Master’s Thesis

Title

Self-Organizing Topology Control based on Reaction-Diffusion Model for Data Gathering in Wireless Sensor Networks

Supervisor
Professor Masayuki Murata

Author
Katsuya Hyodo

February 13th, 2008

Department of Information Networking
Graduate School of Information Science and Technology
Osaka University

Master's Thesis

Self-Organizing Topology Control based on Reaction-Diffusion Model for Data Gathering in Wireless Sensor Networks

Katsuya Hyodo

Abstract

A wireless sensor network (WSN) consists of a large number of small, low-cost, and fragile sensing devices, called sensor nodes, communicating with each other through unreliable and unstable wireless communication. Therefore, control mechanisms for WSNs must be scalable to the number of nodes, adaptive to dynamically changing communication environment, and robust to failures. In addition, due to difficulty in managing sensor nodes in a centralized manner for limited energy and network resources, control mechanisms must be fully distributed and self-organizing. Considering these requirements, it is not surprising that much attention has been paid to application of biological mechanisms in these days, ranging from signal transduction network to swarm intelligence. A variety of biologically-inspired control mechanisms have been proposed for routing, clustering, scheduling, and topology control of WSNs. In this thesis, we focus on topology control for energy efficient and low delay data gathering, which certainly is one of the most important issues of WSNs.

Independently of applications, unless all sensor nodes operate with electric power supply, energy efficiency is a strong requirement for control mechanisms. In addition, it is necessary for sensor nodes to deliver up-to-date sensor information to a sink node of sensor data, e.g. a gateway server, as fast as possible for prompt reaction to detected events and conditions. The amount of consumed energy and the delay in data gathering heavily depend on the topology of a WSN. A variety of topologies have been considered for energy-efficient and low-delay data gathering in a WSN, such as, direct transmission from all sensor nodes to a sink node, a tree topology rooted at a sink node, and multi-tier topology organized by clustering sensor nodes into groups. However, there has never been comprehensive and complete comparison among them so far, to the best

knowledge of us. Therefore, we first need to clarify the best topology for data gathering before considering scalable, adaptive, robust, fully-distributed and self-organizing control mechanism.

Therefore, we begin this thesis with investigation of the best topology for data gathering from viewpoints of energy efficiency and delay. We consider six types of topology, as combinations of with or without-clustering and single or multi-hop transmission for inter or intra-cluster communication, in a variety of scenarios different in the energy consumption model, node density, communication range, and the location of a sink node. Through extensive evaluation, we prove that adopting multi-hop communication in both of inter and intra clusters enables energy-efficient and low delay data gathering in a balanced manner for most of scenarios. Then, we propose a novel topology control mechanism based on a biological self-organizing mechanism, that is, the reaction-diffusion model to organize the best topology in a self-organizing and autonomous way. A reaction-diffusion model is a mathematical model for pattern generation on the surface of body of fishes and mammals. Nodes generate a spatial spot patterns through exchange two virtual morphogens, i.e., activator and inhibitor, among neighboring nodes and calculating a reaction-diffusion equation. The spatial pattern, which is gradient of morphogens, has peaks of activator concentration equally spaced. The node which has a peak of activator concentration is elected as cluster head, and other nodes send own data following the gradient of activator concentration to the cluster head. In addition, we find that the point density of cluster heads near the sink node is larger than that of far form the sink node in optimal topology. Therefore, we extend the mechanism to generate the point density of cluster heads. Through simulation experiments it is shown that a topology organized by our proposal accomplishes as small energy consumption and delay as the best topology derived under an ideal condition does at the difference of about 8.4 % with data fusion, and 30 % without data fusion.

Keywords

Sensor Network

Data Gathering

Topology Control

Biological System

Reaction-Diffusion Model

Contents

List of Figures

List of Tables

1 Introduction

Recent advances in micro-electro-mechanical system (MEMS), wireless communications, and digital electronics technology have made development of small and low-cost sensing devices possible [1]. Wireless sensor networks (WSNs) are composed of a large number of such sensing device, called sensor nodes, which consist of sensing, data processing, and wireless communicating components. WSN enables a broad range of applications, such as environmental monitoring, precision agriculture, health monitoring of human beings and artificial structures, and disaster and crime prevention [2]. WSN is one of the most promising and key technologies for safe, secure, and comfortable society.

In WSNs, since a large number of fragile nodes communicate with each other through unreliable and unstable wireless connections, control mechanisms for WSNs must be scalable to the number of nodes, adaptive to dynamically changing communication environment, and robust to failures. In addition, due to difficulty in managing sensor nodes in a centralized manner for limited energy and network resources, control mechanisms must be fully distributed and self-organizing. Considering these requirements, it is not surprising that much attention has been paid to application of biological mechanisms in these days, ranging from signal transduction network to swarm intelligence [3, 4]. A variety of biologically-inspired control mechanisms have been proposed for routing, clustering, scheduling, and topology control of WSNs [5-8]. For example, ant colony optimization is a well-known biologically-inspired optimization algorithm [9], which models shortest-path establishment and maintenance observed in foraging behavior of ants, ACO has often been applied to routing problems in wired and wireless networks for their similarity. In [5], a routing scheme for many-to-one type of communication was developed based on a distributed ant algorithm to optimize the energy cost for data communication and the utilization of network resources. In this scheme, the robustness can be improved and network life time can be lengthened by learning from a self-organizing biological system. In [6], the authors exploit a mechanism of ant colony in foraging and brood sorting to design a hierarchical and scalable data gathering protocol. In BiSNET (Biologically-inspired architecture for Sensor NETworks) [7], they addressed the autonomy, adaptability, self-healing, and simplicity of MWSNs (multi-model wireless sensor networks). A multi-agent platform called the BiSNET platform is proposed, where an agent can emit pheromones, replicate itself, migrate from one node to another, exchange energy with other agents, and die for energy depletion. Finally, in [8], the authors adopted a biologically-inspired promoter / inhibitor schemes for adaptive parameter control in network security environments. They show that self-regulating techniques based on biological system lead to maximized security and graceful degradation in overload situations.

In this thesis, we focus on topology control for energy efficient and low delay data gathering, which is certainly one of the most important issues of WSNs. Independently of applications, unless all sensor nodes operate with electric power supply, energy efficiency is a strong requirement for control mechanisms. In addition, it is necessary for sensor nodes to deliver up-to-date sensor information to a sink node as fast as possible for prompt reaction to detected events and conditions. The amount of consumed energy and the delay in data gathering heavily depend on the topology of a WSN. There have been many proposals for energy-efficient and low-delay data gathering [10-19] as well as biologically-inspired ones [5, 6, 20, 21]. A variety of topologies were considered in those literatures, such as, direct transmission from all sensor nodes to a sink node, a tree topology rooted at a sink node, and multi-tier topology organized by clustering sensor nodes into groups. However, there has never been comprehensive and complete comparison among them so far, to the best knowledge of us. For example, in PEGASIS [17], they only show the simulation result that indicates PEGASIS performs better than LEACH by about 100 to 200 percent in terms of lifetime of a WSN when 1 percent, 25 percent, 50 percent, and 100 percent of nodes die. Therefore, we first need to clarify the best topology for data gathering before considering scalable, adaptive, robust, fully-distributed and self-organizing control mechanism.

