2015-10

Ant Colony Clustering Routing Protocol for Optimization of Large Scale Wireless Sensor Networks

Li Xinlu
Technological University Dublin

Brian Keegan
Technological University Dublin, brian.x.keegan@tudublin.ie

Fredrick Japhet
Technological University Dublin

Ant Colony Clustering Routing Protocol for Optimization of Large Scale Wireless Sensor Networks

Li Xinlu [1,2], Brian Keegan [2], Fredrick Japhet [2]

[1] Department of Computer Science, Hefei University, Anhui, China
[2]School of Computing, Dublin Institute of Technology, Kevin Street, Dublin 8, Ireland
xinlu.li@mydit.ie brian.x.keegan@dit.ie Fredrick.Mtenzi@dit.ie

Abstract

Wireless sensor networks (WSNs) have broad application prospects in many fields, which include diverse uses such as farming and industrial. The major motivation to improve the routing performance of WSNs is due to the restriction of energy supply, especially for large scale WSNs. In this paper, an ant colony clustering routing protocol (ACCR) is presented. In the clustering phase, we estimate the theoretical energy dissipation as standard energy which results in the energy consumption to fluctuate around the standard energy. In the data transmission phase, we implement a proactive algorithm based on ant colony optimization. The algorithm establishes multiple optimal paths in terms of energy level and link length from cluster head to base station. The simulation results that ACCR is well suited for large-scale network application, achieves better load balancing and lower energy consumption, higher energy efficiency, and then maximize the network lifetime.

Keywords: Ant Colony Optimization, Clustering, Wireless sensor Network, Large Scale Networks

1 Introduction

Due to advances in low-power wireless communications and embedded digital electronics, low cost wireless sensor networks (WSNs) have become available. Such WSNs consist of a large number of distributed sensor nodes and have broad application prospects in military, environmental monitoring, healthcare and other fields[1]. The sensor nodes generally have restrictions including limited energy supply, limited computation and communication abilities, so how efficiently and rationally use energy to extend the network lifetime as much as possible has become the core issues of protocol design.

Hierarchical (clustering) technology is particularly promising and has received much attention in the research community, such as LEACH[2], PEGASIS[3], HEED[4]. The hierarchical model consists of base station (BS), cluster head (CH) and sensor nodes. The CHs collect sensor data from its cluster members, aggregate and send it to the remote BS for the end user. Therefore, the CHs play an important role in a clustering algorithm with CH selection being the key component of a clustering algorithm.

Recently, many routing protocols based on Ant Colony Optimization (ACO)[5] have been developed, and these protocols exploit the metaheuristics behind the shortest path behavior observed in ant colonies to define a nature-inspired metaheuristic for combinatorial optimization. The basic idea behind ACO algorithms for routing is AntNet[6][7], which is for adaptive best-effort routing in IP networks. Zhang et al.[8]found that AntNet does not performs sufficiently well in sensor networks and therefore proposed a basic WSNs routing algorithm directly derived from AntNet. However, the author only focused on building an initial artificial pheromone distribution, and thus, the algorithms are only good at system start-up phase. An energy-efficient ant-based routing algorithm (EEABR)[9] is another proposed ant based algorithm to extend network lifetime by reducing the communication overhead in path discovery. However, the algorithm does not consider data correlation. As a result, the energy consumption of communication is significantly large when multiple network sources exist. Chen et al.[10] have proposed the E-D ANTS algorithm which aims to find a route with minimum energy-delay product in order to maximize network lifetime and to provide a real-time data delivery service. However, the E-D ANTS being flat in nature would be difficult to scale through large topologies.

In this paper, we propose an ant colony clustering routing protocol (ACCR) for large-scale WSN applications. First, ACCR establishes a hierarchical network model by using a novel clustering algorithm. Second, ACCR performs an improved ACO algorithm for route discovery, and establishes multiple optimal paths. Finally, ACCR gives a weight function for dynamically choosing the route. The main goal of our study is to achieve better load balancing and lower energy consumption, higher energy efficiency, and then maximize the network lifetime.
The remainder of the paper is organized as follows. Section 2 introduces the hierarchical model and gives the clustering algorithm. In Section 3, the proposed routing algorithm using ACO is presented. The simulation results are presented in Section 4.

