Contemporary Engineering Sciences, Vol. 9, 2016, no. 17, 809 - 822
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ces.2016.66109

Performance Analysis of Clustering Protocols in Normally and Exponentially Distributed Wireless Sensor Networks

Sangil Choi and Sangman Moh*

Dept. of Computer Engineering, Chosun University
Gwangju, Republic of Korea
*Corresponding author

Abstract

In wireless sensor networks (WSNs), clustering technique is effectively used to reduce energy consumption in battery-operated sensor nodes. In this paper, the performance of typical clustering protocols in normally and exponentially distributed WSNs is analyzed and compared via extensive simulation. Every sensor node in WSNs is powered by batteries with limited energy capacity and, thus, the network lifetime that is directly affected by the energy efficiency of sensor nodes is the most important performance metric. Our performance study shows that the balanced clustering algorithm (BCA) achieves longer network lifetime than the other protocols. Furthermore, the improvement effect is better and better as the non-uniformity is increased.

Keywords: Wireless sensor network, clustering, non-uniform deployment, energy consumption, network lifetime, performance

1 Introduction

Wireless sensor networks (WSNs) are widely used for various applications such as environment monitoring, logistics, target tracking, military fields, home networks, and industrial diagnosis [3]. A WSN consists of many battery-powered sensor nodes that sense their surroundings and send the sensed data to a sink node or base station. In many WSNs, the batteries are difficult to replace and, even if replaceable, the replacement cost is very high [1]. Thus, reducing energy consumption in sensor nodes is very important for prolonging network lifetime.

In WSNs, routing is the process of forwarding data gathered by sensor nodes to the sink or base station. A WSN consists of a lot of sensor nodes, and it is inefficient for all the sensor nodes to send their sensed data to the single sink node or base station directly. Instead, the sensor nodes are grouped as clusters, and every sensor node sends its sensed data to its cluster head (CH). Then, the CHs send the aggregated data to the sink. Such a hierarchical routing is energy-efficient compared to the flat routing that each sensor delivers the sensed data to the sink directly.

The typical hierarchical routing or clustering protocols are low energy adaptive clustering hierarchy (LEACH) [4], low-energy adaptive cluster hierarchy centralized (LEACH-C) [5], hybrid, energy-efficient distributed (HEED) [13], base station controlled dynamic clustering protocol (BCDCP) [11], threshold sensitive energy-efficient sensor network protocol (TEEN) [9], hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks (APTEEN) [10], tree-based clustering (TBC) [6], and balanced clustering algorithm (BCA) [12]. The well-known LEACH is the pioneer clustering protocol in WSNs, and TBC is the most advanced clustering scheme for uniformly deployed WSNs. The recently developed BCA is a single-hop clustering scheme targeted for non-uniformly deployed WSNs. The existing clustering algorithms will be reviewed in more detail in the next section.

In many applications such as environment monitoring, sensor nodes can be non-uniformly deployed due to some limited condition. For example, when the sensors nodes are deployed over a mountain area by a helicopter, there is the possibility that they may be non-uniformly deployed. In such a non-uniformly deployed WSN, the sensing area or coverage area of each cluster varies region by region, i.e., there are many small-area clusters in dense regions and a few large-area clusters in sparse regions. In BCA [12], equal-size clustering is achieved even in non-uniformly deployed WSNs and the excessively redundant nodes are turned into sleep mode to save energy and to prolong network lifetime. In BCA, however, the single-hop transmission from sensor nodes to their CH needs more energy consumption compared to multihop transmission in a cluster because transmission power is exponentially increased with distance. On the other hand, TBC [6] implements a multi-level tree within a cluster enabling multihop transmission, but it does not take the non-uniform deployment into consideration resulting in severely conflicted transmissions and unnecessary energy consumption in dense regions.

In this paper, the typical three clustering protocols of LEACH, TBC and BCA in normally and exponentially distributed WSNs are qualitatively compared first in terms of major features and characteristics and, then, they are quantitatively analyzed via extensive computer simulation in terms of network lifetime. Note again that the network lifetime is the most important performance metric in WSNs. Our simulation results show that BCA is the most energy-efficient clustering protocol among the three protocols. It is also shown that the performance improvement is better and better as the non-uniformity is increased.