We begin this thesis with investigation of the best topology for data gathering from viewpoints of energy efficiency and delay. We consider six types of topology as combinations of with or without-clustering and single or multi-hop transmission for inter or intra-cluster communication, in a variety of scenarios different in the energy consumption model, node density, communication range, and the location of a sink node. We first derive the optimal topology for each of topology type in each of scenarios by solving an optimization problem. Then, the optimal topologies are compared from viewpoints of the total amount of energy consumption and the maximum delay in gathering sensor data from all nodes. Through extensive evaluation which will be shown in Section 3, we prove that a so-called multi-multi topology, i.e. a cluster-based topology where cluster members transmit their sensor data to a cluster head by multi-hop communication and gathered sensor data are sent from cluster heads to a sink node by multi-hop communication, enables energy-efficient and low-delay data gathering in a balanced manner in most of scenarios.

Then, we propose a novel topology control mechanism based on a biological self-organizing mechanism, that is, a reaction-diffusion model, to organize the best topology in a self-organizing and autonomous way. A reaction-diffusion model was first proposed by Alan Turing as a mathematical model for pattern generation on the surface of body of fishes and mammals [22] . In a reaction-diffusion model, through local interactions among molecules of neighboring cells, a variety of patterns of morphogen concentrations emerge in a self-organizing manner. Autonomously generated patterns can be used for routing, clustering, scheduling, and topology control on sensor networks [23-27]. Our early work [23] verified the practicality of reaction-diffusion based pattern formation on a WSN under the influence of loss and delay of message transmission through experiments. We also proposed two acceleration schemes for faster pattern generation. In smart sensor networks for a forest fire application, a stripe pattern is organized from a robot load point to a fire control point through local and mutual interactions among distributed sensor nodes and mobile robots walk along the stripe to fight fire [24]. RDMAC [25] is a reaction-diffusion based MAC protocol, where they noticed the similarity among a scheduling pattern of spatial TDMA and a spot pattern of leopards. A node inhibits packet emission of neighboring nodes in its range of radio signals while encouraging nodes out of the range to send packets for better spatial use of a wireless channel. For camera sensor networks, a cooperative control model for a surveillance system which consists of plural Pan-Tilt-Zoom cameras and having no central control unit is proposed [26]. Each camera adjusts their observation area to decrease blind spots in the whole surveillance area by control algorithms based on a reaction-diffusion model. We also propose an autonomous mechanism based on a reaction-diffusion model for coding rate control in camera sensor networks for remote surveillance and tracking applications [27].

In this thesis, we adapt a reaction-diffusion model to a topology control mechanism for data gathering. Specifically, a spatial spot patterns emerges through autonomous behavior of nodes to exchange two virtual morphogens, i.e. activator and inhibitor, among neighboring nodes. The spatial pattern of heterogeneous distribution of concentration of morphogens, has equally spaced peaks of activator concentration. A node which has a peak of activator concentration is considered a cluster head, and other nodes send their sensor data following the gradient of activator concentration to a cluster head. In addition, we find that the point density of cluster heads near the sink node is larger than that in the region far form the sink node in the optimal topology. To accomplish such heterogeneous distribution of cluster heads, we extend the reaction diffusion-based mechanism to dynamically change the wavelength of a pattern in accordance with the distance from a sink node. Through simulation experiments, it is shown that a topology organized by our proposal accomplishes as small energy consumption and delay as the optimal topology derived under an ideal condition does.

The rest of this thesis is organized as follows. In Section 2, we introduce related work in topology control of WSNs for data gathering application. Next in Section 3, we first categorize six types of topology for data gathering, then derive their optimal topology, and finally compare them in terms of the energy efficiency and delay. In Section 4, we describe our reaction-diffusion based topology control mechanism. We then show and discuss results of simulation experiments in Section 5. Finally, we conclude the thesis in Section 6.

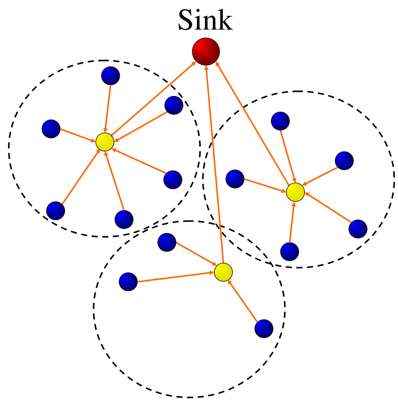


Figure 1: Topology organized in LEACH

2 Related Work

In WSNs, a lot of data gathering or topology control mechanisms have been proposed [5, 6, 10-18, 20, 21]. In this section, we explain LEACH [11], PEGASIS [17], and HIT [16] as typical methods generating cluster-based topology, tree-based topology, and hybrid topology.

2.1 Low-Energy Adaptive Clustering Hierarchy (LEACH)

LEACH (Low-Energy Adaptive Clustering Hierarchy) [11] is the most well known hierarchical routing algorithm for periodic data gathering. The main idea is to cluster sensor nodes into groups and use local cluster heads as routers to a sink node. At the beginning of data gathering cycle called a round, some sensor nodes stand for cluster heads and advertise their candidacy to the whole of a WSN by means of CSMA broadcasting. A node hearing the advertisement chooses the most preferred cluster head, in general based on the distance, and registers itself as a cluster member. A cluster head then assigns time slots to cluster members to collect sensor data from them in a TDMA manner. Finally, aggregated sensor data, ideally shrunk to one size of sensor data, is sent to the sink node from each cluster head by CDMA media access to avoid contention. After repeatedly gather sensor data on the same topology at the regular intervals of data gathering, the next round starts with formation of new clusters. An example of topology organized by LEACH is illustrated in Fig. 1, where yellow circles stand for cluster heads and blue circles correspond to cluster members.

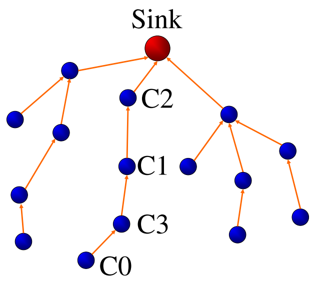


Figure 2: Tree based topology

By clustering, most of nodes can save energy at the sacrifice of cluster heads, which consume much energy in advertisement, reception of sensor data from cluster members, processing of collected sensor data, and transmission of aggregated sensor data to the sink node. To balance energy consumption among sensor nodes for the long lifetime of a WSN, LEACH adopts the probabilistic cluster head rotation mechanism. The node n becomes a cluster head at the round r, if a random number ranging from 0 to 1 is below the threshold $T(n)$

$$T(n) = \begin{cases} \dfrac{p}{1 - p(r \bmod 1/p)} & \text{if } \; n \in G \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where p represents the desired percentage of cluster heads. G is the set of nodes that were not a cluster head in the last $1/p$ rounds. With this mechanism, every node becomes a cluster head once every p rounds and the desired number of clusters are organized. It is shown that LEACH can reduce energy consumption eight times as much as the direct transmission, where all sensor nodes directly transmit their sensor data to the sink node, and the minimum transmission energy routing, where each sensor node transmit sensor data along a path of minimum energy consumption.