2 Hierarchical Model and Clustering Algorithm

We consider a WSN hierarchical model is well suited for large-scale networks, it is based on the model developed in LEACH[2]. We assume that N sensor nodes are randomly distributed within a $M \times M$ rectangle field with the base station located in the center of the field, the network is divided into several clusters, and each cluster has a CH and some senor nodes. After being deployed, all sensor nodes keep less movement, and they also are location-unaware. In order to evaluate the energy dissipation in data communication, we use the First Radio Model described in literature[11].
As discussed previously in Section 1, CHs play an important role in clustering algorithm, and consume more energy than the sensor nodes and run out of energy sooner. Therefore, the remaining energy of sensor node must be taken into account when selecting the cluster heads. Furthermore, we introduce an estimating average energy as standard energy to ensure that the node has a higher energy level than its neighbors and has an increased chance of becoming a CH.

2.1 Estimating Average Energy

Assume that each node dissipates the same energy at each round, until they run out of the energy at almost the same time. If so, theoretically, the network would use the energy resources as much as possible and the lifetime of network would be maximized which is also the target that energy-efficient algorithms should try to achieve.
In the mentioned network and First Radio Model, the energy consumption of each CH consists of receiving data from sensor nodes, data aggregation, and transmitting the fusion data to the BS:

$$E_{ch} = lE_{elec}\left(N/k - 1\right) + lE_{DA}N/k + lE_{elec} + l\varepsilon_{mp}d_{toBS}^{4} \tag{1}$$

Where k is the number of clusters, E_{elec} are the energy dissipation per bit to run transmitter or the receiver circuit, ε_{fs} and ε_{mp} depend on the transmitter amplifier mode we use. The energy dissipation for data aggregation is set as E_{DA}. Each non-cluster head node only needs to transmit data to CH, thus, the energy consumption in each non-cluster head is:

$$E_{non\_ch} = lE_{elec} + l\varepsilon_{fs}d_{toCH}^{2} \tag{2}$$

And the total energy dissipated in a round is:

$$E_{round} = k \cdot E_{cluster} \approx l\left(2NE_{elec} + NE_{DA} + k\varepsilon_{mp}d_{toBS}^{4} + N\varepsilon_{fs}d_{toCH}^{2}\right) \tag{3}$$

Where d_{toBS} is the average distance from CH to BS, and d_{toCH} is the average distance from non-cluster head node to CH. According to literature[2], we assume the nodes are uniformly distributed, we can get:

$$d_{toBS} = 0.765\frac{M}{2} \tag{4}$$

By setting the derivate of E_{round} with respect k to zero,

$$k_{opt} = \sqrt{\frac{n}{2\pi}}\sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}}\frac{M}{d_{toBS}^{2}} \tag{5}$$

Where k_{opt} denote the optimal number of clusters, and we can estimate the total energy dissipated in a round, furthermore, get the theoretical maximum of lifetime:

$$R = \frac{E_{total}}{E_{round}} \tag{6}$$

R is the theoretical amount of round, and the estimating average energy in round r ,

$$\overline{E}(r) = \frac{1}{N} E_{total} \left(1 - \frac{r}{R}\right) \tag{7}$$

2.2 Clustering Algorithm

We consider estimating average energy $\overline{E}(r)$ as the standard energy by giving each node a desired probability of p_i ,

$$p_i = p_{opt} \frac{E(i)}{\overline{E}(r)} \tag{8}$$

Where $E(i)$ is the remaining energy of node i , $p_{opt} = k_{opt} / N$, denotes the optimal percentage of nodes that become CH in each round, the select threshold is given accordingly by

$$T(n_i) = \begin{cases} \dfrac{p_i}{1 - p_i \left(r \bmod 1/p_i \right)} & , n_i \in G \\ 0 & , otherwise \end{cases} \tag{9}$$

Each node determines a random number between 0 and 1, if the number is less than threshold $T(n_i)$, the node will become CH for the current round. After $1/p_i$ rounds, all nodes are expected to have been cluster head once, following which they are all eligible to perform this task in the next sequence of rounds. The details of the clustering algorithm are as follows:

