

# Aggregation Tree Based Data Aggregation Algorithm in Wireless Sensor Networks

<http://dx.doi.org/10.3991/ijoe.v12i06.5408>

Hu Yanhua<sup>1</sup>, Zhang Xincai<sup>2</sup>

<sup>1</sup>Guangxi University of Science and Technology, Liuzhou, Guangxi, China

<sup>2</sup>Zhong yuan University of Technology, Zhengzhou, Henan, China

**Abstract**—In wireless sensor networks, efficient and effective data aggregation algorithms can prolong the network lifecycle by reducing communication of redundant data and improve the security of the networks. Tradition data aggregation algorithms in wireless sensor networks mainly aim to improve the energy utilization, and ignore the security and lifecycle. In order to get a good trade-off between these requirements, we proposed a data aggregation algorithm based on constructing a data aggregation tree. After give a formalism description of the problem, we proposed a data aggregation tree constructing algorithm. By minimize the maximal energy consumption of nodes, the algorithm can prolong the lifecycle. In data aggregation scheduling algorithm, we select the number of communications carefully to get the trade-off between low weighted delay and high network lifecycle. The simulation experiments show that, the proposed data aggregation algorithm consumes less energy while aggregating data from sensor nodes, and thus can prolong the network lifecycle.

**Index Terms**—wireless sensor networks, data aggregation, aggregation tree, aggregation delay

## I. INTRODUCTION

The wireless sensor networks consist thousands of or even tens of thousands of different kinds of sensors are important computing platforms. The sensors in a wireless sensor network are deployed randomly in an area to monitor and collect data of the real world environment [1]. The collected data in the area are usually redundant, and the same data can be collected by multiple sensors. The transmission of data consumes about 70% of the energy in a wireless sensor network, so in order to reduce the communication overhead between sensors and prolong the lifecycle of the underline wireless sensor network, the data collected from different sensors must be aggregated [2]. Until now, researchers have proposed many data aggregation algorithms for data of wireless sensor networks. With respect to the objective of aggregation, these algorithms can be classified into data centered [3], remaining energy based [4], optimized [5] and performance based [6] algorithms. The main idea of these algorithms is that, decrease the collision probability of data in wireless channels by reducing the transmitting data, and thus improve the efficiency of data collecting; check the validity of sensor nodes by comparing data collected from adjacent nodes, discard the data from the disabled nodes to reinforcement the accuracy of data; and aggregate the data sensed from different kinds of sensors to bridge the gap between the high-level user requirements with the low-level raw sensed data [7].

Wireless sensor networks come from the military applications, and have been applied to many fields, such as emergence surveillance [8], environment monitoring [9], target tracking [10], and so on, whereas the secure aggregation of data is still a hot research in wireless sensor networks. In order to assure the security of aggregation of sensed data, the key points can be placed into either the raw data or the aggregation process. In order to improve the security of raw data, data authorization can be used, but it increases the total communication overhead of the network [11]. By secured aggregation algorithms, the aggregation node can reduce the offense of data from invalid nodes, but different applications need different aggregation algorithms, and they cannot make sure the valid redundant data be acquired,

In this paper, we study the problem of data aggregation in wireless sensor networks. In order to prolong the network lifecycle, we aim to select the number of communications of each node carefully to get the trade-off between low weighted delay and high network lifecycle, and thus make the whole network work longer.

## II. RELATED WORKS

Cluster based topology management mechanism [12-13] is commonly used in wireless sensor networks, where a head node is selected in each cluster to manage the nodes in the same cluster, collect data from the intra-cluster nodes, and transmit data between inter-clusters. So, the head node of a cluster must know the semantic of collected data, which makes it impossible to transmit encrypted data between the head node with bases.

ESPDA, proposed by Li et al. [14], is an energy efficient secure data aggregation protocol. In this protocol, sensor nodes generate pattern codes based on the raw data, and send them to the head node. The head node classifies the raw data according to the received pattern codes, selects the pattern code subset via pattern comparing algorithms, and requires the selected sensors to transmit the encrypted data.

SIA, proposed by Przydatek et al. [15], is a secure data aggregation framework for large scale wireless sensor networks. In this framework, they defined a kind of nodes, called aggregators, which aggregate the queried data to reduce the communication cost of the network. In addition, they applied effective random sampling and interactive verifying mechanisms to assure that the data aggregated by aggregators are the maximal approximation of the real values.