2.2 Power-Efficient Gathering in Sensor Information Systems (PEGASIS)

PEGASIS (Power-Efficient GAthering in Sensor Information Systems) [17] is different from LEACH in forming a tree based topology from sensor nodes to a sink node, as shown in Fig. 2, rather than clusters. Since each node communicates with only direct neighbors, they can save energy consumption for data gathering.

In PEGASIS, all nodes are assumed to have global knowledge of a WSN, including location of all nodes and be able to fuse own data together with data received from neighbors. Each node

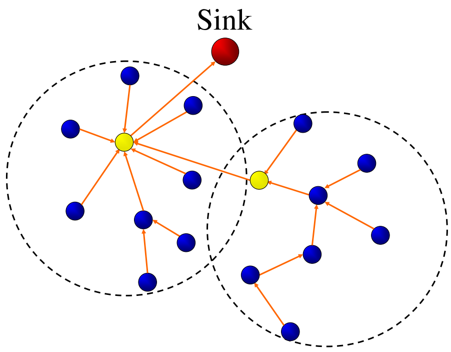


Figure 3: Topology organized in HIT

locally computes the tree using the same greedy algorithm as follows. First, any of the furthest nodes from sink node, such as c0 in Fig. 2, is chosen (if there is a tie, node is selected at random) The closest neighbor to this node will be the next hop node on the tree. By repeating this procedure, a tree is formed from leaf nodes to the sink. In gathering sensor data from nodes, since all nodes build the same tree, a node can determine its time slot so that the gathering delay in terms of the number of slots is minimized.

The simulation results show that PEGASIS outperforms LEACH by about 100 to 200 percent in terms of lifetime of a WSN when 1 percent, 25 percent, 50 percent, and 100 percent of nodes death.

2.3 Hybrid Indirect Transmission (HIT)

If data fusion can be done at each node, tree based topology of PEGASIS leads to less energy consumption than cluster based approach such as LEACH, because tree based topology does not involve long range transmission. However, the number of hops from leaf nodes to the sink node becomes larger and as a result the transmission delay becomes longer. HIT (Hybrid Indirect Transmission) [16] takes the advantages of both topology in order to minimize both energy consumption and delay. HIT has a hybrid topology. Generated topology consists of one or more clusters. In a cluster, cluster members form a tree rooted at a cluster head. Cluster heads also form a tree rooted at a sink node. Therefore, it is a two-tier tree network as shown in Fig. 3. First, one or more nodes are selected as a cluster head by the election scheme like that of LEACH. They broadcast their status including sender ID by using the fixed transmission power. Next, a node receiving this broadcast becomes a cluster member of one of cluster heads in its vicinity. At the same time, a tree of cluster heads is formed. Then, each node including a cluster head and a cluster member makes a blocking set. A blocking set is a set of nodes to which the node should not send sensor data in order to avoid collisions. Based on the blocking list, each node determines a tree to its cluster head. After this, each node computes a TDMA schedule. Finally, sensor data are sent to each cluster head from its cluster members. Each cluster head fuses the collected sensor data together with its own and sends the aggregated sensor data to a sink node traversing an inter-cluster tree. In order to minimize both energy consumption and gathering delay, data transmission is done in parallel in inter and intra cluster communication. For this purpose, HIT has an algorithm for each node to determine the slot assignment of TDMA. HIT assumes that sensor data can be fused together.

By simulation experiments, the authors show that HIT achieve 1.05 – 32.7 times as energy efficiency as LEACH, PEGASIS, and the direct transmission. Moreover, on average, transmission delay in HIT is about 25 % of the delay in the other methods.

3 Analysis on Energy Efficiency and Delay of Network Topology for Data Gathering

The preceding proposals for data gathering in a WSN are all aimed at energy efficiency and/or low delay. A variety of topologies are considered in these literatures in accordance with different assumptions, models, and scenarios. However, there has never been a comprehensive and complete comparison among topologies so far, to the best knowledge of us. Therefore, we will devote this section to the discussion of the best topology for data gathering from the viewpoint of energy efficiency and delay, before considering reaction-diffusion based self-organizing topology control.

3.1 Target System and Application

Before the discussion, we must precisely define the target system and application. The target system in this thesis is WSN which consists of tens of sensor nodes and one sink node. The target application is periodic or aperiodic data gathering from all sensor nodes, such as monitoring natural disaster, such as activity of a volcano, mudslide, and flood, which needs an energy efficient and low delay mechanism in order to observe the environment as long as possible while reporting an unusual event as fast as possible.

100 sensor nodes, operating on batteries, are randomly distributed in an $L \times L$ m^2 two-dimensional monitored region. Sensor nodes are homogeneous in the initial battery power and the equipped transceiver. A sensor node can control the transmission power so that the range of radio signal can be adjusted to the necessary and sufficient level. We consider both cases that a sensor node can fuse multiple sensor data into one or cannot. The sink node has a power supply. Therefore, it is out of account in calculation of energy consumption. The sink node is located at the center of the monitored region or far from the region. Since we consider here the best and optimal topology, we assume that a network topology is built having the complete information about the node distribution. Our research objective of this thesis is to organize the optimal topology in a self-organizing way. In addition, to derive the delay metric on a topology, we assume TDMA-based communication, where ideal slot assignment is performed. It also is possible to estimate the maximum delay in packet transmission assuming CSMA-based communication, but the analysis model largely depends on a specific MAC protocol and the strong assumption on the system configuration.