The rest of this paper is organized as follows: In the following section, the existing clustering protocols are reviewed. In Section 3, the three typical protocols of LEACH, TBC and BCA are qualitatively compared. In Section 4, the performance of the three clustering protocols is quantitatively evaluated via extensive computer simulation and compared. Finally, the paper is concluded in Section 5.

2 Related Works

For more than a decade, many clustering algorithms based on randomness have been studied. Since the pioneer clustering protocol LEACH was introduced [4], more advanced clustering algorithms have been proposed so far [5-6, 9-13]. In this section, they are reviewed with respect to major characteristics and improvements.

2.1. LEACH

In the LEACH protocol [4], each round consists of set-up phase and steady-state phase. Clusters are formed during the set-up phase, and the sensed data are periodically delivered to the sink through CHs during the steady-state phase.

In LEACH, CHs are elected probabilistically every round. Every sensor node generates a random number between zero and one and, then, it becomes a CH if the generated number is less than the calculated threshold value. For a node *n*, the threshold value *T(n)* at the *r*-th round is calculated by

$$\boldsymbol{T(n)} = \begin{cases} \dfrac{\boldsymbol{p}}{1 - \boldsymbol{p}(\boldsymbol{r} \bmod \frac{1}{\boldsymbol{p}})}, & \textbf{if } \boldsymbol{n} \in \boldsymbol{G} \\ 0, & \textbf{otherwise} \end{cases} \tag{1}$$

where the given parameter *p* is the probability that a sensor node becomes a CH and *G* is the set of sensor nodes that have not been chosen as a CH for *1/p* rounds. If a node n has not been chosen as a CH for the last *1/p* rounds, *T(n)* is calculated by (1) and, if the generated random number is less than *T(n)*, the node becomes a CH at the current round; otherwise, *T(n)* is zero and the node n is not elected as a CH at the current round.

Once CHs are chosen according to the above procedure, every CH broadcasts that it has becomes a CH. Then, sensor nodes send a join message to the nearest CH based on the received signal strength of the broadcast messages.

In the steady-state phase after cluster formation, sensor nodes send the sensed data to their CHs periodically in accordance with the TDMA (Time Division Multiple

Access) schedule assigned by their CHs. CHs aggregate the received data and send the aggregated data to the sink node.

Such a series of procedural steps are repeated every round. That is, the CHs are rotated per round because they consume more energy than normal sensor nodes. This makes all the nodes consume energy as evenly as possible, resulting in increased network lifetime. However, when sensor nodes are non-uniformly deployed over the network area, the balanced energy consumption is not possible due to unbalanced clustering.

2.2. TBC

In the TBC protocol, a multi-level tree is constructed in a cluster, in which the CH is the root node [6]. The CH is elected in the same manner as in the LEACH protocol. The broadcast and join messages are also similar to those in LEACH, which are sent by CHs and normal sensor nodes, respectively. Unlike LEACH, however, the location information of the sensor node is included in the join message.

By receiving the join messages from sensor nodes, the CH finds the farthest sensor node, and the distance between the CH and the farthest sensor node is denoted as d_{max}. The maximum distance d_{max} is divided by the tree depth α, where α is also called tree height or the maximum level of the tree. Therefore, the average transmission distance d_{avg} between the node and its parent node in the tree can be represented by

$$d_{avg} = \frac{d_{\max}}{\alpha}. \tag{2}$$

The CH is at level 0 in the tree and member nodes are at the specific level according to the distance from the CH. Fig. 1 shows an example of constructing a tree in TBC when α is 3. Once the cluster is divided into α concentric circles as shown in Fig. 1, each sensor node selects an upper-level node with the minimum distance from the node itself as its parent node. Finally, a single tree is generated.

In a cluster, the multihop transmission through the multi-level tree from sensor nodes to the CH reduces energy consumption in comparison to single-hop transmission because transmission power is exponentially increased with distance. Also, the energy consumption is distributed over the network. As in LEACH, however, the unbalanced clustering causes unbalanced energy consumption over the network if sensor nodes are non-uniformly deployed. In addition, if there is an error or failure at the parent node, the messages from its children nodes cannot be delivered.

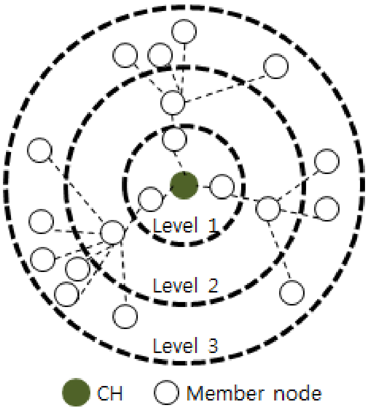


Fig. 1. An example tree in TBC when α = 3.

2.3. BCA

In the BCA protocol [12], every cluster area is almost the same even when sensor nodes are deployed non-uniformly over the network area. The balanced clustering is achieved by electing the CH on the basis of relative node density. For a node *n*, the relative node density *D(n)* is given by dividing node density by network density, where the node density is the ratio of the number of nodes within the node's sensing range over the node's sensing area and the network density is the ratio of the total number of nodes in the network over the network area. Therefore, *D(n)* can be represented by

$$D(n) = \frac{F/(\pi R^2)}{N/A} = \frac{F/N}{\pi R^2/A}, \tag{3}$$

where F is the number of nodes within the node's sensing range, N is the total number of nodes in the network, R is the sensing range, and A is the network area.

The CH is selected according to a new threshold taking the *D(n)* into consideration. That is, for a node n, the new threshold value $\tilde{T}(n)$ at the *r*-th round is calculated by

$$\tilde{T}(n) = T(n) + \frac{mT(n)}{N}\left(\frac{1}{D(n)} - 1\right), \tag{4}$$

where *T(n)* is the same threshold value calculated in (1), N is the total number of nodes in the network, and m is the number of living nodes in the network.

In the region where the node density is high, $\tilde{T}(n)$ is decreased compared to *T(n)* and, thus, a less number of CHs are selected every round. This results in balanced clustering even when sensor nodes are non-uniformly deployed. After cluster formation, if the number of nodes in a cluster exceeds the average number of nodes per cluster in the network, the randomly chosen excessive nodes in the cluster are remained sleep every round. That is, the nodes not included in clusters in dense regions are remained sleep every round. However, when sensor nodes are regularly deployed in the network area, BCA incurs extra overhead for calculating the node density unnecessarily.

2.3. Other Clustering Protocols

LEACH-C [5] is a centralized version of LEACH. That is, the base station elects cluster heads and forms clusters. All nodes in the network send a message including position and residual energy information to the base station. Based on the information, the base station selects cluster heads and divides all nodes to the clusters. Then, the base station broadcasts the information of clusters to all the nodes which are deployed in the network area.

HEED [13] uses some values which take into account the nodes residual energy for cluster formation. A node with more residual energy can be elected as a cluster head for prolonging network lifetime. If candidates for the cluster head have the same residual energy, then their transmission costs are compared.

In BCDCP [11], the complex calculations are assigned to the base station as in LEACHC. In cluster formation, base station elects a candidate set of cluster heads to determine cluster heads. In this scheme, cluster heads send aggregated messages to the base station on a multi-hop basis without direct transmission.

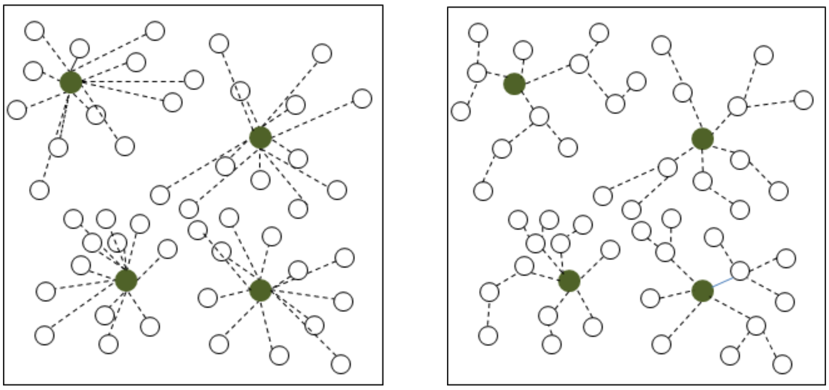
In TEEN [9], sensor nodes manage the threshold data reactively. The process which excludes the threshold value is equal to LEACH. The cluster formation process in TEEN is the same as that in LEACH. After cluster formation, cluster heads transmit the parameters of the data, the hard threshold (HT) value, and the soft threshold (ST) value to their member nodes. All nodes collect and transmit data when the value exceeds the HT value first. After exceeding HT, nodes collect and transmit data only when the measured data exceeds ST.

