# An Adaptive Routing Protocol for Heterogeneous Wireless Sensor Networks

January 2013

Sampath PRIYANKARA

An Adaptive Routing Protocol for Heterogeneous Wireless Sensor Networks January 2013 Sampath PRI

# An Adaptive Routing Protocol for Heterogeneous Wireless Sensor Networks

Submitted to
Graduate School of Information Science and Technology
Osaka University

January 2013

Sampath PRIYANKARA

# List of Publications by the Author

## I. Journals

[1] Sampath Priyankara, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami, "A Clustering Method for Wireless Sensor Networks with Heterogeneous Node Types," *IEICE Transactions on Communications*, Vol. E94-B, No. 8, pp. 2254–2264, August 2011.

[2] Sampath Priyankara, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami, "A Hybrid Routing Method for Wireless Sensor Networks with Heterogeneous Node Types," *IEICE Transactions on Communications*, Vol. J96-B, No. 2, February 2013. (*to be published*) (*written in Japanese*)

## II. International Conferences

[1] Sampath Priyankara, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami, "A Clustering Method for Wireless Sensor Networks with Heterogeneous Node Types," in *Proceedings of the 18th International Conference on Computer Communications and Networks (ICCCN 2009)*, August 2009.

[2] Sampath Priyankara, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami, "Cluster-based Routing Method for Wireless Sensor Networks with Heterogeneous Node Types," in *Proceedings of International Conference on Multimedia, Information Technology and its Applications (MITA 2009)*, Session C-4, pp. 85–88, August 2009.

[3] Sampath Priyankara, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami, "A Clustering/Multi-hop Hybrid Routing Method for Wireless Sensor Networks with Heterogeneous Node Types," in *Proceedings of the IEEE GLOBECOM 2010 Workshop on Heterogeneous, Multi-hop Wireless and Mobile Networks (HeterWMN)*, December 2010.

[4] Sampath Priyankara, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami, "A Generalized Spatial Boundary Analysis Method for Clustering/Multi-hop Hybrid Routing in Wireless Sensor Networks," in *Proceedings of the IEEE Globecom 2011, Workshop on*

*Heterogeneous, Multi-hop, Wireless and Mobile Networks (HeterWMN)*, December 2011.

# III. Domestic Conferences

[1] Sampath Priyankara, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami, "A Clustering Method for Wireless Sensor Networks with Heterogeneous Node Types," in *IEICE Technical Report*, vol. 108, no. 258, NS2008-78, pp. 61–64October 2008 (*written in Japanese*).

[2] Sampath Priyankara, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami, "A Clustering Method for Wireless Sensor Networks with Heterogeneous Node Types," in *Proceedings of the 2009 IEICE General Conference*, BS-4-33, March 2009.

[3] Kana Uekubo, Sampath Priyankara, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami, "An Enhanced DSR Protocol for Heterogeneous Wireless Sensor Networks," in *Proceedings of the 2010 IEICE General Conference*, B-6-136, March 2010. (*written in Japanese*)

[4] Sampath Priyankara, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami, "Hybrid Routing Protocol for Load Balancing around Sink in Wireless Sensor Networks," in *IEICE Technical Report*, vol. 109, no. 448, NS2009-202, pp. 229–234, March 2010. (*written in Japanese*)

[5] Sampath Priyankara, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami, "A Study on Efficient Node Placement for Clustering/Multi-hop Hybrid Routing in Heterogeneous Wireless Sensor Networks," in *Proceedings of the IEICE General Conference*, BS-7-23, September 2010.

[6] Sampath Priyankara, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami, "Spatial Boundary Analysis Method for Clustering/Multi-hop Hybrid Routing in Homogeneous Wireless Sensor Networks," in *Proceedings of the IEICE General Conference*, BS-3-39, March 2012.

[7] Sampath Priyankara, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami, "Hybrid Routing Protocol with Dynamic Reconfiguration for Wireless Sensor Networks," in *IE-*

*ICE Technical Report*, December 2012.(*written in Japanese*)

# Abstract

This dissertation describes the study of an adaptive routing protocol for heterogeneous wireless sensor networks which has been studied since 2007 by the author who was a undergraduate student and has been a Ph. D. student in the Department of Information Networking, Graduate School of Information Science and Technology, Osaka University.

In recent years, Wireless Sensor Networks (WSNs) which consists of a large number of sensor nodes having wireless communication and self-organizing capability receives much attention. Sensor nodes are tiny, battery powered modules with limited on-board processing, storage and radio capabilities. Sensor nodes sense and send their data toward a central processing center which is called "sink" by creating self-organized network. Those nodes are powered by embedded batteries and replacement of those batteries is a very difficult process once those nodes have been deployed. Therefore, the design of protocols and applications for such networks has to be energy aware in order to prolong the lifetime of the network.

WSN can be classified into two types; homogeneous and heterogeneous sensor networks. In homogeneous networks, all the sensor nodes are identical in terms of battery capacity and hardware complexity. On the other hand, in a heterogeneous sensor network, two or more different types of nodes with different battery capacity, communication capability or functionality are used. The motivation is that the more complex hardware and extra battery energy can be embedded in a few cluster head nodes, thereby the hardware cost of the rest of the network and communication cost of the sensing nodes can be reduced.

Clustering and non-clustering communications are most efficient routing methods in WSN. Clustering is one of the promising techniques for sensor networks because of its good scalability and the support for data aggregation. However, since all data are directed to sink, CHs around the sink comes to have high relay traffic. As a result, CH around sink uses much energy in a short time. Specially in heterogenous WSNs, since each high-end node acts as the CH for long duration, nodes around the CH use much energy in a short time to relay the traffic to CH. These non-uniform energy consumption problems around the sink and CHs are two of the main problems which shorten the network lifetime in WSNs.

Non-clustering routing methods are effective in small area WSN and in network with high node density. Non-clustering routing using a large number of nodes enables to balance the relay

traffic all over the WSN. When it is applied to wide area networks, however, the amount of data that has to relay for each node increases as it is near to sink. Furthermore, compared with clustering mode, data aggregation is not efficient in multi-hop routing.

Many routing protocols have been proposed for WSN, which consider energy consumption problem due to battery limitation and/or scalability problem due to large number of sensor nodes. Although most of them suppose homogeneous sensor nodes with clustering routing methods for wide area WSNs and non-clustering routing methods for small area WSNs. However, a few schemes for heterogeneous WSNs which introduce some high-end nodes and select them as the CHs in order to enhance the network performance is focused on recently. Due to lack of the research work focused on heterogeneous WSNs, non-uniform energy consumption problem aroud the CHs and sink are still remain without a efficient solution. In addition, in practical networks nodes are deployed in a non-uniform manner. Also, the residual power of each node decreases with time elapsing. Therefore, as the network time elapsed, the pre-defined routing methods should be reconfigured.

In order to prolong the network lifetime, this dissertation proposes an adaptive routing protocol for heterogeneous WSNs. First, we address the proposed clustering routing method and non-clustering routing method for heterogeneous WSNs. The proposed clustering method successfully avoids non-uniform energy consumption around CHs and extends the network lifetime in wide area WSNs by selecting the cluster heads considering not only transmission power and residual energy of each node but also those of its adjacent nodes. On the other hand, the proposed non-clustering method extends the network lifetime in small area heterogeneous WSNs by considering the residual energy of the nodes to select the optimal path for source to destination. Then, in order to avoid non-uniform energy consumption around the sink, we propose a hybrid routing method which adaptively combines clustering and non-clustering communication methods. In the proposed hybrid routing method, we also propose an analysis method to define the spatial boundary between clustering and non-clustering zones which can flexibly include the constraints on the physical network boundaries and location of the sink. Next, in order to respond to dynamic changes in the practical networks, we propose a network reconfigure method which dynamically re-defines the spatial boundary by considering the variance of the residual energy of a boundary node and its adjacent nodes.

This dissertation evaluates the performance of the proposed protocol through computer simulations, and then confirms its effectiveness.

# Acknowledgments

I would like to express my deepest appreciation to Professor Koso Murakami of the Department of Information Networking in Graduate School of Information Science and Technology, Osaka University for his supervision, continuing encouragement, valuable discussions, academic advice, and various supports throughout my studies and the preparation of this dissertation as a chief investigator.

I would like to express my deep gratitude to Professor Masayuki Murata, Professor Teruo Higashino and Professor Hirotaka Nakano of the Department of Information Networking in Graduate School of Information Science and Technology, Osaka University for serving as members on my dissertation committee.

I would like to express my sincere appreciation to Professor Hideki Tode of the Department of Computer Science and Intelligent Systems in Graduate School of Engineering, Osaka Prefecture University for his helpful comments, valuable discussions, and various assists.

I also would like to express my sincere appreciation Associate Professor Kazuhiko Kinoshita of the Department of Information Networking in Graduate School of Information Science and Technology, Osaka University for his helpful comments, valuable discussions, and also guided me through the academic process, never accepting less than my best efforts.

I must also acknowledge the kind assistance, comments and various assists of Associate Professor Yusuke Hirota of Department of Information Networking in Graduate School of Information Science and Technology, Osaka University and Associate Professor Yosuke Tanigawa of Department of Computer Sciences and Intelligent Systems Graduate School of Engineering, Osaka Prefecture University.

I would like to thank Professor Toshihiro Masaki of the Office for University-Industry Collaboration of Osaka University, Ms. Mariko Tsukuda of the Department of Information Networking in Graduate School of Information Science and Technology, Osaka University for their help and support.

Thanks also goes out to members of Ubiquitous Network team in Murakami Laboratory; Mr. Keisuke Imoto, Mr. Natsuki Inoue, Mr. Ryohei Hashimoto, Mr. Yoshikage Ochi, Ms. Junko Nagata, Mr. Takeshi Hatata, and Ms. Kana Uekubo.

I consider it an honor to work with former and current members of the Murakami Laboratory; Dr. Yusuke Shinohara, Ms. Kanako Uchida, Mr. Kazuki Hamada, Mr. Tatsuya Mori, Mr. Masayuki Morita, Mr. Tsuyoshi Yamada, Mr. Tran Nguyen Trung, Mr. Hiroki Okada, Mr. Wataru Miyazaki, Mr. Makoto Mochiki, Mr. Shinya Yatsuo, Mr. Shusuke Yamazaki, Mr. Yutaka Yamamoto, Mr. Masanori Yoshimoto, Mr. Kohei Ichimura, Mr. Yuichi Kanamori, Mr. Yusuke Hara, Mr. Naosuke Yokoe, Mr. Naotoshi Yoshihara, Mr. Takafumi Sugimoto, Mr. Masashi Nakagawa, Mr. Atsushi Narishige, Mr. Takayuki Hisada, Mr. Eisuke Hirota, Mr. Nguyen Thanh Cong, Mr. Yusuke Harada, Ms. Nima Baidar, Dr. Akihiro Fujimoto, Mr. Tomoki Motohashi, Mr. Naohiro Wakabayashi, Mr. Yuya Inoue, Mr. Ken Katsumoto, Mr. Kentaro Ohishi, Mr. Ryo Kawasumi, Mr. Shiori Kouno, Mr. Daishi Kondo, Mr. Shohei Fujii, Ms. Yukika Maruyama, and Mr. Takayuki Moriguchi.

It gives me great pleasure in acknowledging the support and help from all the members in Sato Yo International Scholarship Foundation (SISF). I would also like to thank ICOM Scholarship Foundation, Nitori Scholarship Foundation and Ministry of Education, Culture, Sports, Science and Technology (Japan) for Monbusho Scholarship.

I also wish to acknowledge all of my friends who have listened to me laugh and cry, and encourage me with friendly advice.

Last but not the least, I am particularly grateful to my parents for their understanding, encouragement and perpetual support.

# Contents


**List of Publications by the Author** **i**

**Abstract** **v**

**Acknowledgments** **ix**

**1 Introduction** **1**

**2 Heterogeneous Wireless Sensor Networks** **7**

2.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7

2.2 Related Work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10

2.2.1 Preceding Studies on Clustering Methods for Heterogeneous WSNs . . 10

2.2.2 Preceding Studies on Non-clustering Methods for Heterogeneous WSNs 11

2.3 Proposed Method . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14

2.3.1 Non-clustering Method . . . . . . . . . . . . . . . . . . . . . . . 14

2.3.2 Clustering Method . . . . . . . . . . . . . . . . . . . . . . . . . . 16

2.4 Performance Evaluation . . . . . . . . . . . . . . . . . . . . . . . . . . 22

2.4.1 Performance Evaluation of the Proposed non-Clustering Method . . . . 22

2.4.2 Performance Evaluation of the Proposed Clustering Method . . . . . . 25

2.5 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 33

**3 Hybrid Routing Method for Wireless Sensor Networks** **35**

3.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 35

3.2 Traditional Routing Methods . . . . . . . . . . . . . . . . . . . . . . . 37

3.2.1 Clustering Methods . . . . . . . . . . . . . . . . . . . . . . . . . . 37

3.2.2 Multi-hop Communication . . . . . . . . . . . . . . . . . . . . . . 38
3.3 Proposed Hybrid Routing Method . . . . . . . . . . . . . . . . . . . . . . 39
3.3.1 Basic Idea of Clustering and Non-clustering Zones . . . . . . . . . . . 39
3.3.2 Routing Methods in Clustering Zone . . . . . . . . . . . . . . . . . . 43
3.3.3 Routing Methods in Non-clustering Zone . . . . . . . . . . . . . . . . 45
3.3.4 Routing Methods Around the Spatial Boundary . . . . . . . . . . . . . 45
3.3.5 Network Formation . . . . . . . . . . . . . . . . . . . . . . . . . . . 45
3.4 Simplified Spatial Boundary Analysis Method for Small Area WSN . . . . . . 46
3.4.1 Network Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46
3.4.2 Mathematical Analysis . . . . . . . . . . . . . . . . . . . . . . . . . 47
3.4.3 Multi-hop Mode . . . . . . . . . . . . . . . . . . . . . . . . . . . . 48
3.4.4 Clustering Mode . . . . . . . . . . . . . . . . . . . . . . . . . . . . 49
3.5 Generalized Spatial Boundary Analysis Method for Wide Area WSN . . . . . . 50
3.5.1 Network Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 50
3.5.2 Definition of the Boundary Function . . . . . . . . . . . . . . . . . . 50
3.5.3 Hexagonal Element Based Analytical Model . . . . . . . . . . . . . . 52
3.5.4 Spatial Boundary for Homogeneous Networks . . . . . . . . . . . . . 55
3.5.5 Spatial Boundary for Heterogeneous Networks . . . . . . . . . . . . . 56
3.5.6 Extension for Non-computable Network Topologies . . . . . . . . . . 57
3.6 Performance Evaluation . . . . . . . . . . . . . . . . . . . . . . . . . . . 57
3.6.1 Simplified Spatial Boundary Analysis Method . . . . . . . . . . . . . 57
3.6.2 Generalized Spatial Boundary Analysis Method . . . . . . . . . . . . . 61
3.7 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 66

**4 Dynamic Reconfiguration of Spatial Boundary** **67**
4.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 67
4.2 Related Work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 67
4.3 Dynamic Reconfiguration of Spatial Boundary . . . . . . . . . . . . . . . . 69
4.3.1 Variance of the Residual Energy . . . . . . . . . . . . . . . . . . . . 70
4.3.2 Handing Over the Role of Boundary Node . . . . . . . . . . . . . . . 71

4.4 Performance Evaluation . . . . . . . . . . . . . . . . . . . . . . . . . . . 73
4.5 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 76

**5 Conclusions** **77**

**A Appendices** **87**

A.1 Proof of Lemma 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . 87
A.2 Proof of Lemma 2 . . . . . . . . . . . . . . . . . . . . . . . . . . . . 88

# Chapter 1

# Introduction

Wireless Sensor Networks (WSN) are networks composed of tiny, battery powered sensor nodes with limited on-board processing, storage and radio capabilities [1]. With recent miniaturization and reduction of energy consumption in wireless communication devices and development of MEMS (Micro Electro Mechanical Systems) technologies, WSN attracts much attention [2] [3].

WSNs are composed of a large number of sensor nodes, which are densely deployed either inside the sense perception (sensing area) or very close to sensing points. Sensor nodes sense and send their data toward a central processing center which is called "sink". WSN may have profound effect on the efficiency of many military and civil applications such as target field imaging, intrusion detection, weather monitoring, security and tactical surveillance, distributed computing, detecting ambient conditions such as temperature, movement, sound, light, or the presence of certain objects, inventory control, disaster management, and so on [7].

Realization of these sensor network applications requires wireless ad hoc networking techniques. Although many protocols and algorithms have been proposed for traditional wireless ad hoc networks, they are not well suited for the unique features and application requirements of sensor networks. Compared with ad hoc networks, the number of nodes in a sensor network can be several orders of higher in magnitude than that in an ad hoc network. Also, sensor nodes are densely deployed and prone to failures. The topology of a sensor network changes very frequently due to node failures and sensor nodes mainly use broadcast communication paradigm where most ad hoc networks are based on point-to-point communications. Prior to above all, sensor nodes are limited in power, computational capacities, and memory and may not have global identification (ID) because of the large amount of overhead and large number of sensors. Since a large number of sensor nodes are densely deployed, neighbor nodes may be very close to each other. Furthermore, the transmission power levels are very low, which is highly desired in covert operations. One of the most important constraints on sensor nodes is the low power consumption requirement. Sensor nodes carry limited, generally irreplaceable, power sources. Therefore, while traditional networks aim to achieve high quality of service (QoS), sensor network protocols must focus primarily on power conservation.

Hence, the design of protocols and applications for such networks has to be energy aware in order to prolong the lifetime of the network. Replacement of the embedded batteries is a very difficult process once these nodes have been deployed. One of the major causes for energy consumption in WSNs is the radio transmission energy, and in general, radio transmission energy is proportional to the 4th power of the transmission distance.

Therefore, many data gathering schemes which shorten the transmitting distance or reduce the number of times of transmission are proposed. For example, Directed Diffusion [40] shorten the transmitting distance by using multi-hop communication and LEACH [41] reduces the number of times of transmission to the base station by using cluster formation and data fusion. Further, HEED (Hybrid Energy-Efficient Distributed clustering) [23] which follows LEACH concept takes residual energy of each node into consideration in order to prolong the network lifetime.

Multi-hop and clustering communications are most efficient routing methods in WSN. Multi-hop routing methods are effective in small area WSN and in network with high node density. It can effectively overcome shadowing and path loss effects. Multi-hop routing using a large number of nodes enables to balance the relay traffic all over the WSN. When it is applied to wide area network, however, the amount of data that has to relay for each node increases as it is near to sink. Furthermore, compared with clustering mode, data aggregation is not efficient in multi-hop routing [42,43].

Clustering [9] is one of the promising techniques for sensor networks because of its good scalability and the support for data aggregation. The data aggregation combines data from multiple sensors to eliminate redundant information and transmissions. It also guarantees that the energy load is well balanced by Cluster Heads (CH) dynamically elected according to a prior optimal mechanism [12]. By rotating the role of CH among all nodes, each node tends to expend the same energy over time. Clustering can be extremely effective in one-to-many, many-to-one, or one-to-all (broadcast) communications. On the other hand, since all data are directed to sink, CHs around the sink comes to have high relay traffic. As a result, CH around sink uses much energy in a short time, and this is one of the main problems that shorten the network life time [21] in clustered WSN.

By the way, WSN can be classified into two types; homogeneous and heterogeneous sensor networks [13]. Many routing protocols have been proposed for WSN, which consider energy consumption problem due to battery limitation and/or scalability problem due to large number of sensor nodes. Although most of them supposes homogeneous sensor nodes, a few schemes which introduce some high-end nodes to enhance the network performance is focused on recently [19, 20].

In homogeneous networks, all the sensor nodes are identical in terms of battery capacity and hardware complexity. By means of clustering in a homogeneous network, it is evident that the CHs will be over-loaded with longer distance transmissions to the remote sink or next CH, and the extra processing is necessary for data aggregation and protocol co-ordination. As a result, CHs expire the battery before other nodes. However, it is desirable to ensure that all the nodes run out of their battery at about the same time, so that very little residual energy remains when the system expires. One way to ensure this is to rotate the role of a CH over all the nodes. However, the downside of using role rotation in homogeneous network is all the nodes should be capable of acting as CHs, and therefore should possess the necessary hardware capabilities [22].

On the other hand, in a heterogeneous sensor network, two or more different types of nodes with different battery capacity, communication capability or functionality are used. The motivation is that the more complex hardware and extra battery energy can be embedded in a few CHs, thereby the hardware cost of the rest of the network and communication cost of the sensing nodes can be reduced. However, since each high-end node acts as the CH for long duration, nodes around the CH use much energy in a short time to relay the traffic to CH.

In order to prolong the network lifetime, this dissertation proposes an adaptive routing protocol for heterogeneous WSNs. First, we propose a new clustering routing method and a non-clustering routing method for heterogeneous WSNs. The proposed clustering method successfully avoids non-uniform energy consumption problem around CHs and extends the network lifetime in wide area WSNs by selecting CHs considering not only transmission power and residual energy of each node but also those of its adjacent nodes. On the other hand, the proposed non-clustering method also extends the network lifetime in small area WSNs by considering the residual energy of the nodes to select the optimal path for source to destination.

Then, in order to avoid non-uniform energy consumption problem around the sink, we propose a hybrid routing method which adaptively combines clustering and non-clustering routing methods. For the proposed hybrid routing method, we also propose an analysis method to define the spatial boundary between clustering and non-clustering zones, which can flexibly include the constraints on the physical network boundaries and location of the sink. Next, to respond to dynamic changes in practical networks, we propose a network reconfiguration method which dynamically re-define the spatial boundary by considering the variance of the residual energy of a boundary node and its adjacent nodes.

This dissertation evaluates the performance of the proposed protocol through computer simulations, and then confirms its effectiveness.

The rest of this dissertation is organized as follows. Chapter 2 presents clustering and non-clustering routing methods for heterogeneous WSNs. First, we discuss some related works and introduce some existing clustering methods for heterogeneous WSNs and some existing non-clustering routing protocols. Then, we propose new clustering and non-clustering methods for WSNs. The proposed clustering method selects CHs considering not only the performance of a certain node but also that of its adjacent nodes. The proposed non-clustering method is an extension of DSR protocol and residual energy of nodes in the path are taken into consideration when selecting an optimal path between source and destination. In addition, the proposed clustering and non-clustering methods have no limitations on the number of node types that can use to crate a WSN. Through the simulation experiments, we verified that our proposed clustering and non-clustering methods increase network lifetime.

Chapter 3 presents the proposed robust clustering non-clustering hybrid routing method for heterogeneous WSN which adaptively combines clustering and multi-hop communication methods. Most of conventional routing methods for WSN are optimized for a specific application. In the proposed method, however, we consider the specifications of the nodes and the characteristics of the networked like outer boundary and location of the sink instead of characteristics of an application to optimize the WSN. Therefore, the proposed method can be easily applied to general-purpose WSN with multiple sensor applications. Also in chapter 3, we propose a spatial boundary analysis method. Through the simulation experiments, we verified that the proposed hybrid methods are able to maintain higher alive node ratio than that of the pure clustering and multi-hop methods in any network topologies.

Chapter 4 proposes a dynamic reconfiguration method of spatial boundary of clustering non-clustering hybrid routing method. In the proposed methods in Chapter 3, we assumed that nodes are uniformly distributed in the network. However, in practical networks nodes are deployed in a non-uniform manner. In addition, the residual power of each node decreases with time elapsing. Moreover, in this method sensor nodes near to the spatial boundary tends to have higher data relay traffic. Therefore, as the network time elapsed, the pre-defined spatial boundary should be reconfigured. In this chapter we propose to dynamically re-define the spatial boundary by considering the variance of the residual energy of a boundary node and its adjacent nodes. Simulation results show that our method increases network lifetime in comparison the existing method with static boundary.

Finally, Chapter 5 concludes this dissertation and describes the future works.

# Chapter 2

# Heterogeneous Wireless Sensor Networks

## 2.1 Introduction

In this chapter, we present clustering and non-clustering routing methods for heterogeneous WSNs. First, we discuss some related work and introduce some existing clustering methods for heterogeneous WSNs, and some existing non-clustering routing protocols for WSNs. Then, we propose clustering and non-clustering methods for WSNs. The proposed clustering method selects CHs considering not only the performance of a certain node but also that of its adjacent node. In the proposed non-clustering method, residual energy of nodes in the path are taken into consideration when selecting the optimal path between source and destination. Remarkably, the proposed clustering and non-clustering methods have no limitations on the number of node types that can be used to create a WSN. Through the simulation experiments, we verified that our proposed clustering and non-clustering methods increased network lifetime.

The design of protocols and applications for heterogeneous WSNs has to be energy aware in order to prolong the lifetime of the network, because the replacement of the embedded batteries is a very difficult process once these nodes have been deployed [8].

We consider clustered WSN [9]. That is one of the promising techniques for sensor networks because of its good scalability and the support for data aggregation [10]. A number of sensor nodes that are in pre-specified area create a virtual group called cluster and dynamically elect a single node as a CH from among the cluster. CHs aggregate data from their cluster members before forwarding them to the sink. The data aggregation combines data from multiple sensors to eliminate redundant information and transmissions [11]. It also guarantees that the energy load is well balanced by CH dynamically elected according to a prior optimal mechanism [12]. By rotating the role of CH among all nodes, each node tends to expend the same amount of energy over time.