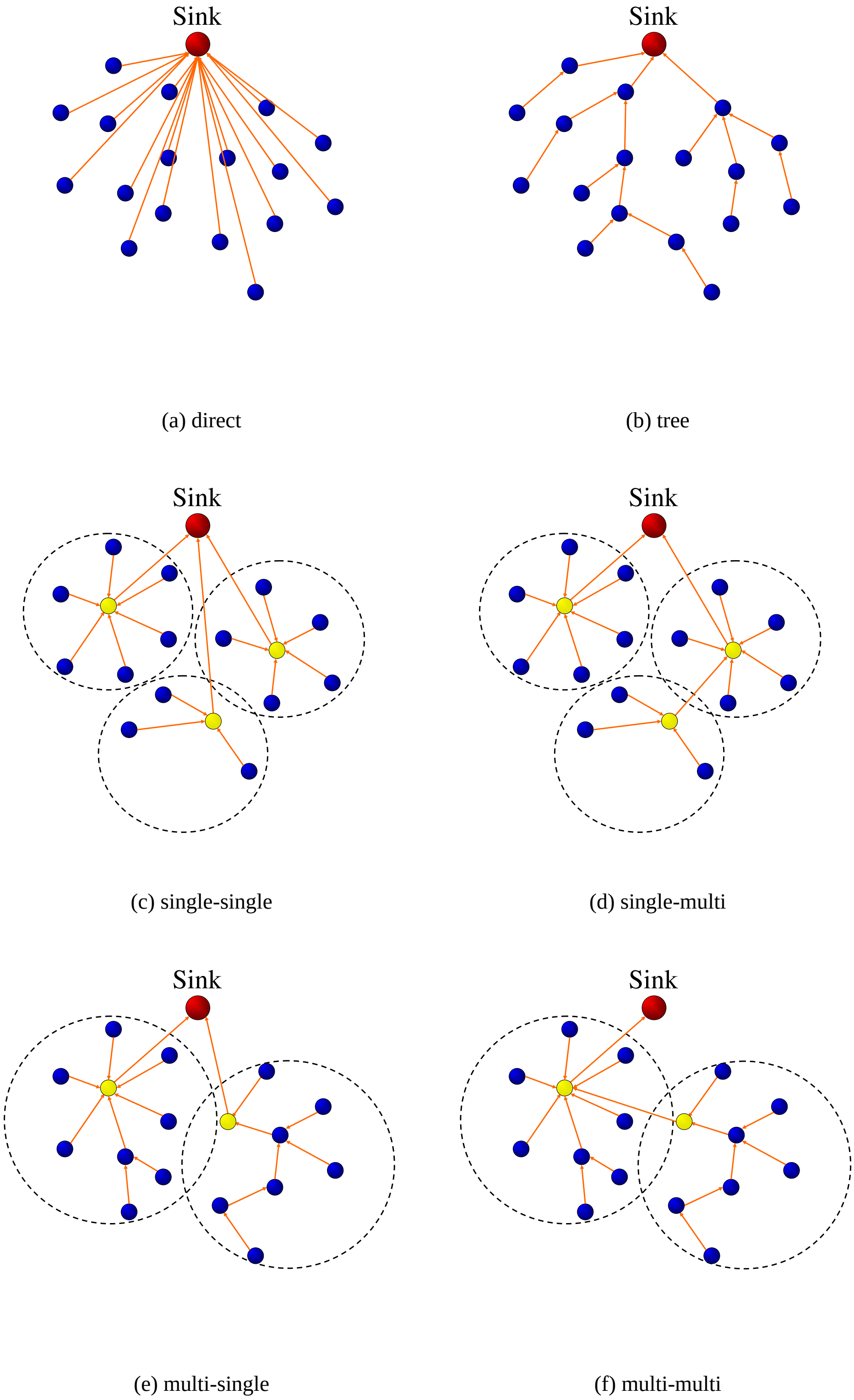
3.2 Classification of Network Topology for Data Gathering

Network topologies for data gathering can be divided in to two major categories: with clustering and without clustering. There are four types of topologies in cluster-based topology and two types in non-clsuter topology. In total, there are six topologies considered in the thesis. Examples are illustrated in Fig. 4.

The first topology, named as direct topology, does not organize any hierarchical structure (Fig. 4(a)). All sensor nodes, shown as blue circles, directly send their sensor data to the sink node, shown as a red circle. Since the electric power consumed in data transmission increases in proportional to the 2nd to 4th power of the distance [11], this leads to the worst case in terms of the energy consumption. The data gathering delay, which is defined as time required to collect sensor data from all sensor nodes, is equal to the number of sensor nodes multiplied by the duration of one time slot in TDMA. It is because that all sensor nodes compete for the single wireless channel for the single destination.

The second topology is the tree topology (Fig. 4(b)). To save energy consumption, sensor nodes limit the transmission power and sensor data are relayed to the sink node by sensor nodes, i.e. multi-hop communication. Usually, sensor nodes form a tree rooted at the sink node [17-19]. Data gathering starts at the edge of a WSN or leaf nodes of the tree towards the sink node. In the tree topology, there is the tradeoff between the energy efficiency and the delay. By limiting the communication range, the amount of energy consumption in data transmission can be reduced. However, it makes a high tree. The delay increases as the height of tree increases. In addition, as the network size, in terms of the number of sensor nodes and the area of the monitored region, increases, the energy consumption and the delay increase.

Clustering or making hierarchy is one of the major approach to make a system scalable. In cluster-based schemes [11-17, 20], nearby sensor nodes form a group, called cluster. One among them becomes a representative of the cluster, called cluster head (yellow circle in figures). A cluster head gathers sensor data from the other sensor nodes in the cluster, called cluster members, directly or forming a tree, fuses the collected sensor data if possible, and sends the aggregated sensor data to the sink node directly or by being relayed by other cluster heads to the sink node. Therefore, there are four combinations of communication types in cluster-based topology. In the single-single topology illustrated in Figs. 4(c), cluster members send their sensor data to a cluster



(a) direct (b) tree

(c) single-single (d) single-multi

(e) multi-single (f) multi-multi

Figure 4: Classification of network topology

head by unicast transmission and then the aggregated data is sent from a cluster head to the sink node by unicast transmission as in LEACH. Since the distance from cluster heads to the sink is farther than that from cluster members to a cluster head, direct transmission from a cluster head to the sink node expenses much energy of the cluster head. Therefore, in the single-multi topology (Fig. 4(d)), cluster heads form a tree as in the tree topology to save energy consumption. One can easily consider the opposite topology, i.e. the multi-single topology, where cluster members form a tree rooted at a cluster head and then cluster heads perform direct transmission to the sink node. In this case, the size of a cluster can be larger than that in the single-single and single-multi topologies. Then, the number of clusters becomes small and less number of nodes appointed as cluster heads sacrifice their energy. Further energy saving can be accomplished in the multi-multi topology illustrated in Fig. 4(f). There are trees of two levels as in HIT. One is for intra-cluster communication and the other is for inter-cluster communication. By establishing two-level trees, the energy consumption in data transmission can be reduced on cluster members as well as cluster heads.

3.3 Derivation of Optimal Topology for Each Topology Type and its Delay

In this section, we give details of derivation of the optimal topology for each of topology types and the maximum data gathering delay on the topology. The amount of energy consumption in gathering sensor data from all nodes is derived as the total of energy consumed in transmission, reception, and processing of sensor data on all nodes. We assume that a sensor node adjusts the range of radio signals to the necessary and sufficient level to reach the next hop node by transmission power control. In deriving the delay, we assume that nodes at the edge or end of a WSN begin transmission of sensor data first, and slot length for TDMA is calculated by dividing data size by transfer rate. Each intermediate node sends the aggregated data as soon as it receives sensor data from all of its child nodes in a tree or its cluster members.

direct topology

The direct topology has only one star-shaped topology as illustrated in Fig. 4(a).

