Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 891321, 12 pages
doi:10.1155/2012/891321

Research Article

Optimal Convergecast Scheduling Limits for Clustered Industrial Wireless Sensor Networks

Xiaoling Zhang, Wei Liang, Haibin Yu, and Xisheng Feng

Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China

Correspondence should be addressed to Haibin Yu, yhb@sia.cn

Received 27 July 2012; Revised 12 October 2012; Accepted 14 October 2012

Academic Editor: Hongwei Zhang

Increased mobility coupled with a possible reduction of cabling costs and deployment time makes wireless communication an attractive alternative for the industrial process monitoring and control. The major obstacles toward the utilization of wireless industrial networks are predominantly the timeliness and reliability requirements. In this paper, orienting to clustered industrial wireless sensor networks, we analyze the performance bounds of the convergecast scheduling, which is a typical many-to-one communication in wireless industrial networks. The analysis aims to the cluster-line and cluster-tree topologies. Each kind is future divided into three scenarios according to an application parameter, named data update rate. Firstly, we establish the lower bounds on the number of timeslots to finish the intracluster and the intercluster convergecast transmissions. Secondly, we establish the lower bounds on the number of channels based on the lower bounds on the number of timeslots and maximum available channels in a multichannel scenario. Lastly, we carry out the extensive analysis-taking packet retransmissions into consideration so as to meet the reliability requirement. Experiment results validate the correctness and tightness of our theory analysis.

1. Introduction

With the success of wireless technologies in consumer electronics, standard wireless technologies are envisioned for the deployment in industrial environments to improve the functionality and the efficiency, which boost the formation of the industrial wireless sensor networks (IWSNs) [1]. Recently, some international organizations are actively promoting the standardized process of IWSNs and have achieved several productions, such as WirelessHART [2] and Wireless network for industrial automation-process automation (WIA-PA) [3]. Convergecast is a typical many-to-one communication pattern in IWSNs. In convergecast, many or all nodes in the network send data to a sink node during a relatively short time period [4]. In these applications all packets generated in the network have to reach a sink node either for record or for computationally intensive analysis.

A guarantee on packet delivery and a bound on convergecast latency are highly desirable in mission critical applications. Guarantee on packet delivery ensures availability of accurate information about the sensing field. An optimal bound on convergecast latency leads to timely detection of the events. Furthermore, a known time bound on convergecast latency can help the sink node to schedule convergecast requests and related computations [4].

As concluded in [4], collisions present a major challenge in convergecast when contention-based MAC protocol like CSMA are employed. Collisions will result in loss of packets. The recovery methods, such as retransmissions method, will increase the latency. Therefore, the convergecast latency incurred by radio interference is far from the optimal [4]. Contention-free MAC protocols like TDMA can be used to eliminate collisions and obtain a bound on the time required to complete convergecast. Hence, we are interested in determining a TDMA schedule such that the entire convergecast can be completed in minimal number of timeslots and in minimal number of channels.

Choi et al. [5] proved that convergecast scheduling problem is NP-hard. Florens et al. [6] consider that the convergecast is the reverse operation of the broadcast scheduling. If node i sends packets to node $i + 1$ by using timeslot t during broadcast scheduling, then node $i + 1$ send packets to node i by using timeslot $T - i + 1$. T is the minimal schedule length and is calculated according to the number of packets that

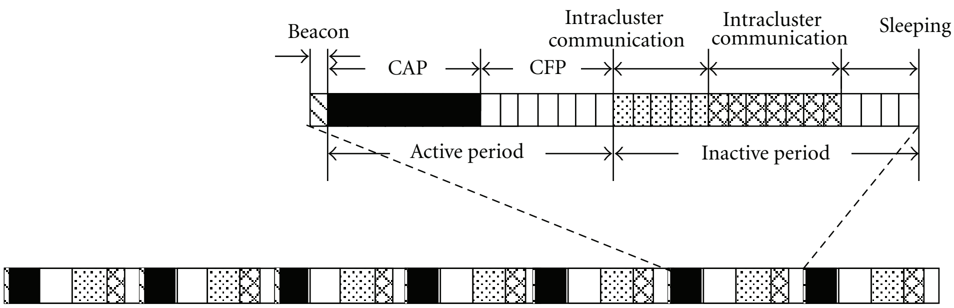

Figure 1: Superframe structure.

sink node sends to node i. Gandham et al. [4] proved that once convergecast scheduling requires at most $3N$ timesltos, where N represents the number of nodes in the network. They also proved that at lease $3N-3$ timeslots are required to complete convergecast in a linear network with N nodes; at least $\max(3n_k-3, N)$ timeslots are required to complete convergecast in a multiline network; at least $\max(3n_k-1, N)$ timeslots are required to complete convergecast in a tree network. Yu et al. [7] consider a network as a disc and proved that the upper bound on the number of timeslots is $24D+6\Delta+16$, where D is the network diameter and Δ is the maximum node degree. Soldti [8, 9] consider the jointly time- and channel-optimal convergecast scheduling, and proved the lower bound on the number of timeslots and channels for convergecast transmission in mesh network. Shang et al. [10] considered how to collect data from sensors deployed in the Euclidean plane in a time-efficient way. Zhang et al. [11] focused on bursty convergecast and proposed a windowless block acknowledgement scheme to achieve reliable and real-time error control and contention control. Liu et al. [12] estimated the probabilistic path delay bounds in an agile manner and designed a delay-sensitive routing protocol.

In summary, there are three problems in previous researches. First, the network topology is ad hoc or mesh and the network deployment is plane or quadrate, which are not suitable for industrial applications. Second, the cyclic data feature is not taken into consideration. Third, previous research is based on one channel, which are inherently inefficient. To solve these three problems, a new result specific to clustered IWSNs, where in particular, cyclic data collection over the network is assumed. In addition, we consider data update rates (DURs) under various assumptions and derive corresponding time and channel performance. This work extends the results in [8], where only mesh topology is assumed and the DURs are not taken into consideration. It should be noted that most of industrial networks use hierarchical topology and faced to different applications with different data features. Therefore, when requirements are more stringent, the results in [8] may no longer be practical. In particular, this paper presents the following contributions.

(i) For a cluster-line IWSN, we analyze and prove the performance limits for completing the convergecast transmission. We establish the lower bounds on the number of timeslots and channels based on three scenarios: all the data update rates (DURs) in a cluster are same, all the DURs in a cluster are different, and the DURs in a cluster are partly same.

(ii) For a cluster-tree IWSN, the same problem as the first item is studied.

The rest of this paper is organized as follows. First, the system model and problem formulation of convergecast scheduling are described in Section 2. The analysis of the time- and channel-optimal convergecast scheduling for networks with cluster-line and cluster-tree topologies are studied in Sections 3 and 4, respectively. This paper is concluded in Section 5.

2. Problem Formulation

Time is equally divided into timeslots. A collection of timeslots repeating constitute one superframe. Our superframe structure extends the IEEE STD 802.15.4-2006 superframe [13], and divides the inactive period into intracluster communication period, intercluster communication period, and sleep period. The extended superframe structure is shown in Figure 1. The superframe length of the gateway device or a routing device can be configured and the basic superframe length $\overline{S_f}$ is also configurable.

One field device initially generates at most one packet destined to the gateway device every DUR time during each convergecast. Packets are routed along $T=(V, E_r)$, where $E_r \subseteq E$ is the set of routing paths. For every device i, $I(i)$ and $O(i)$, respectively, denote the set of links that arrive at or leave device i; $p_t(i)$ denotes the capability of buffer after timeslot t, where we defined that $p_0(i)=1$. Let g_{ij} be the number of packets of a cluster with R_{ij} being the cluster head in one superframe cycle. g_{ij} is decided by the DURs of all field devices in a cluster. Let L_s be the superframe length (in timeslots). Let x_{ijt} be a binary variable indicating whether timeslot t has been allocated to link (i, j). If timeslot t has been allocated to link (i, j), $x_{ijt}=1$; otherwise, $x_{ijt}=0$.

The time- and channel-optimal scheduling to complete convergecast with the objective to minimize both the number of timeslots and the number of channels can be informally formulated as [8].

