Performance Analysis According to the change of Cluster Size in Large Scale Wireless Sensor Networks

Inbo Sim and Jaiyong Lee

Yonsei University, Seoul, Korea

Summary

In large scale sensor networks, scalability and robustness are very important, and sensor nodes are also highly power constrained. Clustering, a method of grouping sensor nodes, can meet these requirements. Sensor networks also need to support fast delivery of important and urgent data. The purpose of this paper is to analyze the effects of cluster size on application's QoS, select proper cluster size based on the analysis, and suggest method to improve performance in clustering environment.

Key words:

Cluster, Scalability, Packet delivery, Delay, Energy, WSN

1. Introduction

Sensor networks have uses in disaster prediction, security, environmental monitoring, and traffic control. A wide variety of network sizes are used in these applications. In environmental monitoring, for example, hundred or thousands of sensor nodes are deployed in a large monitoring region. In such large scale sensor networks, scalability and robustness are very important. In addition, sensor nodes are highly power constrained, and they must work at very low power to prolong the lifetime of the sensor network. Clustering, a method of grouping sensor nodes, can meet these requirements, and it has been the focus of much research on sensor networks.

Also sensor networks need to support fast delivery of important and urgent data. For example, consider a sensor network deployed to sense the temperature in a forest. An abnormally high temperature in a particular location may be an indication of a fire. As a result, such messages have to be transferred to the sink node as fast as possible, not being lost. A sensor network monitoring environmental conditions such as pressure and seismic activity is another example. Similarly, one can think of many situations where packet delivery and delay must be considered. However, considering packet delivery and delay in WSN is still a largely unexplored research field.

The purpose of this paper is to analyze the effects of cluster size on application's QoS (Quality of Service), select proper cluster size based on the analysis, and suggest method to improve performance in clustering environment.

The remainder of the paper is structured in the following manner. In section 2, we describe related studies on clustering and multi-hop communication between clusters. In section 3, we explain our network model. In section 4, we present simulation results and discussion of them. In section 5, we propose method to improve performance in clustering environment, followed by the conclusion in section 6.

2. Related Work

In sensor networks, a low-energy adaptive clustering hierarchy (LEACH) has been proposed as a clustering method for reducing the power consumption of sensor networks [1]. In LEACH, each sensor node decides whether to become a cluster header based on a predetermined percentage. Clusters are constructed with sensor nodes adjacent to cluster headers. Communication from sensors nodes to cluster headers and communication from cluster headers to a sink node is performed via a single hop.

Many clustering methods aimed at improving on LEACH have been suggested. In LEACH, the positions of clusters can be unbalanced, decreasing the network lifetime. Hybrid energy-efficient distributed clustering (HEED) [2] places clusters uniformly over a monitoring region. Furthermore, HEED balances power consumption between sensor nodes. The central controlling algorithm, which provides a regular cluster size, was also proposed [3]. The time complexity of this algorithm is $O(n^3)$, where n is the number of sensor nodes. Although LEACH uses one-hop communication within a cluster, hybrid indirect transmission (HIT)[4] uses multi-hop communication within clusters to limit the interference range and to communication in parallel with as many nodes as possible.

The transmission distance must be reduced to minimize the power consumption of sensor networks. From this standpoint, power-efficient gathering in sensor information systems (PEGASIS) [5] comprises a chain instead of a set of clusters. This chain connects the nearest neighboring sensor nodes, and the distance between sensor nodes is very short. Two-Phase Clustering (TPC) [6] also constructs chains within clusters. Although these studies evaluated various clustering methods by using several metrics such as power consumption or data collection time, most of them focused on small sensor networks in which one-hop communication between an arbitrary sensor node and the sink node can always be achieved. Although multi-hop

---

Manuscript received April 5, 2009
Manuscript revised April 20, 2009

communication between clusters is employed in [6], its performance is not very specifically mentioned.

Studies have also been done on multi-hop communication between clusters in sensor networks. In Connectionless Probabilistic (CoP) routing [7], the monitoring region is divided into square areas, and multi-hop communication takes place between cluster headers positioned at the vertices of these areas. Neander et al. evaluated a sensor network using multi-hop communication between cluster headers mainly through simulation experiments [8] that assumed the sink node can directly communicate with all sensor nodes. The unequal clustering size (UCS) [9] was designed to equalize power consumption among cluster headers. In UCS, a circular monitoring region is split into two concentric circles, called layers. Soro and Heinzelman determined that the size of the cluster in the interior layer should be reduced to equalize the power consumption. Shu et al. divided a monitoring region into multiple layers and derived optimal parameters, such as the cluster radius of each layer and the relay probabilities of cluster headers, to prolong the coverage-time [10].

3. Sensor Network Model

3.1 Network Model

A model of the sensor network under consideration is as follows. Data fusion is not used to reduce the data volume. That is, the data generated by the sensor nodes is transmitted to the sink node without any modification or compression. We assume that both the sink node and the sensor nodes are stationary after deployment. Moreover, they have the ability to control the transmission power depending on the distance between the sensor node and its next-hop node.

Clustering is made by physical partitioning based on geographical node position. Cluster header changes every certain time (T), and the node with the biggest residual energy is selected as the header. The sink node is placed at the center in a 240m×240m square area. The sensor nodes are placed randomly and uniformly throughout the area. The same wireless channel is used in the entire network for the intra-cluster communication and another one is used between cluster headers. Communication in each cluster is one-hop transmission to cluster header. Also, in large sensor networks, one-hop communication between an arbitrary cluster header and the sink node is unrealistic because of sensor node’s physical constraint. Thus, communication between a cluster header and the sink node involves multi-hop transmission with minimum transmission range for connectivity.

3.2 Routing Protocol

In WSN, in addition to energy saving, scalability is another important factor in designing a routing protocol for sensor networks. A good routing protocol has to be scalable and adaptive to the changes in the network topology. Scalability means that the protocol performs well as the network grows larger or as the workload increases. It is best accomplished through decentralized algorithms, where nodes only need local information exchange to make routing decisions. It has been experimentally shown that the protocols which use geographical location of the nodes are more scalable than non-geographical protocols [11].

For these reasons, in this paper we will use geographical routing as a routing protocol for multi-hop transmission between a cluster header and the sink node.

We assume that each cluster header knows the locations of all its neighbors and the location of the destination node (the sink node). Based on this assumption, a transmission strategy can be designed as follows:

Among the neighbor cluster headers closer in distance to the destination node than the source node (the cluster header), the source node sends the packet to the neighbor cluster header closest at itself.

3.3 MAC Protocol

In WSN, S-MAC[12] uses fixed duty cycles. To improve energy efficiency, T-MAC[13] uses a timer to switch to the sleep mode after a certain period of time when it detects that there is no data to send or receive. Also, B-MAC[14] which is asynchronous MAC protocol supports CSMA/CA (Carrier Sense Multiple Access/Collision Avoidance) with LPL (Low Power Listening) where each node periodically wakes up after a sample interval and checks the channel for activity for a short duration of 2.5ms. If the channel is found to be active, the node stays awake to receive the payload following an extended preamble. TDMA protocols such as TRAMA[15] and LMAC[16] have been proposed and these protocols are able to communicate between node pairs in dedicated time slots.

Thus, most of MAC protocols in WSN use CSAM/CA and TDMA methods. TDMA has a natural advantage of collision free medium access. However TDMA method has weak points as follows. It includes clock drift problems and decreased throughput at low traffic loads due to idle slots. The difficulties with TDMA systems are synchronization of the nodes and adaptation to topology changes when these changes are caused by insertion of new nodes, exhaustion of battery capacities, broken links due to interference, the sleep schedules of relay nodes, and scheduling caused by clustering algorithms. The slot assignments, therefore, should be done with regard to such possibilities. However, it is not easy to change the slot

assignment within a decentralized environment for traditional TDMA, since all nodes must agree on the slot assignments. In accordance with common networking lore, CSMA/CA methods have a lower delay and promising throughput potential at lower traffic loads, which generally happens to be the case in wireless sensor networks [17].
For these reasons, in this paper we will use CSMA/CA as a MAC protocol.