As stated before, since each of all sensor nodes has to be assigned one designated time slot in direct topology, the maximum gathering delay is the product of the number of nodes and the slot length in time.

tree topology

The optimal topology of tree topology depends on whether data fusion is possible or not. When multiple sensor data cannot be fused to one, Shortest Path Tree (SPT) is the optimal topology [28]. In this thesis, we use the Dijkstra algorithm for generating SPT [29]. On the other hand, when it is possible to fuse multiple sensor data, Minimum Spanning Tree (MST) is the optimal topology [19]. In this thesis, we use the Prim's MST algorithm to obtain MST [30]. We set the link cost as the sum of energy consumed at the both ends of the link in transmission and reception of sensor data, and energy consumption of data fusion if data fusion is possible. Since the sink node has unlimited power supply, the energy consumed at the sink node for reception is ignored in topology derivation.

In the tree topology, we assume that each sensor node waits for reception of sensor data from all of its child nodes before sending sensor data independently of the possibility of data fusion. Therefore, the furthest sensor node from the sink node dominates the delay. Since all furthest sensor nodes compete for the channel at some senor node or the sink node, the maximum delay can be given as the product of the slot length in time and the sum of the maximum hop count, i.e. the height of the tree and the number of furthest leaf nodes -1.

single-single topology

In cluster-based topology, optimal clustering is known as an NP-hard problem, even with the complete information about node locations [31]. In most of centralized methods [12, 32], cluster heads are chosen by using an SA (simulated annealing) algorithm [33]. In this thesis, we also adopt an SA algorithm to obtain the optimal or near-optimal solution for cluster-based topologies. In problem formulation, we set an objective function as the minimization of energy consumption, delay, or product of them [34], i.e. we generate three kinds of clusters. We derive solutions for different number of clusters and chooses the best by comparing their results.

The delay from cluster members to a cluster head can be calculated by multiplying the number of cluster members and the slot length in time, since communication pattern in a cluster is the same as in the direct topology. The delay from cluster heads to the sink node can be calculated in the similar way. We assume that a cluster head immediately send the aggregated sensor data to the sink node, once it receives sensor data from all of its cluster members Therefore, the maximum delay can be derived as (the maximum number of cluster members + the number of clusters with the same maximum population - 1) $\times$ slot length in time.

single-multi topology

To derive the optimal or near-optimal topology, we first applies an SA algorithm and obtain a set of cluster heads. Then, a tree of cluster heads rooted at the sink node is established by MST or SPT depending on the possibility of data fusion. Finally, remaining nodes are connected to their nearest cluster heads.

By following the same discussions in the above, the maximum delay can be given as (the maximum number of cluster members + the number of hops of the furthest cluster head from the sink node in the cluster head tree + the number of furthest cluster heads in the cluster head tree $-$ 1) $\times$ slot length in time.

multi-single topology

To derive the optimal or near-optimal topology, we first applies an SA algorithm and obtain a set of cluster heads. Then, the sink node is connected to each of cluster heads by a direct link. Finally, MST or SPT is applied to accommodate the remaining nodes, i.e. cluster members by extending the star topology made of cluster heads and the sink.

The maximum delay is obtained by calculating (the maximum number of hops to the sink node from a cluster member + the number of sensor nodes with the same maximum hop distance$-$1) $\times$ slot length in time.

multi-multi topology

To derive the optimal or near-optimal topology, we first applies an SA algorithm and obtain a set of cluster heads. Then, MST or SPT is applied to form a cluster head tree. Finally, MST or SPT is applied to accommodate the remaining nodes, i.e. cluster members by extending the cluster head tree.

The maximum delay is obtained by calculating (the maximum number of hops to the sink node from a cluster member $+$ the number of sensor nodes with the same maximum hop distance-1) $\times$ slot length in time.

3.4 Scenarios of Comparison

The purpose of topology comparison is to clarify the best topology for data gathering in terms of the energy consumption and the gathering delay. We conduct thorough evaluation in a variety of settings in the energy consumption model, node density, communication range, and the location of a sink node. Since we focus here on the characteristics of topology itself, we do not take into account the energy consumption and delay in exchanging control messages to form the topology.

As summarized in Table 1, independently of scenarios, i.e. setting of system configuration, 100 nodes are randomly distributed in the monitored region. The message length for one sensor data is 2000 bits. Without data fusion, n sensor data amounts to 2000n bits. When data fusion is possible, n sensor data can be fused into 2000 bits. The initial energy is identical among sensor nodes and set at 10000 J. The transmission rate is 250 Kbps, taken from IEEE 802.15.4 standard rate. We consider 40 scenarios by combining different settings in the communication range, the energy consumption model, the possibility of data fusion, the location of the sink node, and the area, as summarized in Tables 2 through 5. As the communication range , i.e. the maximum distance that radio signals with the maximum transmission power reach, we consider finite and infinite settings. In the case of infinite communication range, the maximum distance is 400 m in our scenarios, which is long enough for all sensor nodes can directly communicate with the remote sink node. In the case of the finite transmission range, the maximum distance is set at 20 m based on the implementation experiment presented in [35] and the data sheet of well-known commercial node device MICAz [36]. We adopt the energy consumption model of e-LEACH [12], called e-LEACH model, and the practical model we made from the data sheet of MICAz, called MICAz model.

In the e-LEACH model, the amount of energy consumed in transmitting a message of l bits to the distance d m is derived by the following equation.

$$E_{T_x}(l,d) = \begin{cases} lE_{elec} + l\epsilon_{fs}d^2, & d < d_0 \\ lE_{elec} + l\epsilon_{mp}d^4, & d \geq d_0 \end{cases} \quad (2)$$

where $E_{elec} = 50$ nJ/bit is the amount of energy consumed in the transmitter and receiver circuits

Number of node	100
Data size	2000 bit
Initial energy	10000 J (size AA battery)
transmission rate	250 Kbps (Zigbee 2.4GHz)

Table 1: Parameter settings for comparison of topologies

No.	Max range	Energy model	Fusion	Sink	Area
1	∞	e-LEACH ($d_0 = 75$)	$\times$	Remote (120,10)	20×20
2	∞	e-LEACH ($d_0 = 75$)	$\times$	Remote (200,50)	100×100
3	∞	e-LEACH ($d_0 = 75$)	$\times$	Remote (300,100)	200×200
4	∞	e-LEACH ($d_0 = 75$)	$\times$	Center (10,10)	20×20
5	∞	e-LEACH ($d_0 = 75$)	$\times$	Center (50,50)	100×100
6	∞	e-LEACH ($d_0 = 75$)	$\times$	Center (100,100)	200×200
7	∞	e-LEACH ($d_0 = 75$)	$\bigcirc$	Remote (120,10)	20×20
8	∞	e-LEACH ($d_0 = 75$)	$\bigcirc$	Remote (200,50)	100×100
9	∞	e-LEACH ($d_0 = 75$)	$\bigcirc$	Remote (300,100)	200×200
10	∞	e-LEACH ($d_0 = 75$)	$\bigcirc$	Center (10,10)	20×20
11	∞	e-LEACH ($d_0 = 75$)	$\bigcirc$	Center (50,50)	100×100
12	∞	e-LEACH ($d_0 = 75$)	$\bigcirc$	Center (100,100)	200×200