Clustering can be extremely effective in one-to-many, many-to-one, or one-to-all (broadcast) communications. Clustered sensor networks can be classified into two types; homogeneous and heterogeneous sensor networks [13]. Many routing protocols have been proposed for WSN, which consider energy consumption problem due to battery limitation and/or scalability problem due to large number of sensor nodes. Although most of them (e.g. LEACH [14], FLOC [15], ACE [16]) supposes homogeneous sensor nodes, a few schemes (e.g. SEP [17], DEEC [18]) and applications [19, 20] which introduce some high-end nodes to enhance the network performance is focused recently.

In homogeneous networks, all the sensor nodes are identical in terms of battery energy and hardware complexity. By means of clustering in a homogeneous network, it is evident that the CH nodes will be over-loaded with longer range transmissions [21] to the remote sink or next CH, and the extra processing is necessary for data aggregation and protocol co-ordination. As a result, the CH nodes expire battery earlier than other nodes [22].

However, it is desirable to ensure that all the nodes run out of their battery at about the same time, so that very little residual energy is wasted when the system expires. One way to ensure this is to rotate the role of a CH among over all the nodes as proposed in HEED. HEED introduces two parameters to select CHs. First one is the residual energy of each node and the other is intra-cluster communication cost which depends on the number of its neighboring nodes. It achieves lower messaging overhead and balancing the load of CH all over the network. However, the downside of using a homogeneous network and role rotation is that all the nodes should be capable of acting as CHs, and therefore should possess the necessary hardware capabilities.

On the other hand, in a heterogeneous sensor network, two or more different types of nodes with different battery energy, communication capability and functionality are used. High-end nodes which are equipped with more complex hardware, long-range communication capability and extra battery energy can be used as CHs and other low-end nodes can be used as sensing nodes. The motivation is that using high-end CHs, the number of low-end nodes can be reduced to minimum which satisfies the coverage and connectivity requirements. Also, communication cost of the low-end sensing nodes can be reduced. Therefore, compared with homogeneous networks, heterogeneous networks are able to reduce the initial cost of the network for given lifetime requirement or to prolong the network lifetime for given initial cost.

CC (Chessboard Clustering) [24] is a routing protocol for heterogeneous sensor networks consisting of a few powerful high-end sensors in addition to a large number of low-end sensors. In CC, low-end nodes never play the role of CH and physically more powerful high-end nodes become the CHs, while other protocols like HEED, LEACH [14], LRS [26] need an algorithm to select CHs. However, fixing the CH nodes means that the role rotation is no longer possible. Since sensor nodes inside the cluster use multi-hop communication to reach the CH, the nodes around the CH have to relay larger amount of traffic. Therefore, those nodes have highest energy consumption in the cluster and cause rapid exhaust of the batteries. Furthermore, CHs around the sink have the same problem when CHs use multi-hop communication to reach the sink. Therefore, those nodes are called as critical nodes. Also, CHs around the sink are also critical nodes due to high relay traffic around the sink.

Heterogeneous networks lower the hardware cost and reduce the communication cost of the sensing nodes. On the other hand, homogeneous networks achieve uniform energy drainage. However, both features cannot be incorporated in the same network. The objective of this chapter is to design a network architecture which able to maintain above two characteristics in heterogeneous networks.

The rest of the chapter is organized as follows: We first discuss the existing clustering communication methods in section 2.2. In section 2.3, we present our proposed clustering method for WSN with heterogeneous node types, while in section 2.4 we present simulation results on performance evaluation. Finally, in section 2.5, we make some conclusions.

## 2.2 Related Work

### 2.2.1 Preceding Studies on Clustering Methods for Heterogeneous WSNs

#### HEED

HEED is one of the protocols which follow a clustering concept and has four primary goals:

- prolonging network lifetime by distributing energy consumption
- terminating the clustering process within a constant number of iterations/steps
- minimizing control overhead (to be linear in the number of nodes)
- producing well-distributed CHs.

In order to achieve these goals, HEED considers a hybrid of energy and communication cost.

HEED uses two parameters, residual energy and intra-cluster communication cost, in order to form good cluster. Residual energy is used to probabilistically select an initial set of CHs, and communication cost is used to determine which CH the node should belong to when the node has two or more possible CHs.

However, HEED has no consideration to select the optimal sensor nodes for CH in heterogeneous environment. In typical heterogeneous sensor networks, a high-end node tends to be elected as a CH with higher probability, so that nodes around a CH have to relay much heavier traffic and run out of power much earlier than other nodes.

#### CC

CC is a protocol for heterogeneous sensor networks consisting of a few powerful high-end sensors in addition to a large number of low-end sensors. In CC, the sensor network is divided into several small equal-sized cells, and adjacent cells are colored with different colors, as illustrated in Fig. 2.1 (only high-end nodes are shown in Fig. 2.1).

By using given location information, a high-end sensor can determine whether it is in a white cell or a gray cell. During the initialization phase, only the high-end sensors in white cells are active, and the high-end sensors in gray cells turn themselves off. All the low-end

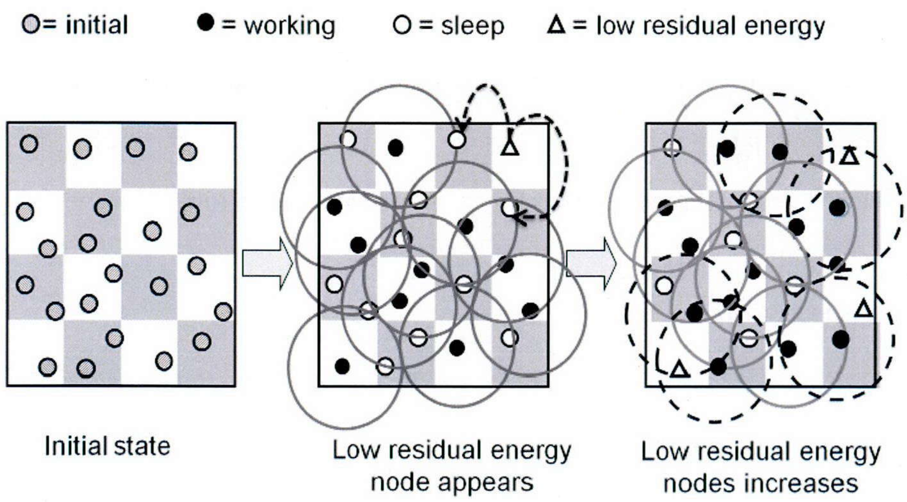



Figure 2.1: Chessboard Clustering

sensors are active. Clusters are formed around the high-end sensors in white cells, and these high-end sensors become CHs. When the high-end sensors in white cells run out of energy later, the high-end sensors in gray cells wake up and form a different set of clusters in the network. Because of the formation of a different set of clusters, previous critical sensors become non-critical sensors, and previous non-critical sensors become critical sensors.

However, CC lacks flexibility. Specifically, it uses just 2 types of nodes and only high-end nodes can become a CH. Power exhaustion of high-end nodes greatly decreases the network connectivity.

### 2.2.2 Preceding Studies on Non-clustering Methods for Heterogeneous WSNs

**Ad hoc On-Demand Distance Vector (AODV)**

There are two types of routing protocols which are reactive and proactive. In reactive routing protocols the routes are created only when source wants to send data to destination whereas proactive routing protocols are table driven. Being a reactive routing protocol AODV uses traditional routing tables, one entry per destination and sequence numbers are used to determine whether routing information is up-to-date and to prevent routing loops.

The maintenance of time-based states is an important feature of AODV which means that a routing entry which is not recently used is expired. The neighbors are notified in case of route breakage. The discovery of the route from source to destination is based on query and reply cycles and intermediate nodes store the route information in the form of route table entries along the route. Control messages used for the discovery and failures of route are as follows:

- Route Request Message (RREQ)
- Route Reply Message (RREP)
- Route Error Message (RERR)
- HELLO Messages.

A route request packet is flooded through the network when a route is not available for the destination from source. The parameters are contained in the route request packet are as follows:

| Source Address | Request ID | Source Sequence Number | Destination Address | Destination Sequence Number | Hop Count |
|---|---|---|---|---|---|

A RREQ is identified by the pair source address and request ID, each time when the source node sends a new RREQ and the request ID is incremented. After receiving of request message, each node checks the request ID and source address pair. The new RREQ is discarded if there is already RREQ packet with same pair of parameters. A node that has no route entry for the destination, it rebroadcasts the RREQ with incremented hop count parameter. A route reply (RREP) message is generated and sent back to source if a node has route with sequence number greater than or equal to that of RREQ.

On having a valid route to the destination or if the node is destination, a RREP message is sent to the source by the node. When a route that is active is lost, the neighborhood nodes are notified by route error message (RERR) on both sides of link. The parameters are contained in the route reply message are as follows:

| Source Address | Destination Address | Source Sequence Number | Hop Count | Life Time |
|---|---|---|---|---|

The HELLO messages are broadcasted in order to know neighborhood nodes. The neighborhood nodes are directly communicated. In AODV, HELLO messages are broadcasted in order to inform the neighbors about the activation of the link. These messages are not broadcasted because of short time to live (TTL) with a value equal to one.

When a source node does not have routing information about destination, the process of the discovery of the route starts for a node with which source wants to communicate. The process is initiated by broadcasting of RREQ. On receiving RREP message, the route is established. If multiple RREP messages with different routes are received then routing information is updated with RREP message of greater sequence number. The reverse path to the node is noted by each node during the transmission of RREQ messages. The RREP message travels along this path after the destination node is found. The addresses of the neighbors from which the RREQ packets are received are recorded by each node. The reverse path is used to send RREP message back to the source but a forward path is setup during transmission of RREP message. This forward path can be called as reverse to the reverse path. The data transmission is started as soon as this forward path is setup. The locally buffered data packets waiting for transmission are transmitted in FIFO-queue.

### Dynamic Source Routing (DSR)

Dynamic Source Routing (DSR) protocol is specifically designed for multi-hop ad hoc networks. DSR allows the network to be completely self-organizing and self-configuring, without the need for any existing network infrastructure or administration. The difference in DSR and other routing protocols is that it uses source routing supplied by packet's originator to determine packet's path through the network instead of independent hop-by-hop routing decisions made by each node. The packet in source routing which is going to be routed through the network carries the complete ordered list of nodes in its header through which the packet will pass. Fresh routing information is not needed to be maintained in intermediate nodes in design of source routing, since all the routing decisions are contained in the packet by themselves. The protocol is composed of the two main mechanisms of "Route Discovery" and "Route Maintenance", which work together to allow nodes to discover and maintain routes to arbitrary destinations in the ad hoc network.

In route discovery mechanism, when a node $S$ wants to send a packet to destination $D$, the route to destination $D$ is obtained by route discovery mechanism. In this mechanism the source node $S$ broadcasts a ROUTE REQUEST packet which in a controlled manner is flooded through the network and answered in the form of ROUTE REPLY packet by the destination node or from the node which has the route to destination. The routes are kept in Route Cache, which to the same destination can store multiple routes. The nodes check their route cache for a route that could answer the request before re-propagation of ROUTE REQUEST. The routes that are not currently used for communication the nodes do not expend effort on obtaining or maintaining them i.e. the route discovery is initiated only on-demand.

Route maintenance mechanism starts when source node $S$ detects if the topology of the network has changed so that it can no longer use its route to destination. If the two nodes that were listed as neighbors on the route moved out of the range of each other and the link becomes broken, the source node $S$ is notified with a ROUTE ERROR packet. The source node $S$ can use any other known routes to the destination $D$ or the process of route discovery is invoked again to find a new route to the destination.

## 2.3 Proposed Method

In this section, we propose a clustering method and a non-clustering method for WSNs with heterogeneous node types. The proposed clustering method selects CHs considering not only the performance of a certain node but also that of its adjacent nodes. In addition, it has no limitations on the number of node types.

### 2.3.1 Non-clustering Method

In DSR, first, a source node sends a RREQ packet to the destination node. The route with earliest reached RREQ to destination is used for the route from the source node to the destination node. Also, relay nodes on the route add their node ID to RREQ header and re-broadcast the RREQ packet. In the proposed method, a node waits time $T$ before it re-broadcasts the RREQ packet. $T$ is decided based on the node's residual energy, communication range, and the number of paths that have been established through the node (hereinafter referred to as the number of active routes). By controlling the waiting time of the RREQ re-broadcasting from the relay node, RREQs via the nodes with large number of active routes and low residual power are delayed to reach the destination node. In other words, it is able to avoid the situation where high-end nodes with large number of active routes or nodes with low residual energy tend to be used as a relay node extremely.

### Number of Active Routes

Every node maintains a source node address list of routing packets (hereinafter referred to as source list). RREP contains the information of the nodes in the route and source node. Therefore, each node is able to calculate the number of the active routes without any additional information. Detailed calculation method of the active routes in arbitrary node $A$ is as follows.

- If the address of the source node is not in the source list of $A$,
  - If node $A$ is in the route, then number of active routes does not change.
  - If node $A$ is not in the route, then decrements the number of active routes.
- If address of the source node is in the source list of $A$,
  - If node $A$ is in the route, then increment the number of active routes.
  - If node $A$ is not in the route, then number of active routes does not change.

### Waiting Time to Re-broadcast

Time $T$ shown in Fig.2.1 is decided based on node's residual energy, communication range, and active roots. We define two time variables $t_1$ and $t_2$ as shown in Eqs. (2.2) and (2.3), respectively. Time $t_1$ is inversely proportional to the communication range of the node. In other words, a node with wider communication range has smaller $t_1$ and a node with a smaller communication range has larger $t_1$. Then, time $t_2$ is set to a small value for the nodes with higher residual energy and less number of active routes. Finally, $T = t_1 + t_2$ is set as a waiting time for a node to re-broadcast a RREQ packet.

$$T = t_1 + t_2 \tag{2.1}$$

$$t_1 = \begin{cases} \alpha & \text{Low-end nodes} \\ 0 & \text{High-end nodes} \end{cases} \tag{2.2}$$

$$t_2 = \frac{\text{Number of active routes}}{\text{Residual Energy}} \times (\text{Random value: 0-1.0}) \times \beta \tag{2.3}$$

### 2.3.2 Clustering Method

Assume that each sensor is aware of its own location. In the proposed method, location information is used to identify adjacent nodes for routing within cluster and inter-cluster routing mechanisms. Therefore, significant location errors are fatal to the proposed method. We assume, however, recent researches for localization on WSN achieve to provide enough accurate information [29]. Sensor nodes can use location services such as [28–31] to estimate their locations, and a Global Positioning System (GPS) receiver is not required for each node. Location awareness is a basic requirement for many sensor networks, since in many cases the sensing data are only meaningful when the location of generating the data is known. And also, global time synchronization is essential in the proposed method.

A sensor node expends more energy in data communication than in sensing or data processing [32], [33]. For example, the energy spent by a mote sensor for transmitting 1-bit data over 20m is equivalent to that of running 1000 CPU instructions [34]. In general, the minimum output power required to transmit a signal over a distance $r$ is proportional to $r^n$, where $2 \leq n < 4$. The exponent $n$ is closer to 4 for low-lying antennae and near-ground channels, as is typical in sensor network communication.

Now, assume that $N$ nodes are uniformly distributed in area $S$. Then, we have the node density $D$ by $D = \frac{N}{S}$. When a node which has communication range $r$ becomes CH, it has $k$ nodes given by $k = \pi \times r^2 \times D = \frac{\pi r^2 N}{S}$. Under the assumption where a node uses its all energy for communication, a node with residual energy $p$ can communicate $t$ rounds until its power exhaustion, where $t = \frac{\alpha p}{r^4}$. Consequently, CH node $i$ can send total amount of data $g_i$ which is given by the following equation;

$$\begin{aligned} g_i &= \beta \times t \times k \\ &= \frac{\alpha \beta N \pi}{S} \times \frac{p}{r^2}, \end{aligned} \tag{2.4}$$

where $\alpha$ and $\beta$ are constants.

Next, we define $G$ as the highest value of $g$ in the WSN, namely $G = max[g_i], i \in [1, N]$. We assume that the sink has detail information of all nodes in the network. Therefore it can find the highest value of $g$ and broadcast it to the network. Finally, we normalize $g_i$ by $G$ and use the result $v_i$ given by Eq. (2.5) as a strength index of the node $i$. This operation is conducted by each node individually.

$$v_i = \frac{g_i}{G} \tag{2.5}$$

Then nodes exchange the $v$ value with their neighbor nodes. Using this value, we can relatively compare the nodes in WSN. By considering $v_i$, all nodes are classified into $C$ classes as $1, 2, 3...C$. In the proposed method, we define two types of classes to each node; individual class and adjacent class. To calculate the Individual class we use only local strength index of the node. Therefore we can consider that it represents the strength of node itself. For adjacent class we use the strength index of adjacent nodes and we can consider that it represents the strength of the adjacent nodes. Those calculation methods are given below.

Node $i$ itself belongs to individual class $I_i$ (individual class value), given by Eq. (2.6).

$$I_i = \lfloor C(1 - v_i) \rfloor \tag{2.6}$$

Example of calculating the individual class value when $C = 4$ is shown in Fig. 2.2. According to Fig. 2.2, node which has $v_i$ value between 0.5 and 0.75 has individual class value 2 ($I_i = 2$).

Next, we calculate the average strength index of the adjacent nodes by $\frac{\sum v_j}{u}, j \in$ {Nodes 1 hop away from node $i$}, where $u$ is the total number of nodes which locate 1 hop away from node $i$. If two nodes are able to establish two-way communication without any relay nodes, we

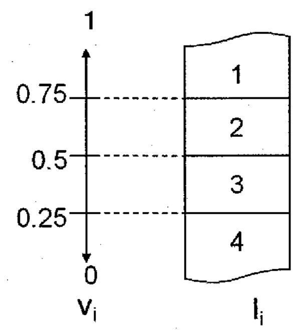



Figure 2.2: Strength index and Individual class

consider they are in one hop away from each-other. Then, node $i$ belongs to adjacent class $A_i$ (adjacent class value), given by Eq. (2.7).

$$A_i = \lfloor C(1 - \frac{\sum v_j}{u}) \rfloor \tag{2.7}$$

Based on these class values, each node becomes the CH and makes a cluster if pre-defined time $T_{wait}$ in Eq. (2.8) passed without receiving clustering messages from any other nodes. We use $\tau$ as a weight in Eq. (2.8) to calculate the waiting time for each node to select or become a CH. $\tau$ is a constant defined as waiting time for a class.

$$T_{wait} = (I_i + \frac{A_i}{u})\tau \tag{2.8}$$

All nodes in a cluster send their data toward its CH. However, adjacent nodes around the CH have to relay all those data to CH. Therefore, we prioritize a node with a higher number of adjacent nodes by dividing the adjacent class value from number of the adjacent nodes, since higher number of adjacent nodes decreases the relay traffic per an adjacent node.

Example of selecting the CHs is shown in Fig. 2.3. In this example we assume that node $k$ which $g_k = G$ has residual energy and communication range 50 $[J]$ and 16 $[m]$, respectively. Each node $i$ has residual energy ($P_i$), communication range ($r_i$) and strength index ($v_i$) as shown in Fig. 2.3. We focus on nodes 7 and 13 to explain the proposed clustering mechanism.

At first, sink broadcasts a message with location of the sink and network start time to network. Then, all the nodes enter to the initial state. Then, each node calculates the strength index according to Eq. (2.5). Applying the result to Eq. (2.6), individual class values of nodes 7 and 13 become $I_7 = 1$ and $I_{13} = 1$. Then, nodes calculate the $\frac{\sum v_j}{u}$; for nodes 7 and 13 they become

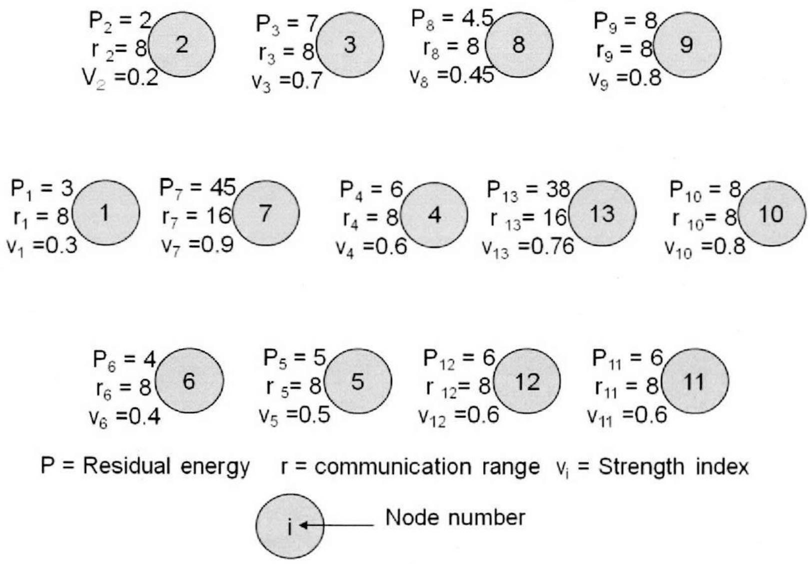



Figure 2.3: Adjacent class value

0.36 and 0.513, respectively. Therefore adjacent class values for nodes 7 and 13 obtained by Eq. (2.7) are $A_7 = 3$ and $A_{13} = 2$. Finally, nodes calculate $T_{wait}$ applying own $I_i$, $A_i$ and $u$ values to Eq. (2.8). Considering above values, $T_{wait}$ for node 13, $1.33\tau$ is smaller than that of node 7, $1.5\tau$.

Clustering schemes which only use residual energy as a parameter might choose node 7 which has weak adjacent nodes as a CH due to its high residual energy. However, the proposed clustering method sets a high priority for node 13 which has strong adjacent nodes to be a CH.

In Eq. (2.8) we consider both node's strength and adjacent nodes' one to determine the waiting time. High-end nodes have lower value of $I_i$ and when the number of adjacent nodes increases, the value of $A_i$ has less effect to $T_{wait}$. On the contrary, a high-end node surrounded by high-end nodes tends to have smaller value of $T_{wait}$ and becomes CH with higher probability. An example in Fig. 2.4 shows that nodes 6 and 15 has the same strength index and lays on the same individual class; $I_6 = I_{15} = 1$. We can show that adjacent class value of nodes 6 and 15 are equal to 2 and 3, respectively by using Eq. (2.7). As a result of calculating the $T_{wait}$ for both nodes, node 6 with higher number of adjacent nodes has smaller $T_{wait}$ and high priority to become CH.

On the other hand, when a node receives messages from other CHs during the waiting time, it keeps them into CH backup list. After waiting $T_{wait}$, it checks the CH backup list for

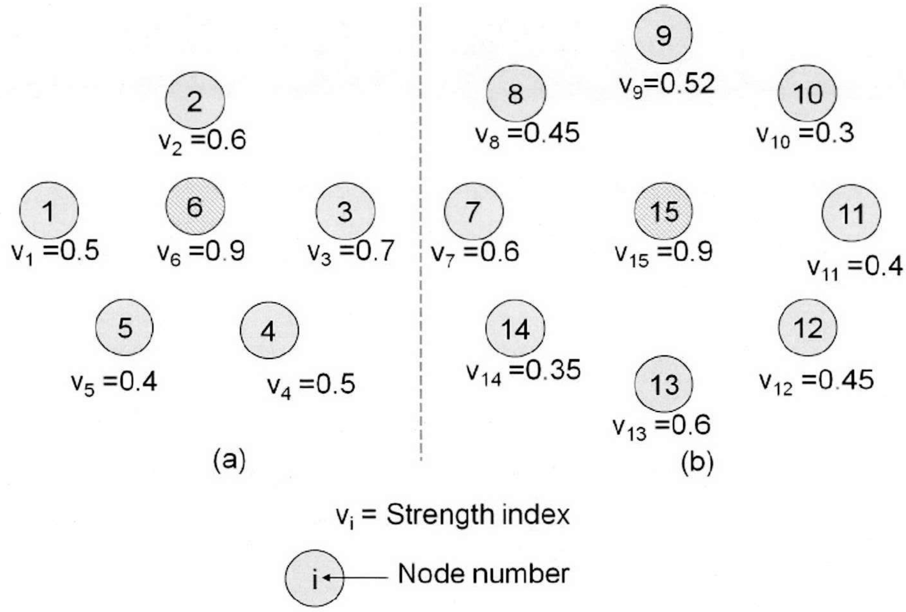



Figure 2.4: Effect of adjacent nodes

accessible CHs and becomes a member of the optimum cluster (i.e. nearest CH to the node), if there are any entries. If there are no accessible CHs in the list, it broadcasts a searching message for accessible CHs. After a certain period, if it cannot find any accessible CHs yet, it becomes CH.

Every node in the WSN holds the strength index of its adjacent nodes, $v$. Therefore, each node can assume the $g$ values of its adjacent nodes by using $v$ and its own $g$ value. From these values, CH re-assumes the $I_{cv}$ and $A_{cv}$ per round, and when at least one of these values drops to lower class, the CH starts to reconstruct the cluster.

### Routing within a Cluster

When a sensor node generates data, it sends packets to its CH. The packets are forwarded to the neighbor node that has the shortest distance to the CH, and the next node relays toward the CH in the same manner. For example, as shown in Fig. 2.5, source node $S$ sends the data packets to destination node $D$ via node $N_i$. Since the nodes are aware of its own location, $z = \overrightarrow{SN_i} \cdot \frac{\overrightarrow{SD}}{|SD|}$ is calculated for each node among adjacent nodes and the node with maximum $z$ is selected as the next hop to destination $D$.