Problem 1. Time-optimal convergecast scheduling:

$$\text{minimize}\, L_s, \tag{1}$$

s.t.

$$\sum_{(i,j)\in I(i)} x_{ijt} + \sum_{(i,j)\in O(i)} x_{ijt} \le 1, \quad \forall i \in V,\ \forall t \in [0, L_s], \tag{2}$$

$$p_{L_s}(\mathrm{GW}) = \sum_{i,j} g_{ij}, \tag{3}$$

$$p_t(i) = p_{t-1}(i) + \sum_{(i,j)\in I(i)} x_{ijt} - \sum_{(i,j)\in O(i)} x_{ijt}, \quad \forall i \in V,\ \forall t \in [0, L_s], \tag{4}$$

$$x_{ijt} \in \{0,1\}, \quad \forall (i,j) \in E_r. \tag{5}$$

Objective (1) guarantees the minimization of the convergecast scheduling length. Constraint (2) is the half-duplex requirement, which restricts devices not to transmit and receive simultaneously. Constraint (3) guarantees that all packets generated by field devices have been sent to the gateway device. Constraint (4) is the memory status. And constraint (5) indicates the link status.

Problem 2. Time- and channel-optimal convergecast scheduling:

$$\text{minimize}\, C_s = \max_{t\in[0,L_s]} \sum_{(i,j)\in E_r} x_{ijt}, \tag{6}$$

s.t.

$$L_s\ \text{calculated by}\ (1) \sim (5), \tag{7}$$

where C_s is the number of channels used for transmission.

3. Time- and Channel-Optimal Convergecast Scheduling for Cluster-Line Topology

Cluster-line topology that is shown in Figure 2 is instrumental for more general routings, and is the preferred topology in certain applications such as pipeline monitoring and unmanned offshore gas production [8]. Without loss of generality, the GateWay device (GW) is placed at the right end of the line, and N routing devices $R_{11} \sim R_{1N}$ are placed from right to left. For routing device R_{1i}, $i \in [1, N]$, it acts as a cluster head and manages several field devices. f_{1i} denotes the number of field devices in the cluster with R_{1i} being cluster head; and F_{1ij} denotes the jth field device, $j \in [0, f_{1i}]$; C_{1ij} denotes the DUR of F_{1ij}. The following content studies three scenarios: all the DURs in a cluster are same, all the DURs in a cluster are different, and the DURs in a cluster are partly same, which are briefly denoted as ADS, ADD, and PDS.

3.1. Scenario I: ADS. Each field device in this cluster initially generates one packet at the begin of each superframe cycle. Considering the single-packet memory, the generated data must be sent out before the end of each superframe. Otherwise, the old data will be overlapped by newly generated data.

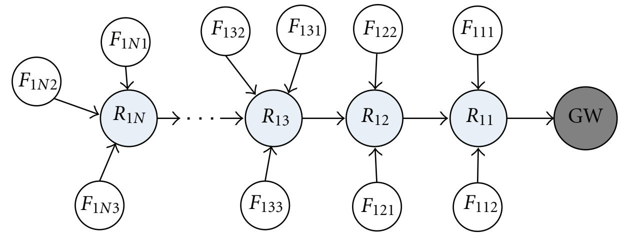

Figure 2: Cluster-line network topology.

3.1.1. Performance Limits on Intracluster Convergecast Scheduling. A routing devices manages one intracluster convergecast scheduling, and assigns timeslots and channels for communications between it and field devices. Because of the half-duplex radio transceiver, the length of intracluster convergecast scheduling should not be less than the number of field devices. For network scalability, we assume that the intracluster communication periods of all clusters are equal, which is defined as $\max_{i\in[1,N]} f_{1i}$. Each field device is assigned one timeslot during the intracluster communication period.

The intracluster convergecast scheduling is simple and intuitive. In one cluster, different field devices using different timeslots, while using one channel can eliminate collisions completely. We only need to consider the interference among clusters. Neighboring clusters using different channels can avoid collisions, which will be an extension in our future work.

3.1.2. Performance Limits on Intercluster Convergecast Scheduling. GW manages the intercluster convergecast scheduling, and allocates timeslots and channels for communications among routing devices and GW. The routing device nearest GW such as R_{11} in Figure 2 is the "bottleneck" of all communications, so that the lower bound on the number of timeslots for intercluster convergecast scheduling is determined by R_{11}.

Theorem 1. *If all the data update rates of field devices in a cluster are same, that is $C_{1i1} = C_{1i2} = \cdots = C_{1if_{1i}}$ and $i = 1, 2 \ldots N$, the lower bound on the number of timeslots used for the intercluster convergecast in a cluster-line routing network is $L_s = 2\sum_{i=2}^{N} f_{1i} + f_{11}$.*

Proof. At the beginning of each intercluster communication period, if $C_{1i1} = C_{1i2} = \cdots = C_{1if_{1i}}$ $(i = 1, 2 \ldots N)$, the number of packets in R_{1i} satisfies $g_{1i} = f_{1i}$. In Figure 2, R_{11} is responsible for receiving and forwarding packets from $R_{12} \sim R_{1N}$ to GW; R_{11} is also responsible for forwarding g_{11} packets in its cluster to GW. Due to the half-duplex constraint, R_{11} needs $2\sum_{i=2}^{N} g_{1i} + g_{11}$ timeslots. Reminded that $g_{1i} = f_{1i}$, then R_{11} needs $2\sum_{i=2}^{N} f_{1i} + f_{11}$ timeslots to forward all packets to GW. Therefore, the lower bound on the number of timeslots used for the intercluster convergecast in a cluster-line network is $L_s = 2\sum_{i=2}^{N} f_{1i} + f_{11}$. □

Theorem 2. *Given any scheduling algorithm S that can complete intercluster convergecast in $L_s = 2\sum_{i=2}^{N} f_{1i} + f_{11}$ timeslots* *in a cluster-line routing network, the lower bound on the number of channels used in S is*

$$\left\lceil \frac{2\sum_{i=2}^{N} f_{1i} + f_{11} + 1}{2} - \sqrt{\left(\frac{2\sum_{i=2}^{N} f_{1i} + f_{11} + 1}{2}\right)^2 - \sum_{i=1}^{N} i f_{1i}} \right\rceil. \tag{8}$$

Proof. Let C_s be the number of channels that can be used in a network, and $\mathrm{PT}_{\max}(t)$ denotes the number of parallel transmissions scheduled in timeslot t ($t \in [1, L_s]$). In order to complete convergecast in L_s timeslots, the following conditions should be satisfied.

(i) When $t = L_s$, there should be only one packet in the network and R_{11} forwards it to GW. That is, only one device could be scheduled to transmitted if $t = L_s$.

(ii) When $t = L_s - 1$, there exist two situations. First, R_{12} sends its last packet to R_{11} in the timeslot t; and R_{11} forwards this packet to GW in timeslot L_s. Second, there exist only two packets in R_{11}; and R_{11} sends these two packets to GW, respectively, in timelsot $L_s - 1$ and L_s. In all, only one device could be scheduled to transmitted if $t = L_s - 1$.

(iii) Deduced as so forth, the maximum number of devices allowed to transmit in timeslot t is $\lceil (L_s - t + 1)/2 \rceil$.