Table 2: Scenarios for comparison of topologies (infinite communication range and e-LEACH energy consumption model)

No.	Max range	Energy consumption	Fusion	Sink	Area
13	∞	MICAz (Open space)	$\times$	Remote (120,10)	20×20
14	∞	MICAz (Open space)	$\times$	Remote (200,50)	100×100
15	∞	MICAz (Open space)	$\times$	Remote (300,100)	200×200
16	∞	MICAz (Open space)	$\times$	Center (10,10)	20×20
17	∞	MICAz (Open space)	$\times$	Center (50,50)	100×100
18	∞	MICAz (Open space)	$\times$	Center (100,100)	200×200
19	∞	MICAz (Open space)	$\bigcirc$	Remote (120,10)	20×20
20	∞	MICAz (Open space)	$\bigcirc$	Remote (200,50)	100×100
21	∞	MICAz (Open space)	$\bigcirc$	Remote (300,100)	200×200
22	∞	MICAz (Open space)	$\bigcirc$	Center (10,10)	20×20
23	∞	MICAz (Open space)	$\bigcirc$	Center (50,50)	100×100
24	∞	MICAz (Open space)	$\bigcirc$	Center (100,100)	200×200

Table 3: Scenarios for comparison of topologies (infinite communication range and MICAz (Open space) energy consumption model)

No.	Max range	Energy consumption	Fusion	Sink	Area
25	20 m	e-LEACH ($d_0 = 75$)	$\times$	Remote (120,10)	20×20
26	20 m	e-LEACH ($d_0 = 75$)	$\times$	Remote (200,50)	100×100
27	20 m	e-LEACH ($d_0 = 75$)	$\times$	Center (10,10)	20×20
28	20 m	e-LEACH ($d_0 = 75$)	$\times$	Center (50,50)	100×100
29	20 m	e-LEACH ($d_0 = 75$)	$\bigcirc$	Remote (120,10)	20×20
30	20 m	e-LEACH ($d_0 = 75$)	$\bigcirc$	Remote (200,50)	100×100
31	20 m	e-LEACH ($d_0 = 75$)	$\bigcirc$	Center (10,10)	20×20
32	20 m	e-LEACH ($d_0 = 75$)	$\bigcirc$	Center (50,50)	100×100

Table 4: Scenarios for comparison of topologies (finite communication range and e-LEACH energy consumption model)

No.	Max range	Energy consumption	Fusion	Sink	Area
33	20 m	MICAz (Indoor)	$\times$	Remote (20,10)	20×20
34	20 m	MICAz (Indoor)	$\times$	Remote (100,50)	100×100
35	20 m	MICAz (Indoor)	$\times$	Center (10,10)	20×20
36	20 m	MICAz (Indoor)	$\times$	Center (50,50)	100×100
37	20 m	MICAz (Indoor)	$\bigcirc$	Remote (20,10)	20×20
38	20 m	MICAz (Indoor)	$\bigcirc$	Remote (100,50)	100×100
39	20 m	MICAz (Indoor)	$\bigcirc$	Center (10,10)	20×20
40	20 m	MICAz (Indoor)	$\bigcirc$	Center (50,50)	100×100

Table 5: Scenarios for comparison of topologies (finite communication range and MICAz (Indoor) energy consumption model)

per bit. d_0 is the threshold in distance and set at 75 m. For the transmission distance d smaller than the threshold, the free space model is adopted, where the power loss of d^2 is applied and the amplifier energy consumption $\epsilon_{fs} = 10$ pJ/bit/m^2 is used. On the other hand, for the distance d larger than the threshold, the multipath model with the power loss of d^4 and the amplifier energy consumption $\epsilon_{mp} = 0.0013$ pJ/bit/m^4 is adopted. The energy consumed in reception of a message of k bits is derived as,

$$E_{R_x}(k) = E_{R_{x-elec}}(k) = kE_{elec}. \tag{3}$$

The energy consumed in data fusion is set at $E_{DA} = 5$nJ/bit/data.

In the MICAz model, we assume that the radio transmitter dissipates the energy of E_{T-elec}=120 nJ/bit to run the transmitter circuit and E_{R-elec}=280 nJ/bit to run the receiver circuit, respectively. As a pathloss model, we use a free space model of d power loss and a multipath model of d^2 power loss. The amplifier energy consumption is set at ϵ_{amp}=240 pJ/bit/m^2 independently of the pathloss model. Therefore, we have the following equation for data transmission

$$E_{T_x}(l,d) = \begin{cases} lE_{T-elec} + ld\epsilon_{amp}, & \text{free space} \\ lE_{T-elec} + ld^2\epsilon_{amp}, & \text{multipath} \end{cases} \tag{4}$$

and for data reception,

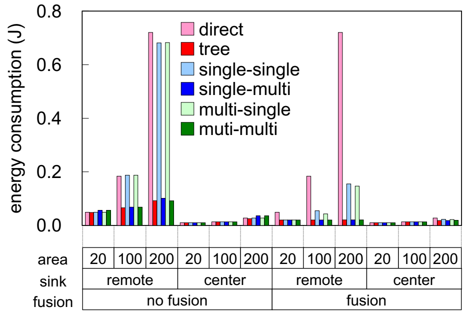
$$E_{R_x}(k) = kE_{R-elec}. \tag{5}$$

We evaluate the topology by using three metrics of evaluation. One is the total energy consumption required for one time data gathering from all nodes in the whole WSNs. The maximum delay in gathering sensor data from all nodes is also considered. For a node with the limited energy capacity, the energy consumption is obviously a major concern. At the same time, several practical applications have limits on acceptable latency, as specified by QoS requirements. For example, sensor data must be gathered as fast as possible even consuming much energy in delay-critical applications such as emergency detection. Generally, higher level of energy saving is accomplished with a penalty of increased delay. When we consider the balance among the energy consumption and the delay, the product of the amount of energy consumption and the delay, i.e. energy$\times$delay is a good measure for optimization as considered in [34]. Although we conducted thorough experiments and obtained a tremendous amount of data, we show some characteristic results in the next section.

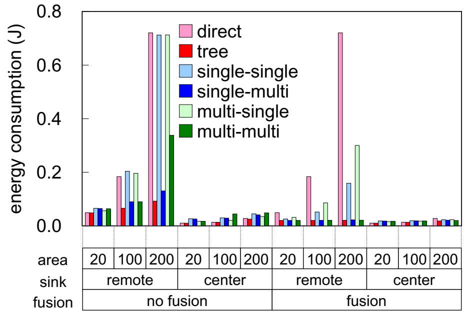
3.5 Results of Comparison

Figure 5 shows the energy consumption when the maximum transmission range is set to infinity and the e-LEACH model is used (scenarios 1 through 12 in Table 2). Figures 5(a), 5(b), and 5(c) are the total amount of energy consumed in gathering sensor data from all nodes in the optimal topology which minimizes energy consumption, delay, and energy$\times$delay, respectively. From these figures, it is apparent that the direct topology, the single-single topology, and the multi-single topology which involve the direct transmission to the sink node consume considerably larger energy than the others in scenarios with the distant sink node. There is no significant difference among the other three topologies, i.e. tree, single-multi, and multi-multi when the optimization is performed to minimize the energy consumption (Fig. 5(a)). When we compare them under the strategy of delay minimization, it can be noticed that the multi-multi topology leads to the highest energy consumption. The reason is that the multi-multi topology forms multi-hop tree in a cluster, which results in large delay. In order to suppress this delay, the multi-multi topology uses the long range transmission between cluster heads by having large clusters, because the long range transmission can shortcut to the sink node. Also in the case of energy$\times$delay minimization, the energy consumption is in order of tree, single-multi, and multi-multi, but the difference is small and 148 % at average, 320 % at a maximum.