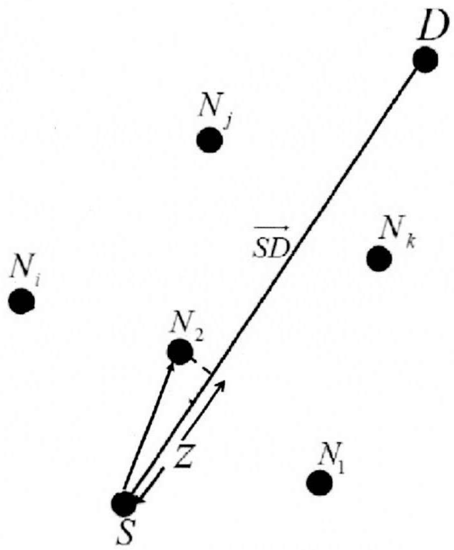



Figure 2.5: Routing within a Cluster

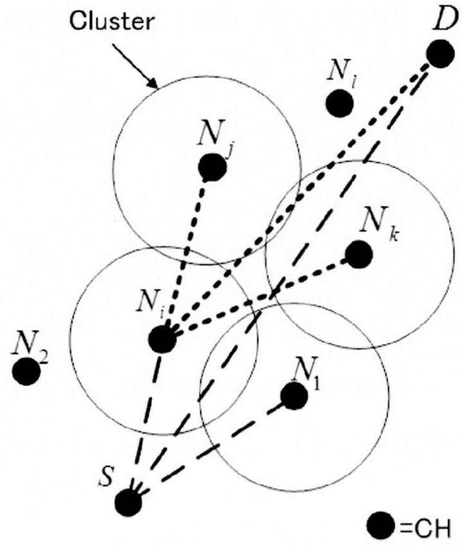



Figure 2.6: Inter-Cluster Routing

**Inter-cluster Routing**

CH has two main tasks; i.e. sending the collected data to sink and relaying the data from other CHs. CHs with lower energy burdened with higher relay traffic could cause early expiration of battery and shorten the network lifetime. We can simply solve this problem by selecting the node with highest residual energy for next hop. After initialization, each CH exchanges location information with neighbor CHs.

All the CHs advertise to adjacent CHs whether they are reachable to the sink or not when they broadcasts the control messages. Once CH has received the control messages from other CHs, it can determine that which adjacent CHs are connected with sink and which are not. For example, in Fig. 2.6, $S$ is source CH and $D$ is destination CH or sink. When a CH wants to send data packets to the sink, it virtually draws a straight line $SD$ between itself and the sink. Then, it searches for neighbor clusters which intersect with line $SD$ ($N_1$ and $N_i$ in Fig. 2.6). Among them, the node with the highest residual energy is finally selected as a next hop to sink.

When line SD does not have any intersections with neighbor clusters, the proposed method selects the CH with highest residual energy from among the adjacent CHs which is reachable to sink as its next hop. Then next hop repeats the same task while it is able to connect with sink.

## 2.4 Performance Evaluation

### 2.4.1 Performance Evaluation of the Proposed non-Clustering Method

Simulation topology was shown in Fig. 2.7(a). Sink was located in the middle of the $100[m] \times 100[m]$ area. Also, 4 high-end nodes and 8 low end nodes are deployed in 1 hop away from the sink. Location of those nodes were shown in Fig. 2.7(b). We deliberately deployed those nodes in order to distribute the relay traffic around the sink.

Simulation parameters were as follows.

- Simulation time: $35000[sec]$
- Number of nodes: 15 high-end nodes, 84 low-end nods, 1 sink
- Initial Energy: $5[mAhr]$ for high-end nodes, $4[mAhr]$ for low-end nodes
- Transmission power: $5[dBm]$ for high-end nodes, -$10[dBm]$ for low-end nodes
- Route cache refresh period: $20[h]$
- Packet generate interval: $30[min]$

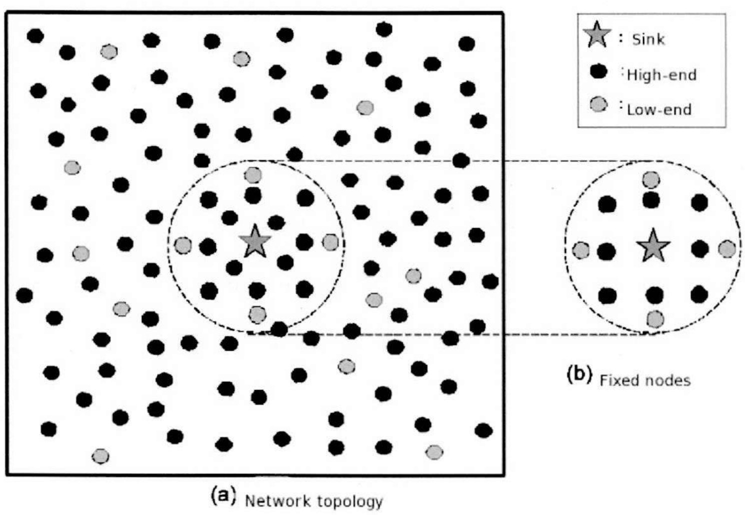



Figure 2.7: Network topology

We evaluated the performance of the proposed method via QualNet simulator and observed the percentage of alive nodes as a performance measure. Definition of a alive node is a node which is able to send the sensed data to the sink successfully. Therefore, disconnected nodes are not considered as alive nodes.

Sink sets a time limit (TTR: Time To Receive) for each source node to send their data to sink. When the sink receives a data from source node, it renovate the TTR for that source node. Calculation method of TTR is shown in Eq. 2.9.

$$\begin{aligned} TTR &= \text{Time to receive data + wait delay time} \\ &\geq \text{Time to receive data + Route cache refresh period} \end{aligned} \quad (2.9)$$

When the route cache refresh period is larger than wait delay time, TTR is not satisfy the Eq. 2.9.

We conducted some simulation experiments to obtain the optimal value for $t_1$ and $t_2$. First, we set the $t_1$ as shown in Eq. (2.10) and changed the $\beta$ value of $t_2$ shown in Eq. (2.3). Simulation results are shown in Fig. 2.8.

$$t_1 = \begin{cases} 0.001(s) \text{ (Low-end nodes)} \\ 0(s) \text{ (High-end nodes)} \end{cases} \quad (2.10)$$

From the results shown in Fig. 2.8, $\beta = 0.01$ has maximum alive node ratio for same $t_1$ value.

Then, we set the $\beta = 0.01$ and changed the $t_1$ for the low-end nodes. Simulation results are shown in Fig. 2.9.

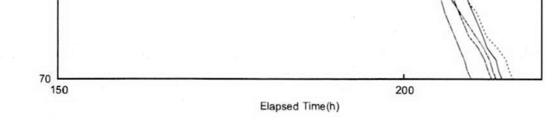



Figure 2.8: Variation of $\beta$

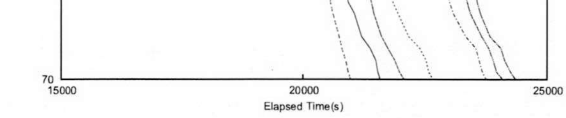



Figure 2.9: Variation of $t_1$

From the results shown in Fig. 2.9, $t_1 = 0.1[s]$ has maximum alive node ratio.

Therefore, from here forward we use the settings shown in Eq. (2.11) as to decide the waiting time for a node.

$$T = \begin{Bmatrix} \alpha \text{ (Low-end nodes)} \\ 0 \text{ (High-end nodes)} \end{Bmatrix} + \frac{\text{Number of active routes}}{\text{Residual Energy}} \times \text{(Random Value: 0-1.0)} \times 0.01 \tag{2.11}$$

In Eq. (2.3), we used a random value between 0 to 1 in order to avoid the prejudices of the wait time $T$.

We re-arranged the $t_2$ in Eq. (2.3) as shown in Eq. (2.4.1) and conducted some simulation experiments to observe the effect of the random value.

- $t_2 = \frac{\text{Number of active routes}}{\text{Residual energy}} \times 0.5 \times \beta$
- $t_2 = \frac{\text{Number of active routes}}{\text{Residual energy}} \times \text{Random value (0.25-0.75)} \times \beta$

Simulation results are shown in Fig. 2.10.

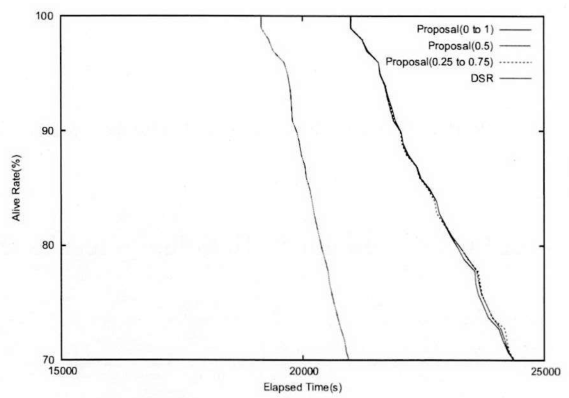



Figure 2.10: Effect of random value

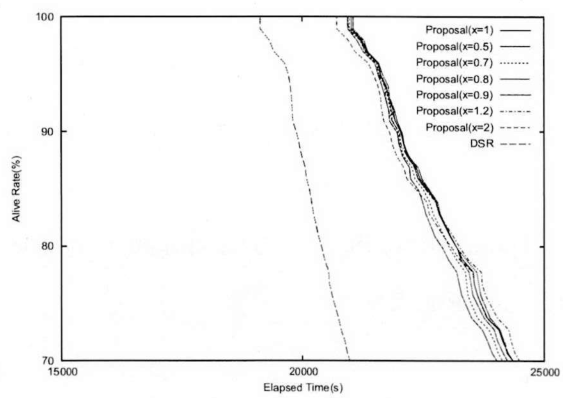



Figure 2.11: Variation of $x$

From the results in Fig. 2.10, the proposed method did not show any significant effect when random value set to 0.5 and 0.25 to 0.75. Therefore, we realized that the proposed method had no effect even when we set a random value in order to preferentially select a nodes with higher residual energy and low number of active routes.

To find out the optimal value for $T$, we conducted some simulation experiments with changing the value of $T$ by $xT$. We changed the $x$ and observed the effectiveness. Simulation results are shown is Fig. 2.11.

Result in Fig. 2.11 show that it has no effect when we increase the value of $T$ by 20%. Therefore, we conclude the values in Eqs. (2.1), (2.2), and (2.3) are optimal.

Those results show that the proposed non-clustering method is able to increase the network lifetime in heterogeneous environment.

### 2.4.2 Performance Evaluation of the Proposed Clustering Method

We evaluated the performance of the proposed clustering method with different $C$ values through simulation experiments and determined optimal $C$ value for the network. After that we conducted further evaluations with the optimized $C$ value and compared it with two other clustering methods; HEED and CC in heterogeneous and homogeneous environments. We have reasonable issues to select HEED and CC as for comparison methods. HEED is dynamically select CHs regardless of node types. It outperforms most of the other clustering methods for homogeneous environment and also work well with heterogeneous environments. CC is specially designed for heterogeneous environment and it uses per-defined CHs.

We used ns2 simulator. To set communication radius, we adjusted the receiving threshold value, RXThresh. This can be obtained by running the program ns/indeputils/propagation/threshold.cc and specifying the propagation model and the desired radius. The parameters used the obtain the RXThresh is shown in Table 2.1. Please refer the ns2 manual for how to set the parameters. We used the same method to set the communication range for nodes in all the simulation experiments in this paper.

High-end nodes are equipped with longer communication range and higher initial battery power. Compare to the high-end nodes, low-end nodes are equipped with shorter communication range and lower initial battery power. For the both high-end and low-end nodes, Initial energy, communication range and number of the nodes are set in each simulation experiment. Basic specification of the nodes are shown in Table 2.2.

Every sensor node generate the data in iteration of 20 minutes. However, every node have a random back off time of maximum 1 minutes before process the sensed data and prepare them for transmission. We used S-MAC protocol as MAC protocol. S-MAC codes are provided by ns2 simulator.

Table 2.1: ns2 settings

| | |
|---|---|
| Radio-propagation model | TwoRayGround |
| Antenna model | OmniAntenna |
| Transmit power | 0.281838 |
| Frequency | 9.14e+08 |
| Transmit antenna gain | 1 |
| Receive antenna gain | 1 |
| System loss | 1 |
| Transmit antenna height | 1.5 |
| Receive antenna height | 1.5 |

Table 2.2: Node specifications for simulations

| | H-node | L-node |
|---|---|---|
| Aggregation ratio | 0.25 | 0.25 |
| Data rate [kbps] | 250 | 250 |
| Calculation cost [J/byte] | $2.73 \times 10^{-6}$ | $2.73 \times 10^{-6}$ |

Due to the large number of nodes in the topology, each simulation experiment takes a long time. Therefore, these simulation results show an average of 50 simulation samples. We assumed lack of simulation samples caused the non smooth graphs. To reveal the cause, we conducted 500 simulation experiments for selected single experiment in each chapter.

Constants $\alpha, \beta$ were equals to 14000 and 1, respectively. The value for $\alpha$ could be obtained by calculations [38] for two-ray ground reflection model. The value for $\beta$ is set to unit value 1 for this simulation. However, as shown in Eq. (2.5), we relatively compare the strength of the nodes. Therefore, value of $\alpha$ and $\beta$ only use in primary stage calculations and has no active effect on CH selection mechanism.

In these experiments, we considered the largest energy consumption factor of node is the transceiver energy consumption. Therefore, we exclude the idle state energy consumption from the simulation experiments. In all simulation experiments, we use $\tau = 15[ms]$.

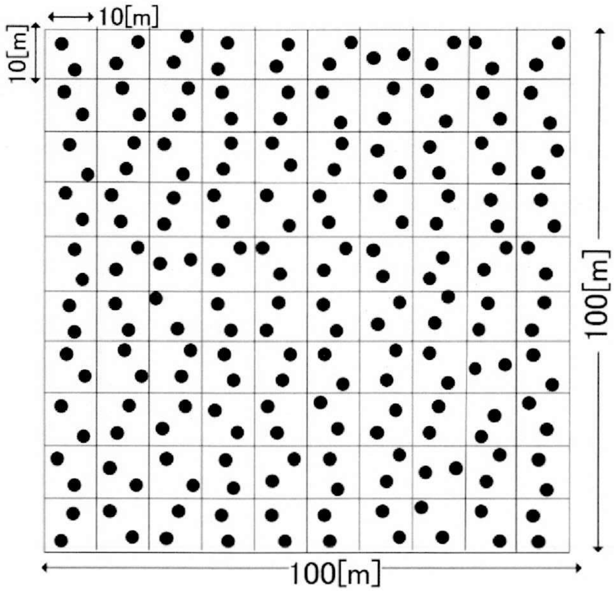



Figure 2.12: Logical division of the coverage area

We observed the percentage of alive nodes and coverage over time.

It is difficult to evaluate the non-uniform energy consumption patterns only by alive node ratio. When non-uniform energy consumption occurs, coverage area dramatically drops with decrements of alive node ratio. On the other hand, when the network is able to maintain higher coverage area in spite of lower alive node ratio, it means to achieve geographically distributed energy consumption pattern which is suitable for WSN. To evaluate the coverage, we logically divided the total area into large number of tiny areas as shown in Fig. 2.12. In each area, when at least one sensor node was alive, we considered it as a covered area. Coverage was defined as the ratio of covered area.

In the first experiment, we used homogeneous network environment and we used single type of node in which initial energy, communication range and amount of nodes were set to $10[J]$, $8[m]$, 200 nodes, respectively. In this experiment, we evaluated the performance of our clustering method with different $C$ values. Simulation results in Fig. 2.13 show alive node ratio as a function of elapsed time in homogeneous environment.

Larger number of classes can rotate the role of CH among all nodes efficiently. However, overhead of reconfiguration increases. Simulation results in Fig. 2.13 show that it has a difference less than 10% between $C = 3$ and $C = 4$ in the homogeneous environment. But $C = 5$

has less alive node ratio than $C = 4$ due to increment of reconfiguration overhead.

In the second experiment, we used heterogeneous network environment and we used 2 type of nodes in which initial energy, communication range and amount of nodes were set to $50[J]$, $16[m]$, 60 high-end nodes and $5[J]$, $8[m]$, 140 low-end nodes, respectively. We set the high-end and low-end node configurations using datasheets of typical products; XE1205a [36] and TR1000 [37], respectively. We evaluated the performance of our clustering method with different $C$ value and simulation results in Fig. 2.14 show alive node ratio as a function of elapsed time in heterogeneous environment.

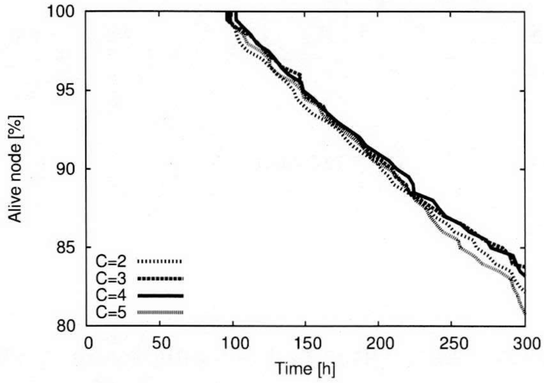



Figure 2.13: Alive node ratio with different C values in experiment 1

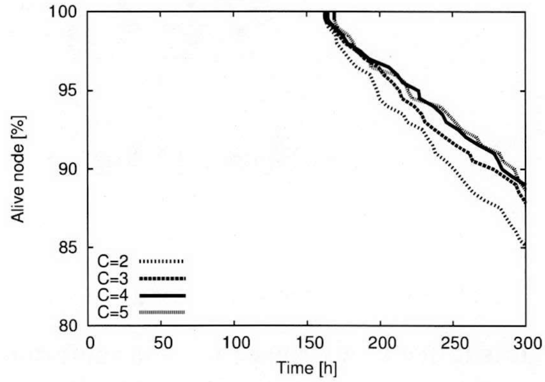



Figure 2.14: Alive node ratio with different C values in experiment 2

In the third experiment, we performed additional simulation experiments in different heterogeneous environment. Total network area was set to $160[m] \times 160[m]$. We used 2 type of nodes

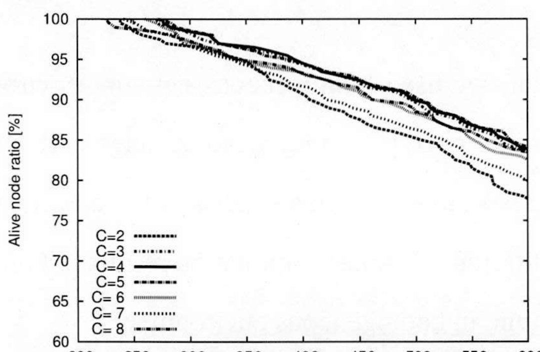



Figure 2.15: Alive node ratio with different C values in experiment 3

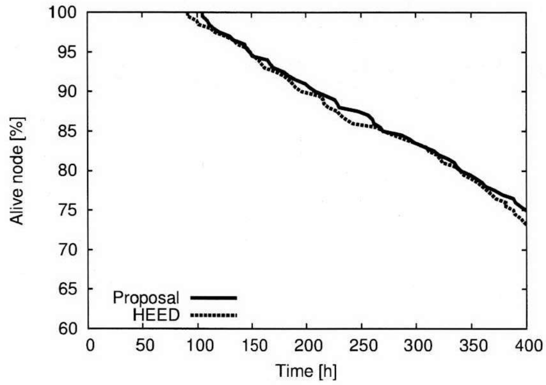
in which initial energy, communication range and amount of nodes were set to $80[J]$, $40[m]$, 64 nodes and $15[J]$, $16[m]$, 256 nodes, respectively. Alive node ratio is shown in Fig. 2.15 with different class values.

In these node specifications, high-end nodes have large batteries and wide communication rage than that of the low end nodes. When C is smaller, once a node has become a CH, it remains as a CH much longer. Low-end nodes around those CHs have high energy consumption and die faster than other nodes. As a result, C=2 has the worst alive node ratio due to rapid drop of critical nodes. On the other hand, larger C causes high frequent re-clustering in the first half of the network lifetime and it has the highest clustering overhead. Because of high clustering overhead, C=8 has lost some nodes in the first half of the network lifetime. However, when it comes to the second half of the network lifetime, it is able to distribute the relay traffic by frequent re-clustering. The conclusion is that smaller C is better for the first half of the network lifetime and larger C is better for the second half of the lifetime. However, proposed protocol does not support dynamic change of C value yet. Therefore, we set the C=4 which has better overall performance.

In the fourth experiment, we use $C = 4$ as the recommended $C$ value and conducted further evaluation for the homogeneous environment described in the first experiment. In this experiment, we compared the result with HEED. We omitted CC from evaluation for homogeneity of the environment. Simulation results in Fig. 2.16 show alive node ratio and Fig. 2.17 shows coverage over time.

In the first half of the network lifetime, the proposed method had more frequent re-clustering than HEED. As a result, the proposed method had lost some coverage area due to re-clustering overhead before HEED as shown in Fig. 2.17. However, in the second half of the network lifetime, HEED had non-uniform energy consumption pasterns due to fixed iteration clustering. On the other hand, the proposed method achieved uniform energy consumption network by dynamic re-clustering and gained some coverage area. Those results show that the proposed clustering method enables to maintain nearly 10% higher alive node ratio than that of HEED.

In the fifth experiment, we evaluated our clustering method with recommended $C = 4$ value and compared it with two other clustering methods; HEED and CC in heterogeneous environment described in the second experiment. Simulation results in Fig. 2.18 and Fig. 2.19



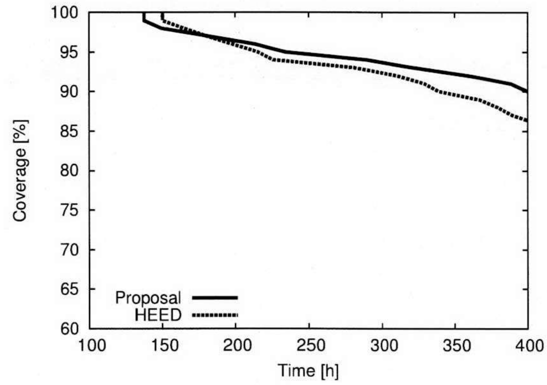
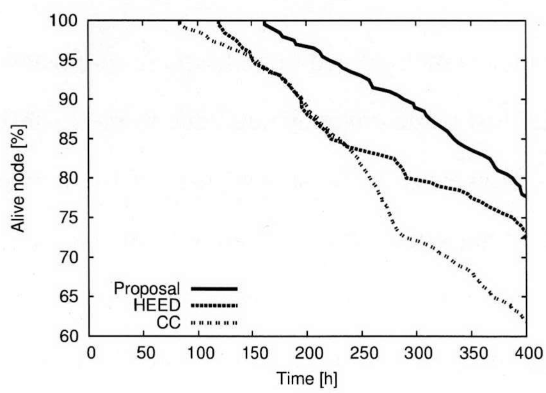




Figure 2.16: Alive node ratio in experiment 4

Figure 2.17: Coverage in experiment 4

show alive node ratio and coverage over time, respectively.

HEED has no mechanism to avoid the non-uniform energy consumption around the CH. On the other hand, the proposed method monitors the class levels of adjacent nodes and performs the re-clustering when those nodes has become weak. CC lacks flexibility and only high-end nodes can become a CH. Power exhaustion of high-end nodes greatly decreases the network connectivity. However, the proposed method dynamically selects the CHs by considering the performance of node itself and its adjacent nodes. Those results showed that the proposed clustering method is able to increase the network lifetime by 80% and 60% more than that of the CC and HEED, respectively in heterogeneous environment and achieve much higher performance than that in homogeneous environment.



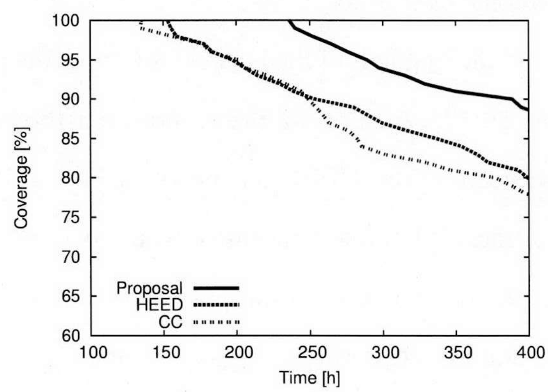




Figure 2.18: Alive node ratio in experiment 5

Figure 2.19: Coverage in experiment 5

In the sixth experiment, we used another heterogeneous network environment optimized for CC. First, we set the total sensing area is to $75[m] \times 75[m]$ square. Then, the whole area is divided into 25 small equal-sized cells ($15[m] \times 15[m]$) as illustrated in Fig. 2.1 (Initial state) and placed a single high-end node in the center of each cell. Low-end nodes are randomly distributed in the whole sensing area. In this environment, we used high-end nodes and low-end nodes in which initial energy, communication range and amount of nodes were set to 43.35[J], 22[m], 25 nodes and 1.83[J], 7.5[m], 175 nodes, respectively. Those values are determined by considering the requirements for minimum node density for a given lifetime constraint and coverage requirement in CC [24].

Fig. 2.20 and Fig. 2.21 show alive node ratio and coverage over time, respectively. Simulation results show that the proposed clustering method achieves more than 10% of coverage than that of CC even in such a CC favorable environment.

