An Energy-efficient UAV-based Data Aggregation Protocol in Wireless Sensor Networks

Qiyue Wu
University of Ottawa
Ottawa, Canada
qwu042@uottawa.ca

Peng Sun
University of Ottawa
Ottawa, Canada
psun044@uottawa.ca

Azzedine Boukerche
University of Ottawa
Ottawa, Canada
boukerch@site.uottawa.ca

ABSTRACT

Energy efficiency is an important issue in Wireless Sensor Networks (WSNs). The energy is mainly consumed by data sensing, data transmission and movement of sensors. The energy consumed by data transmission is much larger than data sensing. A potential solution for energy saving is applying the external devices to collect data to prolong the network lifetime. In this paper, we adopt an unmanned aerial vehicle (UAV) serving as the data mule and propose a novel energy-efficient UAV-based data aggregation protocol in WSNs to reduce the energy consumption of sensors. By considering a clustered WSN, our approach computes an optimal path for data mule through all cluster heads (CHs) while achieving a relatively high system-wide energy efficiency. Moreover, we introduce a genetic algorithm to derive a near-optimal solution for a large-scale WSN while reducing the computing time. We compare our protocol with state-of-the-art approaches, and the simulation results demonstrate that the proposed algorithm improves the network performance on energy efficiency.

KEYWORDS

Wireless sensor network, energy efficiency, routing, data mule, unmanned aerial vehicle

ACM Reference Format:

Qiyue Wu, Peng Sun, and Azzedine Boukerche. 2018. An Energy-efficient UAV-based Data Aggregation Protocol in Wireless Sensor Networks. In *8th ACM International Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications (DIVANet'18), October 28-November 2, 2018, Montreal, QC, Canada.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3272036.3272047

1 INTRODUCTION

Wireless Sensor Networks (WSN) is an ad hoc network consisting of distributed tiny devices called sensor nodes [24]. It is often adopted to military and civil applications, such as target detection and tracking. Sensor nodes can collect, transmit and receive data in order to detect and track the targets in WSNs [28]. Normally, the primary energy source of a sensor node is a battery with the limited energy which is hard to change when the nodes are deployed in the complicated and dangerous environment. The limited available energy budget directly affects the lifetime and performance of the system [3, 6]. Consequently, it is very critical to improve the energy efficiency of WSNs to prolong the battery life and make the WSNs sustainable. Applying solar cells instead of the batteries seems to be a feasible solution, however, it is hard to keep persistent due to varied weather conditions. Nowadays, how to improve energy efficiency is still a crucial issue.

The energy of sensors in WSNs is mainly consumed by three ways, i.e. data transmission, movement and sensing. The huge energy consumed on data transmission causes the sensor runs out of its energy quickly. The nodes closed to the data sink are prone to run out of their energy, since all messages from sensor nodes to the sink have to go through these nodes, which causes a huge amount of energy consumption [20]. Once these nodes deplete their energy, they cannot communicate with others, which possibly cause the network disconnected. It is commonly called energy-hole problem [18]. To avoid this issue, we apply mobile devices for data aggregation, which is a practical way to balance the workload for all sensors in the network.

There are many advantages of data aggregation by adopting mobile devices in WSNs. One of them is ensuring that the data from all sensors can be transmitted to sink. In some particular situations, due to the hazardous environment and cost, the deployed system cannot achieve the system-wide connectivity. For example, in the large-scale military area, we can only deploy sensor nodes at designated positions. However, it may cause the low density of nodes, which means the distance between two nodes might be beyond the transmission range, and the further node cannot transmit messages to the sink through intermediate nodes. Adopting mobile devices to collect data from these nodes can be considered a applicable solution. It can reduce the communication load of all sensors on the one hand [10]. On the other hand, based on mobility, mobile devices can refill their batteries so that they can gather data consistently.

In this paper, we design a novel energy-efficient data aggregation protocol in WSNs by using UAV. We adopt UAV as the data mule to collect data because of its powerful communication and computational abilities. The objectives of the proposed protocol are (1) balancing the system-wide energy consumption and system throughput; (2) reducing the delivery latency from sensors to the data sink. To ensure that the data can be forwarded to the sink within a tolerant delay, we apply genetic algorithm (GA) to search the shortest path taken by the data mule to achieve lower latency [32]. GA considered a heuristic algorithm provides an optimization scheme that identifies near-optimal solutions in order to satisfy the given constraints. Accordingly, the main contribution of our work is solving a joint cluster head (CH) selection problem and

DIVANet'18, October 28-November 2, 2018, Montreal, QC, Canada

ACM ISBN 978-1-4503-5964-1/18/10…$15.00
https://doi.org/10.1145/3272036.3272047

the routing problem of the data mule by a GA-based optimization method. Consequently, we can obtain the optimal route of the data mule, and guarantee lower energy consumption of the WSN in the meantime.

The rest of this paper is organized as follows. Section II discusses related work regarding the energy efficient with data mule. The network model is shown in section III. Section IV proposes the mathematical programming formulation of our protocol. The genetic algorithm of our proposed protocol is described in Section V. Section VI shows the simulation results. The conclusion is made in section VII.