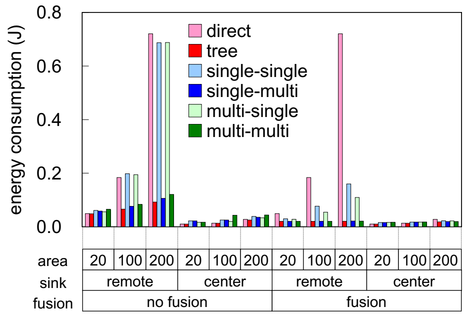
Next in Fig. 6, the maximum delay in data gathering is compared among topologies in the same



(a) Energy consumption of optimal topology minimizing energy consumption

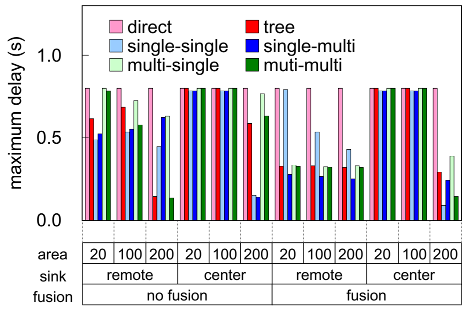


(b) Energy consumption of optimal topology minimizing delay

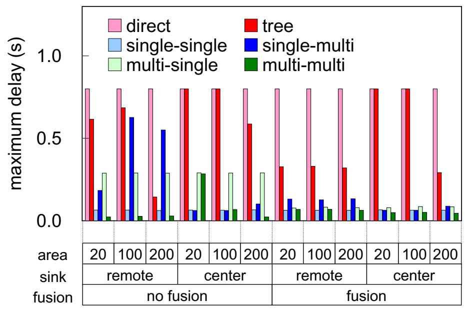


(c) Energy consumption of optimal topology minimizing energy$\times$delay

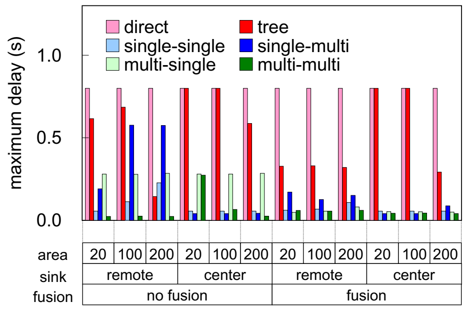
Figure 5: Comparison in energy consumption



(a) Delay of optimal topology minimizing energy consumption



(b) Delay of optimal topology minimizing delay

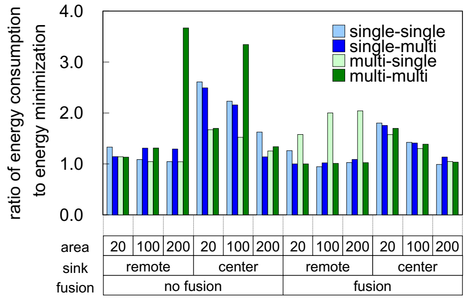


(c) Delay of optimal topology minimizing energy$\times$delay

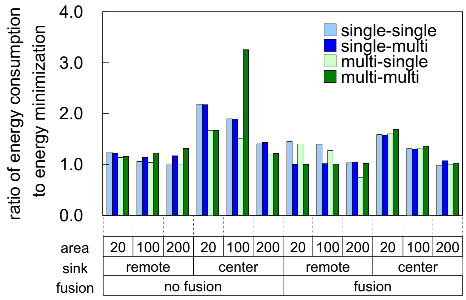
Figure 6: Comparison in delay

12 scenarios. When we want to minimize the energy consumption, the best topology for the small delay changes among the scenarios (Fig. 6(a)). The delay in the direct topology is always the same, because it is the product of the number of nodes and the slot length in time. This is the worst delay in most of the cases. The tree or cluster topology enables nodes which is far from each other to use parallel transmission, but in the direct topology all nodes always exist the interfering area of each other. The tree topology, which can save energy consumption in data gathering (Fig. 5), results in the larger delay, because the height of tree tends to be high to shorten the communication range for energy saving. Next in Fig. 6(b), the multi-multi topology seems the good topology, because it can suppress the delay between cluster heads by using long range transmission and it can also reduce the delay in a cluster than the single-multi and single-single topologies by forming a tree in a cluster. One may notice that the delay of the multi-multi topology is exceptionally high when the sink node is located at the center of 20×20 m^2 region and data fusion is not possible. It is because the nodes near the sink node directly connect to the sink node for suppressing the energy consumption. The same tendency can be observed in Fig. 6(c) where we consider to minimize energy$\times$delay.

Now, we compare the effect of minimization criteria. Especially when we focus on topologies with clustering, minimization of delay or energy$\times$delay increases the amount of consumed energy at about 75 to 370 % than minimization of energy consumption (Fig. 7). The reason why it is 75 % lower than the minimization of energy consumption, is because the both of the topologies is not always the optimal topology. We use the SA algorithm, and it does not always lead to the optimal topology. Moreover, we can see that the energy consumption in delay minimization is higher than that in energy$\times$delay minimization. On the contrary, minimization of energy consumption results in 1.1 to 34 times as long delay as minimization of delay (Fig. 8). On the other hand, minimization of energy$\times$delay results in 0.87 to 3.6 times as long delay as minimization of delay (Fig. 8). Moreover, we can see that the delay in energy$\times$delay minimization is as low as that in delay minimization, except for the single-single topology in a few scenarios. The reason for these exceptional cases is that the delay of the single-single topology in delay minimization can achieve very small delay (Fig. 6(b)) by making large clusters and having long range transmission from cluster heads to the sink node. However, by having long range transmission, such topologies consume much energy as shown in Fig. 5(b). From Figs. 7 and 8, we can conclude that energy$\times$delay is a good metric, because topologies minimizing energy$\times$delay accomplish as low energy consumption as



(a) Topologies minimizing delay

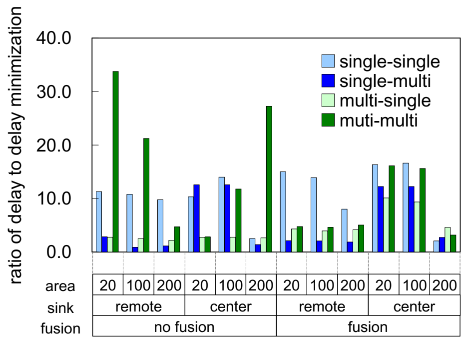


(b) Topologies minimizing energy$\times$delay

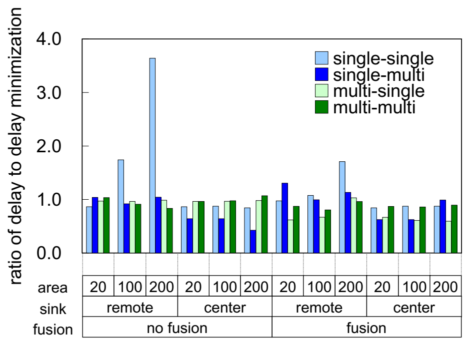
Figure 7: Ratio of energy consumption to energy minimization

topologies minimizing energy consumption and as small delay as topologies minimizing delay except for some scenarios.