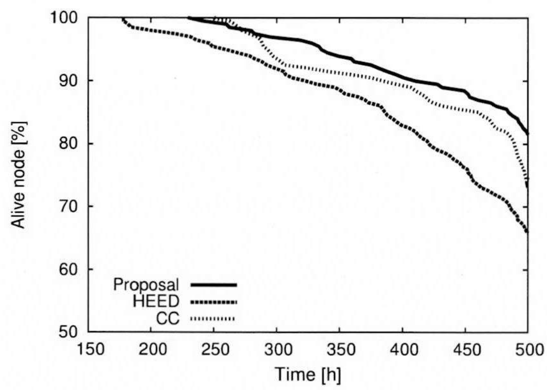



Figure 2.20: Alive node ratio in CC favorable environment

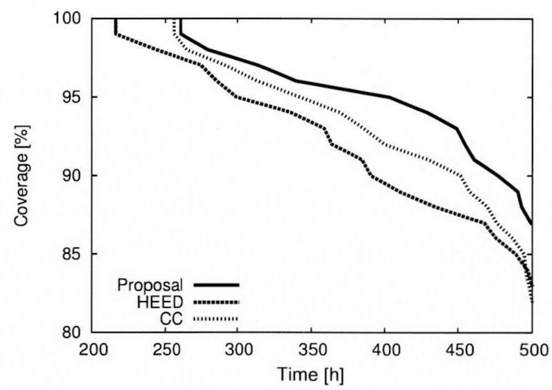



Figure 2.21: Coverage in CC favorable environment

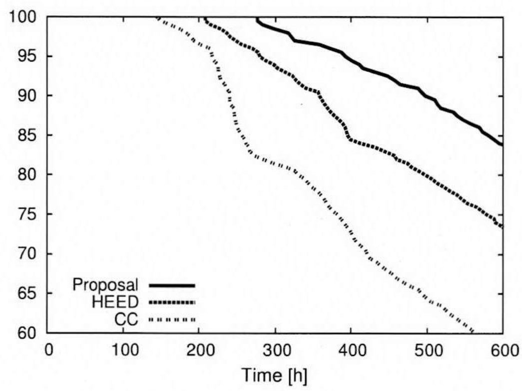



Figure 2.22: Alive node ratio in experiment 7

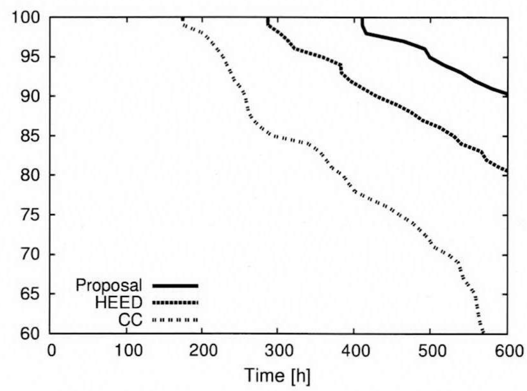



Figure 2.23: Coverage area in experiment 7

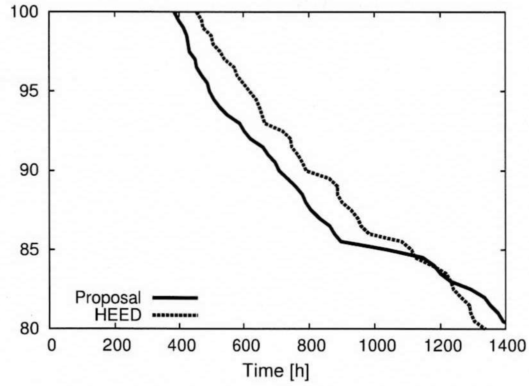



Figure 2.24: Alive node ratio in experiment 8

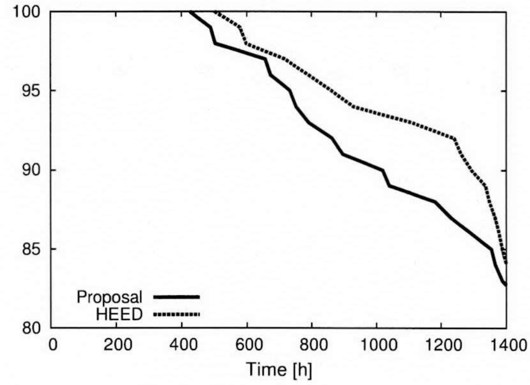



Figure 2.25: Coverage area in experiment 8

In our seventh experiment, we enhanced the heterogeneity by strengthened high-end nodes and weakened low-end nodes. We used 2 type of nodes in which initial energy, communication range and amount of nodes were set to $40[J]$, $80[m]$, 64 nodes and $15[J]$, $16[m]$, 256 nodes, respectively. Total network area was set to $160[m] \times 160[m]$. Then, we logically divided the total area into 256 tiny areas which $10[m] \times 10[m]$ and placed single high-end node in each of those tiny areas. High-end nodes were randomly deployed on the field.

Alive node ratio and coverage area are shown in Fig. 2.22 and Fig. 2.23 respectively. Also, alive node ratio with avarage of 500 samples is shown in Fig. 2.26. In this scenario, HEED always selects the high-end nodes as CH by considering the residual energy of the nodes and sets a long clustering iteration for those CHs. Low-end nodes die before CH and it weakens the connectivity of the network. In CC only high-end nodes are used as CHs. On the other hand, the proposed method considers not only the performance of the node itself but also that of it's adjacent nodes to select the appropriate node for CH. As a result, the proposed method shows the best performance in this environment.

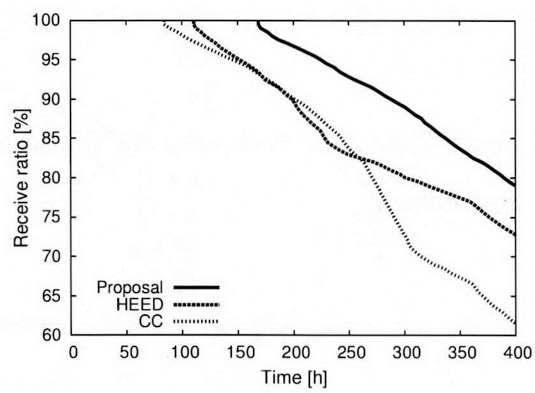



Figure 2.26: Alive node ratio in experiment 7 with 500 samples

In our eighth experiment, we created less heterogeneous environment by increasing the initial energy of low end nodes. Total network area was set to $80[m] \times 80[m]$. We used 2 type of nodes in which initial energy, communication range and amount of nodes were set to $25[J]$, $16[m]$, 30 nodes and $25[J]$, $8[m]$, 220 nodes, respectively. Then, we logically divided the total area into 256 tiny areas each of which ranges $8[m] \times 8[m]$. High-end nodes and low-end were randomly deployed on the field.

In this scenario, the proposed method was misguided by low-end nodes with high initial energy. Alive node ratio and coverage area are shown in Fig. 2.24 and Fig. 2.25 respectively. In initial state, the proposed method chooses low-end nodes as CHs and there is no possibility to select a high-end nodes as a CH. However, for HEED, CHs are randomly selected between low-end nodes and high-end nodes. Therefore, during the first half of the network lifetime, the proposed method creates large amount of CHs and increases the clustering overhead. When it comes to the second half of the network life, the network becomes more heterogeneous than the initial state. As a result, the proposed method enables to gain more alive nodes and larger coverage in the second half of the network lifetime.

Through all the performance evaluation, it is verified that our proposed clustering method increases network lifetime by 80% and 60% more than that of the CC and HEED, respectively in heterogeneous networks. Moreover, even in a homogeneous network, the proposed method has almost the same performance as HEED. It is a useful feature for practical use.

## 2.5 Conclusion

In this chapter, we proposed a new clustering and a new non-clustering routing methods for heterogeneous WSNs. First, we discussed some related work and introduce two existing clustering methods for heterogeneous WSNs and existing non-clustering routing protocols for WSNs. The proposed clustering method selects CHs considering not only the performance of a certain node but also that of its adjacent node. In the proposed non-clustering method, residual energy of nodes in the path are taken into consideration when selecting the optimal path between source and destination. Remarkably, the proposed clustering and non-clustering methods have no limitations on the number of node types that can be used to create a WSN. The simulation experiments validated the result that the proposed clustering method can prolong the lifetime in heterogeneous network environment. Moreover, even in a homogeneous network, the proposed method had almost the same performance as HEED. It is a useful feature for practical use.

Due to uneven manufacturing process, same type of nodes may have variation in their initial energy and/or radio transmission range. Since there is no limit to the number of types of nodes that can be used in the proposed method, such a problem can be addressed by subdividing the nodes into more node types.

In this chapter, we have considered a limited number of aspects of WSNs. The possibility of geographically different networks should also be considered. Another important issue to be explored is a heterogeneous network model where the difference between the sensors is not only the difference in initial energy and communication range, but also in their processing capabilities, and thus the consideration of energy consumption in data processing (compression, fusion, etc.).

# Chapter 3

# Hybrid Routing Method for Wireless Sensor Networks

## 3.1 Introduction

In this chapter, we propose a robust clustering/non-clustering hybrid routing method for heterogeneous WSNs which adaptively combines clustering and multi-hop communication methods. Most of conventional routing methods for WSNs are optimized for a specific application. In the proposed method, however, we considered the specifications of the nodes and the characteristics of the networked like outer boundary and location of the sink instead of characteristics of an application to optimize the WSN. Therefore, the proposed method can be easily applied to a general-purpose WSN, with multiple sensor applications. In this chapter, we propose simplified spatial boundary analysis method for small area WSN and generalized spatial boundary analysis method for wide area WSN. In order to study the energy consumption and traffic patterns around sink and study about efficient bode placement around the Sink, we conducted several preliminary simulation experiments. Through the simulation experiments, we verified that the proposed hybrid methods able to maintain higher alive node ratio than that of the pure clustering and multi-hop methods in any network topology.

In WSN, since all data are directed to sink, CHs or nodes around the sink come to have high relay traffic. As a result, CHs or nodes around sink consume much energy in a short time. As we discussed in chapter 1, non-clustering routing using a large amount of nodes enables to balance the relay traffic all over the WSN. When it is applied to wide area networks, however, the amount of data that has to relay for each node increases as it is near to sink. Furthermore, compared with clustering mode, data aggregation is not efficient in multi-hop routing. On the other hand, data aggregation in clustering combines data from multiple sensors to eliminate redundant information and transmissions. It also guarantees that the energy load is well balanced by CHs dynamically elected according to a prior optimal mechanism. However, CH around the sink come to have high relay traffic.

This non-uniform energy consumption around the sink is one of the main problems that shorten the network lifetime [21] in WSN. Furthermore, in the clustered WSNs, only limited number of CHs around the sink are participate in data relaying process. Therefore, those CHs around the sink uses much energy in a short time. In order to avoid the node failures, those CHs have to perform the frequent re-clustering. Also, specially in wide area WSN, these frequent re-clustering could trigger the total network re-configurations.

The importance of the non-uniform energy consumption problem around the sink differs substantially whether the sensor nodes and/or the sink node are mobile or not. When the sink is mobile, it moves around to collect data from sensor nodes. Therefore, the sink is able to effectively balance the the energy consumption of the WSN. The sensor nodes can transmit the data periodically, or store the data and delay the transmission till the distance between the senor nodes and the sink is minimal in order to decrease the power consumption for data relaying. In the case where sensor nodes are mobile, the nodes can adjust their position and help to balance the energy consumption in areas that have high energy consumption. However, deploying a mobile sinks or mobile nodes will increase the initial cost and management cost of the WSN. Additionally, mobility is impractical for some applications. Therefore, in this research we focused on the WSNs which consisted of stationary nodes and a single sink.

WSN can be classified into two types; homogeneous and heterogeneous WSNs [13]. Heterogeneous networks lower the hardware cost and reduce the communication cost of the sensing nodes. Meanwhile, homogeneous networks achieve uniform energy drainage. However, both features cannot be incorporated in the same network. In Chapter 2, we proposed a new clustering method for WSNs with heterogeneous node types which selects CHs considering not only transmission power and residual energy of each node but also those of its adjacent nodes. It is able to maintain above three characteristics.

However, in heterogeneous WSNs, non-uniform energy consumption problem around the sink still remains without efficient solution. Therefore, in this chapter we propose a hybrid routing method in order to avoid the non-uniform energy consumption problem around the sink.

The rest of the chapter is organized as follows: we first discuss the existing clustering and multi-hop communication methods in section 3.2. In section 3.3, we present our approach to hybrid routing in WSN, while in section 3.6 we present the performance evaluation and simulation results. Finally, in section 3.7, conclusions are drawn.

## 3.2 Traditional Routing Methods

### 3.2.1 Clustering Methods

Many data gathering schemes which shorten the transmitting distance or reduce the number of times of transmission have been proposed.

For example, Directed Diffusion [40] shortens the transmitting distance by using multi-hop communication and LEACH (Low-Energy Adaptive Clustering Hierarchy) [41] reduces the number of times of transmission to the base station by using cluster formation and data fusion. Further, HEED (Hybrid, Energy-Efficient, Distributed Clustering) [23] which follows LEACH concept takes residual energy of each node into consideration in order to prolong the network lifetime and CC (Chessboard Clustering) [24] uses high-end sensor nodes as CHs in order to reduce the data transmission burden on sensor nodes in the network.

In order to balance energy consumption among CHs, Unequal Cluster-based routing protocol (UC) proposed a method by assigning smaller cluster sizes to CHs closer to the sink and larger cluster sizes to CHs faraway from the sink, i.e. cluster size is proportional to the distance between CH and the sink. In UC, sensors are assumed to be deployed in a circular region around a sink in a homogeneous WSN. The smallest cluster size depends on the radius of the circular WSN.

However, CHs that are closer to the sink will relay more traffic and hence will deplete their batteries earlier than CHs that are faraway from the sink. Furthermore, large number of CHs closer to the sink consume more energy to maintain the topology and frequently make re-clustering due to high relay traffic.

In [44], we proposed a clustering method for WSN with heterogeneous node types, which selects CHs considering not only transmission power and residual energy of each node but also those of its adjacent nodes. Brief introduction for our clustering method is as follows.

First, we assume that each sensor is aware of its own location [30, 45]. A sensor node expends more energy in data communication than in sensing or data processing [32]. First, we calculate the total amount of data which each node is able to send until its power exhaustion and use it as a strength index of the node. Then, we normalize the value and use it to relatively compare the nodes in WSN. Using that result, all nodes are classified into $C$ number of classes. Then, we define two types of classes to each node; individual class and adjacent class. Individual class value represents the strength of node itself and adjacent class value represents the strength of its adjacent nodes. Based on these class values, each node becomes the CH and makes a cluster if pre-defined waiting time passed without receiving clustering messages from any other nodes. When a node receives messages from other nodes during the waiting time, it keeps them into CH backup list. After the waiting time expired, it checks the CH backup list for accessible CH and node becomes a member of the optimum cluster (i.e. nearest CH to the node), if there are any entries. If there are no accessible CHs in the list, it broadcasts a searching message for accessible CHs. After a certain period, if it still cannot find any accessible CHs, it becomes CH.

Clustering method we proposed in [44] has no limitations on the number of node types and simulation results show that the proposed clustering method achieves much higher alive node ratio and coverage than those of HEED and CC. Therefore we use [44] as the clustering method in our new hybrid routing method.

### 3.2.2 Multi-hop Communication

Multi-hop wireless networks typically use routing techniques similar to those in wired networks. In a multi-hop WSN, each node plays the dual role of data origination and data router. These traditional routing protocols choose the best sequence of nodes between the source and destination, and forward each packet through that sequence. In this research, we use the shortest path between node and sink. We will discuss further details in section 3.3.2.

## 3.3 Proposed Hybrid Routing Method

### 3.3.1 Basic Idea of Clustering and Non-clustering Zones

#### Efficient Node Placement Around the Sink

As we discussed before, CHs around the sink have to relay heavier traffic. Due to that traffic, those nodes consume more energy than the other CHs. Increasing the low residual energy nodes around sink shortens the network lifetime. Therefore, in our proposed method, we focus on load balancing around sink in order to reduce the number of nodes with low residual energy.

In this study, we divided the network in to two zones; clustering zone and non-clustering zone. Then, we set a all the high-end nodes (H-nodes) in non-clustering zone are placed near to boundary line as shown in Fig. 3.1. Radius of this static non-clustering zone is equal to communication range of the H-node. Furthermore, those H-nodes act as CHs for nodes only in clustering zone. Then, we increase the low-end node (L-node) density in non-clustering zone and observe the effect on network lifetime.

In the clustering zone, we use our proposed clustering method [44], which selects CHs considering not only transmission power and residual energy of each node but also those of its adjacent nodes. First, we assume that each sensor is aware of its own location. A sensor node expends more energy in data communication than in sensing or data processing. First, we calculate the total amount of data which each node is able to send until its power exhaustion and use it as a strength index of the node. Then, we normalize the value and use it to relatively compare the nodes in WSN. Using that result, all

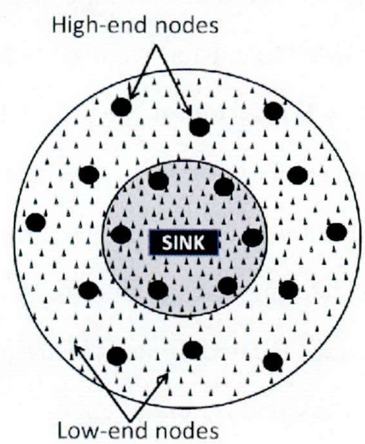



Figure 3.1: Network Model

nodes are classified into $C$ number of classes. Then, we define two types of classes to each node; individual class and adjacent class. Individual class value represents the strength of node itself and adjacent class value represents the strength of its adjacent nodes. Based on these class values, each node becomes the CH and makes a cluster if pre-defined waiting time passed without receiving clustering messages from any other nodes. When a node receives messages from other nodes during the waiting time, it keeps them into CH backup list. After the waiting time expired, it checks the CH backup list for accessible CH and becomes a member of the optimum cluster (i.e. nearest CH to the node), if there are any entries. If there are no accessible CHs in the list, it broadcasts a searching message for accessible CHs. After a certain period, if it still cannot find any accessible CHs, it becomes CH. Simulation results show that the proposed clustering method achieves much higher alive node ratio and more extensive coverage area than those of HEED and CC.

In a non-clustering zone, each node plays the dual roles of data origination and data relay. Multi-hop routing protocols choose the best sequence of nodes between the source and destination, and forward each packet through that sequence. In this research, we use the shortest path between node and sink. The packets are forwarded to the neighbor node that has the shortest distance to the destination, and the next node relays toward the destination in the same manner. Source node $S$ sends the data packets to destination node $D$ via node $N_i$. Since the nodes are aware of its own location, $z = \overrightarrow{SN_i} \cdot \frac{\overrightarrow{SD}}{|SD|}$ is calculated for each node and the node with maximum $z$ is selected as the next hop to destination $D$.

Inside the cluster, we use the above multi-hop communication method to relay the sensing data to CH. In addition we also use the same method as the one for inter-cluster communication between CHs in order to send the aggregated data to non-clustering zone. Especially, among the CHs are in clustering zone, the CHs which are able to directly communicate with nodes in non-clustering zone are referred to as Boundary CHs (BCHs). Each BCH manages a list of next hop nodes located in non-clustering zone and assigns a probability in proportion to its residual energy divided by initial energy to each node. When BCH forwards the data to non-clustering zone, it selects a next node according to that probability.

We evaluate the performance of this model via network simulator ns-2 and observe the percentage of alive nodes. Specifically, we compare it with that of total clustering which adopts the method described in [44] and total multi-hop networks. Then, we increase the node density in non-clustering zone and observe the alive node ratio.

In this simulation model, we set the node configurations using datasheet and information provided by Crossbow. Total network area is set to a circle with radius $R = 80[m]$. We used 2 types of nodes; H-nodes and L-nodes. Their initial energy, communication range and amount of nodes were set to 43.55[J], 30[m], 18 nodes and 5.67[J], 12[m], 180 nodes, respectively.

According to those node specifications, the radius of the non-clustering zone $d$ became $30[m]$.

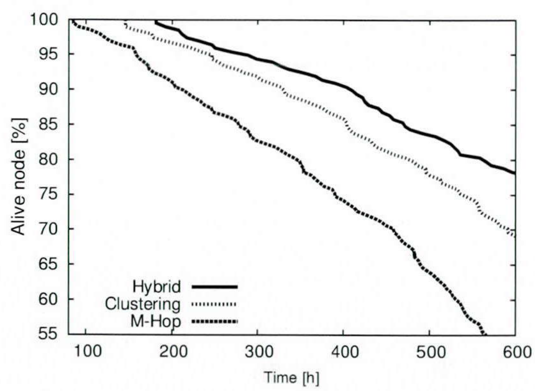



Figure 3.2: Comparison with Others

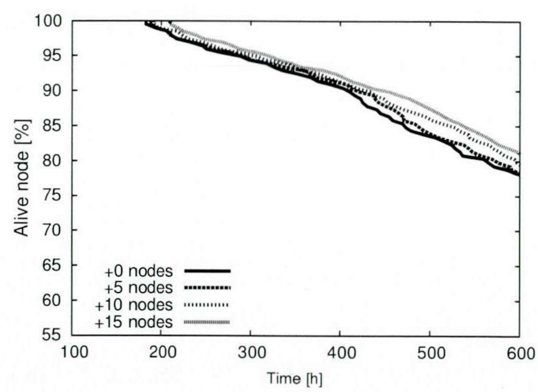
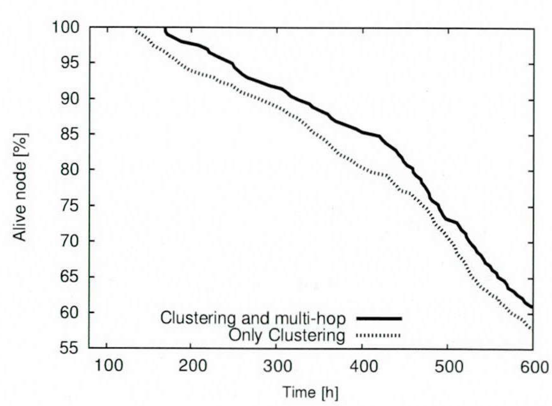



Figure 3.3: With different node density

The simulation result is shown in Fig. 3.2, which is alive node ratio in $d = 30[m]$. It shows that the proposed method achieves higher alive node ratio than that of total clustering and multi-hop methods. Then, in the same environment, we added 5 to 15 L-nodes to non-clustering zone and those simulation results are shown in Fig. 3.3. From Fig. 3.3, higher L-node density in non-clustering zone leads to longer network lifetime.

#### Study of Energy Consumption and Traffic Patterns Around Sink

In order to study the energy consumption and traffic patterns around sink, we conducted several preliminary simulation experiments. In these experiments, we observed the behavior of clustering and multi-hop networks.




In this simulation, we randomly deployed

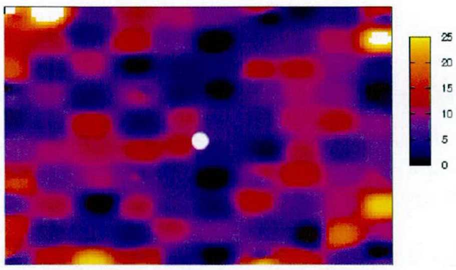



Figure 3.4: Energy distribution on clustered network at 150[h]

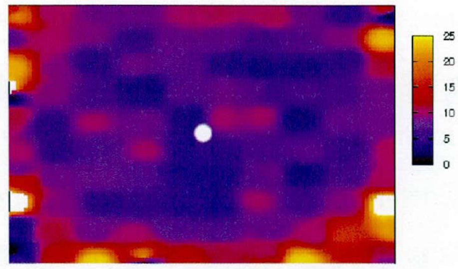



Figure 3.5: Energy distribution on multi-hop/clustering combined network at 150[h]

Table 3.1: Node specifications

| | H-node | L-node |
|---|---|---|
| Communication range [m] | 30 | 12 |
| Initial energy [J] | 43.55 | 5.67 |
| Calculation cost [J/bit] | $2.73 \times 10^{-6}$ | $2.73 \times 10^{-6}$ |
| Aggregation ratio | 0.25 | 0.25 |
| Data transfer rate [kbps] | 250 | 250 |
| Sensing range [m] | 3 | 3 |

50 H-nodes (High-end nodes) and 260 L-nodes (Low-end nodes) into $100[m] \times 100[m]$ area and the sink was placed at the center of the area. Node specifications are shown in Table 3.1.

At first, we conducted a simulation experiment on the above network by creating a clustered network. Secondly, we conducted another experiment on multi-hop/clustering combined network by creating a multi-hop network in the non-clustering zone and clustered network outside of the non-clustering zone. In this network, we defined the non-clustering zone as $50[m] \times 50[m]$ meters squared area around the sink. Also, the center of the non-clustering zone overlapped with sink. We observed alive node ratio in each experiment as shown in Fig. 3.6. It indicates that multi-hop/clustering combined network is able to maintain higher receive ratio than that of the clustered network.

Moreover, we observed the distribution of the residual energy of each node in non-clustering zone. Fig. 3.4 and Fig. 3.5 show the residual energy distribution inside the non-clustering zone of the clustered and multi-hop/clustering combined networks when $150[h]$ elapsed in network time, respectively. Bright areas indicate higher residual energy and dark areas indicate lower residual energy.

Comparing Fig. 3.4 and Fig. 3.5, clustered network has a number of darker spots and multi-hop/clustering combined network has no such a spot. Those darker spots are the results of the unequal energy consumption around the CHs and sink. Also, clustered network consumes more energy to maintain the topology and frequently re-cluster it due to high traffic load. To the contrary, multi-hop network uses large number of nodes which enable to balance the relay traffic all over the network.