SecureDAV, proposed by Mahimkar et al. [16], is a secure data aggregation and verification protocol in wireless

sensor networks. This protocol assigns keys for nodes of the same cluster by applying the elliptic curve based secret sharing scheme. Each normal node computes the average value of the intra-cluster data and gives its partial signature, and the head node collects the signature of other intra-cluster nodes, give a complete signature to the average values and send it to the base. Finally, the base verifies the signature via the corresponding public key.

SRDA, proposed by Sanli et al. [17], is a reference data based secure aggregation protocol. This protocol ascertains the differences by comparing the raw data and the reference data, and the sensor nodes transmit differential data rather than the raw data to reduce the communication overhead. Moreover, Sugandhi et al. [18] proposed a secure and energy-saving data aggregation and authorization protocol with respect to the risk of leaking information while aggregating data. Sun et al. [19] proposed a credible behavior based secure data aggregation and route algorithm by studying the credibility of networks. Wu et al. [20] designed a pattern code based efficient secure data aggregation protocol in wireless sensor networks.

These secure data aggregation algorithms have their specific advantages. However, some of them ignore the limits of resources in wireless sensor networks, and their computation complexity is very high, which consumes much energy while aggregating data; and some of them only focus the security of the data and ignore the security while transmitting data between different sensors.

### III. PROBLEM STATEMENT

Given a sensor network  $G = (V, E)$ ,  $V$  is the set of sensor nodes, where  $|V| = n$  nodes are randomly deployed in a  $L \times L$  region, the communication radius of each node is  $r$ , and the only base node *sink* is in the center of the region, and the link set  $E \subseteq V \times V$  is defined as  $e_{ij} \in E$ , if  $\|v_i - v_j\| \leq r$ .

Given a sensor network  $G = (V, E)$ , we aim to build a data aggregation tree  $T = (V, \hat{E})$ , where for each node  $i \in V$ , its parent node is  $P_i$ , the link between  $i$  with  $P_i$  is denoted as  $e_i$ . The aim of data aggregation is building a sequence  $SCHE = \{E(1), E(2), \dots, E(T)\}$  based on  $\hat{E}$ , where  $E(s)$  ( $1 \leq s \leq T$ ) is the set of scheduled links during time  $s$ . After executing the scheduling sequence, we have

- 1) All collected data from nodes are aggregated to the sink node;
- 2)  $E(1) \cup E(2) \cup \dots \cup E(T) = \hat{E}$ ;
- 3) For each  $E(s)$  ( $1 \leq s \leq T$ ), if  $e_i, e_j \in E(s)$ , then they don't collide.

In this paper, we define the collision as the coordinate collision model [10], and assume that all nodes in the sensor network can receive and send data, but they can't be done at the same time. In addition, we also assume the interference radiuses of all nodes equal to  $r_i$ .

**Definition 1. Link collision.** In the coordinate collision model, two communication links collide, **if and only if** the

receive end of one node is within the radius of another node.

Via constructing an aggregation tree, the constructed aggregation tree should satisfy aggregation delay, average weighted delay and network lifecycle three conditions.

**Definition 2. Aggregation delay.** For one data aggregation process, if the data collected from all nodes arrive the *sink* node after  $t$  time intervals, that is  $SCHE = \{E(1), E(2), \dots, E(T)\}$ , then the delay of the aggregation cycle is  $t$ , i.e.  $d = |SCHE| = t$ .

For an aggregation tree of any structure, the aggregation delay satisfies

$$d \geq \max\{x_i + h_i : i = 1, 2, \dots, n\}, \quad (1)$$

where  $x_i$  represents the number of child nodes of node  $i$ , and  $h_i$  represents the number of hops from the root node to node  $i$ .

As the importance of different regions is different, every node is assigned a weight reflecting its importance. Let the weight of node  $i$  be  $w_i$ . If  $w_i > w_j$ , then we hope that the data collected by node  $i$  arrive the *sink* node earlier than the data collected by node  $j$ . In order to quantify the weighted fairness, we use the following average weighted delay to measure fairness.

**Definition 3. Average weighted delay.** In an aggregation cycle, if the data collected by node  $i$  arrive the *sink* node after  $t_i$  interval, then the average weighted delay of node  $i$  is  $D_i = w_i \times t_i$ , and the average weighted delay of all nodes in the aggregation cycle is

$$D_{avg} = \frac{\sum_{i \in V} w_i \times t_i}{|V|}. \quad (2)$$

**Definition 4. Network lifecycle.** For any scheduling algorithm, if there is a point, where energy runs out, after  $L$  aggregation periods, then the network lifecycle of the aggregation tree constructing process is  $L$ .