3.3.1 CSMA/CA MAC Protocol

Fig.1 shows the operation of CSMA/CA. In this scheme, every DATA communication is preceded by an exchange of control packets when the data packet size exceeds a particular threshold. When a source S wants to transmit to a destination D, it senses its local channel (physical carrier sensing). If the channel is busy, it backs-off after exponentially increasing its back-off timer. Otherwise, the source transmits a request-to-send (RTS) control message to the destination. If the local channel around D is free, D replies with a clear-to-send (CTS) message, which is then followed by the data packet transmission from S to D, and an acknowledgment (ACK) packet transmission from D to S. If the channel around D is busy, S times out waiting for the CTS message, exponentially backs-off its timeout value and retransmits the RTS packet. Both RTS and CTS packets contain the proposed duration of the upcoming data transmission. Nodes located in the vicinity of communicating nodes, that overhear either (or both) of these control packets, must defer transmission for this proposed duration. This is called virtual carrier sensing which is performed in addition to the physical carrier sensing mentioned earlier. It is implemented by means of a variable called the network allocation vector (NAV). A node updates the value of its NAV with the duration field specified in the RTS or CTS. Thus the nodes lying within the transmission range of the transmitter or the receiver do not initiate any transmission while the communication is in progress. The RTS and CTS packets thereby reserve the local channel for the upcoming DATA transmission by silencing the nodes in the vicinity of the transmitter and the receiver. The CSMA/CA MAC protocol uses a back-off interval to resolve channel contention. A source node S, before initiating a transmission chooses a random back-off interval in the range of [0, CW] where CW represents the contention window. The node S then decrements its back-off counter by one after every idle slot time. When the back-off counter reaches 0, node S transmits its packet. If the transmission from S collides with some other transmission, S doubles its CW, and chooses a new random back-off interval from the new range and then attempts retransmission. The contention window is doubled for every collision until it reaches a maximum threshold called the CW_{max}. While in the back-off stage, if a node senses the channel to be busy, then it freezes its back-off counter. When the channel becomes idle once again for duration DIFS (DCF inter-frame spacing), the back-off counter is resumed to count down from its frozen value. A shorter inter-frame space, SIFS is used to separate transmissions pertaining to the same data packet.

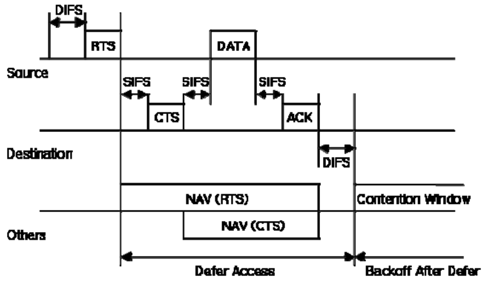

Fig. 1. Operation of CSMA/CA

3.4 Simulation Setting

The simulation was conducted using OPNET. We adopt the transmission strategy explained in section 3.2 for communication between a cluster header in the routing layer and the IEEE 802.11 in the MAC/Physical layer. The transmission data rate is set to 1Mbps, packet size to 128byte, buffer size to 256kbits, long retry number limit to 4, and the packet generated rate is variable from 0.05 to 0.5 packet(s)/sec/node in 0.05 packet(s)/sec/node. The simulations are run on networks with 250 nodes.
We compare the following three parameters for each cluster size(20m×20m, 40m×40m, 60m×60m).
Packet delivery (ratio): ((Total packets generated at each node) – (Dropped packets)) / (Total packets generated at each node)
Delay (sec) (scaled as the log function): The average time taken from when packets were initially sent by the node farthest away from the sink node to when those packets were successfully received at the sink at location (x=120, y=120) in a square area of 240m×240m.
Energy consumption (J): the sum of the energy consumption of each node