APTEEN [10] combines the advantages of LEACH and TEEN. As a hybrid protocol, APTEEN unites the data transmission according to the threshold value of TEEN and the periodic data transmission of LEACH. After cluster formation, the cluster heads transmit the threshold value and parameters that include the TDMA schedule time to the member nodes.

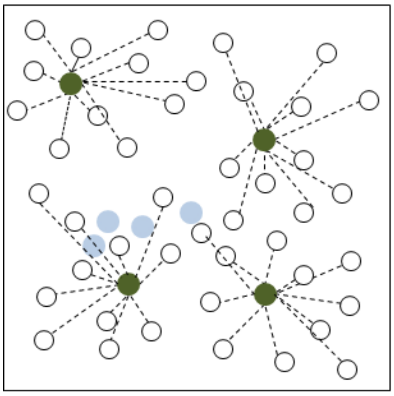
More recently, some works on clustering have been reported in the literature [7-8, 14] even though they do not achieve a major quantum jump. They mainly focus on the improvement of energy efficiency because the energy efficiency is one of the most important design criteria for prolonging network lifetime in battery-operated wireless sensor networks. In addition, they do not take the non-uniform deployment of sensor nodes into consideration yet.

3 Qualitative Comparison of Clustering Protocols

In this section, the three typical clustering protocols are qualitatively compared. The examples of cluster formation in a non-uniformly deployed WSN are shown in Fig. 2, in which the three clustering schemes of LEACH, TBC and BCA are compared schematically. In BCA, some nodes are slept in the densely populated clusters as shown in Fig. 2(c). The sleep nodes are randomly chosen every round.



(a) LEACH (b) TBC



(c) BCA

Fig. 2. Examples of cluster formation in a non-uniformly deployed WSN

The three clustering protocols of LEACH, TBC and BCA are qualitatively compared in terms of various technical aspects in Table 1, in which the major features and characteristics of the three protocols are comparatively summarized.

Table 1. Comparison of Clustering Protocols

Protocol	LEACH	TBC	BCA
CH selection criteria	Probability	Probability	Node density and probability
Performance in non-uniform deployment	Low	Low	Middle
Data delivery latency	Short	Middle	Short
Topology within a cluster	Star	Tree	Star
Impact of node failure	Low	Middle	Low
Energy consumption at CHs	High	Middle	Middle

For non-uniformly deployed WSNs, BCA is better than LEACH and TBC because the node density is additionally considered in selecting CHs. On the other hand, the multi-level tree structure in TBC results in not only the increased end-to-end latency of data delivery but also more loss of data when an upper-level node is failed.

4 Quantitative Performance Study

In this section, the performance of the three clustering protocols in normally and exponentially distributed WSNs is quantitatively evaluated via computer simulation using Matlab and compared with each other. As described earlier, the popular LEACH is a pioneer protocol in clustering for WSNs, TBC is the most advanced clustering scheme for uniformly deployed WSNs, and the recently developed BCA is a single-hop clustering scheme targeted for non-uniformly deployed WSNs.

4.1. Simulation Environment

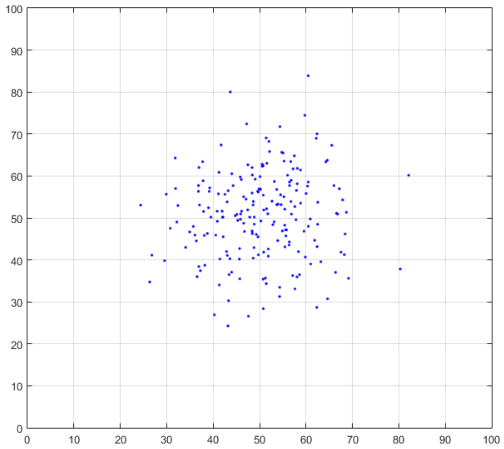
In our simulation, 200 sensor nodes are deployed over the network area of 100 × 100 m^2. The sink node (or base station) is fixed at the location (125, 75), and the initial energy of each senor node is set to 2 J. In our simulation, four non-uniform deployments are experimented as shown in Fig. 3: (1) 200 nodes are deployed according to the normal distribution with mean of (50, 50) and variance of (±10, ±10), (2) 200 nodes are deployed according to the normal distribution with mean of (50, 50) and variance of (±20, ±20), (3) 200 nodes are deployed according to the exponential distribution with mean of (50 ± 10, 50 ± 10), and (4) 200 nodes are deployed according to the exponential distribution with mean of (50 ± 20, 50 ± 20).

In our experiment, the energy consumption model [2] is as follows: The free space (*fs*) model is used if the distance is less than a threshold d_0; otherwise, the multipath (*mp*) model is used. Hence, when transmitting *k* bits of a message along with distance *d*, the energy consumption can be calculated by

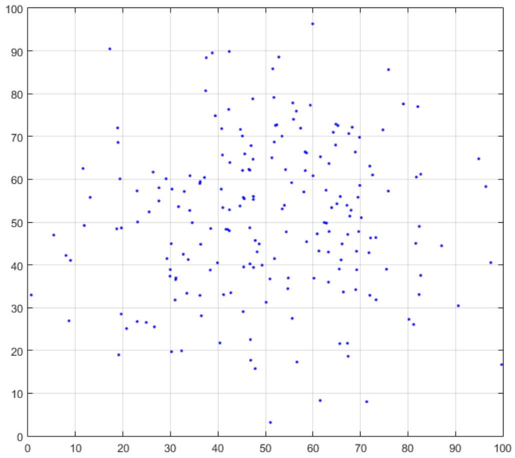
$$\begin{aligned} E_{Tx}(k,d) &= E_{Tx-elec}(k) + E_{Tx-amp}(k,d) \\ &= \begin{cases} kE_{elec} + k\varepsilon_{fs}d^2, & \textbf{if } d < d_0 \\ kE_{elec} + k\varepsilon_{mp}d^4, & \textbf{otherwise} \end{cases} \end{aligned} \tag{5}$$

where d_0 is set to 87 m as in [6]. The energy consumption for receiving *k* bits of data is calculated by

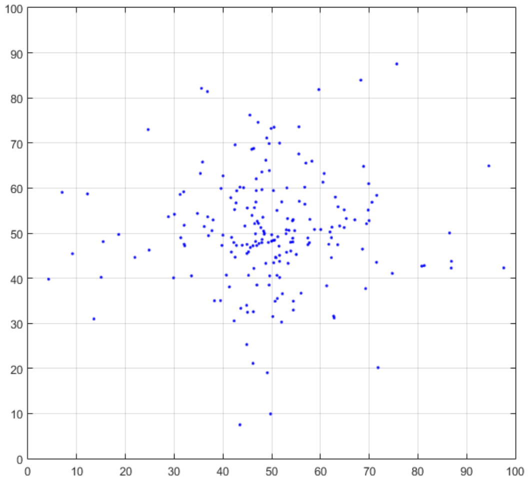
$$E_{Rx}(k,d) = E_{Rx-elec}(k) = kE_{elec} \ . \tag{6}$$