Since $\lceil (L_s - t + 1)/2 \rceil \geq C_s$, when $t \in [1, L_s - 2(C_s - 1)]$ and the upper bound on the number of channels is C_s, we can obtain the following conclusion as in [8]:

$$\mathrm{PT}\max(t) = \begin{cases} C_s, & t \in [1, L_s - 2(C_s - 1)] \\ \left\lceil \dfrac{L_s - t + 1}{2} \right\rceil, & t \in [L_s - 2(C_s - 1) + 1, L_s]. \end{cases} \tag{9}$$

By (9), the maximum number of transmissions that can be scheduled in L_s using C_s channels is

$$\begin{aligned} \sum_{t=1}^{L_s} \mathrm{PT}\max(t) &= \sum_{t=1}^{L_s - 2(C_s - 1)} C_s + \sum_{t=L_s-2(C_s-1)+1}^{L_s} \left\lceil \frac{L_s - t + 1}{2} \right\rceil \\ &= -C_{s^2} + C_s(L_s + 1). \end{aligned} \tag{10}$$

For intercluster convergecast, only routing devices and GW along the line routing transmit packets. Thus, the number of transmissions for completing convergecast is at least $\sum_{i=1}^{N} i f_{1i}$. Combined with (10), the number of transmissions satisfy the following inequality:

$$-C_{s^2} + C_s(L_s + 1) \geq \sum_{i=1}^{N} i f_{1i}. \tag{11}$$

Replacing L_s with $2\sum_{i=2}^{N} f_{1i} + f_{11}$, the lower bound on the number of channels used to complete convergecast in $2\sum_{i=2}^{N} f_{1i} + f_{11}$ timeslots is

$$\left\lceil \frac{2\sum_{i=2}^{N} f_{1i} + f_{11} + 1}{2} - \sqrt{\left(\frac{2\sum_{i=2}^{N} f_{1i} + f_{11} + 1}{2}\right)^2 - \sum_{i=1}^{N} i f_{1i}} \right\rceil. \tag{12}$$

□

From Theorems 1 and 2, we can discover that the results deduced in [8, 9] is special cases of our work when all the data update rates of field devices in a cluster are same.

3.2. Scenario II: ADD. If all the DURs of field devices in a cluster with R_{1i} ($i \in [1, N]$) being cluster head are different, is satisfied the following.

$$C_{1ij} \neq C_{1ik}, \quad \forall j, k \in [1, f_{1i}],\ j \neq k,\ i \in [1, N]. \tag{13}$$

3.2.1. Performance Limits on Intracluster Convergecast Scheduling. Field devices with different DURs can realize the timeslot multiplex. For example, if the superframe cycle is 1 and the DURs of two field device R_{11i} and R_{11j} ($i \neq j$) are 2 and 4, respectively, then the timeslot of R_{11j} can be occupied by R_{11i} when R_{11j} has no data to be transmitted.

Theorem 3. *Given a cluster with more than one field device, if all the data update rates of field devices are different, that is $C_{1ij} \neq C_{1ik}$, for all j, $k \in [1, f_{1i}]$, $j \neq k$ & $i \in [1, N]$, the lower bound on the number of timeslots used for the intracluster convergecast in a cluster-line routing network is 2.*

Proof. Obviously, for a cluster with only one field device, the intracluster convergecast can be completed by using only one timeslot. We consider the cluster with more than one field device below. For a cluster with R_{1i} ($i \in [1, N]$) being cluster head, the number of field devices is f_{1i} ($f_{1i} \geq 2$) and DUR of the jth field device is C_{1ij} ($j \in [1, f_{1i}]$). We assume that the DUR of a field device is 2^n, where n is a natural number. Without loss of generality, we suppose that $C_{1i1} < C_{1i2} \ldots < C_{1if_{1i}} = 2^{\Theta}$ (Θ is a natural number).

Denoted that the superframe length of R_{1i} is $\overline{S_f} \times C_{1i1}$. $\overline{S_f}$ denotes the basis superframe length. F_{1ij} that has the minimal DUR needs one timeslot per superframe cycle to transmit its packet. Other field devices cannot occupy its timeslot and only one timeslot can not finish the intracluster convergecast.

During $C_{1if_{1i}}/C_{1i1}$ superframe cycles, the first timeslot in every intracluster communication period is allocated to F_{1i1} and is denoted as $TC_{1i1} + k_1(C_{1i1} \times \overline{S_f})$, where TC_{1i1} is the index of the timeslot during the first superframe cycle and $k_1 = 0, 1, \ldots (C_{1if_{1i}}/C_{1i1} - 1)$.

For F_{1i2}, the available timeslot is the second timeslot in each intracluster communication period. We choose one timeslot numbered as 2 in the intracluster communication period from the first C_{1i2}/C_{1i1} superframe cycles and denoted as TC_{1i2}, which is to guarantee the packet in F_{1i2} can be sent out before DUR. Since F_{1i2} updates its data every $\overline{S_f} \times (C_{1i2}/C_{1i1})$ timeslots, F_{1i2} will occupied the timeslot indexed by $TC_{1i2} + k_2(C_{1i2} \times \overline{S_f}/C_{1i1})$ during $C_{1if_{1i}}/C_{1i1}$ superframe cycles, where $k_2 = 0, 1, \ldots (C_{1if_{1i}}/C_{1i2} - 1)$. During $C_{1if_{1i}}/C_{1i1}$ superframe cycles, the occupied number of timeslots number as 2 in the intracluster communication period is $C_{1if_{1i}}/(C_{1i1} \times C_{1i2})$ and the remainder number of the second timeslots in the intracluster communication period is $(C_{1if_{1i}}/C_{1i2}) \times (C_{1i2}/C_{1i1} - 1)$.

The timeslot assignment for the rest field devices can be done in the same manner as F_{1i2}. In all, the available timeslot for F_{1ij} $(j \in [1, f_{1i}])$ is the second free timeslot in the intracluster communication period during the first C_{1ij}/C_{1i1} superframe cycles and occupied number is $C_{1if_{1i}}/C_{1ij}$ during $C_{1if_{1i}}/C_{1i1}$ superframe cycles.

For the last field device $F_{1Nf_{1N}}$, the number of idle timeslot numbered as 2 in the intracluster communication period during $C_{1if_{1i}}/C_{1i1}$ superframe cycles is:

$$\begin{aligned}&\frac{C_{1if_{1i}}}{C_{1i1}} - \frac{C_{1if_{1i}}}{C_{1i2}} - \frac{C_{1if_{1i}}}{C_{1i3}} - \cdots C_{1if_{1i}}/C_{1i(f_{1i}-1)} \\ &\quad \geq 2^{\Theta} - 2^{\Theta}/2 - 2^{\Theta}/4 - \cdots 2^{\Theta}/2^{(\Theta-1)} = 2,\end{aligned} \tag{14}$$

which illustrates that the number of free timeslot number as 2 in the intracluster communication period during $C_{1if_{1i}}/C_{1i1}$ superframe cycles is enough for all field devices completing their intracluster convergecast. □

3.2.2. Performance Limits on Intercluster Convergecast Scheduling. The following Corollaries 4 and 5 can be deduced from Theorems 1 and 2.

Corollary 4. *If all the data update rates of field devices in a cluster are different, that is $C_{1ij} \neq C_{1ik}$, for all $j, k \in [1, f_{1i}]$, $j \neq k$ & $i \in [1, N]$, the lower bound on the number of timeslots used for the intercluster convergecast in a cluster-line routing network is $4N - 2$, where N is the number of routing devices in a network.*

Proof. Let g_{1i} be the number of packets in R_{1i} $(i \in [1, N])$ at the beginning of each superframe cycle. The proving approach is similar to Theorem 1. The only difference is that the lower bound of g_{1i} is 2 because of Theorem 3. R_{11} needs $2\sum_{i=2}^{N} g_{1i}$ timeslots for forwarding packets from $R_{12} \sim R_{1N}$ and needs g_{11} timeslots for forwarding own packets to GW. In all, the lower bound on the number of timeslots used for intercluster convergecast scheduling is the sum of these two parts, which can be formulated as $L_s = 2\sum_{i=2}^{N} g_{1i} + g_{11} = 2 \times 2(N-1) + 2 = 4N - 2$. □

Corollary 5. *Given any scheduling algorithm S that can complete intercluster convergecast in $4N - 2$ timeslots in a cluster-line network, the lower bound on the number of channels used in S is $\lceil ((4N-1)/2) + \sqrt{(12N^2 - 12N + 1)/4} \rceil$, where N is the number of routing devices in a network.*

Proof. The proving approach is similar to Theorem 2. The only difference is to replace $L_s = 2\sum_{i=2}^{N} f_{1i} + f_{11}$ in Theorem 2 with $L_s = 4N - 2$. Then, the lower bound on the number of channels satisfy the following inequality:

$$C_s \geq \left\lceil \frac{(4N-2)+1}{2} - \sqrt{\left(\frac{(4N-2)+1}{2}\right)^2 - \sum_{i=1}^{N}(2i)} \right\rceil. \tag{15}$$

□

3.3. Scenario III: PDS. If the DURs of field devices in a cluster with R_{1i} $(i \in [1, N])$ being cluster head are partly same, the following is satisfied:

$$C_{1ij} = C_{1ik}, \quad \exists j, k \in [1, f_{1i}],\ j \neq k,\ i \in [1, N]. \tag{16}$$

3.3.1. Performance Limits on Intracluster Convergecast Scheduling. In this scenario, field devices with different DURs can also realize the timeslot multiplex.