In the aggregation tree based data aggregation, the nodes have three states, i.e. sending, receiving and sleeping, where the power consumption in the sleeping state is far smaller the other two states, so it is ignored in this paper. Then, the power consumption of node  $i$  in each aggregation period is  $E_i = p_i \times t_i + p_r \times r_i$ , where  $p_i$  and  $p_r$  are the power consumptions of one sending and one receiving respectively, and  $t_i$  and  $r_i$  are the numbers of sending and receiving respectively. If all nodes have the same initial energy,  $EN$ , then the lifecycle of a data aggregation tree is

$$L = \min_{i \in N} \left\lfloor \frac{EN}{p_i \times t_i + p_r \times r_i} \right\rfloor. \quad (3)$$

From three definitions defined above we can see that, the performance of data aggregation algorithms mainly depend on the structure of aggregation tree and the scheduling strategy. So, the optimal aggregation scheduling process should construct the sequence of scheduling links,  $SCHE = \{E(1), E(2), \dots, E(T)\}$ , such that  $\min(d)$ ,  $\min(D_{avg})$  and  $\max(L)$ .

IV. CONSTRUCTING AGGREGATION TREE

The construction of aggregation tree is the preparation of the whole data aggregation process, and there are two reasons to build the aggregation tree. On one hand, there is only one base node in the whole sensor network, so the tree structure is suitable to this condition. On the other hand, the tree structure can reduce the overhead of communication in the whole network, and this is necessary for energy limited wireless sensor networks. However, the problem of constructing the optimal aggregation tree has been proved to be NP-hard. Current algorithms of constructing aggregation trees are all approximation algorithms, and they mainly aim to prolong the network lifecycle or reduce the aggregation delay. For the tree structure, the degree of nodes and the depth of the tree are critical for the network lifecycle and aggregation delay.

The algorithm proposed by Li et al. [14] can reduce the aggregation delay, but it doesn't take the balance of energy consumption of sensors into consideration. Then, some sensor nodes may run out their energy very quickly, which makes the network lifecycle be very short. In wireless sensor networks, the energy consumed by CPU operations is much smaller than the energy consumed by data transmission, and the energy consumed by data transmission depends on the number of sending and receiving operations. In order to save energy consuming, it needs to limit the number of operations for sending and receiving data, and then the nodes with the most children nodes will be the bottleneck of the whole network. So, in order to prolong the network lifecycle, the degrees of nodes in the aggregation tree must be balanced. In this paper, we propose a data aggregation tree constructing algorithm. The main idea of the proposed algorithm is adjusting the tree structure according to the consumed energy of each node after the initial aggregation tree is constructed. For predicting the consumed energy for each node, we estimate the consumed energy based on its location in the tree. The detail of the algorithm is in algorithm 1.

Algorithm 1. Aggregation tree constructing algorithm

Input:  $G = (V, E)$ , radius  $r$ ;

Output: Data aggregation tree  $DAT = (V, \hat{E})$

1. Assuming  $v_s$  as root and  $r$  as radius, construct a breadth-first searching tree  $G$ ;
2. For  $\forall v_i \in V$ , let  $C(i) = White$ ,  $Mark(i) = 0$ ,  $R(i) = (level_i, i)$ ;
3. Let  $C_{v_s}(i) = Black$ , broadcast  $Msg(B)$ ;
4. For each node  $v_j$  that has received  $Msg(B)$ ,
5. If  $C(j) = White$ , then
6. let  $C(j) = Grey$  and broadcast  $Msg(G)$ ;
7. If  $C(j) = White$  and " $v_i, \|v_i - v_j\| \leq r, R_j < R_i$ ", then
8. Let  $C(j) = Black$  and broadcast  $Msg(B)$ ;
9. Let  $DS = \{v_i | C(j) = Black\}$ ;
10. For root  $v_s$ , broadcast  $Msg(G_{join})$ ;