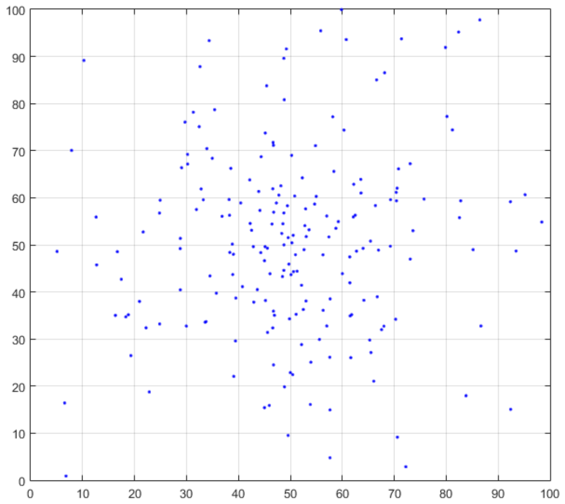
(a) Normal distribution with mean (50, 50) and variance (±10, ±10)



(b) Normal distribution with mean (50, 50) and variance (±20, ±20)



(c) Exponential distribution with mean (50 ± 10, 50 ± 10)



(d) Exponential distribution with mean (50 ± 20, 50 ± 20)

Fig. 3. Four non-uniform deployments of 200 nodes for simulation

In (5) and (6), E_{elec} is the radio electronics energy depending on digital coding, modulation, filtering and spreading of the signal. ε_{fs} and ε_{mp} are constant values for the amplifier energy depending on the distance to the receiver and acceptable bit-error rate.

The parameters used in our simulation are summarized in Table 2. In the table, E_{sense} is the energy consumption required for sensing and E_{da} is the energy consumption for data aggregation. The simulations were performed 100 times for each experiment and the mean value of results was used as the simulation results.

Table 2. Simulation parameters

Parameter	Value
Network area	100 × 100 m^2
Location of sink	(125, 75)
Number of nodes	200
Number of clusters	10
Initial energy	2 J
E_{sense}	5 nJ/bit
E_{da}	5 nJ/bit
E_{elec}	50 nJ/bit
E_{fs}	10 pJ/bit/m^2
E_{mp}	0.00013 pJ/bit/m^4
Sensing range	10 m
Maximum transmission range	136 m

4.2. Simulation Results and Discussion

In our performance study, the network lifetime is extensively evaluated because it is the most important metric in WSNs. The network lifetime in our performance study is defined as the time duration until half of the sensor nodes die due to the energy depletion of battery. So, the number of living nodes is observed with respect to round progress. Figs. 4 to 7 show the number of living nodes along with round for the four scenarios of non-uniform deployment described in Section 4.1.

Figs. 4 and 5 show the number of living nodes along with round for the two normal distribution scenarios of non-uniform deployment. In the first deployment that 200 nodes are deployed according to the normal distribution with mean of (50, 50) and variance of (±10, ±10), the network lifetime is 6 to 85 percent longer than the others. In the second deployment that 200 nodes are deployed according to the normal distribution with mean of (50, 50) and variance of (±20, ±20), the network lifetime is 3 to 31 percent longer than the others. That is, it can be easily inferred that the improvement is better and better as the variance decreases.

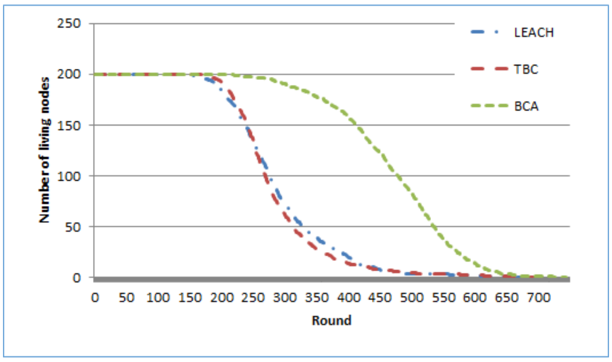


Fig. 4. Network lifetime when 200 nodes are deployed according to the normal distribution with mean of (50, 50) and variance of (±10, ±10)