Theorem 6. *Given a cluster with more than one field device and the data update rates of field devices in a cluster with R_{1i} being cluster head are partly same, that is $C_{1ij} = C_{1ik}$, $\exists j, k \in [1, f_{1i}]$, $j \neq k$, $i \in [1, N]$, the lower bound on the number of timeslots used for the intracluster convergecast in a cluster-line network is $\sum_{j=1}^{Q_{1i}} y_{1ij}$. N_{1ij} is the number of field devices with DUR being C_{1ij} and satisfies $\sum_{j=1}^{Q_{1i}} N_{1ij} = f_{1i}$ $(Q_{1i} \leq f_{1i})$; Q_{1i} is the number of DUR kinds in this cluster; y_{1ij} satisfies:*

$$\begin{aligned}
y_{1i1} &= \left\lceil \frac{N_{1i1} \times C_{1i1}}{C_{1i1}} \right\rceil = \lceil N_{1i1} \rceil; \\
y_{1i2} &= \left\lceil \frac{N_{1i2} \times C_{1i1}}{C_{1i2}} \right\rceil; \\
y_{1ik} &= \begin{cases} 0, & \dfrac{C_{1i1}\left(N_{1i(k-1)}C_{1ik} + N_{1ik}C_{1i(k-1)}\right)}{C_{1i(k-1)}C_{1ik}} < y_{1i(k-1)} \\ \left\lceil \dfrac{N_{1ik} - \left(y_{1i(k-1)}C_{1i(k-1)}/C_{1i1} - N_{1i(k-1)}\right) \times \left(C_{1ik}/C_{1i(k-1)}\right)}{C_{1ik}/C_{1i1}} \right\rceil, & \text{else,} \end{cases}
\end{aligned} \tag{17}$$

where $k = 3, 4, \ldots Q_{1i}$.

Proof. The basic idea of intracluster scheduling is to search free timeslots without increasing the number of total occupied timeslots during each superframe cycle. The scheduling cycle is designated as $C_{1iQ_{1i}}/C_{1i1}$ superframe cycles. We assume that $C_{1i1} \leq C_{1i2} \leq \cdots \leq C_{1iQ_{1i}}$.

The number of field devices that have the minimum DURs being C_{1i1} is N_{1i1}. These field devices can not occupy each other's timeslots. Therefore, the number of timeslots needed by field devices with DURs being C_{1i1} is $y_{1i1} = \lceil N_{1i1} \times C_{1i1}/C_{1i1} \rceil = \lceil N_{1i1} \rceil$.

The field devices with DURs being C_{1i2} need to be allocated the free second timeslot in the intracluster communication period during the first C_{1i2}/C_{1i1} superframe cycles. If these timeslots have already been occupied, the total number of timeslots should be increased. And the increased number of timeslots is $y_{1i2} = \lceil N_{1i2} \times C_{1i1}/C_{1i2} \rceil$.

For N_{1i3} field devices with DUR being C_{1i3}, the available timeslots can be searched from the free second timeslot in the intracluster communication period during the first C_{1i3}/C_{1i1} superframe cycles, which is numbered from 2 to $y_{1i1}+y_{1i2}$. If these timeslots are not enough for the N_{1i3} field devices, the number of timeslots should be increased. The number of free timeslots indexed from 2 to $y_{1i1}+y_{1i2}$ is $(\lceil N_{1i2}\times C_{1i1}/C_{1i2}\rceil \times (C_{1i2}/C_{1i1}) - N_{1i2}) \times (C_{1i3}/C_{1i2})$. The increased number of timeslots is determined by

$$y_{1i3} = \begin{cases} 0, & \dfrac{C_{1i1}}{C_{1i2}C_{1i3}}(N_{1i2}C_{1i3}+N_{1i3}C_{1i2}) < \left\lceil \dfrac{N_{1i2}\times C_{1i1}}{C_{1i2}} \right\rceil \\ \left\lceil \dfrac{N_{1i3} - (\lceil N_{1i2}\times C_{1i1}/C_{1i2}\rceil \times (C_{1i2}/C_{1i1}) - N_{1i2}) \times (C_{1i3}/C_{1i2})}{C_{1i3}/C_{1i1}} \right\rceil, & \text{else.} \end{cases} \tag{18}$$

Generally speaking, for N_{1ik} field devices with DUR being C_{1ik} ($k = 3,4,\ldots Q_{1i}$), the available timeslots are from the free second timeslot to $\sum_{j=1}^{k-1} y_{1ij}$ in the intracluster communication period during the first C_{1ik}/C_{1i1} superframe cycles. If these timeslots are not enough for these N_{1ik} field devices, the number of timeslots should be increased. The number of free timeslots indexed from 2 to $\sum_{j=1}^{k-1} y_{1ij}$ is $(y_{1i(k-1)} \times (C_{1i(k-1}/C_{1i1}) - N_{1i(k-1)}) \times (C_{1ik}/C_{1i(k-1)})$. The increased number of timeslots is determined by

$$y_{1ik} = \begin{cases} 0, & \dfrac{C_{1i1}}{C_{1i(k-1)}C_{1ik}}\left(N_{1i(k-1)}C_{1ik}+N_{1ik}C_{1i(k-1)}\right) < y_{1i(k-1)} \\ \left\lceil \dfrac{[N_{1ik} - (y_{1i(k-1)}\times(C_{1i(k-1)}/C_{1i1}) - N_{1i(k-1)})\times(C_{1ik}/C_{1i(k-1)})]\times C_{1i1}}{C_{1ik}} \right\rceil, & \text{else.} \end{cases} \tag{19}$$

In all, the total number of timeslots used for intracluster convergecast is the sum of the increased number of timeslots, which is $\sum_{j=1}^{Q_{1i}} y_{1ij}$. □

3.3.2. Performance Limits on Intercluster Convergecast Scheduling. The following Corollaries 7 and 8 can be deduced from Theorems 1 and 2.

Corollary 7. *If the data update rates of field devices in a cluster are partly same, that is $C_{1ij} = C_{1ik}$, $\exists j,k \in [1, f_{1i}]$, $j \neq k$, $i \in [1,N]$, the lower bound on the number of timeslots used for the intercluster convergecast in a cluster-line routing network is $L_s = 2\sum_{i=2}^{N}\sum_{j=1}^{Q_{1i}} y_{1ij} + \sum_{j=1}^{Q_{11}} y_{11j}$, where N is the number of routing devices in a network and Q_{1i} is the number of DUR kinds in a cluster with R_{1i} being the cluster head.*

Proof. Let g_{1i} be the number of packets in routing device R_{1i} ($i \in [1,N]$) at the beginning of each superframe cycle. The proving approach is similar to Theorem 1. The only difference is that the lower bound of g_{1i} is $\sum_{j=1}^{Q_{1i}} y_{1ij}$ according to Theorem 6.

R_{11} needs $2\sum_{i=2}^{N} g_{1i}$ timeslots for forwarding packets from $R_{12} \sim R_{1N}$ and needs g_{11} timeslots for forwarding own packets to GW. In all, the lower bound on the number of timeslots used for intercluster convergecast scheduling is the sum of these two parts, which can be formulated as $L_s = 2\sum_{i=2}^{N} g_{1i} + g_{11} = 2\sum_{i=2}^{N}\sum_{j=1}^{Q_{1i}} y_{1ij} + \sum_{j=1}^{Q_{11}} y_{11j}$. □

Corollary 8. *Given any scheduling algorithm S that can complete intercluster convergecast in $2\sum_{i=2}^{N}\sum_{j=1}^{Q_{1i}} y_{1ij} + \sum_{j=1}^{Q_{11}} y_{11j}$ timeslots in a cluster-line routing network, the lower bound on the number of channels used in S is*

$$\left\lceil \frac{2\sum_{i=2}^{N}\sum_{j=1}^{Q_{1i}} y_{1ij} + \sum_{j=1}^{Q_{11}} y_{11j} + 1}{2} - \sqrt{\left(\frac{2\sum_{i=2}^{N}\sum_{j=1}^{Q_{1i}} y_{1ij} + \sum_{j=1}^{Q_{11}} y_{11j} + 1}{2}\right)^2 - \sum_{i=1}^{N}\left[i\sum_{j=1}^{Q_{1i}} y_{1ij}\right]} \right\rceil, \tag{20}$$

where N is the number of routing devices in a network and Q_{1i} is the number of DUR kinds in a cluster with R_{1i} being the cluster head.

Proof. The proof is similar to Theorem 2. The only difference is to replace $L_s = 2\sum_{i=2}^{N} f_{1i} + f_{11}$ in Theorem 2 with $L_s = 2\sum_{i=2}^{N}\sum_{j=1}^{Q_{1i}} y_{1ij} + \sum_{j=1}^{Q_{11}} y_{11j}$. Then, the lower bound on the number of channels used for intercluster convergecast scheduling is

$$\left\lceil \frac{2\sum_{i=2}^{N}\sum_{j=1}^{Q_{1i}} y_{1ij} + \sum_{j=1}^{Q_{11}} y_{11j} + 1}{2} - \sqrt{\left(\frac{2\sum_{i=2}^{N}\sum_{j=1}^{Q_{1i}} y_{1ij} + \sum_{j=1}^{Q_{11}} y_{11j} + 1}{2}\right)^2 - \sum_{i=1}^{N}\left[i\sum_{j=1}^{Q_{1i}} y_{1ij}\right]} \right\rceil. \tag{21}$$

□

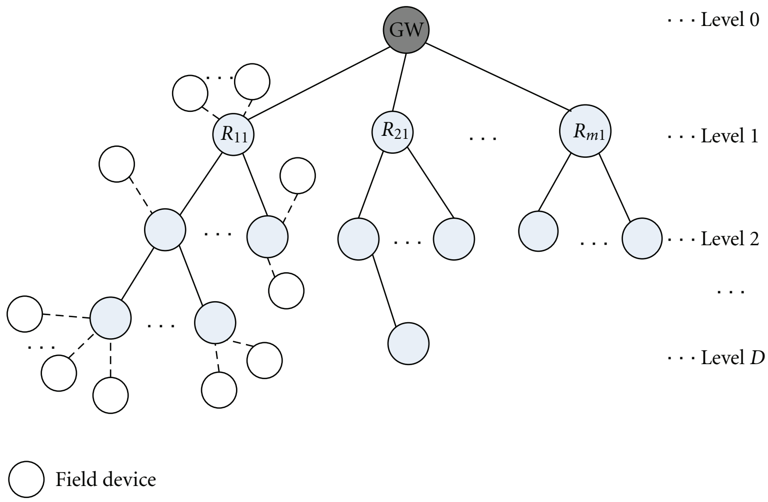

Figure 3: Cluster-tree network topology.

4. Time- and Channel-Optimal Convergecast for Cluster-Tree Topology

Figure 3 is a cluster-tree network. There exist N routing devices in the network, which manage several field devices and constitute clusters. Let D denote the depth of the routing tree. The GW is located at level 0 and has m subtrees $T_1, T_2, \ldots T_m$ rooted at routing devices $R_{11}, R_{21}, \ldots, R_{m1}$, respectively. The number of routing devices in subtrees $T_1, T_2, \ldots, T_m$, respectively, are $n_1, n_2, \ldots n_m$. Without loss of generality, it is assumed that $n_1 \geq n_2 \geq \cdots \geq n_m$. Therefore, T_1 is the largest subtree. Let R_{ij} denote the jth routing device in the ith subtree and f_{ij} be the number of field devices in the cluster with R_{ij} being the cluster head ($i \in [1, m]$, $j \in [1, n_i]$). Next, we will investigate the convergecast scheduling problems in cluster-tree networks from three scenarios: ADS, ADD, and PDS.

4.1. Scenario I: ADS. For a cluster-tree network, if all the DURs in a cluster are same, the number of timeslots for intracluster convergecast with R_{ij} being cluster head is at least equal to the number of field devices in this cluster, that is, the number of timeslots for intracluster convergecast is f_{ij} ($i \in [1, m]$, $j \in [1, n_i]$). For network scalability, we specify that the number of timeslots for intracluster convergecast is $\max_{i\in[1,m], j\in[1,n_i]} f_{ij}$.

The lower bound on the number of timeslots used for intercluster convergecast scheduling is decided by both the largest subtree and the number of packets in network. For ADS, the number of packets in routing devices are equal to the number of timeslots needed for intracluster convergecast. That is to say, the number of packets in routing device R_{ij} is f_{ij} ($j \in [1, n_i]$, $i \in [1, m]$).

Theorem 9. *If all the data update rates of field devices in a cluster are same, the lower bound on the number of timeslots required to complete the intercluster convergecast in a cluster-tree network and N routing devices is* $L_s = \max\{2\sum_{j=2}^{n_1} f_{1i} + f_{11}, \sum_{i=1}^{m}\sum_{j=1}^{n_i} f_{ij}\}$, *where* f_{ij} *is the number of field devices in the cluster with* R_{ij} *being cluster head* ($i \in [1, m]$, $j \in [1, n_i]$).

Proof. For R_{i1} ($i \in [1, m]$), all packets in subtree T_i must go through R_{i1} to GW. Hence, R_{i1} needs at least $2\sum_{j=2}^{n_i} f_{ij} + f_{i1}$ timeslots to forward all packets in T_i. Since the largest subtree is T_1, at least $L_s = 2\sum_{j=2}^{n_1} f_{1j} + f_{11}$ timeslots are needed to complete the convergecast. Meanwhile, since the GW can receive only one packet per timeslot, at least $\sum_{i=1}^{m}\sum_{j=1}^{n_i} f_{ij}$ timeslots are needed to complete the intercluster convergecast in the network having $\sum_{i=1}^{m}\sum_{j=1}^{n_i} f_{ij}$ packets. In all, the lower bound on the number of timeslots needed to complete the intercluster convergecast is $L_s = \max\{2\sum_{j=2}^{n_1} f_{1j} + f_{11}, \sum_{i=1}^{m}\sum_{j=1}^{n_i} f_{ij}\}$. □

Theorem 9 shows that the structure of the routing tree plays a fundamental role in minimizing the convergecast time and dominates the minimum network latency. If $2\sum_{j=2}^{n_1} f_{1i} + f_{11} \leq \sum_{i=1}^{m}\sum_{j=1}^{n_i} f_{ij}$ (i.e., $\sum_{j=2}^{n_1} f_{1j} \leq \sum_{i=2}^{m}\sum_{j=1}^{n_i} f_{ij}$), then the lower bound on the number of timeslots to complete the intercluster convergecast is $L_s = \sum_{i=1}^{m}\sum_{j=1}^{n_i} f_{ij}$, which is equal to the number of packets in the whole network. Otherwise, the lower bound on the number of timeslots to complete the intercluster convergecast is $L_s = 2\sum_{j=2}^{n_1} f_{1i} + f_{11}$, which is equal to the number of packets in the largest subtree T_1. Finding a minimum spanning tree subject to cardinality constraints on the number of routing devices in any subtree, the so called capacitated minimum spanning tree problem, is known to be NP-hard [8], and the effective heuristics algorithm should be carefully designed.

Theorem 10. *Given any scheduling algorithm S that can complete intercluster convergecast in* $L_s = \max\{2\sum_{j=2}^{n_1} f_{1j} + f_{11}, \sum_{i=1}^{m}\sum_{j=1}^{n_i} f_{ij}\}$ *timeslots in a cluster-tree network, the lower bound on the number of channels used in S is*

$$\left\lceil \left(L_s + \frac{1}{2}\right) - \sqrt{\left(L_s + \frac{1}{2}\right)^2 - 2\sum_{d=1}^{D}(d \times n(d))} \right\rceil, \quad (22)$$

where $n(d)$ *denotes the total number of packets in routing devices at level* d.