4. Simulation Results and Discussion

4.1 Packet delivery

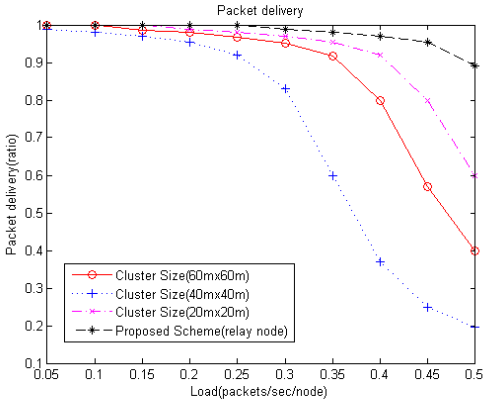

Fig. 2. Packet delivery vs. Load

Fig.2 is a graph of packet delivery. The x-axis shows the load, and the y-axis indicates the packet delivery. We define max load as the maximum load value guaranteeing packet delivery among load values. In WSN, because all nodes transmit packets to the sink node in multi-hop, the nodes which are nearer to the sink node will have to transmit more packets than those nodes which are further away from the sink node[18][19]. Thus, from this fact, we can analogize that max load is decided by the communication between cluster headers neighbor to the sink. When cluster size changes, the traffic in cluster and relayed traffic also changes. This change in traffic affects interference among the cluster headers neighbor to the sink, which have direct effect on max load. According to the increase or decrease in interference, max load also decreases or increases. The result of simulation seen in Fig.2 shows that, in the case where cluster size changes from 20m×20m(40m×40m) to 40m×40m(60m×60m) in this simulation environment, the interference increases(decreases) and max load decreases (increases).

4.2 Delay

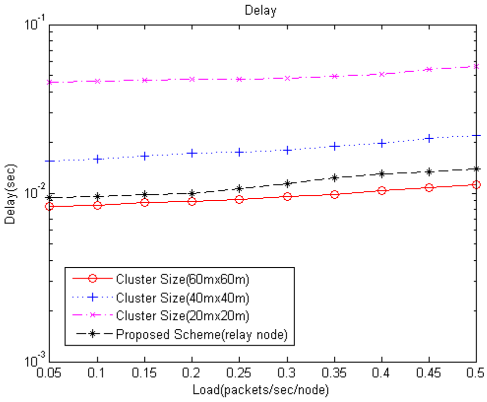

Fig. 3. Delay vs. Load

Fig.3 is a graph of delay. The x-axis shows the load, and y-axis indicates the delay. Delay is important consideration for applications requiring swift data collection in large-scale sensor networks. As mentioned in 4.1, when cluster size changes, traffic in cluster and relayed traffic changes. However, delay has different characteristics than packet delivery in that, despite these changes in traffic, the delay at each hop is not nearly changed in a stable region in which packet delivery is guaranteed [20]. Meanwhile, because there are bigger cluster sizes and lower relayed hop count, when cluster size increases, delay decreases. Also, as a characteristic of 802.11, delay increases dramatically and an increment is determined by the number of interfering neighbors of the node [20]. In this simulation environment with minimum transmission range for connectivity, because the number of interfering neighbors of the cluster header is minimized, an increment is minimized. In addition, the load value where packets start to be dropped is small. For these reasons, delay does not increase suddenly. Based on the results, it can be said that for all loads, bigger cluster size produces better delay characteristics as shown in Fig. 3.

4.3 Energy consumption

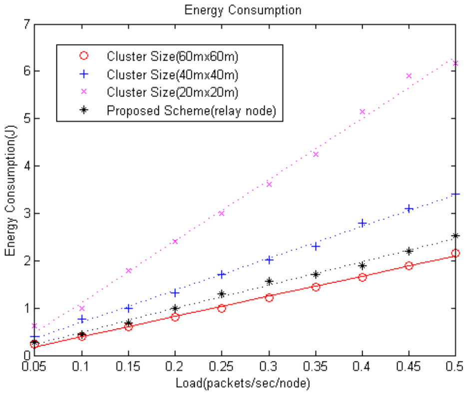

Fig. 4. Energy consumption vs. Load

Fig.4 is a graph of energy consumption. The x-axis shows the load, and y-axis indicates the energy consumption.
In our simulation, we use energy model in [1]. To transmit an l bit message a distance d, the radio expends

$$E_{Tx}(l,d) = E_{Tx-elec}(l) + E_{Tx-amp}(l,d) = lE_{elec} + l\varepsilon_{fs}d^2 \quad (1)$$

And the energy needed to receive l bits of data is

$$E_{Rx}(l) = E_{Rx-elec}(l) = lE_{elec} \quad (2)$$

The electronics energy, E_{elec}, depends on factors such as the digital coding, modulation, filtering, and spreading of the signal, whereas the amplifier energy, $\varepsilon_{fs}d^2$, depends on the distance to the receiver and the acceptable bit-error rate. The communication energy parameters are set as [1]: E_{elec} = 50 nJ/bit, ε_{fs}= 10 pJ/bit/m^2.