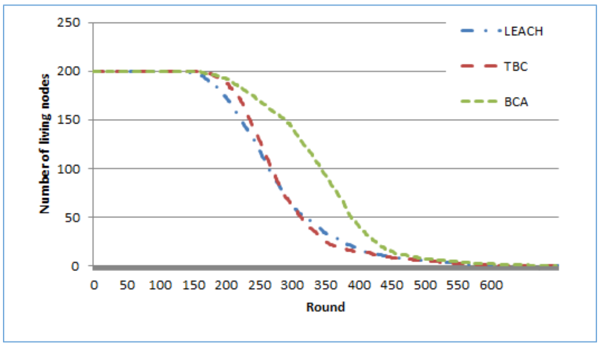


Fig. 5. Network lifetime when 200 nodes are deployed according to the normal distribution with mean of (50, 50) and variance of (±20, ±20)

Figs. 6 and 7 show the number of living nodes along with round for the two exponential distribution scenarios of non-uniform deployment. In the third deployment that 200 nodes are deployed according to the exponential distribution with mean of (50 ± 10, 50 ± 10), the network lifetime is 10 to 76 percent longer than the others. In the fourth deployment that 200 nodes are deployed according to the exponential distribution with mean of (50 ± 20, 50 ± 20), the network lifetime is 6 to 43 percent longer than the others. That is, it can be easily inferred that the improvement is better and better as the mean decreases.

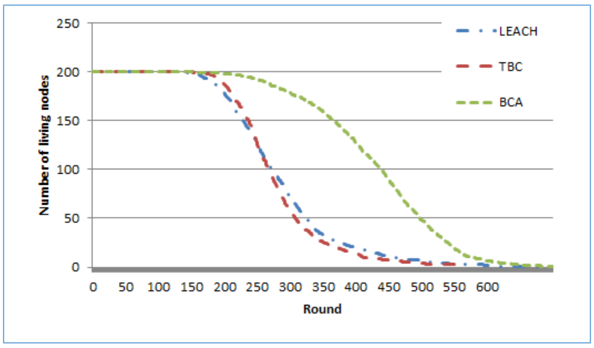


Fig. 6. Network lifetime when 200 nodes are deployed according to the exponential distribution with mean of (50 ± 10, 50 ± 10)

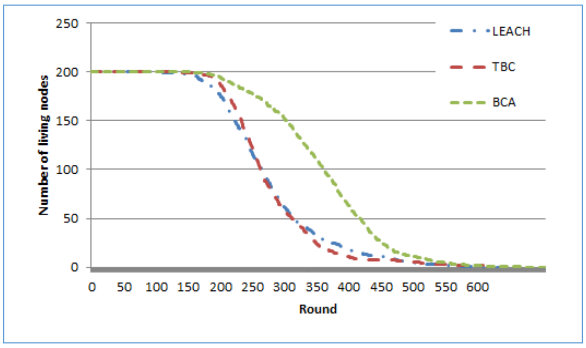


Fig. 7. Network lifetime when 200 nodes are deployed according to the exponential distribution with mean of (50 ± 20, 50 ± 20)

Among the three clustering schemes, LEACH shows the worst performance in our simulation. BCA obviously outperforms LEACH and TBC as shown in all the four graphs. Therefore, BCA can be a primary choice for clustering in normally and exponentially distributed WSNs. In BCA, the network lifetime is remarkably prolonged. This is mainly due to that CHs are distributed evenly over the network area and every cluster has almost the same coverage area. Moreover, excessively redundant nodes are turned into sleep mode to save energy. That is, the unnecessary redundant sensing and transmissions are significantly reduced.

5 Conclusions

In this paper, the three typical clustering protocols in normally and exponentially distributed WSNs have been qualitatively and quantitatively analyzed and compared with each other. As the most important performance metric of

WSNs, the network lifetime has been evaluated for the four different scenarios of sensor distribution. According to the simulation results, BCA outperforms the other protocols and the performance improvement is better and better as the non-uniformity is increased. In BCA, a less number of cluster heads are selected every round in the highly populated region. This enables balanced clustering even when sensor nodes are normally and exponentially distributed. Furthermore, the nodes not included in clusters in the highly populated region are remained sleep for energy saving.