This experiment shows multi-hop/clustering combined network is more efficient than the clustered network in smaller network area with high relay traffic. Therefore, we can apply this method to WSN to avoid the non-uniform energy consumption around the sink. In other words, it is efficient to create a small multi-hop network around the sink and make a cluster network outside. However, it is essential to determine the optimal boundaries for multi-hop network around the sink. In this chapter we propose two methods to determine the optimal boundaries. Simplified spatial boundary analysis method for small are WSNs is explained in section 3.4 and generalized spatial boundary analysis method for wide are WNSs is explained in section 3.5.

### 3.3.2 Routing Methods in Clustering Zone

CH has two main tasks; i.e. sending the collected data to sink and relaying the data from other CHs. CHs with lower energy burdened with higher relay traffic could cause early expiration of battery and shorten the network life-time. We can simply solve this problem by selecting the node with highest residual energy for next hop. After initialization, each CH exchanges location information with neighbor CHs. All the CHs advertise to adjacent CHs whether they are reachable to the non-clustering zone or not when they broadcasts the control messages. Once CH has received the control messages from other CHs, it can determine which adjacent CHs are connected with non-clustering zone and which are not. For example, in Fig. 3.8, $S$ is source CH and $D$ is the sink. When a CH wants to send data packets to the non-clustering zone,

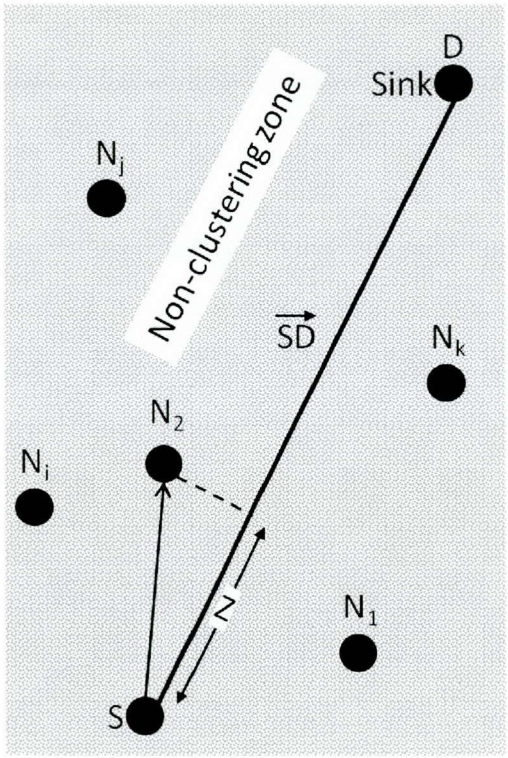



Figure 3.7: Routing inside the non-clustering zone

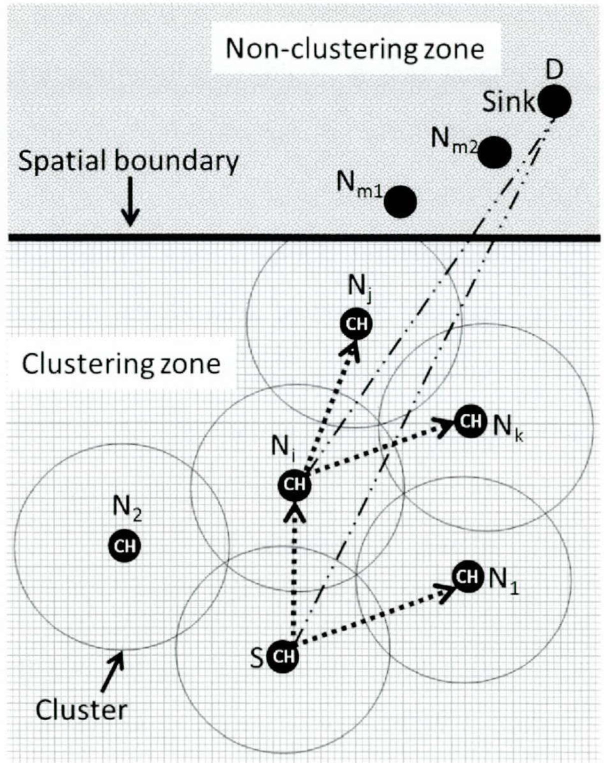



Figure 3.8: Routing between CHs in clustering zone

it virtually draws a straight line $SD$ between itself and the sink. Then, it searches for neighbor clusters which intersects with line $SD$ ($N_1$ and $N_i$ in Fig. 3.8. Among them, the node with the highest residual energy is finally selected as a next hop to non-clustering zone. Then next hop repeats the same task while it is able to connect with non-clustering zone.

When line SD does not have any intersections with neighbor clusters, the proposed method selects the CH with highest residual energy from among the adjacent CHs which is reachable to non-clustering zone as its next hop. Then next hop repeats the same task while it is able to connect with non-clustering zone.

Inside the cluster, we use multi-hop communication method to relay the sensing data to CH. In addition we also use multi-hop communication for inter-cluster communication between CHs for sending the aggregated data to sink. Each node inside the cluster can know the CH's location by specific message that is broadcast by CH.

### 3.3.3 Routing Methods in Non-clustering Zone

When a sensor node generates data, it sends packets to sink. The packets are forwarded to the neighbor node that has the shortest distance to the sink, and the next node relays toward the sink in the same manner. For example, as shown in Fig. 3.7, source node $S$ sends the data packets to destination node $D$ (which is sink) via node $N_2$. Since the nodes are aware of its own location, $z = \overrightarrow{SN_i} \cdot \frac{\overrightarrow{SD}}{|SD|}$ is calculated for each node and the node with maximum $z$ is selected as the next hop to destination $D$.

### 3.3.4 Routing Methods Around the Spatial Boundary

When nodes in the non-clustering zone receive the data, they forward the data to the sink by using multi-hop communication. Furthermore, CH which is able to directly communicate with nodes in non-clustering zone is referred to as Boundary CH (BCH). Each BCH manages a list of next hop nodes located in non-clustering zone and assigns a probability in proportion to its residual energy divided by initial energy to each node. When BCH forwards the data to non-clustering zone, it selects a next node according to the probability.

### 3.3.5 Network Formation

First, the radius of non-clustering zone $d$ is determined by using the method explained in section 3.4.2. Then, sink broadcasts that information to network. Each node in the network enables to receive that message and evaluates whether it is in the non-clustering zone or clustering zone upon its location information and $d$. Then, the nodes in the clustering zone form clusters by using the method explained in [44] and forward the data to non-clustering zone. When nodes in the non-clustering zone receive that data, they forward the data to the sink by using multi-hop communication. Furthermore, CH which is able to directly communicate with nodes in non-clustering zone is referred to as Boundary CH (BCH). Each BCH manages a list of next hop nodes located in non-clustering zone and assigns a probability in proportion to its residual energy divided by initial energy to each node. When BCH forwards the data to non-clustering zone, it selects a next node according to that probability.

## 3.4 Simplified Spatial Boundary Analysis Method for Small Area WSN

As we introduced in previous section, CHs around the sink have high relay traffic. Due to that traffic, those nodes consume more energy than the other CHs. Increasing the low residual energy nodes around sink shortens the network lifetime. Therefore, in our proposal method, we focus on load balancing around sink in order to reduce the number of nodes with low residual energy.

In the proposed method, nodes placed near to sink perform a multi-hop network (non-clustering zone) and nodes far away from sink perform clusters (Clustering Zone). Before performing them, operation boundary for non-clustering zone must be determined. Therefore, first we analyze the network and define the spatial boundary for non-clustering zone. Outside of that zone, we define as clustering zone. Next section explains the calculations of average energy consumption of a node in area $C(d)$ ( Notification $C(x)$ represents the circle area with radius of $x$ $[m]$) when doing multi-hop communication ($E_m$) and clustering ($E_c$). Then considering $E_m \leq E_c$, we calculate the maximum value of $d$ which is the radius of multi-hop communication range.

### 3.4.1 Network Model

First, we assume that whole area of the network is circle $C(R)$ with radius $R$ and sink is located in the center of that circle. It is assumed that the non-clustering zone $C(d)$ is a concentric circle with $C(R)$ and $d(< R)$. Then we uniformly distribute $n_H$ number of High-end nodes (H-nodes) and $n_L$ number of Low-end nodes (L-nodes) over area $C(R)$. Specifications of the nodes are shown in Table 3.2.

We consider node heterogeneity in terms of communication range and initial energy but not in sensing capabilities. Sensing range for all types of nodes are circle with radius $r_s$ $[m]$ where, $R > r_s\sqrt{3}/2$.

For the clustering, we use the clustering formation method which is specifically described in [44]. We assume that the size of sensing data collected in CH becomes $a$ times smaller

Table 3.2: Notation of node specifications

| | H-node | L-node |
|---|---|---|
| Number of nodes | $n_H$ | $n_L$ |
| Communication range [m] | $r_H$ | $r_L$ |
| Initial energy [J] | $b_H$ | $b_L$ |
| Transmission power [W] | $e_{TH}$ | $e_{TL}$ |
| Receiving power [W] | $e_R$ | $e_R$ |

$(0 < a \leq 1)$ after aggregation. Energy consumption for aggregation is $e_c$ $[J/byte]$. Radius of cluster depends on the maximum transmission range of the node and clustering factor $\phi$ is defined as $\frac{Clustering\,radius}{Maximum\ communication\,radius}$. To simplify the analysis, we assume that no data aggregation is performed in multi-hop mode and data packets are simply relayed from other nodes to sink. However, there are complex methods for data aggregation in multi-hop WSN [42, 43] and we will consider them in our future works.

### 3.4.2 Mathematical Analysis

Considering 100% coverage area, we use circles of radius $r_s$ $[m]$ to pack the plane as shown in Fig. 3.9. Around each circle, suppose to draw a concentric circle of radius $r(\sqrt{3}r_s/2)$. Then we can pack 3 circles of radius $r_s$ next to each other, and those three centers of circles can form an equilateral triangle. Also, three corresponding concentric circles of radius $r$ can intersect at the middle of the equilateral triangle. The optimum packing of circles in the plane has density $\pi/\sqrt{12}$, which is that of the hexagonal packing (Kepler's Conjecture [49]). For the connectivity, $2r_s \leq r_L$. For this model we claim that *Lemma 1.* and 2. are true.

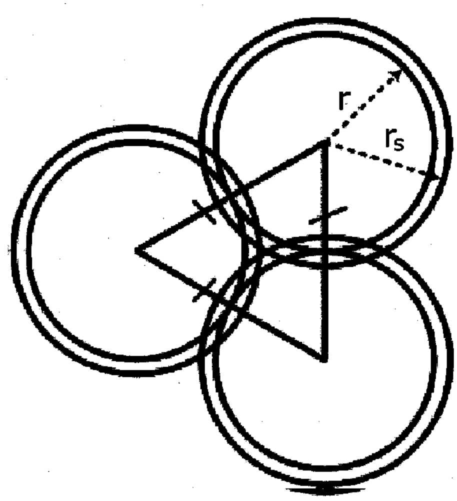



Figure 3.9: Positions of 3 circles

A.1: *Lemma 1.*

It takes maximum $\frac{2\pi}{\sqrt{12}}\left(\frac{2d}{\sqrt{3}r_s}+\frac{2}{\sqrt{3}}\right)^2-1$ necessary hops to deliver a packet from a source to a destination.

A.2: *Lemma 2.*

The number of L-nodes skipped by single H-node is approximately equal to $\lfloor\frac{r_H}{r_L}\rfloor$.

Refer the A.1 and A.2 for further details and the proofs of above.

### 3.4.3 Multi-hop Mode

Non-clustering zone is a circle with radius $d$ and concentric with sink. Specifications of the nodes are shown in Table 3.2. Let us assume that amount of packets from the out-side of the non-clustering zone is equal to $P$. Data-transfer rate is $b$ $[kbps]$ for all nodes. Because of uniform distribution of H-nodes and L-nodes, amount of data for H-nodes, $P_H$ is, $P_H=\frac{n_H P}{n_H+n_L}$ and amount of data for L-nodes, $P_L$ is, $P_L=\frac{n_L P}{n_H+n_L}$. Energy consumption for data relay in H-nodes, $V_H$ is, $V_H=P_H\left(e_R+e_{TH}\right)/b$ and energy consumption for data relay in L-nodes, $V_L$ is, $V_L=P_L\left(e_R+e_{TL}\right)/b$.

The probability that a route consists of $x$ number of H-nodes and $y$ number of L-nodes, $F_m(x,y)$, is

$$F_m(x,y)={}_{X_{max}}\mathrm{C}_x\cdot{}_{Y_{max}}\mathrm{C}_y\left(\frac{n_H}{n_H+n_L}\right)^x\left(\frac{n_L}{n_H+n_L}\right)^y, \tag{3.1}$$

where $X_{max}=\frac{\pi}{\sqrt{12}}\left(\frac{d}{2r_H}+\frac{2}{\sqrt{3}}\right)^2-1$, when the route consists with only H-nodes. In the same manner, $Y_{max}=\frac{\pi}{\sqrt{12}}\left(\frac{d}{2r_L}+\frac{2}{\sqrt{3}}\right)^2-1$, when the route consists with only L-nodes. Each node in the route has to relay the data from Clustering area and send its own $p$ packets of data to sink. Therefore, energy consumption is,

$$\begin{aligned}E_m=\sum_{y=Y_{min}}^{Y_{max}}\sum_{x=0}^{\lceil y/\lfloor\frac{r_H}{r_L}\rfloor\rceil}&\left(F_m(x,y')\frac{xV_H+y'V_L}{x+y'}\right.\\&\left.+\frac{xpe_{TH}+y'pe_{TL}}{b(x+y')}\right),\end{aligned} \tag{3.2}$$

where $y'=y-x\lfloor\frac{r_H}{r_L}\rfloor$, $Y_{min}=\left\lceil\frac{R}{r_L}\right\rceil$.

Note that energy consumption in calculation unit is significant only in aggregation mode. Otherwise it can be negligible.

### 3.4.4 Clustering Mode

Let us assume that clusters are performed in area $C(d)$. The calculation of average energy consumption in clustering area $C(d)$, $E_c$, is given below. The numbers of H-node CHs and L-node CHs are equal to $h_H$ and $h_L$, respectively. Nodes' specifications are shown in Table 3.2. All the CHs are uniformly distributed in area $C(d)$. Therefore, amount of data for H-node CHs, $P_H$, is $P_H = \frac{Ph_H}{h_H+h_L}$ and amount of data for L-node CHs $P_L$, is $P_L = \frac{Ph_L}{h_H+h_L}$.

CH performs 2 major operations in the network. First, CH collects the sensing data from sensor nodes in the cluster including CH itself. Then, it aggregates them and sends to sink by multi-hop communication between CHs. Second, CH relays the traffic from other CHs to sink. Radius of clusters, $r_H^s = \phi r_H$ if H-node becomes a CH and $r_L^s = \phi r_L$ if L-node becomes a CH. CH node has to collect the data from its own cluster and aggregate them before sending to sink. If H-node becomes a CH, energy consumption for its own cluster process, $V_{H1}$, is $V_{H1} = \frac{(n_H+n_L)}{\pi R^2}\pi(\phi r_H)^2 p[(e_R + ae_{TH})/b + e_c]$, and if L-node becomes a CH, energy consumption for its own cluster process, $V_{L1}$, is $V_{L1} = \frac{(n_H+n_L)}{\pi R^2}\pi(\phi r_L)^2 p[(e_R + ae_{TL})/b + e_c]$. CH node has to relay the data from other CHs. If CH is a High-end CH node, $V_{H2} = P_H(e_R + e_{TH})/b$ and if CH is a Low-end CH node, $V_{L2} = P_L(e_R + e_{TL})/b$.

The probability that a route consists of $x$ number of High-end CH nodes and $y$ number of Low-end CH nodes, $F_c(x, y)$, is

$$F_c(x, y) = {}_{X_{max}}C_x \cdot {}_{Y_{max}}C_y \left(\frac{n_H}{n_H + n_L}\right)^x \left(\frac{n_L}{n_H + n_L}\right)^y, \tag{3.3}$$

where $X_{max} = \frac{\pi}{\sqrt{12}}\left(\frac{d}{2r_H} + \frac{2}{\sqrt{3}}\right)^2 - 1$, when the route consists of only High-end CH nodes. In the same manner, $Y_{max} = \frac{\pi}{\sqrt{12}}\left(\frac{d}{2r_L} + \frac{2}{\sqrt{3}}\right)^2 - 1$, when the route consists of only Low-end CH nodes. Therefore, energy consumption in clustering mode in area $C(d)$ is,

$$E_c = \sum_{y=Y_{min}}^{Y_{max}} \sum_{x=0}^{y/\lfloor \frac{r_H}{r_L} \rfloor} \left( F_c(x, y') \frac{xV_{H2} + y'V_{L2}}{x + y'} + \frac{xV_{H1} + y'V_{L1}}{x + y'} \right), \tag{3.4}$$

where $y' = y - x\lfloor \frac{r_H}{r_L} \rfloor$.

Finally, from Eq. (3.4) and Eq. (3.2), we calculate the maximum $d$ which satisfies the $E_m \leq E_c$.

## 3.5 Generalized Spatial Boundary Analysis Method for Wide Area WSN

### 3.5.1 Network Model

In the proposed method, nodes placed near to sink perform the operation of a multi-hop network (non-clustering zone) and nodes far away from sink perform clustering operation (Clustering Zone). Before performing them, operation boundary for non-clustering zone must be determined. Therefore, first we analyze the network and define the boundary for non-clustering zone. Outside of that zone, we define as clustering zone. Next section explains the calculations of average energy consumption for a node in non-clustering zone; $E_m$ when performing multi-hop communication in non-clustering zone and $E_c$ when clustering in non-clustering zone. Then considering $E_m \leq E_c$, we calculate the boundary function for non-clustering zone $B_{m-hop}$ which defines the area for non-clustering zone.

We assume that every node can be aware of its location information. It is a reasonable assumption, since recent researches have proposed efficient methods to obtain the location information with high accuracy. For example, [30] proposes to perceive both the anchors' positions and neighbor node information by Neighbor Position-based Localization Algorithm that can greatly enhance the positioning accuracy compared with conventional overlapping connectivity localization algorithms. In addition, [45] proposed a method in which a probabilistic model of wireless propagation characteristics in a real environment is constructed, and the node position is estimated by maximum-likelihood estimation using the model.

### 3.5.2 Definition of the Boundary Function

In our analytical method, we consider the physical network boundaries and location of the sink to define the spatial boundary between clustering and non-clustering zones. We use Circular Polar Coordinates (CPC) system and shift the origin of CPC to location of the sink. First, we

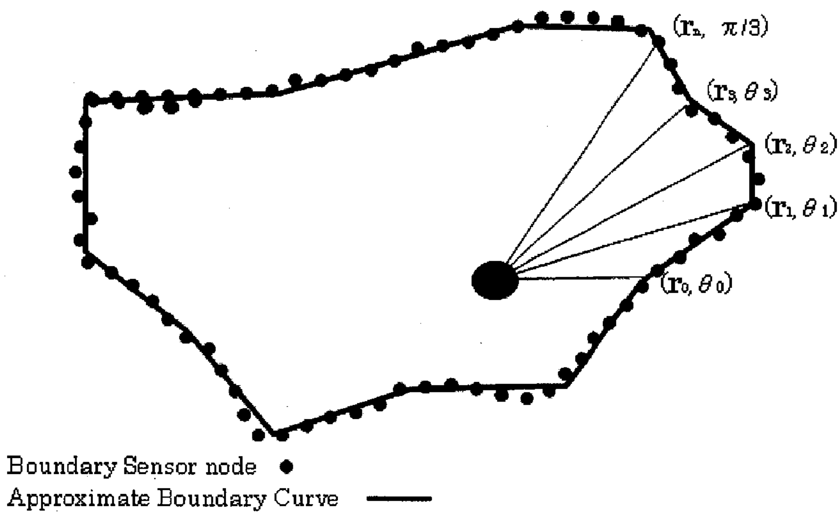



Figure 3.10: Network boundary

gather the location details of random points on the physical network boundary. The coordinates of the a arbitrary node $i$ in the system is given by $(r_i, \theta_i)$, where $r_i$ is the radial coordinate and $\theta_i$ is the angular coordinate. Then, we manually divide the whole boundary into several sub-boundaries by considering the characteristics of the total network area. Then, we apply Least Squares Method (LSM) for arbitrary sub-boundary $i$ and obtain the approximate equation of sub-boundary function $b_i$. As shown in Fig. 3.10, we obtain $n$ sub-boundaries $b_0$, $b_1$, ..., $b_n$ in the range of $0 \leq \theta < 2\pi$ ($\theta_0 = 0$ in CPC). By combining all the sub-boundary functions, we obtain boundary function shown in Eq. (3.5) in the range of $0 \leq \theta < 2\pi$. Each sub-boundary is a curve segment, with fixed beginning and ending points defined by angular coordinates.

Boundary Function of the network is given by $B(r, \theta)$.

$$B(r,\theta) = \begin{cases} b_0 \ (\ \theta_0 \leq \theta < \theta_1) \\ b_1 \ (\ \theta_1 \leq \theta < \theta_2) \\ . \\ . \\ . \\ b_n \ (\ \theta_{n-1} \leq \theta < 2\pi) \end{cases} \tag{3.5}$$

**Conditions**

- Sink must be located in the area encircled by the boundary or on the boundary which is defined by function $B(r, \theta)$.

- An arbitrary line that goes through the sink must only have less than 2 intersection points with all the sub-boundaries in $\mathbb{B} - \mathbb{B}_0$, where $\mathbb{B}$ and $\mathbb{B}_0$ are defined as follows.

  Let's assume that the arbitrary line going through the sink is $L = \alpha$ where $0 \leq \alpha < 2\pi$. And, $\mathbb{B}$ is a set of all the sub-boundaries and $\mathbb{B}_0$ is a set of sub-boundaries that goes through the sink.

  $\mathbb{B} = \{b_0, b_1, \ldots, b_n\}$

  $\mathbb{B}_0 = \{b_k | b_k \in \mathbb{B}$ : and $b_k(0,0) = 0$, where $0 \leq k \leq n\}$, $\mathbb{B}_0 \subset \mathbb{B}$ and $|\mathbb{B}_0| \leq 2$

  $|\mathbb{B}_0| = 0$: Sink is located in the area encircled by the boundary $B(r, \theta)$.

  $|\mathbb{B}_0| = 1$: Sink is located on one of the sub-boundaries in $\mathbb{B}$.

  $|\mathbb{B}_0| = 2$: Sink is located on the intersection point of two sub-boundaries in $\mathbb{B}$.

When $|\mathbb{B}_0| > 2$, we cannot directly use our analytical method to compute the boundary function of the non-clustering zone. However, we use the concept of power line which is described in Section 3.5.6 to obtain a approximate computable model of the network.

**Indicator function for boundary**

We use indicator function $I_b$ which is shown in Eq. (3.6) to decide whether the given node is inside the boundary $B(r, \theta)$ or outside. For arbitrary node $i$ at $(r_i, \theta_i)$ is given by,

$$I_b(r_i, \theta_i) = \begin{cases} 1 & i \text{ is inside the boundary } B(r, \theta) \\ 0 & i \text{ is outside the boundary } B(r, \theta) \end{cases} \tag{3.6}$$

### 3.5.3 Hexagonal Element Based Analytical Model

In our previous research [46], we used circular elements to create a model for the network. In this paper, we use hexagonal elements as shown in Fig. 3.11 to simplify the calculations.

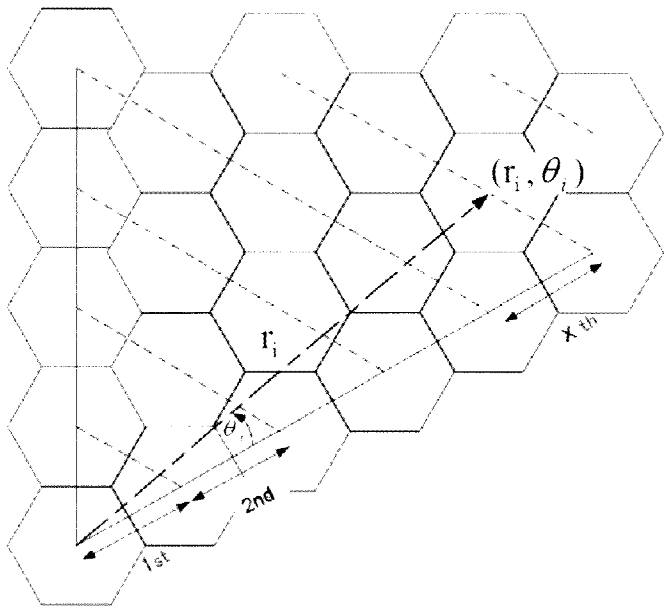



Figure 3.11: Hexagonal elements

The main advantage of using CPC with hexagonal elements is that methods used to define the boundary for non-clustering zone in 1/6 of the network ($0 \leq \theta < \pi/3$ on CPC) are repeatable for remaining 5 areas of the network. Therefore, hereinafter we only consider the area $0 \leq \theta < \pi/3$ and finally repeat the same methods for other areas. In our model, we use 2 different types of nodes which are high-end nodes and low-end nodes. Specifications of those nodes are shown in Table 3.3.