11. For each node  $v_i$  that has received  $Msg(G_{join})$
12. If  $v_i \notin DS$ ,  $Mark(i) = 0, C(i) = Grey$ , then
13. Let  $DS = DS \cup \{v_i\}$ , broadcast  $Msg(B_{join})$ ,  $Mark(i) = 1$ ;
14. If  $v_i \notin DS$ ,  $Mark(j) = 0, C(j) = Black$ , then
15. Let  $DS = DS \cup \{v_i\}$ , broadcast  $Msg(G_{join})$ ,  $Mark(i) = 1$ ;
16. Assuming  $DS$  as non-leaf nodes and other nodes as leaf nodes, construct a tree;
17. For all non-root nodes  $v_i \in V$ , let its degree be  $Deg(i)$ ,  $H(i) = hops(v_i, v_s)$ ;
18. Solve the following optimal problem:
 
$$P_i = \arg \min_j (Deg(j))$$

$$s.t. \|v_i - v_j\| \leq t$$

$$H(i) - H(j) = 1$$
19. Return  $DAT = (V, \hat{E})$ ;

In the algorithm, we first construct a breadth-first searching tree for  $G$ . If the level of node  $i$  is  $level_i$ , we tag  $R(i) = (level_i, i)$ . After the first 8 steps, all black nodes build a dominant set  $DS = \{v_i | C(j) = Black\}$ ; for lines 9-16, we construct a connected tree; and finally for all non-root nodes, we select the up level node with the least children nodes as its parent node, and build the final aggregation tree  $DAT = (V, \hat{E})$ .

V. DATA AGGREGATION SCHEDULING ALGORITHM

The aim of data aggregation scheduling algorithm is to assign intervals for nodes. In this paper, we proposed an approximating maximal weighted independent set based scheduling algorithm. Firstly, we select the link set, where each link is able to communicate, and construct the link collision matrix according to the collision relationships in the data aggregation tree. Then, based on the collision matrix, we ascertain the communication links in each interval according to constructing an approximate maximal weighted independent set.

In order to reduce the energy consumption of nodes, the state-of-the-art algorithms schedule each link only once, but they can't make sure the data with high priorities are scheduled first. If we don't limit the number of communications for each node, then the energy consumption of the network will increase greatly. So, we need to select the number of communications gently to get the trade-off between low weighted delay and high network lifecycle.

**Definition 5. Marginal link set.** For the tree  $T = (V, \hat{E})$ , if all children nodes of the node  $v_i$  has been scheduled after  $t - 1$  intervals, then the link  $e_i$ , whose sending end is  $v_i$ , is the marginal link of  $T$  at time  $t$ . We denote  $FLS(T)_t$  as the set of marginal link at time  $t$ .

**Definition 6. Accumulated weight.** For  $e_i \in \hat{E}$  and in the interval  $t$ , the sum of remaining weight that hasn't been used to transmit data is the accumulated weight in this link, i.e.  $ACW(e_i)_t$ .

**Definition 7. Active link set.** In the interval  $t$ , the link set that can be used to schedule, including marginal link and non-marginal link with accumulated weighted greater than threshold  $d$ , is the active link set at time  $t$ , i.e.

$$ALS(T)_t = \{e_i | e_i \in FLS(T)_t, \vee ACW(e_i) \geq \delta\} . \quad (4)$$

**Definition 8. Collision link set.** For " $e_i \in E_t$ ", if  $e_i$  collides with  $n$  links in  $E_t$ , then these  $n$  links construct the collision link set. The sum of all weights in the collision link set is the collision set weight for  $e_i$  to  $E_t$ . In the scheduling algorithm that limits the number of communications, the marginal links are the objects to schedule, and even there exists high priority links, they should be scheduled after they become marginal links.

In this paper, we assume the active links as the objects to schedule, add the intermediate links, whose priorities are higher than the predefined threshold  $d$ , into the scheduling queue, and then make a balance between weighted delay and energy consumption. The detail of the proposed data aggregation scheduling algorithm is in algorithm 2.

Algorithm 2. Data aggregation scheduling algorithm

Input: Link weight set  $W = \{w_1, w_2, \dots, w_n\}$ , data aggregation tree  $DAT$ , threshold  $d$ ;

Output:  $SCHE = \{E(1), E(2), \dots, E(t)\}$ ;

1. Initialize  $SCHE = \Delta$ ,  $t = 1$ , and tag all links to be unscheduled;
2. For each unscheduled link  $e_i$
3. If  $e_i \in FLS(T)$  or  $ACW(e_i) \geq d$ , then
4. Let  $FLS(T)_t = FLS(T)_t \cup \{e_i\}$ ;
5. Construct the link collision matrix  $IM_t$  based on  $ASL(T)_t$ : if links  $i$  and  $j$  collide in  $T$ , then let  $IM_t(i, j) = 1$ ; and  $IM_t(i, j) = 0$  otherwise;
6. Construct the maximal independent set  $MIS_t$  based on the adjacent matrix  $IM_t$ ;
7. For each active  $e_j$  ( $e_j \notin MIS_t$ )
8. If  $e_i$  has not been scheduled, then
9. If  $ACW(e_i)$  is bigger than the collision set weight of  $MIS_t$ , then
10. Substitute the link collision set in  $MIS_t$  with  $e_j$ , and construct the approximate maximal weighted independent set  $MWIS_t$ ;
11. Let  $E(t) = MWIS_t$ ,  $SCH = SCH \cup \{E(t)\}$ ;
12. If  $e_i \in MWIS_t \cap FLS(T)_t$  then
13. Tag  $e_i$  be already scheduled;
14. If there is still unscheduled links in  $DAT$ , then

15. Let  $t = t + 1$ , and go to step 2.
16. Return  $SCHE = \{E(1), E(2), \dots, E(t)\}$ ;

The output  $SCHE$  of algorithm 2 satisfies the following three conditions:

- $\hat{E} = E(1) \cup E(2) \cup \dots \cup E(t)$ ;
- The data collected by all nodes arrive at *sink* node;
- If links  $e_i, e_j \in E(s)$  ( $1 \leq s \leq t$ ), then there is not a collision between  $e_i$  and  $e_j$  in  $T$ .

Based on the above conditions, we can conclude that the proposed algorithm is an effective and secure data aggregation algorithm.

## VI. EXPERIMENTS

In the experiments, we denote the proposed algorithm as *aggreTree*, and compare it with two state-of-the-art data aggregation algorithms, DAS and IAS. In the experimental setting, we randomly deploy a number of sensor nodes in an area of  $200m \times 200m$ , and the sink node is at the center of the area. The parameters of the simulated wireless sensor network are the same as [18].

Firstly, we compare the aggregation delay, the average weighted delay and the network lifecycle of the three algorithms, and the results are in figures 1-3 respectively. In figure 1, the proposed algorithm has the least aggregation delays under different node densities, which means that the proposed algorithm consumes less time while aggregating data from sensor nodes. In figure 2, the proposed algorithm also has the least average weighted delay, and the reason is that by adjusting the tree structure, the proposed algorithm can have a good trade-off between low weighted delay and high network lifecycle. From figures 1 and 2, we can infer that the proposed algorithm will result longer network lifecycle, and this assumption is verified in figure 3.

Next, we let  $d = 0$ ,  $d = 40$  and  $d \rightarrow \infty$ , and compare the five different results. Figure 4-6 are the comparison results aggregation delay, the average weighted delay and the network lifecycle. From these figures we can see that, as we increase  $d$  from 0 to 40 to  $\infty$ , the aggregation delay reduces, the network lifecycle increases, but the average weighted delay reduces and then increases again. So, in order to get good performance, we must carefully select the parameter threshold  $d$  for different performance metrics.

Finally, we let  $N = 200$ , 600 and 1000, and observe the effect of  $d$  on the aggregation delay, the average weighted delay and the network lifecycle. The results are in figures 7-9 respectively. From these figures we can see that, as  $d$  increases the aggregation delay decreases, the network lifecycle increases, but for average weighted delay, it doesn't change very much. The reason is that, the average weighted delay takes the aggregation tree reconstruction into consideration, so the communication cost reduces, and thus the average weighted delay doesn't have an obvious relationship with parameter threshold  $d$ .