Begin
For each node in network
1: caculate $\overline{E}(r)$, p_i and $T(n_i)$ according to the Equation(7)(8)(9)
2: **if** rand(0,1) < $T(n_i)$ **then**
3: broadcast cluster head status
4: wait for request messages
5: broadcast Time Division Multiple Access (TDMA) schedule to members
6: **else**
7: wait for cluster head to broadcast message
8: select the nearest CH as subordinate cluster
9: send the join information to become cluster members
10: **end if**
End

Table 1 Algorithm: Clustering Alogrithm basedon remaining energy and average energy

This improvement makes sure that nodes have more remaining energy than average energy and increases their chance of being chosen as a CH. The remaining energy would fluctuate around the standard energy, correspondingly the network energy evenly dissipates and prolongs the network lifetime. In the evolving progress, each node just needs localized communication, not the global information.

3 Routing Algorithm based on ACO

In ACCR, we use a proactive protocol for multi-hop model in inter-cluster communication base on improved ACO to establish multiple paths in terms of lower energy consumption and shortest path. There are three kinds of control packets in ACCR: HELLO packet, Forward Ant (FANT) and Backward Ant (BANT). When the CH is ready to transmit the data to BS, an HELLO packet is broadcasted to its neighbors in its communication range to form the neighborhood table. FANT is generated by CH, and is used to find the route from the CH to BS according to the selected probability. Once FANT reaches the BS, ACCR calculates the amount of pheromone value in this path, meanwhile, BANT is created and used to update the artificial pheromone value along the reverse path. By

performing this process with several iterations, each node will be able to know which the best neighbors to send data towards BS.
In order to reduce the overhead of the control message, we define the message format of FANT (Fig. 1) and BANT (Fig. 2) separately.

hp_type	seqno	M^k	E_{avg}	$E_{\min}$	pkt_src	F_len

Fig. 1: The message format of FANT

- hp_type indicates the type of the massage
- seqno is the number of a FANT
- pkt_src is the CH node which create the FANT, the 2-tuple of (pkt_src, seqno) could define a FANT uniquely.
- E_{avg} and $E_{\min}$ are the average energy and minimum energy level till the current node in the path respectively.
- F_len is the length of the path (hops).
- The identifier of every nodes is saved into M^k which could avoid the loop route, and this field is similar with literature[9], only stored the last two visited nodes to reduce the length of the message.

hp_type	seqno	phe_value	pkt_dst	B_len

Fig. 2: the message format of FANT

The field is similar with FANT with the following exceptions:

- phe_vlaue denote the the amount of pheromone value which calculated in BS to update the pheromone value for each intermediate node.
- B_len is the path length from BS to current node. pkt_dst is the destination of BANT, actually, corresponding to the pkt_src field of FANT.

Since the FANT or BANT does not save the whole visited nodes, each CH to store a visited list, which saves information regarding the previous node, the forward node, the ant identification, the ant original node and a timeout value. Besides, CH needs to establish a routing table to store information about its neighborhoods, such as energy level and pheromone trail.

nb_addr	nb_enegy	phe_value	hops

Fig. 3: the routing table

- nb_addr denotes the identifier of the neighbor node
- nb_enegy is the remaining energy of the neighbor node which is obtained by the HELLO message
- phe_value indicates the pheromone trail between the current node with its neighbor node
- Hops field denotes the reachable hops to BS. The record is updated when receiving a BANT.

Each FANT selects the next hop with a given probability $p_{ij}(t)$,

$$p_{ij}(t)=\begin{cases}\dfrac{\tau_{ij}^{\alpha}(t)\times\eta_{ij}^{\beta}(t)}{\sum_k \tau_{ij}^{\alpha}(t)\times\eta_{ij}^{\beta}(t)}, & \forall j\in N_i \text{ and } j\notin M^k \\ 0, & \text{otherwise}\end{cases} \tag{10}$$

Where $\tau_{ij}(t)$ is the pheromone trail on link (i,j), and $\eta_{ij}(t)$ is the local heuristic value of the link (i,j). α and β are two parameters that control the relative importance of pheromone trail and heuristic value. Here we redefine the $\eta_{ij}(t)$, which should be calculated as a function of both parameters: the remaining energy level and the average energy in current round.