2 RELATED WORK

To cope with the energy efficient problem in WSNs, many approaches have been proposed in recent years. Currently, the existing approaches can be roughly categorized into three classes [3]: node scheduling (e.g. dynamic duty-cycle adjustment topology [5, 19, 21]), data-driven approaches (e.g. energy-efficient data acquisition [25]), and mobility-based approaches. In this paper, we develop a mobility-based protocol since it can improve the energy conservation in WSNs [26]. For example, a connected network might become disconnected due to energy depletion of sensor nodes. The connectivity of a deployed WSN can be recovered if nodes have the capacity of mobility. Additionally, if a mobile device is adopted to gather data, the average length of data forwarding path from the sensor node to the sink can be reduced. Accordingly, the accumulated forwarding and link errors can be reduced. Consequently, the network lifetime can be prolonged as well. There are two common types of schemes based on mobility, i.e., mobile-sink-based schemes and mobile-relay-based schemes. Many mobile-sink-based protocols have been proposed [8, 15, 16, 32]. For example, Wang et al. [32] introduced a mobile sink into WSNs. In the presented work, the mobile sink could move to some designated locations to communicate with sensor nodes. The mobile sink stayed at a place for a certain period. Nodes covered by the mobile sink could send data to the sink directly. The rest of nodes sent data to the mobile sink through paths selected between the mobile sink and themselves on the basis of some given routing protocols. A rendezvous-based approach with a mobile sink was introduced in [15]. The authors adopted rendezvous nodes which gathered the data from clusters and sent the data to the mobile sink once the sink appeared in its communication range. The advantage of this protocol was minimizing the overall network overhead and energy expenditure. Furthermore, it can balance energy consumption among sensors to prolong network lifetime. However, the crucial problem for mobile-sink-based protocols was the shortest path routing protocol. Because of the movement of the sink, the nodes have to rebuild the route for transmission according to the location of the sink frequently. The overhead of constructing the new topology may be huge.

To overcome the problem mentioned above, the researchers also provided mobile-relay-based schemes [4, 13, 14, 22, 23, 27, 34]. One kind of mobile-relay-based schemes is applying mobile devices to gather data. The mobile devices traverse around in a network, communicate with sensor nodes to collect data, and eventually forward the collected data to the sink. Kong et al. [14] proposed a GA-based routing protocol for WSNs. Compared with traditional WSNs composed of sensor nodes and the data sink, a new component, the middle layer, was introduced in their work. The middle layer consisted of several stations which are for receiving data from sensor nodes and forwarding data to the sink. Therefore, for each sensor, the average of energy consumed by data transmission can be reduced. However, extra construction cost was introduced for building stations, which was a main drawback of their work. A load balanced and energy-efficient data collection scheme with a data mule was proposed in [23]. The path of data mule was a fixed vertical path in the middle of the whole network. The sink divided the network into grid-shaped clusters with different sizes. The size of each cluster was determined by the distance from the cluster to the path of data mule, that meant the size of further clusters was bigger than the closed ones for balancing the workload of all sensors. The node closest to the centroid was selected as the CH. The similar work could be found in [27]. The approaches mentioned previously share common drawback, which is, the load of the CHs closed to the path of data mule was higher than others that shorten the lifetime of these nodes. The authors in [34] presented a two-phase cluster-based genetic algorithm (CBGA). They introduced a new notation, waypoint, which is the centroid of the overlapped communication area of a set of sensor nodes. Therefore, instead of visiting actual CHs, the data mule can communicate with every node in the system by visit a set of waypoints. For more details, the first phase was to configure small disjoint clusters based on the proposed clustering algorithm, and the second phase was to design the path of data mule by using GA. However, the path length of the data mule is considerable, which causes the high delivery latency.

In this paper, we develop a mobile-relay-based approach by adopting GA, and our proposed protocol improves the performance of WSNs (i.e. delivery latency and energy efficiency) by shortening the path length of data mule.

3 SYSTEM MODEL

In this paper, all the sensor nodes are randomly deployed in a uniform-distributed large-scale area. The set $S = \{S_1, S_2, S_3, ..., S_N\}$ denotes the set of N sensor nodes. We make several assumptions for the network model, which are listed as the follows:

- The field of interest (FoI) is divided into multiple grids with the identical size. The nodes in one grid form a cluster. The number of nodes is on the basis of [30] in order to guarantee the connectivity in each cluster [29].
- Each node is stationary after deployment and has a unique id ID_i.
- All nodes have identical sensing range R_s. Meanwhile, they have same transmission range R_c.
- The sink is deployed outside of the FoI. It knows the location of each node.
- The UAV considered as a data mule has computational and transmission capacities. Its transmission range is donated by R_{DM}. The UAV can communicate with sensor nodes when the distance between them is $min\{R_c, R_{DM}\}$.

The network model in the 2D platform is shown in Fig. 1.