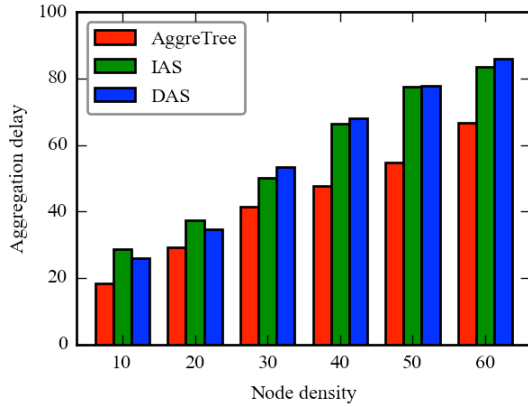


Figure 1. Comparison of aggregation delay

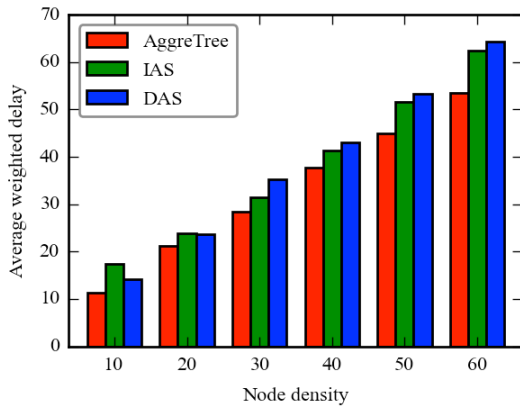


Figure 2. Comparison of average weighted delay

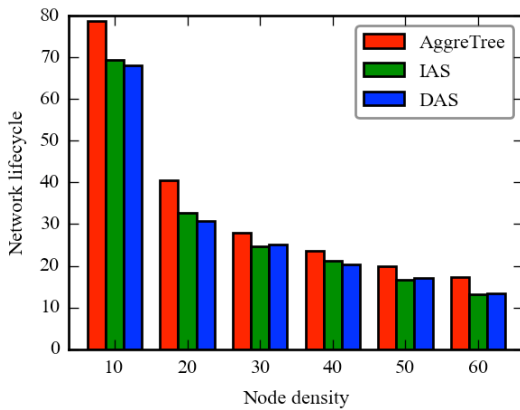


Figure 3. Comparison of network lifecycle

VII. CONCLUSION

Data aggregation is a hotspot in the research of wireless sensor networks. While randomly deployed in a region, sensors usually collect redundant data, which will consume more energy. For the energy-limited sensors, this is a critical problem. Tradition data aggregation algorithms in wireless sensor networks mainly aim to improve the energy utilization, and ignore the security and lifecycle. In order to get a good trade-off between these requirements, we proposed a data aggregation algorithm based on constructing a data aggregation tree. After give a formalism description of the problem, we proposed a data aggregation tree constructing algorithm. By minimize the maximal energy consumption of nodes, the algorithm can prolong the lifecycle. In data aggregation scheduling algo-