Proof. Let C_s be the number of channels that can be used in a network, and $\mathrm{PT}_{\max}(t)$ denotes the number of parallel transmissions scheduled in timeslot t ($t \in [1, L_s]$).

To complete convergecast in L_s timeslots, the following conditions should be satisfied [8].

(i) For $t = L_s$, there should be only one packet in the network and R_{11} is required to send it to GW. That is, only one device could be scheduled to be transmitted if $t = L_s$.

(ii) For $t = L_s - 1$, two devices could be scheduled to be transmitted in the same timeslot.

(iii) Deduced as so forth, the maximum number of devices allowed to transmit in timeslot t is $L_s - t + 1$.

Since $\lceil (L_s - t + 1)/2 \rceil \geq C_s$ when $t \in [1, L_s - 2(C_s - 1)]$ and the upper bound on the number of channels is C_s, we can obtain the following conclusion [8]:

$$\mathrm{PT}\max(t) = \begin{cases} C_s, & t \in [1, L_s - C_s] \\ L_s - t + 1, & t \in [L_s - C_s + 1, L_s]. \end{cases} \tag{23}$$

By (23), the maximum number of transmissions that can be scheduled in L_s using C_s channels is

$$\begin{aligned} &\sum_{t=1}^{L_s} \mathrm{PT}\max(t) \\ &\quad = \sum_{t=1}^{L_s - C_s} C_s + \sum_{t=L_s-C_s+1}^{L_s} L_s - t + 1 \\ &\quad = (L_s - C_s) \times C_s + C_s \times (C_s + 1) - \frac{C_s \times (C_s + 1)}{2} \\ &\quad = -\frac{1}{2} C_{s^2} + C_s \left(L_s + \frac{1}{2} \right). \end{aligned} \tag{24}$$

For the intercluster convergecast, only routing devices and GW along the routing tree transmit packets. Thus the number of transmissions for completing convergecast is at least $\sum_{d=1}^{D} dn(d)$. Combined with (24), the number of transmissions satisfy the following inequality:

$$-\frac{1}{2} C_{s^2} + \left(L_s + \frac{1}{2} \right) C_s \geq \sum_{d=1}^{D} dn(d). \tag{25}$$

The inequality (25) can rewritten as

$$C_s \geq \left\lceil \left(L_s + \frac{1}{2} \right) - \sqrt{\left(L_s + \frac{1}{2} \right)^2 - 2 \sum_{d=1}^{D} (d \times n(d))} \right\rceil, \tag{26}$$

where $L_s = \max\{2 \sum_{j=2}^{n_1} f_{1j} + f_{11}, \sum_{i=1}^{m} \sum_{j=1}^{n_i} f_{ij}\}$. □

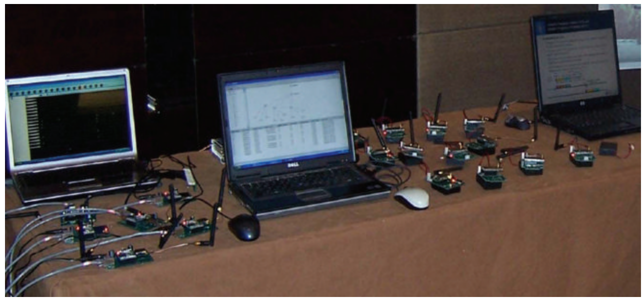

Figure 4: Experiment platform.

4.2. Scenario II: ADD. In Scenario II, the DURs of field devices in the cluster with R_{ij} ($i \in [1, m]$, $j \in [1, n_i]$) being cluster head satisfy the following relationship:

$$C_{ijk} \neq C_{ijl}, \quad \forall k, l \in \left[1, f_{ij}\right], \ k \neq l. \tag{27}$$

Theorem 11. *Given a cluster with more than one field device, if all the data update rates of field devices in a cluster are different, that is $C_{ijk} \neq C_{ijl}$, for all k, $l \in [1, f_{ij}]$, $k \neq l$ & $i \in [1, m]$, $j \in [1, n_i]$, the lower bound on the number of timeslots used for the intracluster convergecast in a cluster-tree network is 2.*

Proof. Similar to Theorem 3. □

For ADD, the number of packets in a routing device is equal to the number of timeslots for intracluster convergecast. According to Theorem 11, at the beginning of the intercluster convergecast, the number of packets in a routing device is 2.

Corollary 12. *If all the data update rates of field devices in a cluster are different, that is $C_{ijk} \neq C_{ijl}$, for all k, $l \in [1, f_{ij}]$, $k \neq l$ & $i \in [1, m]$, $j \in [1, n_i]$, the lower bound on the number of timeslots used for the intercluster convergecast in a cluster-tree network is $L_s = \max\{4n_1 - 2, \sum_{i=1}^{m} n_i\}$, where n_i is the number of routing devices in subtree T_i ($i \in [1, m]$).*

Proof. Let g_{ij} be the number of packets in R_{ij} during one superframe cycle. Reminded that the number of packets in a routing device is 2 at the beginning of the intercluster communication period, we can conclude that $g_{ij} \geq 2$ ($i \in [1, m]$, $j \in [1, n_i]$). Similar to the proof of Theorem 9, the lower bound on the number of timeslots used for the intercluster convergecast in a network with cluster-tree routing is $L_s = \max\{4n_1 - 2, \sum_{i=1}^{m} n_i\}$. □

Corollary 13. *Given any scheduling algorithm S in a network with cluster-tree routing that can complete intercluster convergecast in $L_s = \max\{4n_1 - 2, \sum_{i=1}^{m} n_i\}$ timeslots, the lower bound on the number of channels used in S is*

$$\left\lceil \left(L_s + \frac{1}{2} \right) - \sqrt{\left(L_s + \frac{1}{2} \right)^2 - 2 \sum_{d=1}^{D} (d \times n(d))} \right\rceil, \tag{28}$$

where $n(d)$ denotes the total number of packets in routing devices at level d.

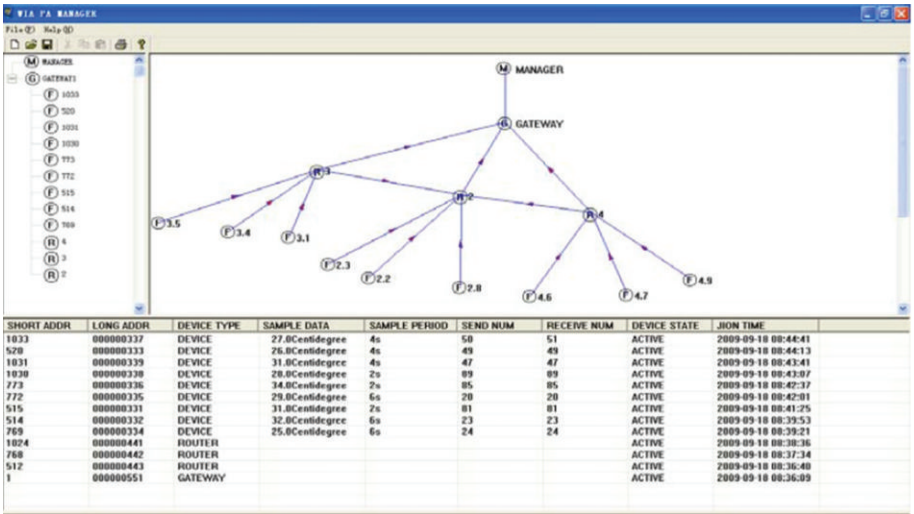

Figure 5: Network topology and monitoring data.

Proof. The proof of Corollary 13 is similar to Theorem 10. The only difference is that $L_s = \max\{4n_1 - 2, \sum_{i=1}^{m} n_i\}$ according to Corollary 13. □

4.3. Scenario III: PDS. In Scenario III, the DURs of field devices in the cluster with R_{ij} $(i \in [1, m],\ j \in [1, n_i])$ being cluster head satisfy the following relationship:

$$C_{ijk} = C_{ijl}, \quad \exists k, l \in \left[1, f_{ij}\right],\ k \neq l. \tag{29}$$

Corollary 14. *Given a cluster with R_{ij} $(i \in [1, m], j \in [1, n_i])$ being cluster head, the data update rates of field devices in a cluster are partly same, that is $C_{ijk} = C_{ijl}$, $\exists k, l \in [1, f_{ij}]$, $k \neq l$. Supposing that the number of field devices with DUR being C_{ijk} is N_{ijk} and $\sum_{j=1}^{Q_{ij}} N_{ijk} = f_{ij}$ $(Q_{ij} \leq f_{ij})$, the lower bound on the number of timeslots used for the intracluster convergecast in a cluster-tree network is $\sum_{j=1}^{Q_{ij}} y_{ijk}$, where $i \in [1, m]$, $j \in [1, n_i]$, $k \in [1, Q_{ij}]$. Q_{ij} is the number of DUR kinds in the cluster with R_{ij} being cluster head, and y_{ijk} is defined as:*

$$\begin{gathered} y_{ij1} = \left\lceil \frac{N_{ij1} \times C_{ij1}}{C_{ij1}} \right\rceil = \left\lceil N_{ij1} \right\rceil; \\ y_{ij2} = \left\lceil \frac{N_{ij2} \times C_{ij1}}{C_{ij2}} \right\rceil, \\ y_{ijk} = \begin{cases} 0, & \dfrac{C_{ij1}\left(N_{ij(k-1)} C_{ijk} + N_{ijk} C_{ij(k-1)}\right)}{C_{ij(k-1)} C_{ijk}} < y_{ij(k-1)} \\ \left\lceil \dfrac{N_{ijk} - \left(y_{ij(k-1)} \times \left(C_{ij(k-1)}/C_{ij1}\right) - N_{ij(k-1)}\right) \times \left(C_{ijk}/C_{ij(k-1)}\right)}{C_{ijk}/C_{ij1}} \right\rceil, & \text{else}, \end{cases} \end{gathered} \tag{30}$$

where $k = 3, 4, \ldots Q_{ij}$.

Proof. similar to Theorem 6. □

The performance limits on the intercluster convergecast scheduling are determined by the performance limits on the intracluster convergecast scheduling, which have been established in Corollary 14.

Corollary 15. *If the data update rates of field devices in a cluster are partly same, that is $C_{ijk} = C_{ijl}$, $\exists k, l \in [1, Q_{ij}]$, $k \neq l$, the lower bound on the number of timeslots used for the intercluster convergecast in a cluster-tree network is $L_s = \max\{2\sum_{j=2}^{n_1}\sum_{k=1}^{Q_{1j}} y_{1jk} + \sum_{k=1}^{Q_{11}} y_{11k}, \sum_{j=1}^{n_i}\sum_{k=1}^{Q_{ij}} y_{ijk}\sum_{i=1}^{m}\}$, where n_i is the number of routing devices in subtree T_i $(i \in [1, m]$, $j \in [1, n_i]$, $k \in [1, Q_{ij}])$.*

Proof. Let g_{ij} be the number of packets in routing device R_{ij} during one superframe cycle. Reminded that the number of packets in the routing device is $\sum_{j=1}^{Q_{ij}} y_{ijk}$ at the beginning of the intercluster convergecast, we can conclude that $g_{ij} \geq \sum_{j=1}^{Q_{ij}} y_{ijk}$ ($i \in [1, m]$, $j \in [1, n_i]$, $k \in [1, Q_{ij}]$). Similar to the proof of Theorem 9, the lower bound on the number of timeslots used for the intercluster convergecast in a network with cluster-tree routing is $L_s = \max\{2\sum_{j=2}^{n_1}\sum_{k=1}^{Q_{1j}} y_{1jk} + \sum_{k=1}^{Q_{11}} y_{11k}, \sum_{i=1}^{m}\sum_{j=1}^{n_i}\sum_{k=1}^{Q_{ij}} y_{ijk}\}$. □

Corollary 16. *Given any scheduling algorithm S that can complete intercluster convergecast in* $L_s = \max\{2\sum_{j=2}^{n_1}\sum_{k=1}^{Q_{1j}} y_{1jk} + \sum_{k=1}^{Q_{11}} y_{11k}, \sum_{i=1}^{m}\sum_{j=1}^{n_i}\sum_{k=1}^{Q_{ij}} y_{ijk}\}$ *timeslots in a cluster-tree network, the lower bound on the number of channels used in S is*

$$\left\lceil \left(L_s + \frac{1}{2}\right) - \sqrt{\left(L_s + \frac{1}{2}\right)^2 - 2\sum_{d=1}^{D}(d \times n(d))} \right\rceil, \qquad (31)$$

where $n(d)$ denotes the total number of packets in routing devices at level d.

Proof. The proof of Corollary 16 is similar to Theorem 10. The only difference is that $L_s = \max\{2\sum_{j=2}^{n_1}\sum_{k=1}^{Q_{1j}} y_{1jk} + \sum_{k=1}^{Q_{11}} y_{11k}, \sum_{i=1}^{m}\sum_{j=1}^{n_i}\sum_{k=1}^{Q_{ij}} y_{ijk}\}$ according to Corollary 16. □

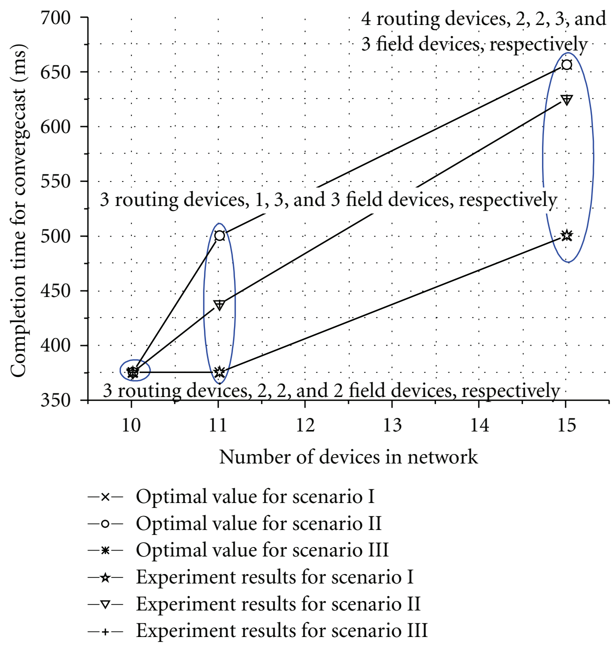

Figure 6: Convergecast time by experiment in network with cluster-line topology.

5. Extensive Analysis

The theoretical analysis of the lower bounds on the number of timeslots and the number of channels gives deep insights into the network planning and reliable scheduling algorithm designing. In order to increase the reliability for IWSNs, a retransmission scheme is introduced.

We consider the general cluster-tree network. According to the theoretical analysis in Section 4, the lower bound on the number of timeslots for intracluster convergecast in Scenario I is $\max_{i\in[1,m], j\in[1,n_i]} f_{ij}$ and the lower bound on the number of timeslots for intercluster convergecast is $\max\{2\sum_{j=2}^{n_1} f_{1i} + f_{11}, \sum_{i=1}^{m}\sum_{j=1}^{n_i} f_{ij}\}$. If we consider x retransmissions for each packet (default value of *mac Max Frame Retries* in IEEE STD 802.15.4-2006 [10]), the lower bounds on the number of timeslots for intracluster and intercluster convergecast in Scenario I are $x+1$ times of $\max_{i\in[1,m], j\in[1,n_i]} f_{ij}$ and $\max\{2\sum_{j=2}^{n_1} f_{1i} + f_{11}, \sum_{i=1}^{m}\sum_{j=1}^{n_i} f_{ij}\}$, respectively. Meanwhile, the number of timeslots during the beacon, CAP and CFP periods are 16 according to the IEEE STD 802.15.4-2006 [11]. In all, the superframe length can be calculated as $(16 + (x+1) \times \max_{i\in[1,m], j\in[1,n_i]} f_{ij} + (x+1) \times \max\{2\sum_{j=2}^{n_1} f_{1i} + f_{11}, \sum_{i=1}^{m}\sum_{j=1}^{n_i} f_{ij}\}) \times duration\ of\ timeslot$. For a network having 1000 field devices, that is $\sum_{i=1}^{m}\sum_{j=1}^{n_i} f_{ij} = 1000$, and the duration of the timeslot being 10 ms, if $x = 3$ and $2\sum_{j=2}^{n_1} f_{1j} + f_{11} < \sum_{i=1}^{m}\sum_{j=1}^{n_i} f_{ij}$, the maximum superframe length is 80.16 s, which is allowable in most industrial process applications. In Scenario II and Scenario III, the superframe lengths are less than 80.16 ms because of smaller number of timeslots.

The analysis of superframe length indicates that the network with hierarchical topology can accommodate more devices and can guarantee better timeliness. For applications required large network and tight timeliness, if we assign field devices with different DURs in same cluster, the requirements can be well satisfied.

6. Experiment

We validate the tightness from a real implementation on the SIA2420 modules [14]. The SIA 2420 module is completely compliant to the IEEE STD 802.15.4-2006 physical layer and built in multiple specifications, such as WIA-PA, WirelessHART and Zigbee. It is a high reliable, low energy-consumption and long-distance wireless product of original equipment manufacturer (OEM), which is controlled by a 16-bit microcontroller and is installed by a MultiMedia Communication eXchange (MMCX) antenna connector. The transmission radius of SIA 2420 is 25 meters. The user interface is built on the Microsoft Visual C++ 6.0. Figures 4 and 5 exhibits the experiment platform, network topology and user operation interface, respectively.

Based on our experiment system, we investigate the bounds of the number of timeslots and channels with cluster-line and cluster-tree routing topologies, respectively. We use three parameters to represent the network with cluster-tree topology, which are M, D, O. M denotes the number of children of GW; D denotes the depth of the tree; and O denotes the number of field devices in the cluster. Each simulation setting is represented as $\{M, D, O\}$. Different combination of these three parameters will generate different topology. Figure 6 compares the experiment convergecast time with the optimal results for three scenarios in the network with the cluster-line routing topology. The experiment results

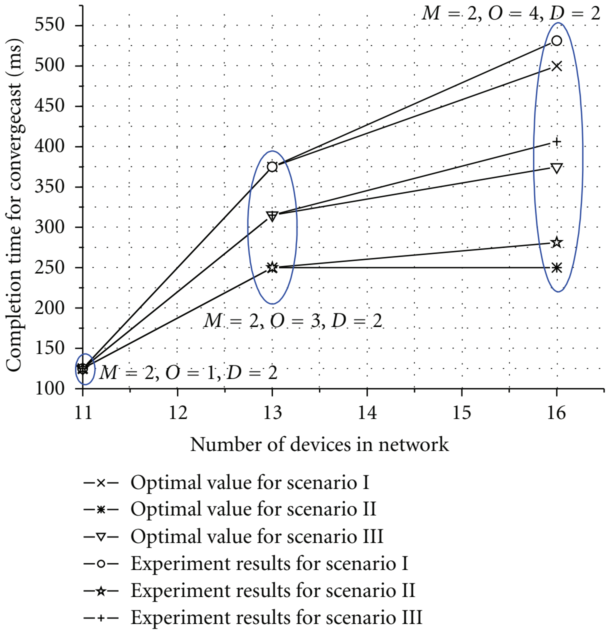

Figure 7: Convergecast time by experiment in network with cluster-tree topology.

show that there exist the optimal strategies for convergecast in the network with cluster-line topology, the conclusion of which is same as the proof results. Figure 7 compares the experiment convergecast time with the optimal results for three scenarios in the network with the cluster-tree routing topology. The experiment results of the network with cluster-tree routing topology show that our two-stage scheduling algorithms are near-optimal and the maximum degree of approximation is 99.9%.

7. Conclusion

In this paper, orienting to clustered networks including cluster-line and cluster-tree, we analyze the performance limits of the convergecast scheduling. Our analysis is divided according to the network topology (cluster-line or cluster-tree) and the application scenarios (ADS, ADD, and PDS). We establish the lower bounds on the number of timeslots to finish the intracluster and the intercluster convergecast transmissions. In addition, we establish the lower bounds on the number of channels based on the lower bounds on the number of timeslots and available channels in multichannel scenario. The analysis results are future extended to take packet retransmissions into consideration so as to meet the reliability requirement.

Working with this framework has opened up several possibilities of future research. Work currently in progress is under assumption of reliable link-level transmissions. In the future, an extension to the framework is planned to analyze the performance of the networks in presence of intermittent connectivity and channel errors. The channel-constrained convergecast scheduling when the number of available channels is less than the lower bound required for optimal convergecast is a great challenge for research. As comparison with the convergecast transmission in this paper, another interesting issue to be investigated in the future work is the dissemination communications from the gateway device to other devices in the network.

Acknowledgments

This work was supported by the Natural Science Foundation of China under Contract 61202418 and 61174026, the National High Technology Research and Development Program of China under Contract 2011AA040101, the Special Program for Key Basic Research Founded by MOST under Contract 2010CB334705, and the Foundation of Chinese Academy of Sciences under Contract (KGCX2-EW-104).

References

[1] A. Willig, "Recent and emerging topics in wireless industrial communications: a selection," *IEEE Transactions on Industrial Informatics*, vol. 4, no. 2, pp. 102–122, 2008.

[2] HART Communication Foundation, "Wire- lessHARTTM technical data sheet," May 2009.

[3] Industrial communication network—Field-bus specifications—WIA-PA communication network and communication profile, http://www.iec.ch/.

[4] S. Gandham, Y. Zhang, and Q. Huang, "Distributed minimal time convergecast scheduling in wireless sensor networks," in *Proceedings of the 26th IEEE International Conference on Distributed Computing Systems (ICDCS '06)*, pp. 462–466, Washington, DC, USA, 2006.

[5] H. Choi, J. Wang, and E. A. Hughes, "Scheduling on sensor hybrid network," in *Proceedings of the 14th International Conference Computer Communications and Networks (ICCCN '05)*, pp. 503–508, San Diego, Calif, USA, October 2005.

[6] C. Florens, M. Franceschetti, and R. J. McEliece, "Lower bounds on data collection time in sensory networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1110–1120, 2004.

[7] B. Yu, J. Li, and Y. Li, "Distributed data aggregation scheduling in wireless sensor networks," in *Proceedings of the 28th IEEE Conference on Computer Communications (INFOCOM '09)*, pp. 2159–2167, Orlando, Fla, USA, April 2009.

[8] P. Soldti, *On cross-layer design and resource scheduling in wireless networks [Ph.D. thesis]*, KTH, School of Electrical Engineering, Stockholm, Sweden, 2009.

[9] P. Soldati, H. Zhang, and M. Johansson, "Deadline-constrained transmission scheduling and data evacuation in wirelesshart networks," in *Proceedings of the European Control Conference (ECC '09)*, Budapest, Hungary, September 2009.

[10] W. Shang, P. Wan, and X. Hu, "Approximation algorithm for minimal convergecast time problem in wireless sensor networks," *Wireless Networks*, vol. 16, no. 5, pp. 1345–1353, 2010.

[11] H. Zhang, A. Arora, Y. R. Choi, and M. G. Gouda, "Reliable bursty convergecast in wireless sensor networks," *Computer Communications*, vol. 30, no. 13, pp. 2560–2576, 2007.

[12] X. Liu, H. Zhang, Q. Xiang, X. Che, and X. Ju, "Taming uncertainties in real-time routing for wireless networked sensing and control," in *Proceedings of the 13th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '12)*, Hilton Head, SC, USA, 2012.

[13] IEEE SA Standards Board, "Part 15.4: wireless medium access control (MAC) and physical layer (PHY) specification for low-rate wireless personal area networks (LR- WPANs)," Nmuber 802.15.4, 2006.

[14] SIA 2420, http://www.industrialwireless.cn/03.asp?pd=dl&nc-las-sid=628&anclassid=14.