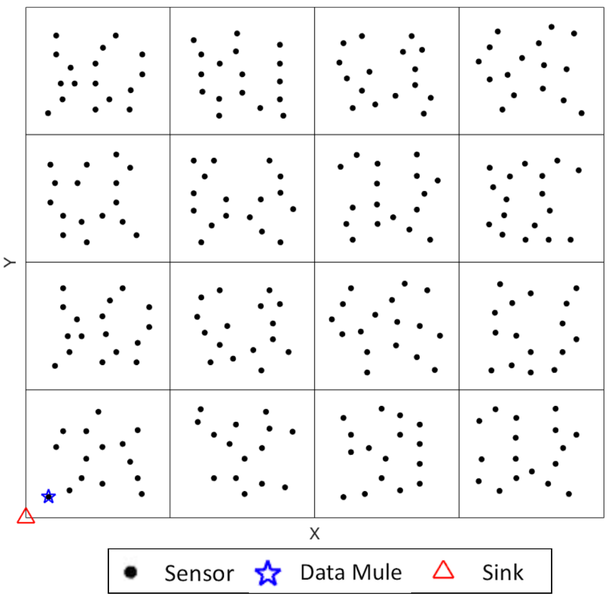


Figure 1: Network Model

4 PROBLEM FORMULATION

In this section, we introduce the formulation for minimizing the energy consumed by sensor nodes and the data mule in WSNs.

For the sensor nodes, we only consider the energy consumption of data transmission. We adopt the energy model introduced in [11], which is also adopted in [7]. For a given distance d between two nodes, the energy consumption for transmitting a k-bit message is given by [11],

$$E_t = E_{elec} \times k + \epsilon_{amp} \times k \times d^2, \tag{1}$$

where E_{elec} is the energy cost of transmitter electronics with unit J/bit. ϵ_{amp} is the energy cost of transmit amplifier with unit $J/bit/m^2$, which is proportional to the square of distance of data transmission [17]. The first part of this formulation presents the energy consumption of radio dissipation, and the second is the energy consumption of amplifying radio. In each cluster, cluster members send the sensed data to CH. Hence, the total energy consumption in the cluster j when each cluster member forwards one packet to CH is

$$E_T^j = \sum_{u \in C_j} E_{elec} \times k + \epsilon_{amp} \times k \times d^2_{(CH,u)}, \tag{2}$$

where CH is the cluster head of cluster j, u is the cluster member in cluster j.

For data mule, the energy is mainly consumed by movement. According to the dynamics, the energy E_{DM} required for data mule to move form current position to next CH in cluster j is

$$E_{DM}^j = P \times \frac{D_{(current,CH_j)}}{v}, \tag{3}$$

where P and v represent the power rate and the speed of data mule, respectively. $D_{(current,CHj)}$ is the distance between the current location of data mule and next CH.

In order to balance the energy expenditure between the sensor nodes and data mule, we introduce two user-configurable coefficients α, β. The coefficients can be adjusted based on the network requirements. For example, when the sensor nodes are deployed in some remote and hazardous area, it is hard to change their batteries. In this case, compared to the energy consumption in data mule, the energy consumption in sensor nodes is more critical. Therefore, the coefficient of energy consumption for sensor nodes should be higher. Moreover, we need normalize E_T^j and E_{DM}^j to make these two values on different scales to a common scale. E_T^j is normalized by E_T^{max} which is the maximum value of E_T^j among all clusters in the deployed WSN. Because P and v in (3) are constants, E_{DM}^j is proportional to the moving distance of data mule. Thus, E_{DM}^j is normalized by E_{DM}^{max} that is the energy consumed by data mule when it travels between the pair of CHs with the longest distance.

In general, the objective of our work is to minimize the energy consumption of the whole network. The corresponding optimization functions for the overall energy consumption is given by,

$$min \quad E_{total} \tag{4}$$

$$E_{total} = \sum_{j=1}^{M} (\alpha \times \frac{E_T^j}{E_T^{max}} + \beta \times \frac{E_{DM}^j}{E_{DM}^{max}}) \tag{5}$$

$$s.t. \quad \alpha + \beta = 1, \tag{6}$$

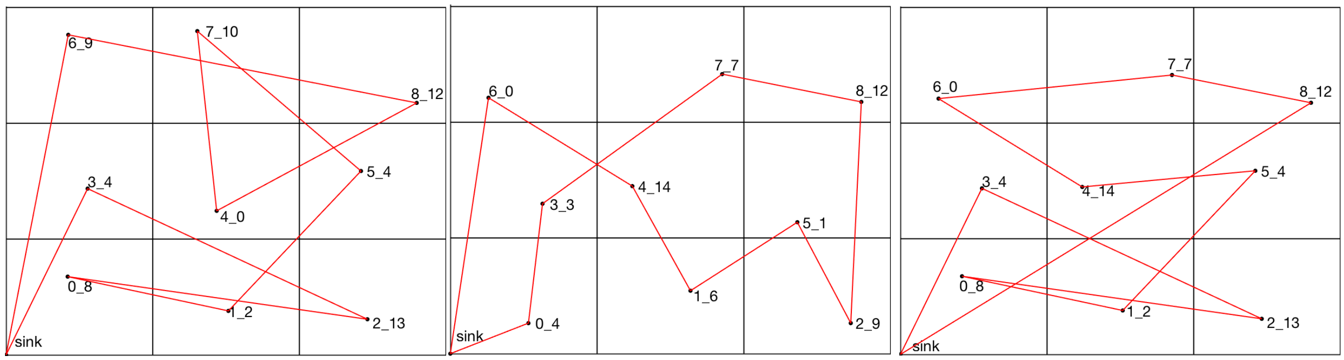
where M is the number of clusters, and E_{total} is the total energy consumption of the entire system.