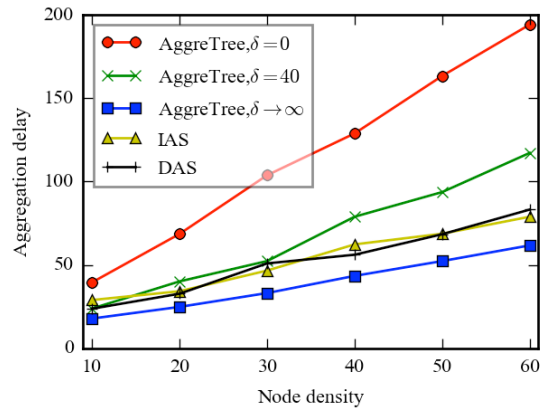


Figure 4. Comparison of aggregation delay for  $d$

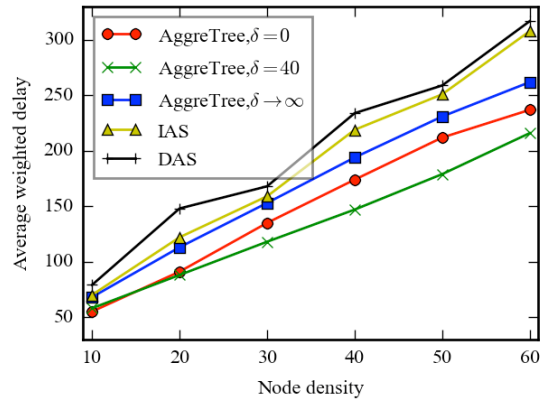


Figure 5. Comparison of average weighted delay for  $d$

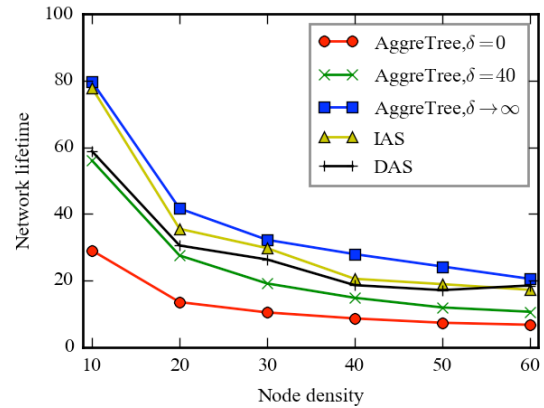


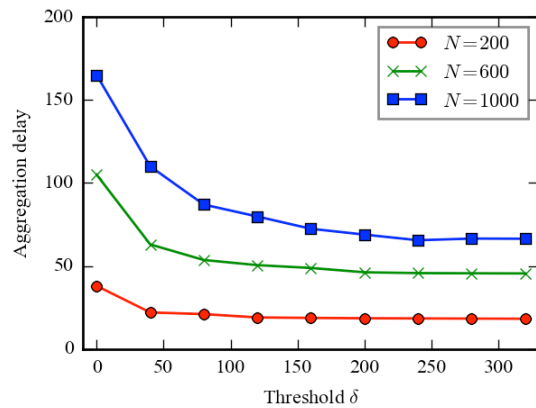
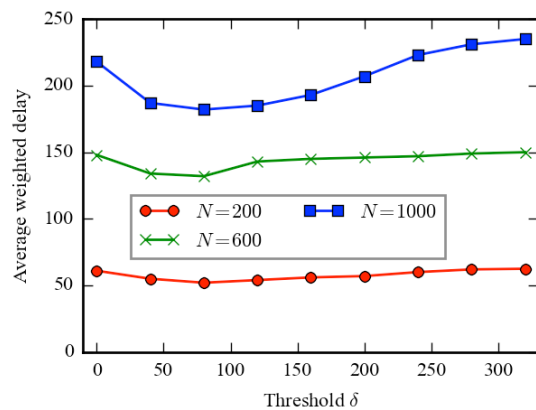
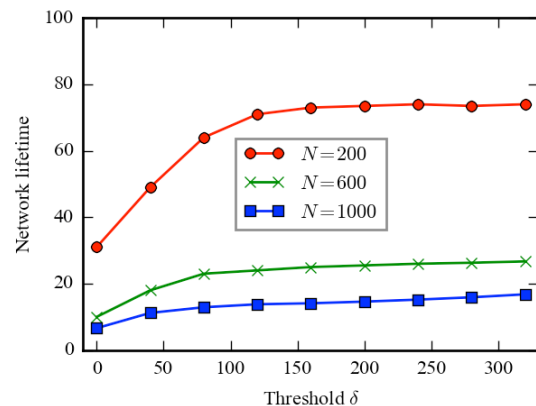
Figure 6. Comparison of network lifecycle for  $d$

rithm, we select the number of communications carefully to get the trade-off between low weighted delay and high network lifecycle. The simulation experiments show that, the proposed data aggregation algorithm consumes less energy while aggregating data from sensor nodes, and thus can prolong the network lifecycle.

REFERENCES

[1] Younis M, Senturk I F, Akkaya K, et al. Topology management techniques for tolerating node failures in wireless sensor networks: A survey[J]. Computer Networks, 2014, 58: 254-283. <http://dx.doi.org/10.1016/j.comnet.2013.08.021>

[2] Madhumathy P, Sivakumar D D. Power Efficient Data Aggregation in Wireless Sensor Networks[C]//Special Issue, 4th National Conference on Advanced Computing, Applications & Technologies. 2014.

Figure 7. Comparison of aggregation delay for  $N$ Figure 8. Comparison of average weighted delay for  $N$ Figure 9. Comparison of network lifecycle for  $N$ 

- [3] Yuan F, Zhan Y, Wang Y. Data density correlation degree clustering method for data aggregation in WSN[J]. Sensors Journal, IEEE, 2014, 14(4): 1089-1098. <http://dx.doi.org/10.1109/JSEN.2013.2293093>
- [4] Gupta G P, Misra M, Garg K. Energy and trust aware mobile agent migration protocol for data aggregation in wireless sensor networks[J]. Journal of Network and Computer Applications, 2014, 41: 300-311. <http://dx.doi.org/10.1016/j.jnca.2014.01.003>
- [5] Bagaa M, Challal Y, Ksentini A, et al. Data aggregation scheduling algorithms in wireless sensor networks: Solutions and challenges[J]. Communications Surveys & Tutorials, IEEE, 2014, 16(3): 1339-1368. <http://dx.doi.org/10.1109/SURV.2014.031914.00029>
- [6] Li H, Wu C, Hua Q S, et al. Latency-minimizing data aggregation in wireless sensor networks under physical interference model[J]. Ad Hoc Networks, 2014, 12: 52-68. <http://dx.doi.org/10.1016/j.adhoc.2011.12.004>

- [7] Lowry R, Leadbetter A. Semantically supporting data discovery, markup and aggregation in the European Marine Observation and Data Network (EMODnet)[C]//EGU General Assembly Conference Abstracts. 2014, 16: 13693.
- [8] Sert S A, Onur E, Yazici A. Security attacks and countermeasures in Surveillance Wireless Sensor Networks[C]//Application of Information and Communication Technologies (AICT), 2015 9th International Conference on. IEEE, 2015: 201-205.
- [9] Guo X, Wang Y, Wei X. Design of WSN-based environment monitoring system on repair of gas leakage[C]//Control and Decision Conference (CCDC), 2015 27th Chinese. IEEE, 2015: 3340-3344.
- [10] Demigha O, Hidouci W K, Ahmed T. On energy efficiency in collaborative target tracking in wireless sensor network: a review[J]. Communications Surveys & Tutorials, IEEE, 2013, 15(3): 1210-1222. <http://dx.doi.org/10.1109/SURV.2012.042512.00030>
- [11] Goel P K, Sharma V K. Data Aggregation Security in Wireless Sensor Network[J]. International Journal of Advanced Research in Computer Science, 2012, 3(5).
- [12] Katiyar V, Chand N, Soni S. Clustering algorithms for heterogeneous wireless sensor network: A survey[J]. International Journal of Advanced Networking and Applications, 2011, 2(4): 745-754.
- [13] Boyinbode O, Le H, Takizawa M. A survey on clustering algorithms for wireless sensor networks[J]. International Journal of Space-Based and Situated Computing, 2011, 1(2-3): 130-136. <http://dx.doi.org/10.1504/IJSSC.2011.040339>
- [14] Li H, Lin K, Li K. Energy-efficient and high-accuracy secure data aggregation in wireless sensor networks[J]. Computer Communications, 2011, 34(4): 591-597. <http://dx.doi.org/10.1016/j.comcom.2010.02.026>
- [15] Przydatek B, Song D, Perrig A. SIA: Secure information aggregation in sensor networks[C]//Proceedings of the 1st international conference on Embedded networked sensor systems. ACM, 2003: 255-265.
- [16] Mahimkar A, Rappaport T S. SecureDAV: A secure data aggregation and verification protocol for sensor networks[C]//Global Telecommunications Conference, 2004. GLOBECOM'04. IEEE. IEEE, 2004, 4: 2175-2179.
- [17] Cai D, He X, Han J. SRDA: An efficient algorithm for large-scale discriminant analysis[J]. Knowledge and Data Engineering, IEEE Transactions on, 2008, 20(1): 1-12. <http://dx.doi.org/10.1109/TKDE.2007.190669>
- [18] Sugandhi N, Manivannan D. Analysis of Various Deterioration Factors of Data Aggregation in Wireless Sensor Networks[J]. IJET, Volume5, No1, ISSN, 2013: 0974-4024.
- [19] Sun Y, Luo H, Das S K. A trust-based framework for fault-tolerant data aggregation in wireless multimedia sensor networks[J]. Dependable and Secure Computing, IEEE Transactions on, 2012, 9(6): 785-797. <http://dx.doi.org/10.1109/TDSC.2012.68>
- [20] Wu J Q, Guo J H. Improved pattern-based encrypted data aggregation scheme for clustered wireless sensor networks[C]//Design, Manufacturing and Mechatronics: Proceedings of the 2015 International Conference on Design, Manufacturing and Mechatronics (ICDMM2015). World Scientific, 2015: 475-481.

## AUTHORS

**HU Yanhua** is with the Department of Electrical and Computer Engineering, Lushan College of Guangxi University of Science and Technology, Liuzhou, Guangxi 545616, China (huyanhuals@sina.com).

**ZHANG Xincui** is with the Department of Information Technology, College of Information & Business, Zhongyuan University of Technology, Zhengzhou, Henan 451191, China.

Submitted 27 December 2015. Published as resubmitted by the authors 25 March 2016. This work was supported in part by the National Significant Science and Technology Special Project of China and the National Natural Science Foundation of China.