### Node expression and hop count

As shown in Fig. 3.11, CPC of arbitrary node $i$ is given by $(r_i, \theta_i)$, and $I_b(r_i, \theta_i) = 1$ then,
$(r_i, \theta_i) = \left( \frac{x r_s}{\cos(\alpha) + \frac{\sin(\alpha)}{\sqrt{3}}}, \alpha \right)$,
where $k = 0, 1, 2, .., x$ and $\alpha = \frac{\pi k}{3x}$.
Minimum hop count $x$ to sink $(0, 0)$ from $(r_i, \theta_i)$ is, $x = r_s \left( \cos(\theta) + \frac{\sin(\theta)}{\sqrt{3}} \right) / r_{H/L/S}$, where $r_{H/L/S} = r_H$ for high-end nodes, $r_{H/L/S} = r_L$ for low-end nodes, and $r_{H/L/S} = r_s$ for maximum number of hops in given distance.

Table 3.3: Description of the parameters

| | H-node | L-node |
|---|---|---|
| Number of nodes | $n_H$ | $n_L$ |
| Communication range [m] | $r_H$ | $r_L$ |
| Sensing range[m] | $r_s$ | $r_s$ |
| Initial energy [J] | $b_H$ | $b_L$ |
| Transmission power [W] | $e_{TH}$ | $e_{TL}$ |
| Receiving power [W] | $e_{RH}$ | $e_{RL}$ |
| Calculation cost [J/bit] | $e_c$ | $e_c$ |
| Aggregation ratio | $\alpha$ | $\alpha$ |
| Data transfer rate [kbps] | $\rho$ | $\rho$ |

### Number of the nodes in given area

By using Eq. (3.6), we calculate the number of the nodes within the given hop count $X_{min}$ to $X_{max}$.

For low-end nodes,

$$n_L(X_{min}, X_{max}) = \sum_{x=X_{min}}^{X_{max}} \sum_{k=0}^{x} I_b \left( \frac{x r_L}{\cos(k\alpha) + \frac{\sin(k\alpha)}{\sqrt{3}}}, \frac{k\pi}{3x} \right) \tag{3.7}$$

For high-end nodes,

$$n_H(X_{min}, X_{max}) = \sum_{x=X_{min}}^{X_{max}} \sum_{k=0}^{x} I_b \left( \frac{x r_H}{\cos(k\alpha) + \frac{\sin(k\alpha)}{\sqrt{3}}}, \frac{k\pi}{3x} \right) \tag{3.8}$$

### Average number of nodes inside a cluster

It is assumed that cluster $C_0$ is surrounded by other 6 clusters. In $C_0$, sensor nodes that are in last edge of the cluster also partially belong to other 6 clusters located around $C_0$. Therefore, total average number of nodes inside the cluster $n_C$ with clustering radius $r_C = jr_s$ is equal to,

$n_C(j) = 6(1+2+3+\ldots+j-1) + \frac{j}{2} \times 6 \ = \ 6(j-1)j/2 + \frac{j}{2} \times 6$,

$n_C(j) = 3j^2$.

Cluster radius $r_C$ for CH is, $r_C = \frac{r_{HorL}}{\sqrt{3}}$. For Low-end node CH, $j = r_L/\sqrt{3}r_s$ and for High-end node CH, $j = r_H/\sqrt{3}r_s$.

### 3.5.4 Spatial Boundary for Homogeneous Networks

In homogeneous networks, we use single type of sensor nodes. Let's assume that only L-nodes are in the network. Specifications of those nodes are shown in Table 3.3.

By using Eq. (3.6), we calculate the number of the nodes $n_L$within the given hop count $X_{min}$ to $X_{max}$.

Next, we calculate the average number of nodes inside a cluster. As mentioned in Section 3.5.3, total average number of nodes inside the cluster equals to $n_C = 3(\frac{r_L}{r_s})^2$.

We assume that non-clustering zone is in the area of hop $x = 1$ to $x_b$ whereas clustering zone is $x = x_b$ to $X_{MAX}$, where, in arbitrary point $(r_i, \theta_i)$ in $0 \leq \theta_i < \pi/3$, $X_{MAX} = max\left[\lceil\frac{r_i}{r_L}\rceil \cdot I_b(r_i, \theta_i)\right]$.

First, we calculate the total amount of data generated in clustering zone. We assume that every node generates $p[kbit]$ of data. Then, amount of data generated in single cluster $P_{cluster}$ is $P_{cluster} = \left(\lfloor\frac{r_L}{r_s}\rfloor^2 + 1\right)\alpha p$. Therefore, total amount of data generated in clustering zone $P_{Czone}$ is $P_{Czone} = P_{cluster} \times n_L\left(x_b, X_{MAX}\right)$.

For multi-hop communications in non-clustering zone, nodes does not aggregate the data. Each node in the route has to relay the data from clustering area in addition to sending its own $p$ packets of data to sink. Since non-clustering zone is in the area of hop $x = 1$ to $x_b$, average energy consumption $E_m$ is,

$$E_m = \sum_{x=1}^{x_b}\Bigg(\rho\frac{P_{Czone}}{x}(e_{RL} + e_{TL}) + \rho p e_{TL} x + \rho p(x-1)(e_{RL} + e_{TL})\Bigg)\frac{1}{n_L(1, \lceil\frac{x_b \cdot r_L}{r_s}\rceil)}. \tag{3.9}$$

Note that energy consumption in calculation unit is significant only in aggregation mode. Otherwise it can be negligible.

Then, we calculate the average energy consumption for clustering in non-clustering zone. Average energy consumption of a single node when performing clustering operation in non-clustering zone $E_c$ is,

$$E_c = \sum_{x=1}^{x_b} \Bigg( \rho \frac{P_{Czone}}{x}(e_{RL} + e_{TL}) + \rho P_{cluster} e_{TL} x \\ + \rho P_{cluster}(x-1)(e_{RL} + e_{TL}) \Bigg) \frac{1}{n_L(1, x_b)}. \tag{3.10}$$

Finally, from Eq. (3.9) and Eq. (3.10), we calculate the maximum $x_b$ which satisfies the $E_m \leq E_c$ and add the equation of $x_b$ to the boundary function of non-clustering zone $B_{m-hop}$ in range $0 \leq \theta < \pi/3$. By repeating the same procedure, we can calculate the other 5 factors of the $B_{m-hop}$ in range $\pi/3 \leq \theta < 2\pi$.

### 3.5.5 Spatial Boundary for Heterogeneous Networks

We assume that non-clustering zone is in the area of hop $x = 1$ to $x_b$ whereas clustering zone is $x = \frac{r_L x_b}{r_H}$ to $X_H$. In arbitrary point $(r_i, \theta_i)$, $X_H = \lceil \frac{max(r_i)}{r_H} \rceil$ and $I_b(r_i, \theta_i) = 1$.

First, we calculate the total amount of data generated in clustering zone. We assume that every node generates $p[kbit]$ of data. Then, amount of data generated in single cluster $P_{cluster}$ is $P_{cluster} = \left(3\lfloor \frac{r_H}{\sqrt{3} r_L} \rfloor^2 + 1\right) \alpha p$. Therefore, total amount of data generated in clustering zone $P_{Czone}$ is $P_{Czone} = P_{cluster} \times n_H \left( \lceil \frac{r_L x_b}{r_H} \rceil, X_H \right)$.

For, multi-hop communication in non-clustering zone, we assume that non-clustering zone only contains low-end nodes and does not aggregate the data. Each node in the route has to relay the data from clustering area and also sends its own $p$ packets of data to sink. Since non-clustering zone is in the area of hop $x = 1$ to $x_b$, average energy consumption $E_m$ is,

$$E_m = \sum_{x=1}^{x_b} \Bigg( \rho \frac{P_{Czone}}{x}(e_{RL} + e_{TL}) \\ + \rho p e_{TL} x + \rho p(x-1)(e_{RL} + e_{TL}) \Bigg) \frac{1}{n_L(1, x_b)}. \tag{3.11}$$

Note that energy consumption in calculation unit is significant only in aggregation mode. Otherwise it can be negligible.

Then, we calculate the average energy consumption for clustering in non-clustering zone. For this calculation, we assume that non-clustering zone contains high-end nodes and low-end nodes. Therefore, average energy consumption of a single node when performing clustering operation in non-clustering zone $E_c$ is,

$$E_c = \sum_{x=1}^{\lceil \frac{x_p r_L}{r_H} \rceil} \Bigg( \rho \frac{P_{Czone}}{x}(e_{RH} + e_{TH}) + \rho P_{cluster} e_{TL} x \\ + \rho P_{cluster}(x-1)(e_{RL} + eTL) \Bigg) \frac{1}{n_L(1, x_b) + n_H(1, \lceil \frac{x_p r_L}{r_H} \rceil)}. \quad (3.12)$$

Finally, from Eq. (3.11) and Eq. (3.12), we calculate the maximum $x_b$ which satisfies the $E_m \leq E_c$ and add the equation of $x_b$ to the boundary function of non-clustering zone $B_{m-hop}$ in range $0 \leq \theta < \pi/3$. By repeating the same methods, we can calculate the other 5 factors of the $B_{m-hop}$ in range $\pi/3 \leq \theta < 2\pi$.

### 3.5.6 Extension for Non-computable Network Topologies

When $|\mathbb{B}_0| > 2$, we cannot directly compute the spatial boundary of the non-clustering zone. Fig. 3.12 shows an example of non-computable network topology. Line segments AS, BS, and CS have 1, 3, and 3 intersection points with boundary function, respectively. According to the conditions of the boundary function, the area colored in gray cannot be included into the computation. Therefore, we assume that equivalent amount of data from gray area is generated on line segment DB named power line. Then, we calculate the total amount of data generated in the gray area and consider it when we calculate the boundary of the non-clustering zone in 1/6 of the network which includes the power line.

## 3.6 Performance Evaluation

### 3.6.1 Simplified Spatial Boundary Analysis Method

We evaluated the performance of our proposed method via network simulator ns-2 [47] and observed the percentage of alive nodes as a performance measure. Definition of a alive node is a node which is able to send the sensed data to the sink successfully. Therefore, disconnected nodes are not considered as alive nodes.

Every sensor node generate the data in iteration of 20 minutes. However, every node have a random back off time of maximum 1 minutes before process the sensed data and prepare them

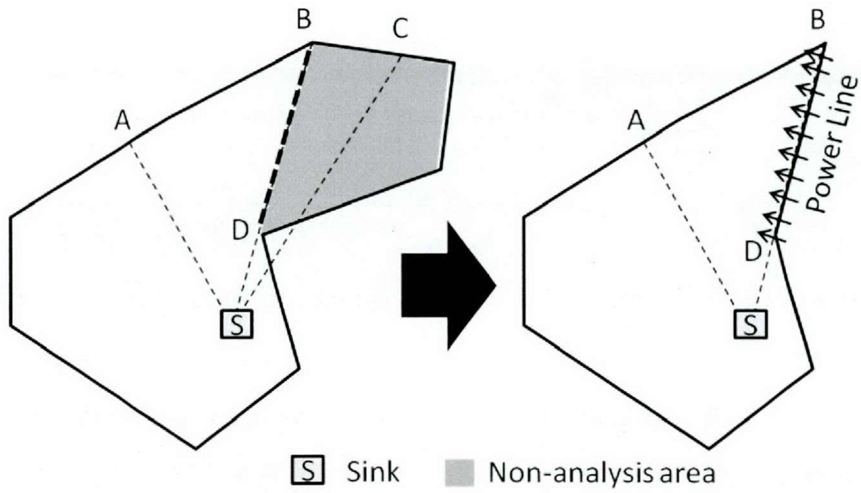



Figure 3.12: Power line

Table 3.4: Node specifications for simulations

| | H-node | L-node |
|---|---|---|
| Number of nodes | 18 | 180 |
| Communication range [m] | 30 | 12 |
| Initial energy [J] | 43.55 | 5.67 |
| Aggregation ratio | 0.11 | 0.11 |
| Data rate [kbps] | 250 | 250 |
| Calculation cost [J/byte] | $2.73 \times 10^{-6}$ | $2.73 \times 10^{-6}$ |

for transmission. We used S-MAC protocol as MAC protocol. S-MAC codes are provided by ns2 simulator. We used same setting for all the simulation experiments in this chapter.

In this simulation model, we set the node configurations using datasheet and information provided by Crossbow [48]. Total network area was set to a circle with radius $R = 80[m]$, clustering factor was $\phi = 0.75$. Other node specifications were shown in Table 3.4.

According to those node specifications, the radius of the non-clustering zone $d$ became $38.07[m]$. We conducted the performance evaluation in our proposed method and compared it with clustering network which adapted the method described in [44] and multi-hop network. Each simulation result is average over 50 different network topologies.

The simulation result is shown in Fig. 3.13, which is alive node ratio in $d = 38.07[m]$. It shows that the proposed method achieves to prolong the lifetime of WSN with heterogeneous node types.

In addition, to evaluate the validity of our analysis model, we observed the alive node ratio with different values of the radius of multi-hop area. In Fig. 3.14, "$d + 0$" means the case using the value derived by our analysis. From the graph, "$d + 0$" outperforms other cases. In other words, it proves that our analysis is valid.

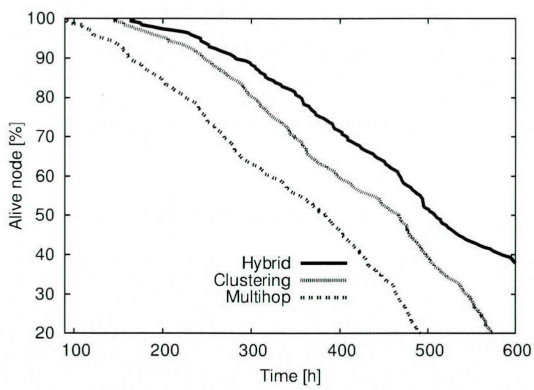



Figure 3.13: Network lifetime

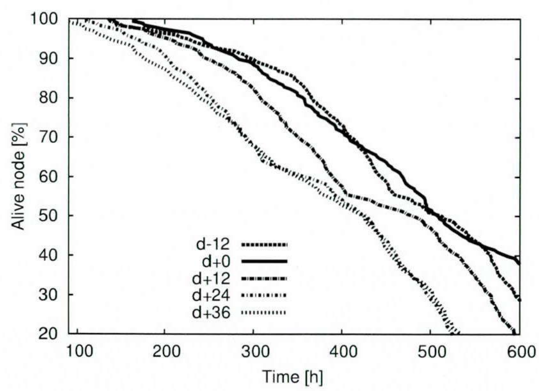



Figure 3.14: Validity of analysis model

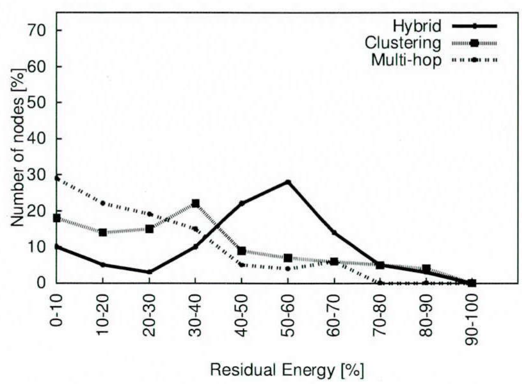



Figure 3.15: Residual energy when alive node ratio is reached to 99.9%

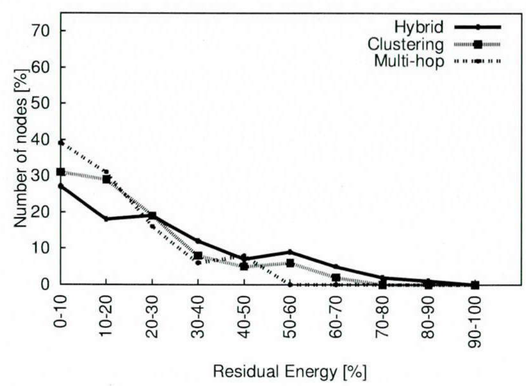



Figure 3.16: Residual energy when alive node ratio is reached to 80%

We also observed the residual energy of the nodes in certain alive node ratio points. When the alive node ratio reached to 99.9%, 80%, 60%, we measured the $\frac{Residual\ Energy}{Initial\ Energy}[\%]$ for each nodes. The simulation results for alive node ratio reached to 99.9%, 80%, 60% are shown in Fig. 3.15, Fig. 3.16, and Fig. 3.17, respectively. Compared with clustering and multi-hop

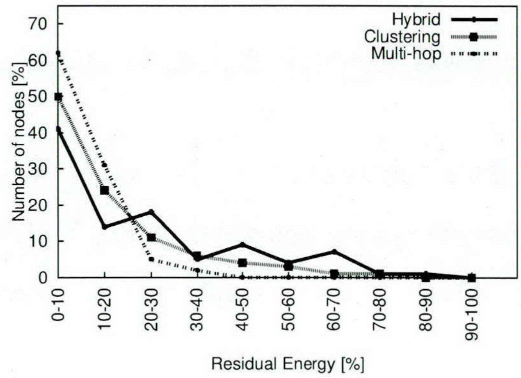



Figure 3.17: Residual energy when alive node ratio is reached to 60%

methods, the proposed method is able to maintain higher number of nodes with medium level of residual energy and lower number of nodes with low level of residual energy. This result confirms that our proposal method can maintain balanced energy consumption in the network.

In our next experiment, we increase the the number of high-end nodes to 60 and that of the low-end nodes to 540. Other node specifications are same as in Table 3.4. We used two topologies; topology 1 is shown in Fig. 3.18 and topology 2 in Fig. 3.19.

In our proposed method, we assume that total network area is circle and used the radius of the circle $R$ to evaluate the spatial boundary for non-clustering zone. However, in topology 1 and 2, we cannot obtain a direct value for the $R$. Therefore, we used two methods to obtain an approximate vale for $R$.

**Method 1**

We calculate the average distance from each node to sink and use it as $R$. In Figs. 3.18 and 3.19, non-clustering zone calculated by the method 1 marked as $R = Node\_Avg.$. Also in Figs. 3.20 and 3.21, network lifetime marked as $R = Node\_Avg.$ for the method 1.

**Method 2**

We consider $R = max\{distance\ from\ node_i\ to\ sink\}$, where $node_i$ is an arbitrary node in the network. In Figs. 3.18 and 3.19, non-clustering zone calculated by the method 2 marked as $R = Node\_Max$. Also in Figs. 3.20 and 3.21, network lifetime marked as $R = Node\_Max$ for method 2.

Network lifetime for topology 1 and 2 are shown in Figs. 3.20 and 3.21, respectively. Each simulation result is average over 50 different node deployments.

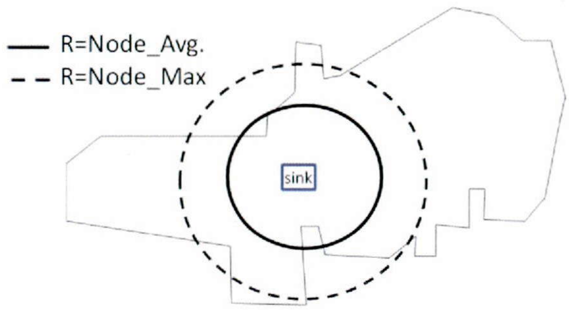



Figure 3.18: Topology 1

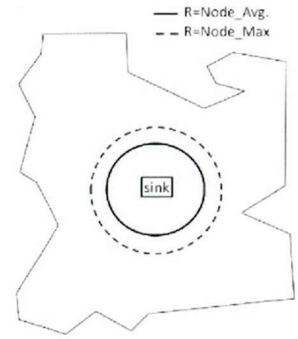



Figure 3.19: Topology 2

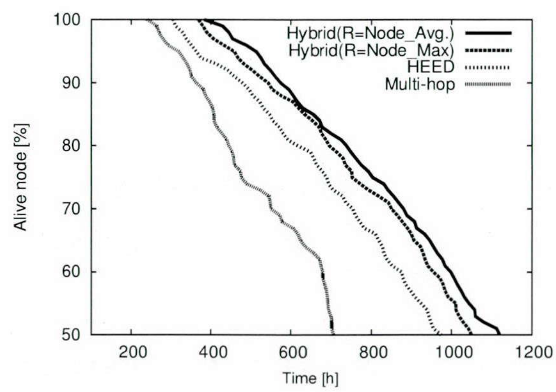



Figure 3.20: Network lifetime in topology 1

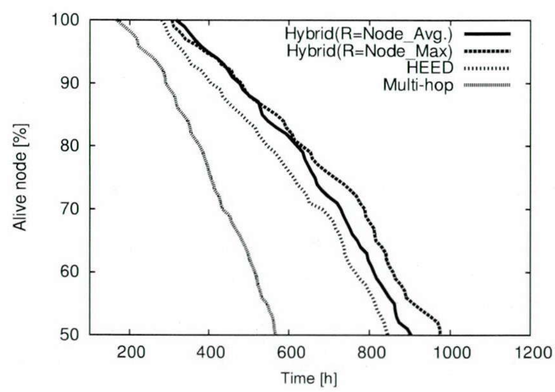



Figure 3.21: Network lifetime in topology 2

Through these experiments, it is verified that the proposed hybrid methods able to maintain higher alive node ratio than that of the pure clustering and multi-hop methods in any network topology.

### 3.6.2 Generalized Spatial Boundary Analysis Method

We evaluate the performance of the proposed method via network simulator ns-2 and observe the percentage of alive nodes. Definition of an alive node is a node which is able to send the sensed data to the sink successfully. Therefore, disconnected nodes are not considered as alive nodes. Specifically, we compare it with those of total clustering and total multi-hop networks. Each simulation result is averaged over 50 different node deployments in a given network topology. In this simulation, we set the node configuration using datasheet and information provided

Table 3.5: Specifications of the nodes

| | H-node | L-node |
|---|---|---|
| Communication range [m] | 30 | 12 |
| Initial energy [J] | 43.55 | 5.67 |
| Calculation cost [J/bit] | $2.73 \times 10^{-6}$ | $2.73 \times 10^{-6}$ |
| Aggregation ratio | 0.25 | 0.25 |
| Data transfer rate [kbps] | 250 | 250 |
| Sensing range [m] | 3 | 3 |

by Crossbow [48] and those values are shown in Table 3.5.

In our first experiment, we used heterogeneous network. In this experiment, we use 30 H-nodes and 220 L-nodes. Node specifications are shown in Table 3.5. Total network area is set to a circle with radius $R = 96[m]$ and center(48,0), sink located at (0,0). By method mentioned in Section 3.5.5, the boundary function of total network area is $B(r, \theta) = r^2 - 96r\cos(\theta - \pi) + 96^2 - r^2$.

According to those node specifications, the boundary function of the non-clustering zone $B_{m-hop}$ is shown in Eq. (3.13). We conduct the performance evaluation in the proposed method and compare it with clustering network which adopts the method described in [44], and total multi-hop network.

$$B_{M-hop} = \begin{cases} L[(24,0),(36,\pi/3)] \ (\ 0 \le \theta < \pi/3) \\ L[(36,\pi/3),(48,2\pi/3)] \ (\ \pi/3 \le \theta < 2\pi/3) \\ L[(48,2\pi/3),(60,\pi)] \ (\ 2\pi/3 \le \theta < \pi) \\ L[(60,\pi),(48,4\pi/3)] \ (\ \pi \le \theta < 4\pi/3) \\ L[(48,4\pi/3),(36,5\pi/3)] \ (\ 4\pi/3 \le \theta < 5\pi/3) \\ L[(36,5\pi/3),(24,2\pi)] \ (\ 5\pi/3 \le \theta < 2\pi) \end{cases} \tag{3.13}$$

The simulation result, shown in Fig. 3.22 indicates the alive node ratio of the WSN. It shows that the proposed method achieves to prolong the lifetime of WSN with heterogeneous node types for longest time compared to the other methods.

In our second and third experiment, we used topologies with irregular network boundaries.