References

[1] Asaduzzaman and H. Y. Kong, Energy Efficient Cooperative LEACH Protocol for Wireless Sensor Networks, *Jounal of Communications and Networks*, **12** (2010), no. 4, 358-365. http://dx.doi.org/10.1109/jcn.2010.6388472

[2] W. Bo, H. Y. Hu and F. Wen, An Improved LEACH Protocol for Data Gathering and Aggregation in Wireless Sensor Networks, *Proc. of 2008 Int. Conf. on Computer and Electrical Engineering*, (2008), 398-401. http://dx.doi.org/10.1109/iccee.2008.59

[3] L. Dan, K.D. Wong, H.H. Yu and A. M. Sayeed, Detection, Classification, and Tracking of Targets, *Proc. of IEEE Signal Processing Magazine*, **19** (2002), no. 2, 17-29. http://dx.doi.org/10.1109/79.985674

[4] W.R. Heinzelman, A. Chandrakasan and H. Balakrishnan, Energy-efficient Communication Protocols for Wireless Microsensor Networks, *Proc. of the Hawaii International Conference on Systems Sciences*, Vol. 2, (2010), 10-19. http://dx.doi.org/10.1109/hicss.2000.926982

[5] W.B. Heinzelman, A.P. Chandrakasan and H. Balakrishnan, An Application-Specific Protocol Architecture for Wireless Microsensor Networks, *IEEE Transactions on Wireless Communications*, **1** (2002), no. 4, 660-670. http://dx.doi.org/10.1109/twc.2002.804190

[6] K. T. Kim, C. H. Lyu, S. S. Moon and H. Y. Yoon, Tree-Based Clustering (TBC) for Energy Efficient Wireless Sensor Networks, *Proc. of IEEE 24th Int. Conf. on Advanced Information Networking and Applications Workshop*, (2010), 680-685. http://dx.doi.org/10.1109/waina.2010.62

[7] J.-S. Lee and W.-L. Cheng, Fuzzy-Logic-Based Clustering Approach for Wireless Sensor Networks Using Energy Predication, *IEEE Sensors Journal*, **12** (2012), no. 9, 2891-2897. http://dx.doi.org/10.1109/jsen.2012.2204737

[8] K. Li and K. A. Hua, Mobility-Assisted Distributed Sensor Clustering for Energy-Efficient Wireless Sensor Networks, *Proc. of 2013 IEEE Global Communications Conference*, (2013), 316-321. http://dx.doi.org/10.1109/glocom.2013.6831090

[9] A. Manjeshwar and D. Agrawal, TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks, *Proc. of 15th Int. Parallel and Distributed Processing Symposium*, (2001), 2009-2015. http://dx.doi.org/10.1109/ipdps.2001.925197

[10] A. Manjeshwar and D. P. Agrawal, APTEEN: A Hybrid Protocol for Efficient Routing and Comprehensive Information Retrieval in Wireless, *Proc. of Int. Parallel and Distributed Processing Symposium*, (2002), 195-202. http://dx.doi.org/10.1109/ipdps.2002.1016600

[11] S. D. Muruganathan, D. C. F. Ma, R. I. Bhasin, and A. O. Fapojuwo, A Centralized Energy-Efficient Routing Protocol for Wireless Sensor Networks, *IEEE Communications Magazine*, **43** (2005), no. 3, S8-S13. http://dx.doi.org/10.1109/mcom.2005.1404592

[12] H. Shin, S. Moh, I. Chung, and M. Kang, Equal-Size Clustering for Irregularly Deployed Wireless Sensor Networks, *Wireless Personal Communications*, **82** (2014), no. 2, 995-1012. http://dx.doi.org/10.1007/s11277-014-2262-5

[13] O. Younis and S. Fahmy, HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks, *IEEE Transactions on Mobile Computing*, **3** (2004), no. 4, 366-379. http://dx.doi.org/10.1109/tmc.2004.41

[14] L. Xu, G.M.P. O'Hare and R. Collier, A Balanced Energy-Efficient Multihop Clustering Scheme for Wireless Sensor Networks, *Proc. of 7th IFIP Wireless and Mobile Networking Conference*, (2014), 1-8. http://dx.doi.org/10.1109/wmnc.2014.6878886

Received: June 27, 2016; Published: August 11, 2016