$$\eta_{ij}(t)=\frac{E_{init}-\overline{E}(r)}{E_{init}-E(j)},\ \forall j\in N_i \text{ and } j\notin M^k \tag{11}$$

where E_{init} is the initial energy for each node, $E(j)$ denote the remaining energy of node j. $\overline{E}(r)$ is the estimating average remaining energy in round r. $E_{init}-E(j)$ indicates the energy dissipation of node j, and $E_{init}-\overline{E}(r)$ indicates the average energy dissipation of each node in round r. As can be

seen from the equation(11), the neighbor node with relatively less energy dissipation than average energy dissipation has a higher probability to be the next hop. This method is beneficial to balance the energy consumption of the network and prolong the lifetime.

When the FANT moves hop-by-hop and reaches the BS, the amount of pheromone value $\Delta\tau(t)$ attached the path that FANT just traveled along is calculated. We calculate $\Delta\tau(t)$ in the following manner:

$$\Delta\tau(t) = \frac{1}{E_{init} - \left[\left(E_{\min} - F_d(t)\right) / \left(E_{avg} - F_d(t)\right)\right]} \tag{12}$$

$E_{\min}$ and E_{avg} are the minimum energy level and average energy in current path towards BS respectively and $F_d(t)$ denotes the length of the path. The BANT carries $\Delta\tau(t)$ at the start of its journey following the reverse path. Each node in the path updates the routing table, including hops and pheromone value. Each pheromone value field is updated by following:

$$\tau_{ij}(t) = (1-\rho)\tau_{ij}(t) + \Delta\tau_{ij}(t) \tag{13}$$

Where a control coefficient $\rho \in (0,1)$ is used to determine the weight of evaporation for each tour, and the $\Delta\tau_{ij}(t)$ is given by

$$\Delta\tau_{ij}(t) = \frac{E(j)}{E_{init} B_d(t)} \Delta\tau(t) \tag{14}$$

Where $B_d(t)$ carried by BANT denotes the path length from BS to current node. Following this manner, the pheromone value is the function of both energy levels and length of the path, as a result, the link with more energy will have more pheromone trail, and the nodes near the BS will have more pheromone, this would lead the data flow in the direction of the BS and accelerate the convergence rate. The details of the algorithm as follows:

FANT for the proposal	BANT for the proposal
Begin	**Begin**
For each CH node in network	***For Base Station***
1: Broadcast HELLO to neighbor CH node, initialize routing table(nb_enegy,nb_addr), create FANT	1: **if** recieve FANT **then**
2: FANT choose the next hop node j according to equation(10).	2: Caculate the amount of pheromone value $\Delta\tau(t)$ according to equation(12)(13)(14).
3: **while** the BS is not reached	3: Create BANT according the information of FANT
4: **if** (pkt_src, seqno) in visited list **then**	4: Delete FANT message
5: back to the previous hop node	5: **end if**
6: goto step 2	***For each CH node in network***
7: **else**	6: **while** the orignal node is not reached
8: update visited list in node j	7: BANT moves back along opposite direction
9: update M^k, E_{avg}, $E_{\min}$,F_len in FANT	8: BANT using formular to update phermone value in routing table and hops
10: **end if**	9: Delete the visited record in intermediate node
10: **end while**	10: **end while**
End	**End**

Table 2 Algorithm: Optimal routing based on ACO

After a number of iterations, which depends on specific application and network scale, we can establish multiple paths from cluster head to base station. In the steady-state phase of data transmission, each CH choose an optimal or suboptimal path to transmit the aggregation data towards BS with the probability.

$$P_i = \frac{\tau_i^k E_{avg\_i}^m d_i^n}{\sum_{i \in N} \tau_i^k E_{avg\_1}^m d_i^n} \tag{15}$$

Where τ_i denote the average pheromone trail of each path i, E_{avg_i} denotes the average remaining energy of each nodes in path i, and the d_i is the length of path i. k, m, n is weight values,

$k+m+n=1$. As a result, the path could be chosen dynamically, thus load balancing among the paths could be achieved.