5 THE PROPOSED PROTOCOL

As mentioned previously, data mule has to visit the sink and CHs of all clusters. In each cluster, a CH has to be chosen among all nodes according to the total energy consumption of all nodes in its cluster calculated by (2) and the distance that data mule needs to move to it. However, the computational complexity for solving the aforementioned formulation to achieve the global optimization could be huge. To reduce the complexity, we design a novel protocol based on GA. GA is an applicable method to derive the local optimum, and it is suitable for optimization problems in a large-scale WSN [9]. GA firstly generates the initial population (i.e., initial solutions) and stores the population in a list based on their fitness value. In each iteration, on the basis of the initial population(or the subsequent population generated by the previous iteration), new solutions would be generated through three main operations, i.e., selection, crossover, and mutation. After the predetermined number of iterations, the solution with the best fitness value is selected to be the final solution. To define an appropriate fitness function to measure the quality of the solution is essential to implement GA. In what follows, we will introduce the proposed protocol in details.

5.1 Topology Construction Phase

Initially, sensor nodes are deployed randomly in the large-scale area. The FoI is divided into multiple grids with the identical size. The sensors in each grid form a cluster. Recall that, the global position information is known by the sink, and every node knows its location information. For each CH, it gathers data from all nodes in the same cluster and then forwards gathered data to the data mule. In each



(a) Parent1: Fitness=0.214751

(b) Parent2: Fitness=0.21253

(c) Offspring: Fitness=0.262065

Figure 2: The offspring generated by SCX with the pair of parents chromosomes

cluster, to reduce overall energy consumed by sensor nodes for transmitting data to the CH, the aggregated length of the path from cluster members to the CH has to be shortened. Accordingly, the CH in each cluster is chosen by the data sink based on the sum of energy consumed by other nodes in its cluster when these nodes transmit data to it, and we introduce E_T^i to represent it.

5.2 Initial Population Generation

After E_T^i of all sensors is calculated, the data sink begins to figure out the path of the data mule by using GA. GA uses chromosomes to encode solution of data [2]. In our work, each unique chromosome represents an optional path of data mule. Every gene in the chromosome is denoted by the sink or a sensor node in the system, which is encoded by *ID* of each node or the sink. In this paper, we use a new designed random initial algorithm to generate initial population. To generate a new chromosome, the system randomly chooses candidate genes, and then adds them to the chromosome. In our design, for each chromosome, only one sensor node can be selected from a cluster. Let H be the population size. The value of H should be set large enough, which is based on the size of FoI. Assuming that there are M clusters in the area, and each cluster has at least n sensors, the number of optional trajectories for data mule is $n^M \times M!$. Thus, the probability that we can obtain the optimal solution increases as the value of H increases.

5.3 Fitness

The fitness function evaluates the quality of each solution for deriving the optimum. The better chromosome with higher fitness value should be chosen. As the mentioned in [1], for the minimization problem, the fitness function is set to be the reciprocal of the objective function. Therefore, the fitness value of our work is defined as

$$F = \sum_{j=1}^{M} \left(\frac{1}{\alpha} \times \frac{E_T^{max}}{E_T^j} + \frac{1}{\beta} \times \frac{E_{DM}^{max}}{E_{DM}^j}\right), \tag{7}$$

where E_T^{max} and E_{DM}^{max} is calculated by sink in advance.

5.4 Selection and Crossover

Based on generated initial population, chromosomes are sorted according to their fitness value. The new chromosome is generated by two steps: 1) Selection operation: Two chromosomes randomly are chosen from the population as two parents; 2) Crossover operation: Based on the chosen parents, one offspring is produced through crossover operation.

In GA, the crossover is an essential operation. The search space of GA is improved by crossover because new offsprings are generated constantly. Thus, the crossover probability P_C should be set high. Similar to [34], we also adopt sequential constructive crossover (SCX) [1] as crossover operator, which guarantees that the quality of the offspring generated by SCX is higher than its parents. In other words, the offspring is closer to the optimal solution.

Let's introduce the SCX through an example which is based on the assumption in our protocol. The pair of chromosomes as parents are chosen by selection operation among the population, which are shown in Fig. 2-(a) and Fig. 2-(b), respectively. Here, the node is defined by i_j, where i indicates the cluster number, and j is the serial number of nodes in its cluster. We set $\alpha = 0.8$ and $\beta = 0.2$ in this example. The candidate gene with a smaller value derived by (5) should be added to the offspring. Recall that, it is necessary to guarantee that only one sensor in each cluster has to be visited by the data mule. The offspring is empty at first. Since the data aggregation procedure always starts from the sink, the sink has to be the initial point of all chromosomes. Starting from the sink, the candidate of the next node of the offspring in parent 1 is node 3_4, and the candidate in parent 2 is node 6_0. The value of node 3_4 is smaller than node 6_0, which are 0.3614 and 0.483, respectively. So, node 3_4 is added to the offspring and set to be the current node now. Furthermore, The cluster 3 which has been visited need to be recorded. In next step, the candidate is 2_13 in the parent 1, and the current node is not existing in the parent 2. But node 3_3 in the parent 2 is from cluster 3, so the candidate for the parent 2 is node 7_7 which is the next node of node 3_3 in its path. As the end of this step, node 2_13 with small value is added to the offspring. To ensure only one node in each cluster should be visited, all nodes in the visited cluster will not be considered as candidate genes anymore. In the parent, the candidate genes can only be selected from one of the unvisited clusters. For example, the current node is node 0_8. For the parent 2, the sink which is the next node of node 0_4 that is from cluster 0 should be chosen for the candidate. However, it has been visited in the beginning, and there are some clusters have not been visited yet. Hence, according to the order of unvisited clusters (i.e., cluster 1, cluster 2, cluster 4, cluster 5, cluster 6, cluster 7, cluster 8), the candidate in parent 2 is node 1_6. After the offspring is generated, its fitness value should be calculated. In this example, the offspring (Fig. 2-(c)) has higher fitness value than its parents, which means the solution represented by the offspring is better than its parents. Then, the parents with the lower fitness value will be replaced by the offspring.

5.5 Mutation

Mutation is another important operation of GA by which, new chromosomes can be generated by replacing some genes in original chromosomes. Consequently, it is helpful to improve the diversity of population and search space. The mutation probability P_M would be set low in case the optimal search is broken. In this paper, we adopt the Gaussian mutation introduced in [12]. In detail, the i-th gene is mutated on the basis of an offset generated by the Gaussian distribution $N(\mu, \sigma^2)$ with the mean μ and the variance σ^2. Furthermore, new gene has to be selected from the cluster of the replaced gene. Gaussian mutation is effective for GA to converge towards a better solution.

5.6 Steady Phase

By running GA in a predetermined number of generations, the chromosome with the highest fitness value is chosen as the final solution, which is the optimal path for the data mule. The sink forwards a message which contains the information about the path and topology of each cluster to the data mule. Then, the data mule goes along the given route to initialize deployed system. The data mule visits CHs and informs them about corresponding topology information. Once a CH receives the message from the data mule, it broadcasts this message to its cluster members. When cluster members receive the message from the CH, they start to transmit the sensed data to the CH. The CH aggregates the data and forwards to the data mule once the data mule enters its transmission range. From the second round, the data mule visits all CHs to gather data and forwards to the sink in the end. We provide an example of the steady phase in Fig. 3. The solid line demonstrates the path of the data mule.

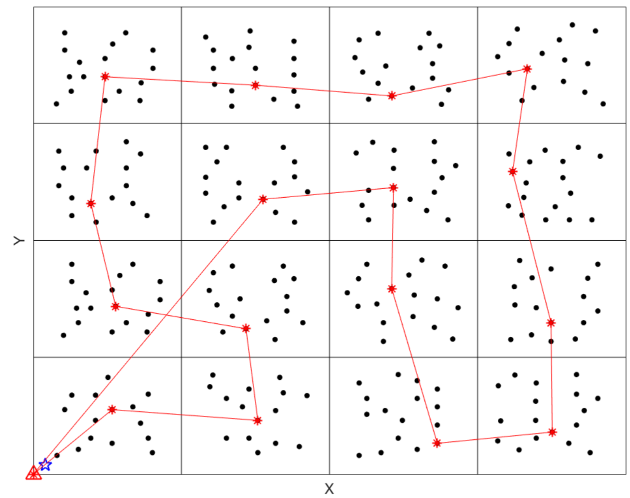


Figure 3: Steady Phase

6 PERFORMANCE EVALUATION

To evaluate the proposed protocol, we implemented a set of experiments on the OMNeT++ simulator. The FoI was divided into multiple grids with the identical size 250m×250m, and we varied the size of the FoI by increasing the number of clusters M from 3×3 to 6×6. To satisfy the connectivity between nodes in one cluster, the number of nodes in each cluster was calculated by $N_s \geq \frac{2\sqrt{3}A}{3R_s^2}$ [30], where A is the area of the cluster, and R_c is the communication range which is set to 50m in our simulation. Moreover, the speed and power rate of the data mule were set to 20m/s and 178W [31]. The transmission range of the data mule is far more than that of the nodes, hence it was not considered. The energy consumption was characterized by (1), and the value of E_{elec} and ϵ_{amp} were introduced in [11]. The parameters of Gaussian mutation were set according to [35], while the mean and the variance of the Gaussian number were changed based on the size of the FoI. The simulation parameters are listed in Table 1.

In this paper, we compared our work with three other different schemes, i.e., centre-based, greedy-based, clustering-based genetic algorithm (CBGA) [34], regarding *data update period*, *system throughput* and *system-wide energy consumption*. In the centre-based protocol, the node located closest to the centroid of each cluster is chosen to be the CH. In this scenario, the data mule visits all CHs according to an S-shaped route that is top-to-bottom and left-to-right. In addition, we developed a greedy-based algorithm as a heuristic algorithm for the data mule routing problem. The CBGA is implemented based on one-hop data collection scheme [34]. We will analyze the experiment results in the following.

The first experiment compared data update period of different schemes, which contains two main components, i.e., moving delay of data mule, data forwarding latency. Due to the limited speed of

Table 1: The list of simulation parameters

Parameter	Description	Value
R_c	Communication range	50m
E_{elec}	The energy consumption dissipation of transmitter circuit	50nJ/bit [11]
ϵ_{amp}	The energy consumption dissipation of transmitter circuit	100$pJ/bit/m^2$ [11]
v	The speed of data mule	20m/s [31]
P	The power of data mule	178W [31]
α	The coefficient for energy consumption of sensor nodes	0.8
β	The coefficient for energy consumption of the data mule	0.2
P_C	Crossover probability	0.8
P_M	Mutation probability	0.2
K	Packet size	4000 bits [33]

Table 2: The moving distance of the data mule in different size of the FoI

	The number of nodes			
SCHEMES	150	260	410	590
Proposed protocol	$2310m$	$3987m$	$6661m$	$8465m$
Centre-based protocol	$3029m$	$4775m$	$7653m$	$10120m$
Greedy-based protocol	$3264m$	$5548m$	$9878m$	$13716m$
CBGA	$4673m$	$8024m$	$12461m$	$22892m$

the data mule, the moving delay is dominated factor of the update period, which is mainly related to the moving distance of the data mule. This criterion is essential for data aggregation by using the data mule. The objective of deployed WSNs is monitoring the FoI or detecting the intruder. Therefore, the update period should be as short as possible in order to track the target. The results of the moving distance of the data mule and the data update period in different protocols are shown in Table 2 and Fig. 4, respectively. We can see that our proposed protocol has much shorter update period than contrast approaches. The results for the first three protocols are similar. But CBGA has the longest update period among them because the number of waypoints is nearly six times more than other works in the same size of the FoI. It causes much longer moving distance in CBGA. Although the limitation of the operation time of the data mule is not considered in our experiment, if the operation time is too long, the data mule would run out of its energy before returning to the data sink. Consequently, the shorter update period of the system is an advantage for our proposed protocol.

In the second experiment, we compared the system throughput. Fig. 5 shows the results of the system throughput with unit *Mbps*. It depicts that the system throughput increases as the number of

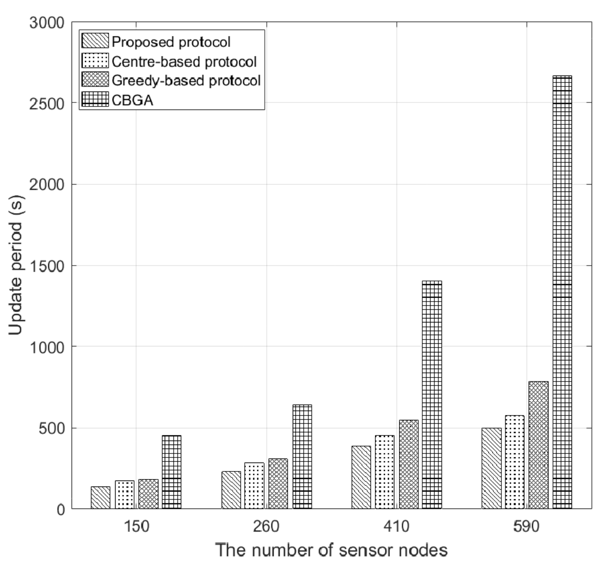


Figure 4: Comparison of data update period

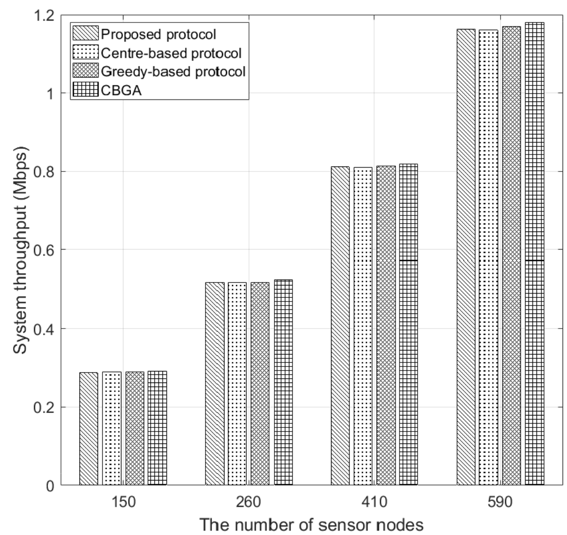


Figure 5: Comparison of system throughput

sensor nodes increases in all protocols. This is because the amount of data generated by deployed sensors monotonically increases as the increasing size of the FoI.

The last experiment was conducted to compare the performance in terms of system-wide energy consumption, which measures the energy consumed by all sensors per bit. As shown in Fig. 6, our proposed protocol provided 1%-28.4% improvement compared with other three protocols. The reason is that the energy consumed by sensors in each cluster is taken into consideration as the objective function in our work. The sum of energy consumed by cluster members is less than centre-based protocol and greedy-based one. For CBGA, it is based on one-hop and without forwarding data by intermediate nodes, but the distance between the waypoint and its sensor nodes might be longer, which introduces extra energy consumption in each sensor. On the contrary, in our work, the transmission distance for each node was minimized. Hence, the energy consumption of each node is lower, and the system-wide energy consumption decreases by 1% to 8% comparing with CBGA.