From the viewpoint of the whole network, when cluster size increases, distance *d* increases and relayed traffic decreases. From equation (1), because *d* does not give a dominant impact on the energy consumption, as the cluster size increases, relayed traffic decreases, resulting in decrease in energy consumption. However, if *d* increases and gives a dominant impact on the energy consumption, though the increase of the cluster size brings about the decrease of the relayed traffic, because *d's* increase is relatively large, as a result, the energy consumption increases when the cluster size increases. But, in large scale sensor networks with sensor node's physical constraint, relayed traffic is many and *d* isn't large enough to give a dominant impact on the energy consumption. In such environment, when cluster size increases, since the amount of decreased energy consumption is much greater than the amount of increased energy consumption, the energy consumption decreases as a whole, such as Fig. 4.

5. Performance Enhancement Scheme

The simulation results show that when cluster size increases, energy consumption and delay become better and packet delivery can become better or worse. In this chapter, a scheme is proposed in order to maintain delay and energy characteristics of maximum cluster size, and to improve packet delivery in WSN based on clustering.

As mentioned in 4.1, the max load of network is decided by the cluster headers neighbor to the sink. This is because that is where traffic gathers. In addition, when cluster size increases to maximum size, the transmission range also increases to guarantee communication connectivity among cluster headers. If this causes more interference with traffic gathering in the sink, then packet delivery decreases. The decrease of packet delivery can be improved most effectively by positioning relay nodes between the sink and cluster headers that are one-hop close to the sink. Because these relay nodes decrease transmission range of cluster headers that are one-hop close to the sink, the interference of cluster headers that are one-hop close to the sink decreases, resulting improvement of packet delivery. To have communication connectivity, maximum cluster size is limited by the maximum transmission range. In this paper, the maximum cluster size was assumed to be 60m×60m, and a simulation was carried out after positioning four relay nodes at the each vertex of the square of 60m×60m centering the sink in the same simulation environment that has a 60m×60m cluster size. The results seen in Fig.2-4 show that, although the scheme does not guarantee maximum packet delivery, it improves packet delivery significantly by sacrificing the least increase of energy consumption and delay.

6. Conclusion

It was shown in this paper that when cluster size increases, energy consumption and delay become better and packet delivery can become better or worse. Based on these results, it can be concluded that by allocating relay nodes between the sink and cluster headers that are one-hop close to the sink, the proposed scheme improves packet delivery significantly by sacrificing the least increase of energy and delay.

Acknowledgment

This research was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Assessment) (IITA-2009-C1090-0902-0038)

References

[1] W.R. Heinzelman, A. Chandrakasan, H. Balakrishnan, An application-specific protocol architecture for wireless microsensor networks, IEEE Transactions on Wireless Communications, October 2002, Vol.1, pp. 660-670

[2] O. Younis and S. Fahmy, Distributed clustering in ad-hoc sensor networks: A, hybrid, energy-efficient approach, IEEE INFOCOM, March 2004, pp.629-640

[3] S. Ghiasi, A. Srivastava, X. Yang, and M. Sarrafzadeh, Optimal energy aware clustering in sensor networks, Sensor Magazine, July 2002, pp. 258-269

[4] B.J. Culpepper, L. Dung, and M. Moh, Design and analysis of hybrid indirect transmissions (HIT) for data gathering in wireless micro sensor networks, ACM Mobile Computing and Communications Review, Jan. 2004, vol.8, no.1, pp. 61-83

[5] S. Lindsey, C. Raghavendra, and K.M. Sivalingam, Data gathering algorithms in sensor networks using energy metrics, IEEE Trans. Parallel Distrib. Syst., Sept.2002, vol.13, no9, pp. 924-935