4 Performance Evaluation

We compare the ACCR protocol with other protocols mentioned in Section 1. We evaluate these protocols in different scenarios, which consist of a different number of nodes from 100 to 500. In each scenario, the sensor nodes are randomly distributed on an $M \times M$ rectangle region with $M = 500m$. Sensor nodes are responsible for monitoring and sending the relevant data to the BS. To reduce the impact caused by the randomness, we run these algorithms ten times to obtain the average results. The network model is described in Section 2. Since the parameters in ACO greatly influence the performance of the protocol, we adopt the values 1, 5 and 0.5 for α, β, ρ by default in accordance with best practice from literature[5].

The main parameters we used are as follows:

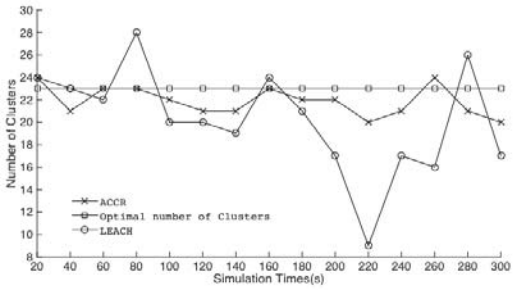
Parameter	Value	Parameter	Value
Date packet size	512 *bytes*	E_{DA}	50nJ / *bit*
Control packet size	20 *bytes*	E_{init}	2.0 J
E_{DA}	50 nJ / *bit*	α	1
ε_{mp}	0.0013 pJ / *bit* / m^4	β	5
ε_{fs}	10 pJ / *bit* / m^2	ρ	0.5

Table 3 Experiment Metrics

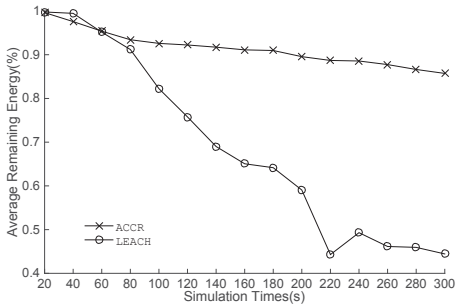
We simulate these protocols under Matlab, and the main metrics that we have been considered are Characteristics of Clusters, Average Energy, Energy Efficiency, and Network Lifetime.

4.1 Characteristics of Clusters

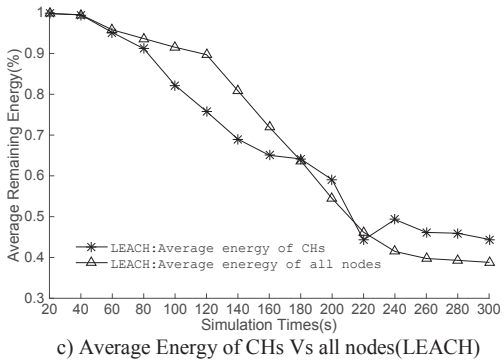
We investigate the number of clusters and average residual energy per cluster head as the main characteristics. If the number of clusters k converge to the expected value k_{opt}, this indicates the clustering mechanism is steady and optimal[2]. As can be seen from the Fig. 4a, the k of ACCR and LEACH also fluctuate around the k_{opt}, however, the fluctuation range of the ACCR is smaller than that of LEACH, basically, the k of LEACH is random fluctuation. The reason is that ACCR defines desired probability function(8)(9) for each node, while LEACH utilizes the stochastic selection simply.



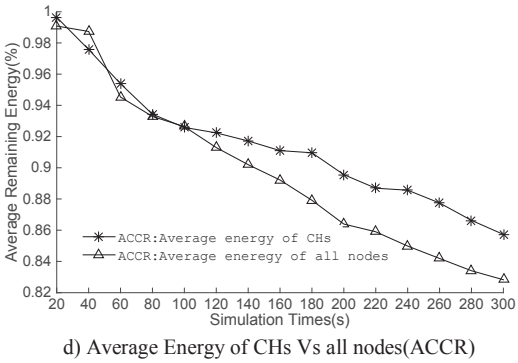
a) Number of Clusters



b)Average Energy of CHs



c) Average Energy of CHs Vs all nodes(LEACH)



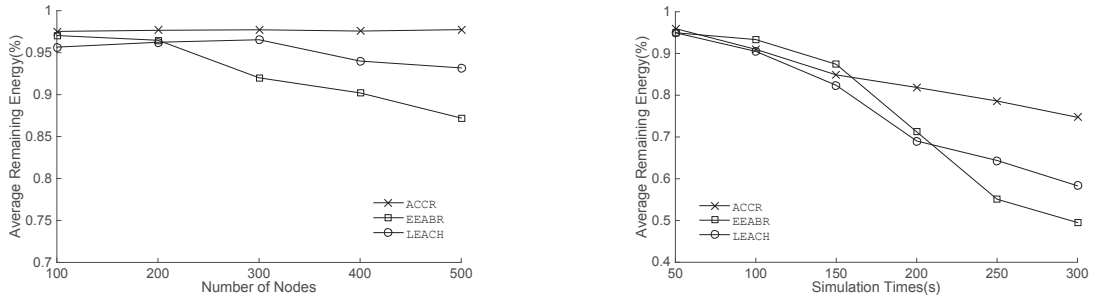
d) Average Energy of CHs Vs all nodes(ACCR)

Fig. 4 Characteristics of Clusters

Fig. 4b presents the average energy per CHs comparing ACCR with LEACH, Fig. 4c and Fig. 4d illustrate the average energy per CHs versus that of all nodes. In this context, we define average energy as the average residual energy after clustering, data collection and data aggregation for verifying the characteristics of cluster algorithm. These figures indicate that the CHs determined by ACCR not only have more residual energy than LEACH, but also more than the residual energy of sensor nodes in ACCR. While the residual energy of CHs and sensor nodes in LEACH is not related due to random selection strategy.

4.2 Average energy

Average energy is defined as the ratio between the total remaining energy and the number of sensor nodes at the end of simulation. It is proportional with network lifetime in the case of load balancing.



a) Average Energy Vs Number of Nodes

b) Average Energy Vs Simulation Times

Fig. 5: Average Energy

Fig. 5a presents the results of different number of nodes (from 100 to 500), while Figure 4b) presents the results of different simulation times (from 50s to 300s). As we can see from Fig. 5a, It is obvious that ACCR consumes less energy than the others, and ACCR gives the better results in the large-scale application, especially, when the density of nodes reach the top (500 nodes in $M \times M$ region), the average energy of ACCR is 14% and 5% more than EEABR and LEACH respectively. After the number of nodes exceeds 200, the LEACH outperforms EEABR, this proves the clustering protocols is more suitable for large-scale network than flat routing protocols. Overall, ACCR has more remaining energy level, and this performance is particularly obvious when the simulation time reach 150s. This is because ACCR consider both energy levels and link length when update the local heuristic value and pheromone trail. After several iterations, ACCR could find the optimal with higher energy and shorter path quickly.

4.3 Energy Efficiency

Energy Efficiency denotes the ratio between total consumed energy and the number of packets received by the BS. The metric shows the energy dissipation of node when transmitting a data packet to BS. As can be seen from Fig. 6, ACCR presents a significant performance improvement over other protocols. When the number of nodes reaches 500, the Energy Efficiency of ACCR increases by about 15% and 50% more than EEABR and LEACH respectively. As the network scale increasing, the density of nodes also increases, correspondingly, the communication distance of the nodes would decrease，this would lead to reduce the energy consumption between the nodes, and hence, increase energy efficiency.

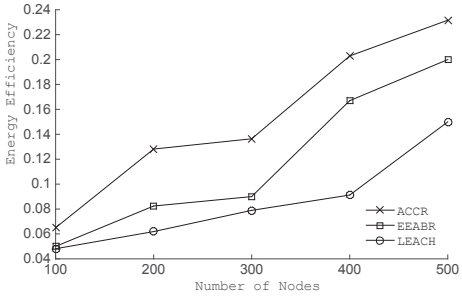


Fig. 6: Energy Efficiency

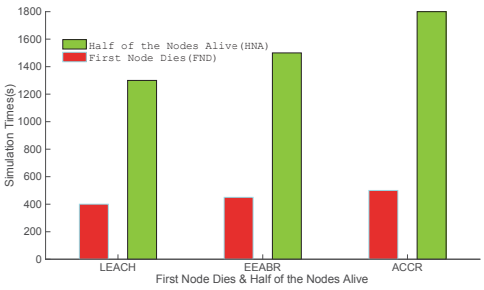


Fig. 7: Network Lifetime

4.4 Network Lifetime

Without loss of generality, we use First Node Dies (FND) and Half of the Nodes Alive (HNA) to measure the network lifetime respectively. In practical applications, these two metric are more meaningful than the time of last node dies. Fig. 7 presents the simulation result of network lifetime. We can see that the FND of three protocols are approximate, however, the HNA of them are significantly different, and the HNA of ACCR improves distinctly by 38% and 20% than LEACH and EEABR respectively. The result indicates the stable period of the network using ACCR is longer than that obtained by other protocols. The reason is that clustering and routing based on ACO can greatly contribute to reducing energy consumption and achieving load balancing among the sensor nods.

5 Conclusion and Future Work

In this paper, we present an ant colony clustering routing protocol (ACCR) for large scale WSNs application. ACCR produces load balance clustering and optimal multiple paths by using the localized communication and ACO metaheuristic. ACCR introduces the theoretical energy dissipation based on clustering network model, and consider it as a standard energy in the deterministic selection. Furthermore, in inter-cluster communication, ACCR establishes multiple optimal paths from CH to BS based on improved ACO algorithm. ACCR considers the energy levels and path length as key metrics when updating the local heuristic value and pheromone trail. The simulation results show that ACCR is well suited for large-scale network application, achieves better load balancing and lower energy consumption, higher energy efficiency, and then maximize the network lifetime.

All of our work here is focused on the CH selection and data transmission from CH to BS. Further performance improvement could be gained through investigation of the intra-cluster communications. Moreover, we plan to implement ACCR protocol in a large size testbed, which would test the protocol performance impact of the real implementation.

References

[1] Zungeru, A. M., Ang, L. M., & Seng, K. P. (2012). Classical and swarm intelligence based routing protocols for wireless sensor networks: A survey and comparison. Journal of Network and Computer Applications, 35(5), 1508-1536.

[2] Heinzelman, W. B., Chandrakasan, A. P., & Balakrishnan, H. (2002). An application-specific protocol architecture for wireless microsensor networks. Wireless Communications, IEEE Transactions on, 1(4), 660-670.

[3] Lindsey, S., & Raghavendra, C. S. (2002). PEGASIS: Power-efficient gathering in sensor information systems. In Aerospace conference proceedings, 2002. IEEE (Vol. 3, pp. 3-1125).

[4] Younis, O., & Fahmy, S. (2004). HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. Mobile Computing, IEEE Transactions on, 3(4), 366-379.

[5] Dorigo, M., Maniezzo, V., & Colorni, A. (1996). Ant system: optimization by a colony of cooperating agents. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 26(1), 29-41.

[6] Di Caro, G., & Dorigo, M. (1998). AntNet: Distributed stigmergetic control for communications networks. Journal of Artificial Intelligence Research, 317-365.

[7] Di Caro, G. (2004). Ant Colony Optimization and its application to adaptive routing in telecommunication networks (Doctoral dissertation, PhD thesis, Faculté des Sciences Appliquées, Université Libre de Bruxelles, Brussels, Belgium).

[8] Zhang, Y., Kuhn, L. D., & Fromherz, M. P. (2004). Improvements on ant routing for sensor networks. In Ant Colony Optimization and Swarm Intelligence (pp. 154-165). Springer Berlin Heidelberg.

[9] Camilo, T., Carreto, C., Silva, J. S., & Boavida, F. (2006). An energy-efficient ant-based routing algorithm for wireless sensor networks. In Ant Colony Optimization and Swarm Intelligence (pp. 49-59). Springer Berlin Heidelberg.

[10] Wen, Y. F., Chen, Y. Q., & Pan, M. (2008). Adaptive ant-based routing in wireless sensor networks using Energy* Delay metrics. Journal of Zhejiang University SCIENCE A, 9(4), 531-538.

[11] Heinzelman, W. R., Chandrakasan, A., & Balakrishnan, H. (2000, January). Energy-efficient communication protocol for wireless microsensor networks. In System sciences, 2000. Proceedings of the 33rd annual Hawaii international conference on (pp. 10-pp). IEEE.