7 CONCLUSION

Energy efficiency is a crucial problem for WSNs since the limited energy of battery-constrained WSNs that affects the lifetime and performance of systems. In this paper, we introduced a novel energy-efficient UAV-based data aggregation protocol to improve the system-wide energy efficiency of WSNs. We applied a UAV as the data mule to gather data from sensors. We formalized an optimization problem for minimizing the total energy consumption of the entire system. Due to high computing complexity of this problem, a GA-based optimization protocol is also developed to derive the optimal solution. Our protocol had three phases. Firstly, we constructed the topology of the network. Then, the sink calculated the route for data mule and selected the CHs of each cluster by executing GA. After that, the system entered the steady phase,

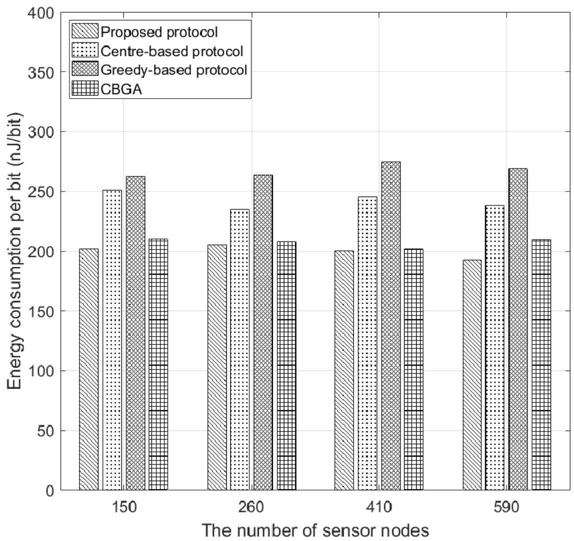


Figure 6: Comparison of system-wide energy consumption

and the data mule traversed designated path and gathered data from each cluster. The simulation results showed that the proposed protocol could achieve shorter data update period and much less energy consumption in WSNs.

REFERENCES

[1] Zakir H. Ahmed. 2010. Genetic Algorithm for the Traveling Salesman Problem using Sequential Constructive Crossover Operator. 3 (2010), 96–105.

[2] M. Alnuaimi, K. Shuaib, K. Alnuaimi, and M. Abdel-Hafez. 2015. Data Gathering in Delay Tolerant Wireless Sensor Networks Using a Ferry. *Sensors* 15, 10 (2015), 25809–25830.

[3] G. Anastasi, M. Conti, M. Di Francesco, and . Passarella. 2009. Energy Conservation in Wireless Sensor Networks: A Survey. *Ad Hoc Netw.* 7, 3 (2009), 537–568.

[4] Constantinos Marios Angelopoulos, Julia Buwaya, Orestis Evangelatos, and José Rolim. 2015. Traversal strategies for wireless power transfer in mobile ad-hoc networks. In *Proc. MSWiM*. 31–40.

[5] J. Aranda, H. Carrillo, and D. Mendez. 2017. Enhanced multimodal switching mechanisms for node scheduling and data gathering in wireless sensor networks. In *Proc. COLCOM*. 1–6.

[6] Rodolfo W.L. Coutinho, Azzedine Boukerche, Luiz F.M. Vieira, and Antonio A.F. Loureiro. 2014. Transmission power control-based opportunistic routing for wireless sensor networks. In *Proc. MSWiM*. 219–226.

[7] R. Dasgupta and S. Yoon. 2017. Energy-Efficient Deadline-Aware Data-Gathering Scheme Using Multiple Mobile Data Collectors. *Sensors* 17, 4 (2017), 742.

[8] S. R. Gandham, M. Dawande, R. Prakash, and S. Venkatesan. 2003. Energy efficient schemes for wireless sensor networks with multiple mobile base stations. In *Proc. IEEE GLOBECOM*, Vol. 1. 377–381.

[9] Pramod D. Ganjewar, S. Barani, Sanjeev J. Wagh, and Santosh S. Sonavane. 2018. Survey on Data Reduction Techniques for Energy Conservation for Prolonging Life of Wireless Sensor Network. *IEEE Wireless Commun.* 10, 2 (2018).

[10] Chirihane Gherbi, Zibouda Aliouat, and Mohamed Benmohammed. 2016. An adaptive clustering approach to dynamic load balancing and energy efficiency in wireless sensor networks. *Energy* 114 (2016), 647 – 662.

[11] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. 2000. Energy-efficient communication protocol for wireless microsensor networks. In *Proc. HICSS*. 1–10.

[12] R. Hinterding. 1995. Gaussian mutation and self-adaption for numeric genetic algorithms. In *Proc. IEEE CEC*, Vol. 1. 384–389.