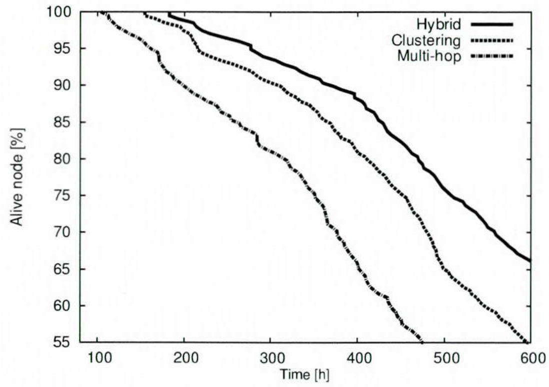



Figure 3.22: Simulation results in heterogeneous network

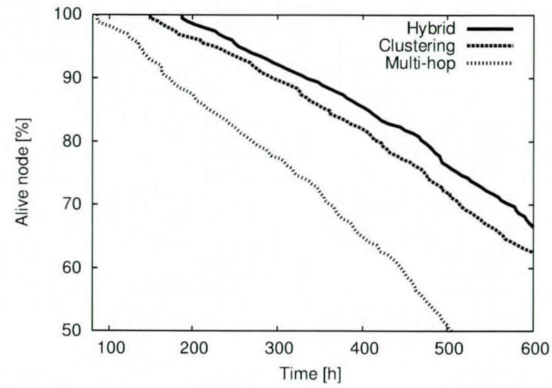



Figure 3.23: Simulation results in homogeneous network

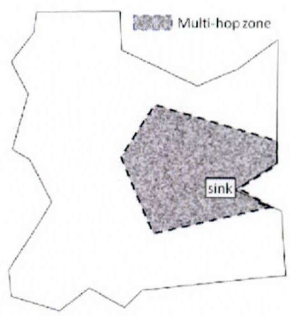



Figure 3.24: Simulation model 1

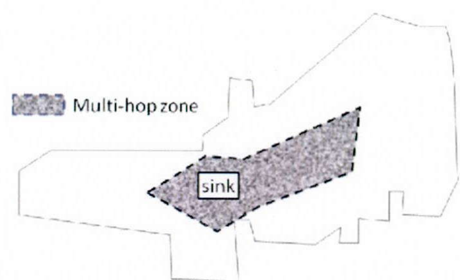



Figure 3.25: Simulation model 2

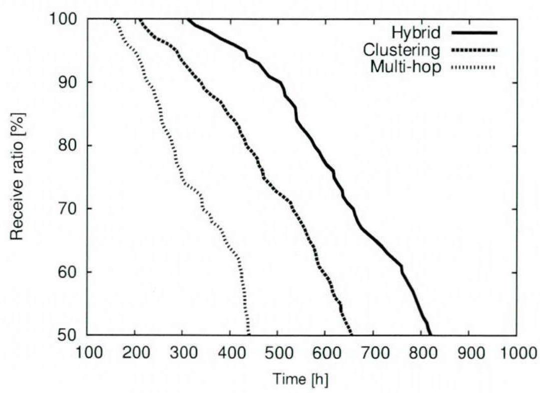



Figure 3.26: Alive node ratio on model 1

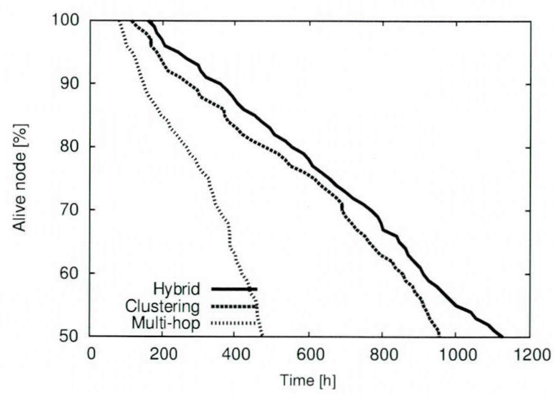



Figure 3.27: Alive node ratio on model 2

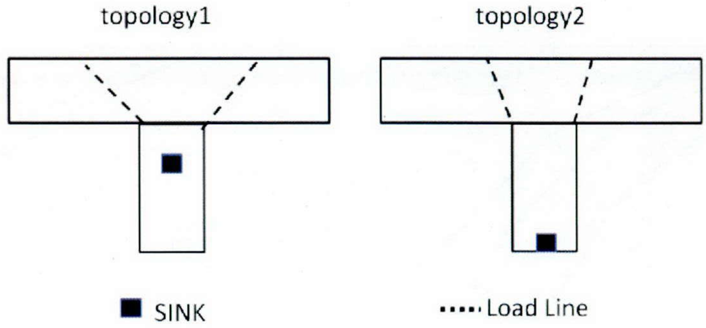



Figure 3.28: Simulation model 3 topology 1 and 2

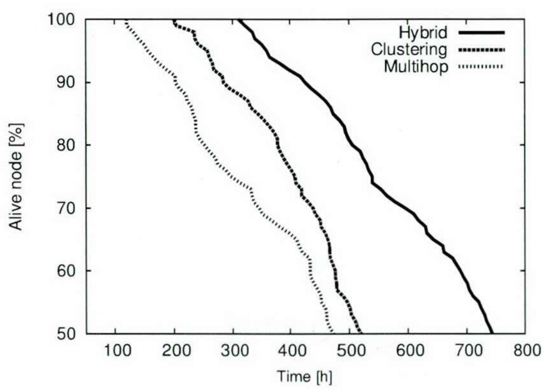



Figure 3.29: Alive node ratio on model 3 topology 1

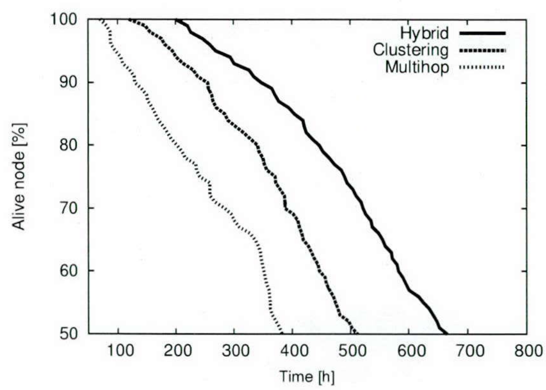



Figure 3.30: Alive node ratio on model 3 topology 2

For our second experiment, we used topology shown in Fig. 3.24 and topology shown in Fig. 3.25 used in third experiment. Figs. 3.24 and 3.25 also show the shape of the boundary functions on corresponding network topologies. In these experiments, we used 60 H-nodes and 540 L-nodes.

Sink is located on the edge of the network in simulation model 1 shown in Fig. 3.24. Comparing with the simulation model 2 in Fig. 3.25, simulation model 1 has uniform relay traffic from wide side on the network. Simulation model 2 is a narrow long network and sink is located on the center of the network. However, the relay traffic from the horizontal direction is much larger than that of the vertical direction. In order to handle such relay traffic, the proposed method forms a narrow long non-clustering zone around the sink. This indicates the accuracy of analytical method to estimate the optimal spatial boundary for non-clustering zone around the sink.

The simulation result, shown in Figs. 3.26 and 3.27 indicates the alive node ratio of the WSN with irregular network boundaries. Also, Fig. 3.31 shows the same simulation results as Fig. 3.26 with average taken by 500 simulation samples. From there graphs, partialized traffic patterns in simulation model 2 shortened the network lifetime compared with simulation model 1. Overall result, however, shows that the proposed method achieves to prolong the lifetime of WSN with heterogeneous node types for longest time compared to the other methods.

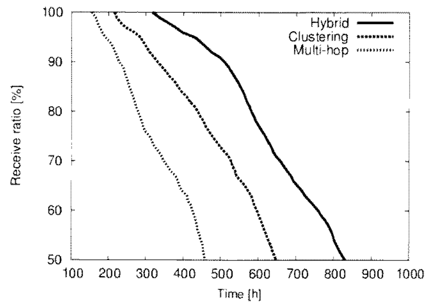



Figure 3.31: Alive node ratio on model 1 with 500 samples

In addition to our second experiment, we evaluated the performance of homogeneous network on topology shown in Fig. 3.24. However, the shape of the boundary function was a different one from the one in Fig. 3.24. We used 600 single type of nodes and specifications are shown under L-node column in Table 3.5. We derived the boundary function by using the method mentioned in Section 3.5.4.

The simulation result, shown in Fig. 3.23, indicates the alive node ratio as a function of elapsed time. It shows that the proposed method achieves to prolong the lifetime of WSN with homogeneous node types for longest time compared with pure clustering [44] or pure multi-hop methods. However, superiority of the heterogeneous network scenario is more significant than in homogeneous network scenario.

In our forth and fifth experiments, we evaluate the validity of power line concept described in Section 3.5.6. We used two topologies; topology 1 and topology 2 in Fig. 3.28. In these experiments, we used 60 H-nodes and 540 L-nodes. Alive node ratio of the topology 1 and topology 2 is shown in Figs. 3.29 and 3.30 respectively.

In topology 1 and topology 2 in Fig. 3.28, we cannot directly compute the spatial boundary of the non-clustering zone. Therefore, we used the power line concept to estimate the optimal spatial boundary for the non-clustering zone. Proposed method with the power line concept indicates the higher alive node ratio than that of the total clustering and total multi-hop networks.

Through all the simulation results shown above indicate that the proposed method achieves to prolong the lifetime of WSN with homogeneous or heterogeneous node types for longest time compared to the other methods.

## 3.7 Conclusion

In this paper, in order to prolong the lifetime of heterogeneous WSNs, we proposed a robust hybrid routing method which adaptively combines clustering and multi-hop communication methods.

We described the network model and mathematical analysis method to define the optimal area of the non-clustering zone. Then, we explained inter-cluster, intra-cluster routing methods and routing in the boundary of non-clustering zone and clustering zone. Furthermore, we showed the excellent performance of our proposed method compared with pure clustering network and multi-hop network. In addition, we fluctuated the radius of multi-hop area to confirm robustness of the proposed method.

The proposed method has the similar network topology with UC when UC has size 1 clusters (cluster with a 1 node) near to the sink. However, UC performs a single hop non-clustering zone and it may not be the optimal size of non-clustering zone for each and every network. Furthermore, UC is designed for homogeneous WSNs and it has no efficient way to use the advantage of the node's performance heterogeneity. On the other hand, the proposed method provides the optimal area for the non-clustering zone by considering the specifications of the nodes and the characteristics of the networked like outer boundary and location of the sink. The proposed method efficiently avoids the non-uniform energy consumption around the sink in heterogeneous WSNs.

# Chapter 4

# Dynamic Reconfiguration of Spatial Boundary

## 4.1 Introduction

In this chapter, we propose a dynamic reconfiguration method for the proposed spatial boundary of clustering non-clustering hybrid routing method. In the proposed methods in chapter 3, we assumed that nodes were uniformly distributed in the network. However, in practical networks nodes are deployed in a non-uniform manner. In addition, the residual power of each node decreases with time elapsing. Therefore, pre-defined spatial boundary is not optimal for all practical networks. Moreover, in this method sensor nodes near to the spatial boundary tend to have higher data relay traffic. Therefore, as the network time elapsed, the pre-defined spatial boundary should be reconfigured. In this chapter we propose to dynamically re-define the spatial boundary by considering the variance of the residual energy of a boundary node and its adjacent nodes. Simulation results show that our method increases network lifetime in comparison to the existing method with static boundary.

## 4.2 Related Work

As we introduced in previous section, CHs around the sink have high relay traffic. Due to that traffic, those nodes consume more energy than the other CHs. Increasing the low residual energy nodes around sink shortens the network lifetime. Therefore, our previous method [46], we focused on load balancing around sink in order to reduce the number of nodes with low residual energy.

We assumed that every node can be aware of its location information. It is a reasonable assumption, since recent researches have proposed efficient methods to obtain the location information with high accuracy.

In the proposed method, nodes placed near to sink perform the operation of a non-clustering network (Non-clustering zone) and nodes far away from sink perform clustering operation (Clustering Zone). Before performing them, operation boundary for non-clustering zone must be determined. Therefore, first we analyzed the network and defined the boundary for non-clustering zone. Outside of that zone, we define as clustering zone. First, we calculate the average energy consumption for a node in non-clustering zone; $E_m$ when performing multi-hop communication in non-clustering zone and $E_c$ when clustering in non-clustering zone. Then considering $E_m \leq E_c$, we calculate the boundary function for non-clustering zone $B_{m-hop}$ which defines the spatial boundary between non-clustering and clustering zones.

First, the boundary function of non-clustering zone $B_{m-hop}$ is determined by using the method explained in above. Then, sink broadcasts the information to network. Each node in the network enables to receive the message including the boundary information and evaluates whether it is in the non-clustering zone or clustering zone upon its location information and $B_{m-hop}$. Then, the nodes in the clustering zone form clusters by using the method explained in [44] and forward the data to non-clustering zone. When nodes in the non-clustering zone receive the data, they forward the data to the sink by using multi-hop communication. Furthermore, CH which is able to directly communicate with nodes in non-clustering zone is referred to as Boundary CH (BCH). Also, inside the non-clustering zone,a node which is on the edge of the non-clustering zone is referred as boundary node. Each BCH manages a list of next hop nodes located in non-clustering zone and assigns a probability in proportion to residual energy divided by initial energy to each node. When BCH forwards the data to non-clustering zone, it selects a next node according to this probability.

We conducted simulation experiments to study the behavior of the static spatial boundary method in [46]. In this experiment, we randomly deployed 180 low-end nodes and 20 high-end

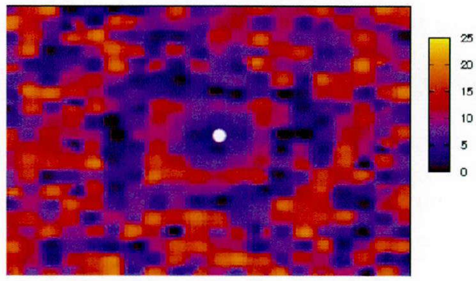



Figure 4.1: Energy distribution on hybrid network at 135[h]

nodes into $60[m] \times 60[m]$ area. Specifications of those nodes are shown in Table 4.1. Fig. 4.1 shows the residual energy distribution of the network in network time at $150[h]$. Bright areas represent the high residual energy and dark areas represent low residual energy. Through performance evaluation, it is verified that our proposed hybrid method increases network lifetime by $150\%$ and $60\%$ more than that of the non-clustering and, clustering respectively in heterogeneous networks.

In this method, we assumed that nodes are uniformly distributed in the network. However, in practical networks nodes are deployed in a non-uniform manner. In addition, the residual power of each node decreases with time elapsed. Therefore, pre-defined spatial boundary not be the optimal for all practical networks. Moreover, in this method sensor nodes near to the spatial boundary tends to has higher data relay traffic. In Fig. 4.1, dense dark spots near to the spatial boundary are the boundary nodes which relay the data from the clustering zone to non-clustering zone. Those nodes residual energy become critical as the network time elapse due to high relay traffic. Therefore, as the network time elapsed, the pre-defined spatial boundary should be reconfigured.

## 4.3 Dynamic Reconfiguration of Spatial Boundary

As we introduced in previous section, nodes around the spatial boundary have high relay traffic. Due to that traffic, those nodes consume more energy than the nodes in the non-clustering zone. Increasing the low residual energy nodes around spatial boundary shortens the network lifetime. Also, pre-defined spatial boundary not be the optimal for all practical networks. Therefore, in this chapter we propose to dynamically re-define the spatial boundary by considering the variance of the residual energy of a boundary node and its adjacent nodes.

We assume that every node can be aware of its location information. It is a reasonable assumption, since recent researches have proposed efficient methods to obtain the location information with high accuracy. For example, [30] proposes to perceive both the anchors' positions and neighbor node information by Neighbor Position-based Localization Algorithm that can greatly enhance the positioning accuracy compared with conventional overlapping connectivity localization algorithms. In addition, [45], proposed a method in which a probabilistic model of wireless propagation characteristics in a real environment is constructed, and the node position is estimated by maximum-likelihood estimation using the model.

At first, we use hybrid routing method explained in [46] to create WSN.In this research we focus on heterogeneous WSN. In this Nodes placed near to sink perform the operation of a non-clustering network (non-clustering zone) and nodes far away from sink perform clustering operation (Clustering Zone). We assume that after the time $T_{static}$ from the network start, dynamic spatial boundary mechanism will start. We also assume that, every node keeps a record of average amount of data relay per one iteration ($P_j$) and the probability of get selected as a relay node($k_j$) in one iteration.

### 4.3.1 Variance of the Residual Energy

We focus on a single boundary node along the spatial boundary. Example of the focused boundary node and its adjacent nodes are shown in Fig. 4.2. Inside the non-clustering zone, nodes which are on the edge of the non-clustering zone is referred as boundary nodes. We focus on of the boundary nodes $n_B$ as shown in Fig. 4.2. Node $n_{m1}, n_{m2}, ...., n_{mi}$ are all the nodes in the non-clustering zone which are able to directly communicate with $n_B$ except the other boundary nodes. Node $n_{c1}$ and $n_{c2}$ are the adjacent nodes of $n_B$ in the clustering zone.

Let $N$ denote the set of these nodes where $N = \{n_{m1}, n_{m2}, ...n_{mi}, n_{c1}, n_{c2}\}$. We define the area which covered by all the nodes in $N$ as region of $n_B$ represent with the notification $R_{n_B}$. In Fig. 4.2, $R_{n_B}$ is indicated by broken lines.

When the network time is $T_{static}$, node $n_B$ calculates the variance of the residual energy ratio in the region $R_{n_B}$. Also, $n_i$ ($n_i \in N$) is an arbitrary node in the region $R_{n_B}$ and $i$th row node in the set $N$.

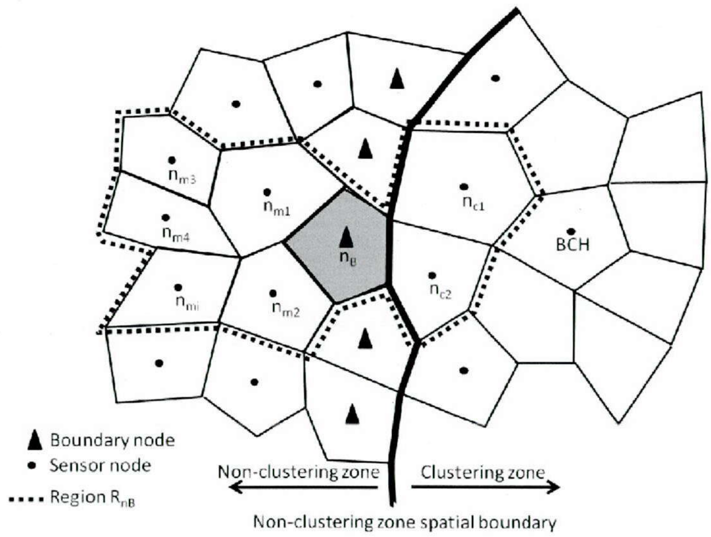



Figure 4.2: Boundary node and its region

$$\sigma_{nB}^2 = \frac{1}{\mid N \mid} \sum_{i=1}^{|N|} (E_i - E_{avg.})^2. \tag{4.1}$$

$\mid N \mid$ is the number of nodes in set $N$ and $E_i$ is ratio of the residual energy to initial energy of the node $n_i$. Average residual energy of the nodes in set $N$ is $E_{avg.}$ and $\sigma_{nB}^2$ in Eq. (4.1) is the variance of the residual energy of the region $R_{n_B}$.

Also, node $n_B$ requests to its adjacent nodes to calculate the variance of residual energy of their own regions and report back to node $n_B$. However, other boundary nodes adjacent with node $n_B$ are not considered as adjacent nodes of node $n_B$.

### 4.3.2 Handing Over the Role of Boundary Node

Node $n_B$ considers the variance of the residual energy that acquired from adjacent nodes to select the best node to hand over the role as a boundary node. As described in 4.2, the non-clustering zone, residual energy of the one hop away nodes are used to probabilistically select the next hop. Therefore, region which consist of nodes with similar residual energy levels should have more data relay traffic for load balancing. As a result, node $n_B$ hand over its role as a boundary with node which has the region with lowest variance of residual energy. When node $n_B$ has handed over the role as a boundary node, node $n_B$ assigned to clustering zone or non-clustering zone.

For example, in Fig. 4.2, if node $n_B$ decides to hand over its role as a boundary node to node $n_{m1}$, then node $n_B$ is assigned to clustering zone and becomes a member of the optimal cluster by following the method in [44]. In addition, $n_{m1}$ becomes a new boundary node. If node $n_B$ decides to hand over its role as a boundary node to node $n_{c1}$, then the node $n_B$ assign to the non-clustering zone and follow the multi-hop routing method described in previous section. In addition, node $n_{c1}$ is left from the clustering zone, join with non-clustering zone. Then, node $n_{c1}$ becomes a new boundary node.

When node $n_B$ has selected the optimal node for hand over the role of the boundary node, it informs the value of $\sigma^2_{nB}$ to the selected node.

Next, newly selected boundary node estimates the number of iterations till the next boundary node reconfiguration process. Every node keeps a record of average amount of data relay per one iteration ($P_j$) and the probability of be selected as a relay node($k_j$) in one iteration. Newly selected boundary node obtains those information and uses them to estimate the variance of the residual energy of the region after $I$ number of iterations.

Nodes in the region $R_{n_B}$ can be categorized as nodes in the non-clustering zone and nodes in the clustering zone. For nodes in the clustering zone, those nodes are in the edge of the cluster and only have to send the sensed data to CH. Therefore, for a arbitrary node $j$, residual energy ratio $E_j^I$ after $I$ iteration is,

$$E_j^I = E_j - I\alpha\rho e_T. \tag{4.2}$$

$E_j$ is the current residual energy and $\alpha$ is the data transfer rate, $\rho$ is the size of the single data packet and $e_T$ is the energy consumption per transmitting a packet.

For the nodes in the non-clustering zone has to send the sensed data to sink and relay the data traffic from clustering zone to sink. Therefore, for arbitrary node $n_j$, residual energy ratio $E_j^I$ after $I$ iteration is,

$$E_j^I = E_j - I\alpha\rho e_T - Ik_iP_i\alpha(e_R + e_T). \tag{4.3}$$

$n_j$, $P_j$ and $k_J$ means arbitrary node, average amount of data relay per one iteration and probability of get selected as a relay node in one iteration, respectively. $e_R$ is the energy consumption per receiving a packet. By using Eqs. (4.2) and (4.3), node estimates the variance of the residual energy $\sigma_I^2$ of the region after $I$ iterations. Then, the estimated $\sigma_I^2$ is compared with $\sigma_{n_B}^2$ and determines the maximum $I$ for $\sigma_I^2 \leq \tau\sigma_{n_B}^2$.

After a node completes the role of the boundary node for $I$ iterations, it starts the process from section4.3.1 to select a new node to hand over the role of the boundary node.

In order to select a best node to hand over the role as a boundary node, node has to exchange the residual energy information between adjacent nodes. However, size of the control data packet which use to exchange the residual energy information are much smaller than the size of the actual data packet. Furthermore, the control packet exchange occurs only when a boundary node wants to handing over the role of boundary node to another node. Therefore, energy consumption for control packet exchange is not significant.

## 4.4 Performance Evaluation

Table 4.1: Specifications of the nodes

| | H-node | L-node |
|---|---|---|
| Communication range [m] | 30 | 12 |
| Initial energy [J] | 43.55 | 5.67 |
| Calculation cost [J/bit] | $2.73 \times 10^{-6}$ | $2.73 \times 10^{-6}$ |
| Aggregation ratio | 0.25 | 0.25 |
| Data transfer rate [kbps] | 250 | 250 |
| Sensing range [m] | 3 | 3 |

We evaluated the performance of the proposed method via network simulator ns-2 and observed the percentage of alive nodes. Definition of a alive node is a node which is able to send the sensed data to the sink successfully. Therefore, disconnected nodes are not considered as alive nodes. Specifically, we compare it with those of total clustering and total non-clustering networks and hybrid routing method [46]. Each simulation result is averaged over 50 different node deployments in a given network topology.

Every sensor node generate the data in iteration of 20 minutes. However, every node have a random back off time of maximum 1 minutes before process the sensed data and prepare them for transmission. We used S-MAC protocol as MAC protocol. S-MAC codes are provided by ns2 simulator. We used same setting for all the simulation experiments in this chapter. In this simulation, we set the node configuration using datasheet and information provided by Crossbow [48].

In our simulation experiment, we used heterogeneous network. In this experiment, we used 49 H-nodes and 400 L-nodes. Node specifications are shown in Table 4.1. Total network area is $100[m] \times 100[m]$, sink located at center of the topology. We assumed, sensing range is circle with $3[m]$ radius of for both H-nodes and L-nodes. Time till start the dynamic configuration $T_{static}$ is set to $100[h]$.

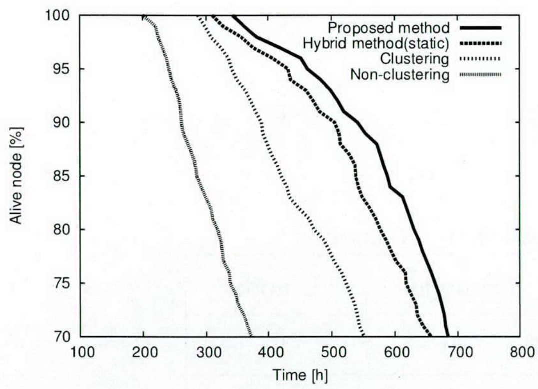



Figure 4.3: Alive node ratio in heterogeneous network

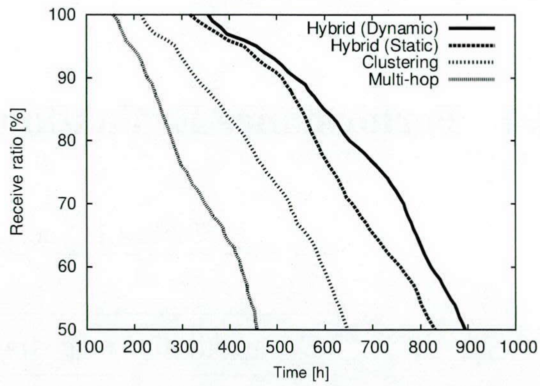



Figure 4.4: Alive node ratio in simulation model 1 with 500 samples

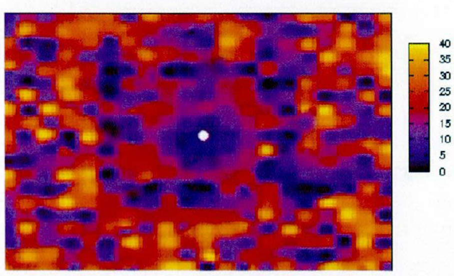



Figure 4.5: Without dynamic reconfiguration

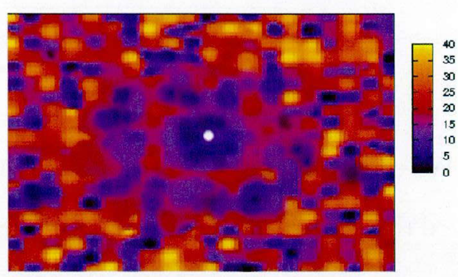



Figure 4.6: With dynamic reconfiguration

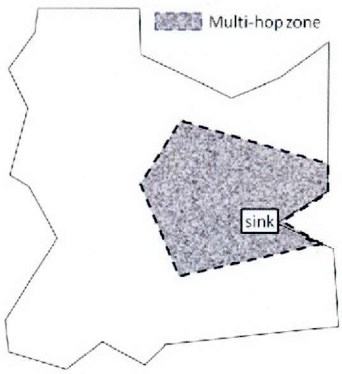



Figure 4.7: Simulation model 1

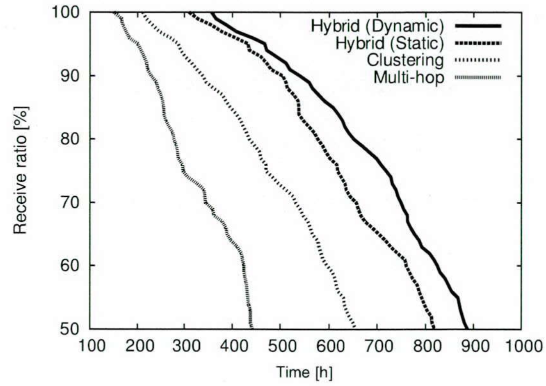



Figure 4.8: Alive node ratio in simulation model 1

The simulation result, shown in Fig. 4.3 indicates the alive node ratio as a function of elapsed time. It shows that the proposed method achieves to prolong the lifetime of sensor network with heterogeneous node types for longest time compared to the other methods.