[6] W. Choi, P.Shah, and S.K. Das, A framework for energy-saving data gathering using two-phase clustering in wireless sensor networks, Proc. First Annual International Conference on Mobile and Ubiquitous System: networking and Services, Aug. 2004, pp. 203-212

[7] A.A. Papadopoulos and J.A. McCann, Connectionless probabilistic (CoP) routing: An efficient protocol for mobile wireless ad-hoc sensor networks, Proc. 24th International Performance Computing and Communications Conference, April 2005, pp. 73-77

[8] J. Neander, E. Hansen, M. Nolin, and M. Bjorkman, Asymmetric multihop communication in large sensor networks, Proc. International Symposium on Wireless Pervasive Computing, Jan. 2006

[9] S. Soro and W.B. Heinzelman, Prolonging the lifetime of wireless sensor networks via unequal clustering, Proc. 19th

IEEE International Parallel and Distributed Processing Symposium, April 2005

[10] T. Shu, M. Krunz, and S. Vrudhula, Power balanced coverage-time optimization for clustered wireless sensor networks, Proc. ACM MobiHoc, May 2005, pp. 111-120

[11] D. Braginsky and D. Estrin, Rumor Routing Algorithm for Sensor Networks, ACM WSNA, 2002, pp. 22–31

[12] W. Ye, J. Heidemann, D. Estrin, Medium Access Control with Coordinated Adaptive Sleeping for wireless Sensor networks, IEEE/ACM Transaction on Networking, June 2004.

[13] T.V. Dam, K. Langendeon, An adaptive Energy-Efficient MAC protocol for wireless Sensor networks, Sensys'03, November 2003.

[14] J. Polastre, J.Hill, and D. Culler, Versatile low media access for wireless sensor networks, SenSys, November 2005.

[15] V. Rajendram, K. Obraczka, and J. J. Garcia-Luna-Aceves, Energy-efficient, collision-free medium access control for wireless sensor networks, SenSys, 2003.

[16] L.F.W. van Hoesel and P.J.M. Havinga, A lightweight medium access protocol for wireless sensor networks, INSS, 2004

[17] I. Demirkol, C. Ersoy, and Fatih Alagöz, MAC Protocols for Wireless Sensor Networks: A Survey, IEEE Communications magazine, April 2006, pp. 115-121

[18] Dali Wei, H.Anthony Chan, and Kevin V. N Kameri, Circular-Layer Algorithm for Ad Hoc Sensor Networks to Balance Power Consumption, IEEE SECON, September 2006, Vol.3, pp.945-950

[19] Kumar Padamanabh, Rajar Roy, Doughnut Effect in wireless Sensor Network and its Solution, IEEE SECON, September 2006, Vol.3, pp.957-963

[20] H. Zhai, Y. Kwon and Y. Fang, Performance analysis of IEEE 802.11 MAC Protocols in wireless LANs, Wirel. Commun. Mob. Comput., 2004, pp. 917-931

Inbo Sim received the B.S. degree in electronic engineering from ROK Air Force Academy, in 1995 and the M.S. degrees in electrical and electronics engineering from Yonsei, Republic of Korea, in 2000, where he is currently pursuing the Ph.D. degree. He has been with Korea Air Force as an officer from 1995. His research interest is in MAC/routing protocol design in WSNs

Jaiyong Lee received Ph.D. degree in Computer engineering from Iowa State University, USA in 1987. He was with ADD (Agency for Defense Development) as a research engineer during 1977 - 1982, and with Computer Science Dept. of POSTECH as an associate professor during 1987 - 1994. Since 1994, he has been a professor in School of EE, Yonsei University. He was the Executive Vice President of OSIA (Open standards and Internet Association) during 2005 and an IT professional of MIC (Ministry of Information and Communication Republic of Korea) during 2003 - 2007. He is also actively serving in the IT areas such as Executive Director of KICS (Korea Institute of Communication Sciences), Advisory Board Member of Korean Association of RFID/USN and Editor of JCN and ETRI journal. Recently, he is performing research works in the areas of QoS and mobility management for supporting 4G architecture, multicast protocol design for wireless access network and sensor MAC/routing protocol design