[13] DT. Ho, E. I. Gr Otli, P. B. Sujit, and TA. Johansen. 2015. Optimization of Wireless Sensor Network and UAV Data Acquisition. *J. Intell. Robot. Syst.* 78, 1 (2015), 159–179.

[14] L. Kong, J. Pan, V. Snášel, P. Tsai, and T. Sung. 2018. An energy-aware routing protocol for wireless sensor network based on genetic algorithm. *Telecommun. Syst.* 67, 3 (2018), 451–463.

[15] C. Konstantopoulos, G. Pantziou, D. Gavalas, A. Mpitziopoulos, and B. Mamalis. 2012. A Rendezvous-Based Approach Enabling Energy-Efficient Sensory Data Collection with Mobile Sinks. *IEEE Trans. Parallel Distrib. Syst.* 23, 5 (2012), 809–817.

[16] C. Konstantopoulos, G. Pantziou, N. Vathis, V. Nakos, and D. Gavalas. 2014. Efficient Mobile Sink-based Data Gathering in Wireless Sensor Networks with Guaranteed Delay. In *Proc. MobiWac*. 47–54.

[17] J. S. Leu, T. H. Chiang, M. C. Yu, and K. W. Su. 2015. Energy Efficient Clustering Scheme for Prolonging the Lifetime of Wireless Sensor Network With Isolated Nodes. *IEEE Commun. Lett.* 19, 2 (2015), 259–262.

[18] J. Li and P. Mohapatra. 2005. An analytical model for the energy hole problem in many-to-one sensor networks. In *Proc. IEEE VTC*, Vol. 4. 2721–2725.

[19] Kin Sum Liu, Jie Gao, Shan Lin, Hua Huang, and Brent Schiller. 2016. Joint Sensor Duty Cycle Scheduling with Coverage Guarantee. In *Proc. MobiWac*. 11–20.

[20] Reem E. Mohemed, Ahmed I. Saleh, Maher Abdelrazzak, and Ahmed S. Samra. 2017. Energy-efficient routing protocols for solving energy hole problem in wireless sensor networks. *Comput. Netw.* 114 (2017), 51 – 66.

[21] Sarwar Morshed and Geert Heijenk. 2014. TR-MAC: An Energy-efficient MAC Protocol Exploiting Transmitted Reference Modulation for Wireless Sensor Networks. In *Proc. MSWiM*. 21–29.

[22] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah. 2016. Mobile Internet of Things: Can UAVs Provide an Energy-Efficient Mobile Architecture?. In *Proc. IEEE GLOBAECOM*. 1–6.

[23] B. Prince and S. Gupta. 2016. Load balanced and energy efficient data collection scheme using data mule for WSNs. In *Proc. IEEE RTEICT*. 1651–1655.

[24] P. Rawat, K. D. Singh, H. Chaouchi, and J. M. Bonnin. 2014. Wireless sensor networks: a survey on recent developments and potential synergies. *J. Supercomput.* 68, 1 (2014), 1–48.

[25] M. Sartipi and R. Fletcher. 2011. Energy-Efficient Data Acquisition in Wireless Sensor Networks Using Compressed Sensing. In *Proc. DCC*. 223–232.

[26] R. Silva, J. S. Silva, and F. Boavida. 2014. Mobility in wireless sensor networks – Survey and proposal. *Comput. Commun.* 52, Supplement C (2014), 1 – 20.

[27] S. K. Singh, P. Kumar, and J. P. Singh. 2016. An energy efficient Odd-Even round number based data collection using mules in WSNs. In *Proc. WCSP*. 1255–1259.

[28] Éfren L. Souza, Eduardo F. Nakamura, and Richard W. Pazzi. 2016. Target Tracking for Sensor Networks: A Survey. *ACM CSUR* 49, 2 (2016), 30:1–30:31.

[29] P. Sun and A. Boukerche. 2016. Integrated connectivity and coverage techniques for wireless sensor networks. In *Proc. MobiWac*. 75–82.

[30] P. Sun and N. Samaan. 2015. Random Node Failures and Wireless Networks Connectivity: Theoretical Analysis. *IEEE Wirel. Commun. Lett.* 4, 5 (2015), 461–464.

[31] SZ DJI Technology Co., Ltd. 2018. Phantom 4 PRO specs. [Online].Available: https://www.dji.com/phantom-4-pro/info#specs

[32] Z. M. Wang, S. Basagni, E. Melachrinoudis, and C. Petrioli. 2005. Exploiting Sink Mobility for Maximizing Sensor Networks Lifetime. In *Proc. HICSS*. 287.1–287.9.

[33] D. Wu, J. He, H. Wang, C. Wang, and R. Wang. 2015. A hierarchical packet forwarding mechanism for energy harvesting wireless sensor networks. *IEEE Commun. Mag.* 53, 8 (2015), 92–98.

[34] S. Y. Wu and J. S. Liu. 2014. Evolutionary path planning of a data mule in wireless sensor network by using shortcuts. In *Proc. IEEE CEC*. 2708–2715.

[35] Y. Yoon and Y. H. Kim. 2013. An Efficient Genetic Algorithm for Maximum Coverage Deployment in Wireless Sensor Networks. *IEEE Trans. Cybern.* 43, 5 (2013), 1473–1483.