Finally, we evaluate topologies which minimize the energy$\times$delay in terms of the energy$\times$delay in Fig. 9. To have a closer look, the figure is magnified in the range of 0.00 to 0.05. In the figure, the multi-multi topology leads to the lowest energy$\times$delay in most scenarios. When the sink is located at the center of a small monitored region, i.e. 20×20 m^2 or 100×100 m^2, and the data fusion is not possible, the energy$\times$delay of the multi-multi topology is higher than that of the single-single and single-multi topologies. The reason for this is that the size of sensor data increases as being relayed in a two-tier tree topology. On the contrary, in the single-single and single-multi topologies, there is less chance to aggregate multiple sensor data in data gathering. Since the single-single topology results in the larger delay for contention among cluster heads, the



(a) Topologies minimizing energy consumption



(b) Topologies minimizing energy$\times$delay

Figure 8: Ratio of delay to delay minimization

single-multi topology is the best for these two scenarios.

Before proposing our topology control mechanism, we examine the best topology in more detail. Figure 10 shows examples of the best topology with or without data fusion on the same distribution of sensor nodes. The minimization metric is the energy$\times$delay. Red open circles represent cluster heads, blue open circles correspond to cluster members, and red filled circles denote a sink node. The region size is 200$\times$200 m^2, the transmission range is infinity, the sink node is located at the center of the monitored region, and the e-LEACH model is used (scenarios 6 and 12). The multi-multi topology is the best for the case with data fusion. In Fig. 10(a), cluster

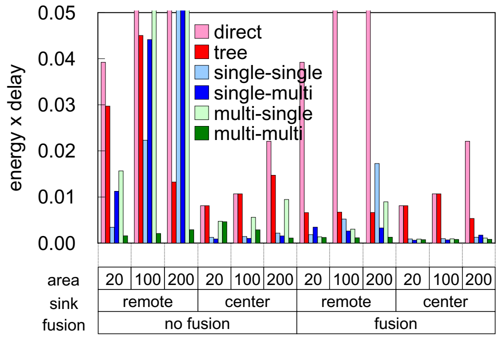
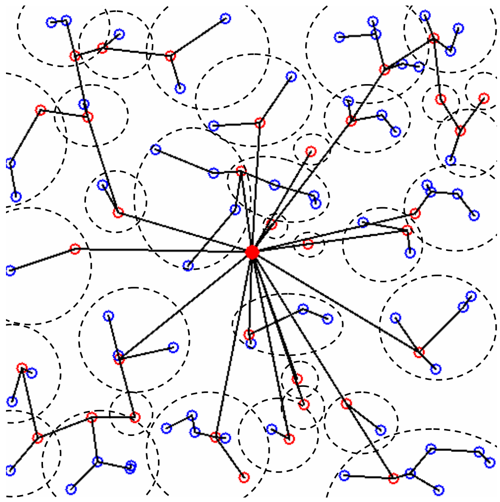
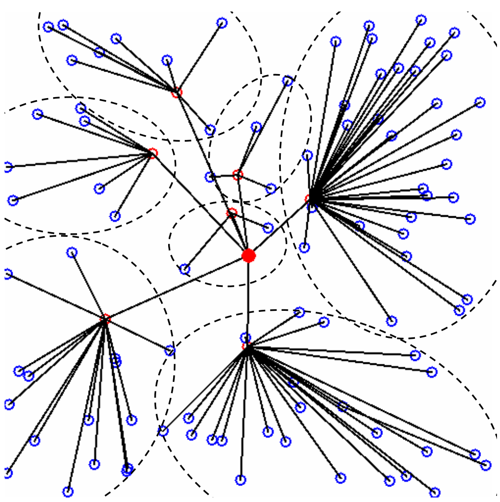


Figure 9: Evaluation on energy$\times$delay



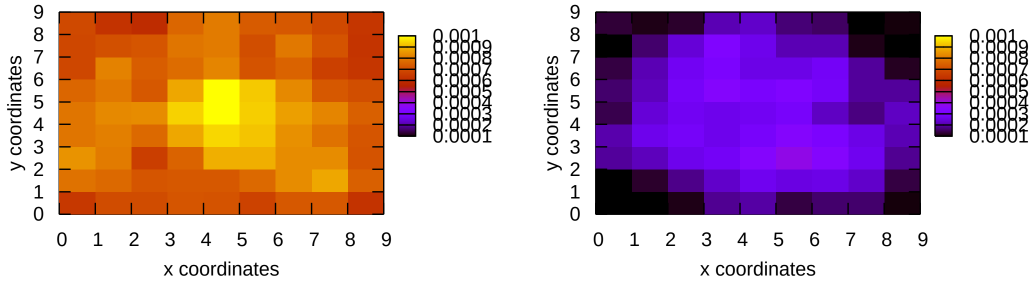
(a) Best topology with data fusion



(b) Best topology without data fusion

Figure 10: Best topology for data gathering

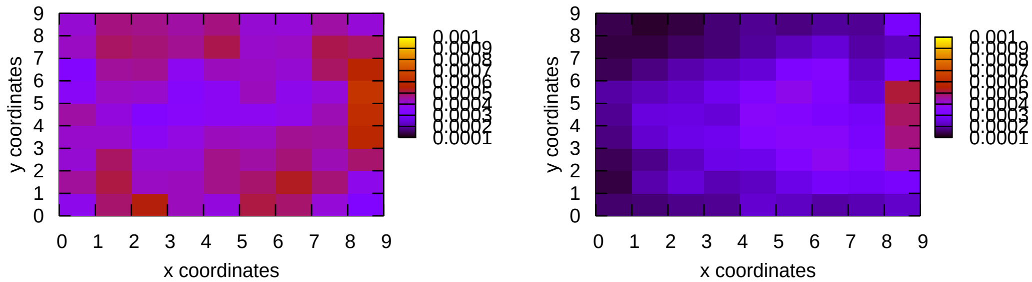
members first transmit sensor data to a cluster head by multi-hop communication. