators by MSP-G do not meet the reliability requirement, some additional concentrators are required to ensure each smart meter has the reliability property. Thus, MSP-G selects more concentrators than MSPLRG-G and MSPLRG-PSO. This results also confirms the conjecture that if we construct and select the clusters of concentrators without addressing the reliability guarantee of smart meters, some smart meters would not have the reliability property. Second, MSPLRG-G selects less concentrators than MSPLRG-PSO. This implies a greedy algorithm usually outperforms a randomized algorithm in the selection of concentrators. As anticipated, the greater the required latency the smaller the number of the selected concentrators. In addition, it can be seen the result for the number of selected concentrators is similar to that for the number of assigned channels.

The minimum and maximum sizes of the DODAGs: As can be seen in Fig. 5, NAIVE has the worst performance in terms of the minimum size of the DODAGs among NAIVE, MIN-FIRST, and MSPLRG-G, which is reasonable. Besides, MIN-FIRST locally identify the best-fit smart meter to be included in the cluster of a concentrator by a greedy strategy while MSPLRG-G globally identify the best-fit one using a linear program, resulting in that MSPLRG-G outperforms MIN-FIRST. The result for the maximum size of the DODAGs is similar to that for the minimum size of the DODAGs.

VI. Conclusion

In this paper, we study the scenario of deploying concentrators in the establishment of an AMI network, model the scenario as the Minimum Sink Placement with Latency and Reliability Guarantee (MSPLRG) problem in an LWSN, and show the MSPLRG problem is NP-hard. Then, we propose an algorithm for the MSPLRG problem. To ensure that the proposed algorithm meets the latency and reliability requirements in a real AMI network, we jointly considers the routing protocol RPL [1], the MAC protocol TSCH [2], and the scheduling approach DeTAS [3].

Our algorithm first constructs a cluster for each sink, then selects the sinks, and finally establish the disjoint DODAGs of selected sinks, where the DODAG of a sink is used to route the data from the sensors in the DODAG to the sink based on RPL. In the phase of cluster construction, we ensure each sensor has the reliability, schedule, path, and subset properties in a cluster to ensure disjoint DODAGs of selected sinks can be established in the phase of DODAG establishment and any sensor in each DODAG can meet the latency and reliability requirements. In the phase of sink selection, we enumerate all slotframe sizes to select sinks using the greedy strategy for the Minimum Unweighted Set Cover problem to ensure the number of selected sinks is theoretically bounded by the optimal number, where the slotframe consists of consecutive time slots and continuously repeats over time based on TSCH. In the phase of DODAG establishment, to keep the balanced sizes of the DODAGs, we evaluate the worth of a sensor in the DODAG of a sink by formulating a linear program.

We conduct simulations on the real data of the locations of premises (as the locations of smart meters) from the Google Maps and the locations of poles (as the locations of concentrators) from the government open data [18], where all smart meters and concentrators are assumed to be equipped with CC2420 RF transceivers [19]. The simulation results show our algorithm in the phases of cluster construction and sink selection has a better performance in terms of the numbers of selected concentrators and assigned channels, as compared to MSP-G [7], [8] and MSPLRG-PSO [9], [10]. The simulation results also show our algorithm in the phase of DODAG establishment has a better performance in terms of the balanced sizes of the DODAGs, as compared to two naive heuristics NAIVE and MIN-FIRST.

References

[1] T. Winter, P. Thubert, A. Brandt, J. Hui, and R. Kelsey. (2012) RFC 6550: RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks. [Online]. Available: https://tools.ietf.org/html/rfc6550

[2] "IEEE Standard for Local and metropolitan area networks–Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs) Amendment 1: MAC sublayer," *IEEE Std 802.15.4e-2012 (Amendment to IEEE Std 802.15.4-2011)*, pp. 1–225, April 2012.

[3] N. Accettura, E. Vogli, M. R. Palattella, L. A. Grieco, G. Boggia, and M. Dohler, "Decentralized traffic aware scheduling in 6TiSCH networks: Design and experimental evaluation," *IEEE Internet of Things Journal*, vol. 2, no. 6, pp. 455–470, 2015.

[4] R. R. Mohassel, A. Fung, F. Mohammadi, and K. Raahemifar, "A survey on advanced metering infrastructure," *International Journal of Electrical Power & Energy Systems*, vol. 63, pp. 473–484, 2014.

[5] P. Kulkarni, S. Gormus, Z. Fan, and F. Ramos, "AMI mesh networks – A practical solution and its performance evaluation," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1469–1481, 2012.

[6] SG Network System Requirements Specification Interim Release 3. [Online]. Available: https://energy.gov/gc/downloads/sg-network-system-requirements-specification-interim-release-3

[7] L. Sitanayah, K. N. Brown, and C. J. Sreenan, "Planning the deployment of multiple sinks and relays in wireless sensor networks," *Journal of Heuristics*, vol. 21, no. 2, pp. 197–232, 2015.

[8] Y. Bejerano, "Efficient integration of multihop wireless and wired networks with QoS constraints," *IEEE/ACM Transactions on Networking*, vol. 12, no. 6, pp. 1064–1078, 2004.

[9] H. Safa, W. El-Hajj, and H. Zoubian, "A robust topology control solution for the sink placement problem in WSNs," *Journal of Network and Computer Applications*, vol. 39, pp. 70–82, 2014.

[10] C. Zhao, C. Wu, X. Wang, B. W.-K. Ling, K. L. Teo, J.-M. Lee, and K.-H. Jung, "Maximizing lifetime of a wireless sensor network via joint optimizing sink placement and sensor-to-sink routing," *Applied Mathematical Modelling*, vol. 49, pp. 319–337, 2017.

[11] E. I. Oyman and C. Ersoy, "Multiple sink network design problem in large scale wireless sensor networks," in *IEEE ICC*, vol. 6, 2004, pp. 3663–3667.

[12] N. A. Pantazis, S. A. Nikolidakis, and D. D. Vergados, "Energy-efficient routing protocols in wireless sensor networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 551–591, 2013.

[13] J. A. Sanchez, P. M. Ruiz, J. Liu, and I. Stojmenovic, "Bandwidth-efficient geographic multicast routing protocol for wireless sensor networks," *IEEE Sensors Journal*, vol. 7, no. 5, pp. 627–636, 2007.

[14] K. Sohrabi, J. Gao, V. Ailawadhi, and G. J. Pottie, "Protocols for self-organization of a wireless sensor network," *IEEE Personal Communications*, vol. 7, no. 5, pp. 16–27, 2000.

[15] X. Yang, J. Guo, P. Orlik, K. Parsons, and K. Ishibashi, "Stability metric based routing protocol for low-power and lossy networks," in *IEEE ICC*, 2014, pp. 3688–3693.

[16] J. Ko, A. Terzis, S. Dawson-Haggerty, D. E. Culler, J. W. Hui, and P. Levis, "Connecting low-power and lossy networks to the internet," *IEEE Communications Magazine*, vol. 49, no. 4, 2011.

[17] S. Gormus, F. Tosato, Z. Fan, Z. Bocus, and P. Kulkarni, "Opportunistic RPL for reliable AMI mesh networks," *Wireless Networks*, vol. 20, no. 8, pp. 2147–2164, 2014.

[18] [Online]. Available: https://data.gov.tw/dataset/33305

[19] Single-Chip 2.4 GHz IEEE 802.15.4 Compliant and ZigBee Ready RF Transceiver. [Online]. Available: http://www.ti.com/product/CC2420

[20] M. Souryal, C. Gentile, D. Griffith, D. Cypher, and N. Golmie, "A methodology to evaluate wireless technologies for the smart grid," in *IEEE SmartGridComm*, 2010, pp. 356–361.

[21] D. S. De Couto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multi-hop wireless routing," *Wireless Networks*, vol. 11, no. 4, pp. 419–434, 2005.

[22] W. T., P. M., and G. L. (2015) Using IEEE 802.15.4e Time-Slotted Channel Hopping (TSCH) in the Internet of Things (IoT): Problem Statement. [Online]. Available: https://tools.ietf.org/html/rfc7554

[23] R. M. Karp, "Reducibility among combinatorial problems," in *Complexity of Computer Computations*. Springer, 1972, pp. 85–103.

[24] D. S. Johnson, "Approximation algorithms for combinatorial problems," *Journal of Computer and System Sciences*, vol. 9, no. 3, pp. 256–278, 1974.

[25] V. Chvatal, "A greedy heuristic for the set-covering problem," *Mathematics of Operations Research*, vol. 4, no. 3, pp. 233–235, 1979.

[26] S. Ramanathan, "A unified framework and algorithm for channel assignment in wireless networks," *Wireless Networks*, vol. 5, no. 2, pp. 81–94, 1999.

[27] [Online]. Available: http://dclab.cs.nthu.edu.tw/˜yushu/grids.tar.gz