For further study, we examined a one sample of the above 50 simulations and visualized the distribution of the residual energy ratio. Figs. 4.5 and 4.6 shows the distribution of the residual energy ratio in hybrid routing method with static spatial boundary (without dynamic reconfiguration) and the hybrid routing method with dynamic reconfiguration of spatial boundary in network time at $300[h]$, respectively. Bright areas represent the high residual energy and dark areas represent low residual energy.

In Fig. 4.5, dark spots near to the spatial boundary are the boundary nodes which relay the data from the clustering zone to non-clustering zone. Those nodes residual energy become critical as the network time elapse due to high relay traffic. On the other hand, in Fig. 4.6 shows that proposed method has no darker spots near to spatial boundary.

In our next experiment, we used topology with irregular network boundaries as shown in Fig. 4.7. Fig. 4.7 also show the shape of the boundary functions on corresponding network topology. In this experiments, we used 60 H-nodes and 540 L-nodes.

Sink is located on the edge of the network in simulation model 1 shown in Fig. 4.7 and Fig. 4.4 shows the same simulation results with average of 500 samples. Simulation model 1 has uniform relay traffic from wide side on the network. However, the relay traffic from the horizontal direction is much larger than that of the vertical direction. In order to handle such relay traffic, the proposed method forms a narrow long non-clustering zone around the sink.

The simulation result, shown in Fig. 4.8 indicates the alive node ratio of the WSN with irregular network boundaries. According to the alive node ratio shown in Fig. 4.8, until the network time reaches to $500[h]$, the proposed method and hybrid routing method with static spatial boundary have only a slight different between alive node ratios. However, when after the alive node ratio dropped below the 90% (at about $500[h]$), the proposed method ables to maintain much higher alive node ration than the hybrid routing method with static spatial boundary. This result shows that our proposed method is more effective for extend the network lifetime in low alive node ratio situations.

Overall result, however, shows that the proposed method achieves to prolong the lifetime of sensor network with heterogeneous node types for longest time compared to the other methods.

The simulation result, shown above indicates that the proposed method achieves to well distributed relay traffic load by dynamic re-configuring the spatial boundary.

## 4.5 Conclusion

In this chapter, in order to maximize the network lifetime of WSN, we introduced a new dynamic reconfiguration method for spatial boundary in clustering and non-clustering hybrid WSN.

We described the dynamic reconfiguration of the spatial boundary by considering the variance of the residual energy of the boundary node and its adjacent nodes. Then, we described the method of handing over the role of a boundary node to adjacent node. Furthermore, we evaluated the performance of the proposed method and compared it with hybrid network with static spatial boundary, clustering network and non-clustering network.

As a future work, we enhance the proposed method by considering the effect of dynamic changes of adjacent boundary nodes.

# Chapter 5

# Conclusions

This dissertation has studied an adaptive routing protocol for heterogeneous wSNs.

In recent years, Wireless Sensor Networks (WSNs) which consists of a large number of sensor nodes having wireless communication and self-organizing capability receives much attention. Sensor nodes are tiny, battery powered modules with limited on-board processing, storage and radio capabilities. Sensor nodes sense and send their data toward a central processing center which is called "sink" by creating self-organized network. Those nodes are powered by embedded batteries and replacement of those batteries is a very difficult process once those nodes have been deployed. Therefore, the design of protocols and applications for such networks has to be energy aware in order to prolong the lifetime of the network.

In order to prolong the network life time, this dissertation proposes an adaptive routing protocol for heterogeneous WSNs. First, we address the proposed clustering routing method and non-clustering routing method for heterogeneous WSNs. Proposed clustering method successfully avoid non-uniform energy consumption around CHs and extend the network lifetime in wide area WSNs by selecting the cluster heads considering not only transmission power and residual energy of each node but also those of its adjacent nodes. On the other hand, proposed non-clustering method extend the network lifetime in small area WSNs by considering the residual energy of the nodes to select the optimal path for source to destination. Then, in order to avoid non-uniform energy consumption around the sink, we presets the proposed hybrid routing method which adaptively combined clustering and non-clustering communication methods. In proposed hybrid routing method, we proposed a analysis method which can flexibly include the constraints on the physical network boundaries and location of the sink to define the spatial boundary between clustering and non-clustering zones. Next, to responds to dynamic changes in the practical networks, we presents the proposed network reconfigure method which dynamically re-define the spatial boundary by considering the variance of the residual energy of a boundary node and its adjacent nodes.

Chapter 2 presented clustering and non-clustering routing methods for heterogeneous WSNs. First, we discussed some related works and introduce you to two existing clustering methods for heterogeneous WSNs, which were HEED (Hybrid Energy-Efficient Distributed clustering) and CC (Chessboard Clustering) and two existing non-clustering routing protocols which were Ad hoc On-Demand Distance Vector (AODV) and the Dynamic Source Routing(DSR) protocol. Then, we presented our proposed clustering and non-clustering methods for WSNs. Proposed clustering method, selects Cluster Heads (CH) considering not only the performance of a certain node but also that of its adjacent node. The proposed non-clustering method is extension for DSR protocol and residual energy of nodes in the path are taken under consideration when select a optimal path between source and destination. In addition, proposed clustering and non-clustering methods had no limitations on the number of node types that can use to crate a WSN.

Through the simulation experiments, we verified that smaller number of classes $C$ is better for the first half of the network lifetime and larger $C$ is better for the second half of the lifetime.. Simulation results show that the proposed clustering method is able to increase the network lifetime by 80% and 60% more than that of the CC and HEED, respectively in heterogeneous environment and achieve much higher performance than that in homogeneous environment. Also, simulation results show that the proposed clustering method achieves more than 10% of coverage than that of CC even in such a CC favorable environment. Through all the performance evaluation, it is verified that our proposed clustering method increases network lifetime than CC and HEED, in heterogeneous networks. Moreover, even in a homogeneous network, the proposed method has almost the same performance as HEED. It is a useful feature for practical use.

Chapter 3 proposed a robust clustering non-clustering hybrid routing method for heterogeneous WSN which adaptively combines clustering and multi-hop communication methods.

Most of conventional routing methods for WSN are optimized for a specific applications. In order to study the energy consumption and traffic patterns around sink and study about efficient bode placement around the Sink, we conducted several preliminary simulation experiments. In the proposed method, however, we considered the specifications of the nodes and the characteristics of the networked like outer boundary and location of the sink instead of characteristics of an application to optimize the WSN. Therefore, the proposed method can be easily applied to general-purpose WSN with multiple sensor applications. Also in chapter 3, we proposed simplified spatial boundary analysis method for small area WSN and generalized spatial boundary analysis method for wide area WSN. In order to calculate the optimal spatial boundary in non-computable networks, we introduced the power line concept. Through the simulation experiments, we verified that the proposed hybrid methods able to maintain higher alive node ratio than that of the pure clustering and multi-hop methods in any network topology.

Chapter 4 proposed the dynamic reconfiguration method of spatial boundary of clustering non-clustering hybrid routing method. In our proposed methods in chapter 3, we assumed that nodes were uniformly distributed in the network. However, in practical networks nodes are deployed in a non-uniform manner. In addition, the residual power of each node decreases with time elapsing. Therefore, pre-defined spatial boundary is not the optimal for all practical networks. Moreover, in this method sensor nodes near to the spatial boundary tends to has higher data relay traffic. Therefore, as the network time elapsed, the pre-defined spatial boundary should be reconfigured. The proposed method considers the variance of the residual energy of a boundary node and its adjacent nodes. Simulation results show that our method increases network lifetime in comparison the existing method with static boundary.

# Bibliography

[1] I F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," IEEE Communications Magazine, vol. 40, no. 8, pp 102–114, 2002.

[2] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the World with Wireless Sensor Networks," International Conference on Acoustics, Speech, and Signal Processing (ICASSP2001), May 2001.

[3] G. J. Pottie and W. J. Kaiser, "Wireless Integrated Network Sensors," Communications of the ACM, vol. 43, no. 5, pp. 51–58, May 2000.

[4] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," Computer Networks, Elsevier, vol. 52, no. 12, pp. 2292–2330, 2008.

[5] AKARI Architecture Design Project, "AKARI Architecture Conceptual Design for New Generation Network version 2.0," National Institute of Information and Communications Technology (NICT), pp 163–186, August 2009.

[6] N. Baker, M. Zafar, B. Moltchanov, and M. Knappmeyer, "Context-Aware Systems and Implications for Future Internet," Towards the Future Internet (Citeseer), pp 335–345, 2009.

[7] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, "A survey on wireless multimedia sensor networks," Computer Networks, vol. 51, pp. 921–960, 2007.

[8] I. Dietrich and F. Dressler, "On the lifetime of wireless sensor networks," ACM Transactions on Sensor Networks (TOSN), vol. 5, no. 1, pp. 5, 2009.

[9] A. A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," Computer Communications, vol. 30, pp. 2826–2841, 2007.

[10] J. L. Al-Karaki, R. Ul-Mustafa, and A. E. Kamal, "Data aggregation in wireless sensor networks-exact and approximate algorithms," Workshop on High Performance Switching and Routing (HPSR), pp. 241–245, 2004.

[11] W. R. Heinzelman, J. Kulik, and H. Balakrishnan, "Adaptive protocols for information dissemination in wireless sensor networks," in ACMMobiCom'99, pp. 174–185, 1999.

[12] G. Gupta and M. Younis, "Performance evaluation of load-balanced clustering of wireless sensor networks," 10th International Conference on Telecommunications, vol. 3, pp. 1577–1583, March 2003.

[13] V. Mhatre and C. Rosenberg, "Homogeneous vs heterogeneous clustered sensor networks: a comparative study," IEEE International Conference on Communications, vol. 6, pp. 3646–3651, Jun 2004.

[14] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "Application specific protocol architecture for wireless microsensor networks," IEEE Transactions on Wireless Networking, vol. 1, no. 4, pp. 660–670, 2002.

[15] M. Demirbas, A. Arora, and V. Mittal, "FLOC: a fast local clustering service for wireless sensor networks," Proceedings of Workshop on Dependability Issues in Wireless Ad Hoc Networks and Sensor Networks (DIWANS04), June 2004.

[16] H. Chan and A. Perrig, "ACE: an emergent algorithm for highly uniform cluster formation," Proceedings of the 1st European Workshop on Sensor Networks (EWSN), January 2004.

[17] G. Smaragdakis, I. Matta, and A. Bestavros, "SEP: A stable election protocol for clustered heterogeneous wireless sensor networks," Second International Workshop on Sensor and Actor Network Protocols and Applications (SANPA 2004), August 2004.

[18] L. Qing, Q. Zhu, and M. Wang, "Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks," Computer Communications, Elsevier, vol. 29, no. 12, pp. 2230–2237, 2006.

[19] N. Wang, N. Zhang, and M. Wang, "Wireless sensors in agriculture and food industry - Recent development and future perspective," Computers and electronics in agriculture, vol. 50, no. 1, pp. 1–14, 2006.

[20] X. Liu, P. Kulkarni, P. Shenoy, and D. Ganesan, "Snapshot: A Self-Calibration Protocol for Camera Sensor Networks," 3rd International Conference of Broadband Communications, Networks and Systems (BROADNETS 2006), October 2006.

[21] X. Du, Y. Xiao, and F. Dai, "Increasing network lifetime by balancing node energy consumption in heterogeneous sensor networks," Wireless Communications and Mobile Computing, vol. 8, pp. 125–136, September 2006.

[22] O. Younis, M. Krunz, and S. Ramasubramanian, "Node clustering in wireless sensor networks: Recent developments and deployment challenges," IEEE Network, vol. 20, pp. 20–25, 2006.

[23] O. Younis and S. Fahmy, "HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks," IEEE Transactions on Mobile Computing, vol. 3, pp. 366–379, October 2004.

[24] X. Du and Y. Xiao, " Energy efficient Chessboard Clustering and routing in heterogeneous sensor networks," International Journal of Wireless and Mobile Computing, vol. 1, pp. 121–130, February 2006.

[25] T. Shu, M. Krunz, and S. Vrudhula, "Power balanced coverage-time optimization for clustered wireless sensor networks," Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing, 2005.

[26] S. Lindsey, C. S. Raghavendra, and K. Sivalingam, "Data Gathering in Sensor Networks using the Energy Delay Metric," Proceedings of the IPDPS Workshop on Issues in Wireless Networks and Mobile Computing, April 2001.

[27] W. R. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, " Energyefficient Communication Protocol for Wireless Microsensor Networks," 33rd Hawaii International Conference on System Sciences (HICSS-33), January 2000.

[28] L. Doherty, L. E. Ghaoui, and K. S. J. Pister, "Convex position estimation in wireless sensor networks," IEEE Infocom 2001, April 2001.

[29] A. Savvides, C. Han, and M. Strivastava, "Dynamic fine-grained localization in ad-hoc networks of sensors," ACM MOBICOM01, July 2001.

[30] L. Lazos, R. Poovendran, and S. Čapkun, "ROPE: robust position estimation in wireless sensor networks," 4th international symposium on Information processing in sensor networks (IPSN), April 2005.

[31] J. Blumenthal, D. Timmermann, C. Buschmann, S. Fischer, J. Koberstein, and N. W Luttenberger, "Minimal transmission power as distance estimation for precise localization in sensor networks," Proceedings of the 2006 international conference on Wireless communications and mobile computing, July 2006.

[32] J. Hill, R. Szewczyk, and A. Woo, "System architecture directions for network sensors," 9th International Conference on Architectural Support for Programming Languages and Operating Systems, November 2000.

[33] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: A survey," Ad Hoc Networks, Elsevier, vol. 7, no. 3, pp. 537–568, 2009.

[34] P. Levis, N. Patel, D. Culler, and S. Shenker, "A self-regulating algorithm for code propagation and maintenance in wireless sensor networks," First USENIX/ACM Symposium on Networked Systems Design and Implementation (NSDI), March 2004.

[35] The Network Simulator ns-2: Documentation, http://www.isi.edu/nsnam/ns/ns-documentation.html

[36] Datasheet of XE1205a : www.semtech.com/images/datasheet/xe1205.pdf

[37] Datasheet of TR1000 : www.rfm.com/products/data/tr1000.pdf

[38] D. Kotz, C. Newport,R. S. Gray, J. Liu, Y. Yuan, and C. Elliott, "Experimental evaluation of wireless simulation assumptions," 7th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems, October 2004.

[39] S. Priyankara, K. Kinoshita, H. Tode, and K. Murakami, "A Clustering Method for Wireless Sensor Networks with Heterogeneous Node Types," 18th International Conference on Computer Communications and Networks (IEEE ICCCN 2009), August 2009.

[40] U. Rawat and M. Sharma, "Directed Diffusion: Features, Current Developments, Issues and Analysis," International Journal of Computer Applications, vol. 49, no. 12, pp. 31–35, 2012.

[41] W. R. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," IEEE Transactions on Wireless Communications, vol. 1, pp. 660–670, 2002.

[42] X. H. Xu, S. G. Wang, X. F. Mao, S. J. Tang, and X. Y. Li, "An improved approximation algorithm for data aggregation in multi-hop wireless sensor networks," Proceedings of the 2nd ACM international workshop on Foundations of wireless ad hoc and sensor networking and computing, August 2009.

[43] P. J. Wan, S. C. H. Huang, L. Wang, Z. Wan, and X. Jia, "Minimum-latency aggregation scheduling in multi-hop wireless networks," Proceedings of the tenth ACM international symposium on Mobile ad hoc networking and computing, May 2009.

[44] S. Priyankara, K. Kinoshita, H. Tode, and K. Murakami, "A Clustering Method for Wireless Sensor Networks with Heterogeneous Node Types," IEICE Transactions on Communications, vol. 94, no. 8, pp. 2254–2264, 2011.

[45] M. Takashima, D. Zhao, K. Yanagihara, K. Fukui, S. Fukunaga, S. Hara, and K.I. Kitayama, "Location estimation using received signal power and maximum likelihood estimation in wireless sensor networks," Electronics and Communications in Japan (Part I: Communications), vol. 90, no. 12, pp. 62–72, 2007.

[46] S. Priyankara, K. Kinoshita, H. Tode, and K. Murakami, "Clustering/Multi-hop Hybrid Routing Method for Wireless Sensor Networks with Heterogeneous Node Types," IEEE Globecom 2010 Workshop on Heterogeneous, Multi-hop Wireless and Mobile Networks, December 2010.

[47] The network Simulator ns-2, http://www.isi.edu/nsnam/ns/doc/ns_doc.pdf.

[48] Crossbow SN9040CA data sheet,http://www.xbow.com/Products/Product_pdf_files/Wireless_pdf/Sensor_Node_Datasheet.pdf

[49] Jie Wu, " Handbook on Theoretical and Algorithmic Aspects of Sensor, Ad Hoc Wireless, and Peer-to-Peer Networks," pp. 41–52, Auerbach Publications.

# Appendix A

# Appendices

## A.1 Proof of Lemma 1

*It takes maximum* $\frac{2\pi}{\sqrt{12}}\left(\frac{2d}{\sqrt{3}r_s}+\frac{2}{\sqrt{3}}\right)^2-1$ *necessary hops to deliver a packet from a source to a destination.*

*Proof:*

Assume that all the sensor nodes are within an area of a circle with radius $d$, $d > r$. The transmission range of each mobile station is $r$. Denote the graph generated by connecting all pairs of vertices within each others' transmission range as G; that is, two vertices are connected if and only if their geographic distance is less than or equal to 1. Then an upper bound for the diameter of G is, $\frac{2\pi}{\sqrt{12}}\left(\frac{d}{r}+\frac{2}{\sqrt{3}}\right)^2-1$. In other words, it takes maximum $\frac{2\pi}{\sqrt{12}}\left(\frac{d}{r}+\frac{2}{\sqrt{3}}\right)^2-1$ hops necessary to deliver a packet from a source to a destination.

Let $D$ be the diameter of the graph G. Choose two vertices $u$, $v$ such that the distance in G between $u$ and $v$, $d_G(u,v)$, is $D$; that is, $d_G(u,v) = D$. Then there exist distinct vertices $u_i, 1 \leq i \leq D+1$ such that,

$$u = u_1 \leftrightarrow u_2 \leftrightarrow u_3 \leftrightarrow ... \leftrightarrow u_D \leftrightarrow u_{D+1} = v$$

is a shortest path of length $D$ between $u$ and $v$. Let us define a set $I = \{u_i : \text{i is odd}\}$. Then the size of $I$, denoted by $\mid I \mid$, is $\lceil (D+1)/2 \rceil$. Then $I$ becomes an independent set of vertices.

*Claim 1* : $I$ is an independent set of vertices in graph $G$.

Suppose that $I$ is not an independent set; that is, there exists at least one edge between some pair of vertices in $I$. Without loss of generality, we assume that there are two vertices $u_{2j+1}, u_{2k+1}$ in $I$ such that $j < k$ and $u_{2j+1} \leftrightarrow u_{2k+1}$. By the definition of $I$, the two vertices are on the shortest path from $u$ to $v$. Then the shortest path can be represented as,

$$u = u_1 \leftrightarrow ...u_{2j} \leftrightarrow u_{2j+1} \leftrightarrow u_{2k+1} \leftrightarrow u_{2k+2}... \leftrightarrow u_{D+1} = v$$

The length of this path, as represented, is $D - [(2k+1) - (2j+1)] + 1 = D + 2j - 2k + 1$. It is less than $D$. This contradicts to $d_G(u, v) = D$. Therefore, there is no edge between any pair of vertices in $I$; in other words, $I$ is an independent set in $G$. By *Claim 1* and definition of $G$, we have *Claim 2*.

*Claim 2* : The geographic distance between any pair of vertices in $I$ is larger than $2r$.

Then, for each vertex $u_i \in I$, we define a circle $S_i$ with a center at $u_i$ and with a radius of $r_s$. By *Claim 2*, two circles $S_j, S_k$ are disjoint if $j \neq k$. Because all the vertices in $I$ are covered by the circle $C(d)$, all the disjoint circles $S_i$ ($i$ is odd) can be covered by a larger circle named $C(d + r_s)$, which has the same center as the circle $C(d)$ and has a radius of $d + r_s$ where $r_s$ is sensing radius of the node. Now we can relate this diameter problem to the circle packing problem, that is, how to effectively pack these non-overlapping circles $S_i$ ($i$ is odd) as many as possible into the large circle $C(d+r_s)$. By the remark following the proof of *Lemma 1* we have,

$$\frac{\sum_i (Area\, of\ S_i)}{Area\, of\ C(d + 2r_s/\sqrt{3})} \leq \frac{\pi}{\sqrt{12}}$$

$\frac{|I|\pi r^2}{\pi\left(d+\frac{2r}{\sqrt{3}}\right)^2} \leq \frac{\pi}{\sqrt{12}}, \mid I \mid \leq \frac{\pi}{\sqrt{12}}\left(\frac{d}{r} + \frac{2}{\sqrt{3}}\right)^2$

$\mid I \mid = \lceil (D+1)/2 \rceil \geq \frac{D+1}{2}$

$$D \leq 2 \mid I \mid -1 \leq \frac{2\pi}{\sqrt{12}}\left(\frac{d}{r} + \frac{2}{\sqrt{3}}\right)^2 - 1.$$

## A.2 Proof of Lemma 2

*Number of L-nodes skipped by single H-node is approximately equal to* $\lfloor \frac{r_H}{r_L} \rfloor$.

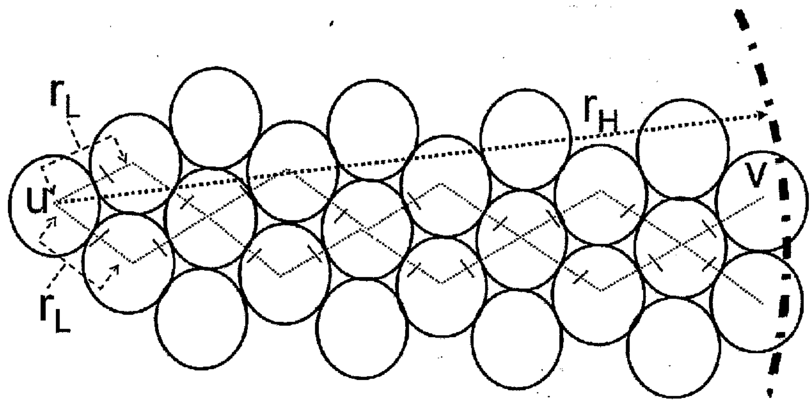



Figure A.1: Number of L-node hops skipped by H-node

*Proof:*

Node packing is optimum when it creates hexagonal packing. Considering that topology, we can assume a topology between source node $u$ and destination node $v$ as shown in Fig.A.1. Let the communication range of H-node is $r_H$ and that of L-node is $r_L$.

There are $n$ nodes between $u$ and $v$. When $n$ is odd,

$$\sqrt{(nr_L cos(\pi/6))^2 + (r_L sin(\pi/6))^2} \leq r_H \leq (n+1)r_L cos(\pi/6)$$

$$\frac{\sqrt{3n^2+1}}{2} \leq \frac{r_H}{r_L} \leq \frac{\sqrt{3}(n+1)}{2}.$$

On the other hand, when $n$ is even,

$$nr_L cos(\pi/6) \leq r_H \leq \sqrt{((n+1)r_L cos(\pi/6))^2 + (r_L sin(\pi/6))^2}$$

$$\frac{\sqrt{3}n}{2} \leq \frac{r_H}{r_L} \leq \frac{\sqrt{3(n+1)^2+1}}{2}.$$

Definition of the $\left\lfloor \frac{r_H}{r_L} \right\rfloor$ is,

$$\left\lfloor \frac{r_H}{r_L} \right\rfloor = k \;\; if\, and\, only\, if \;\; k \leq \frac{r_H}{r_L} < k+1$$

$$\begin{aligned} n < \frac{\sqrt{3}n}{2} < \frac{\sqrt{3n^2+1}}{2} \leq \frac{r_H}{r_L} \\ \leq \frac{\sqrt{3}(n+1)}{2} < \frac{\sqrt{3(n+1)^2+1}}{2} < n+1 \end{aligned} \tag{A.1}$$

Therefore,

$$n = \left\lfloor \frac{r_H}{r_L} \right\rfloor .$$

In other words, the mean number of L-nodes skipped by single H-node is equal to $\lfloor \frac{r_H}{r_L} \rfloor$.

An Adaptive Routing Protocol for Heterogeneous Wireless Sensor Networks January 2013 Sampath PR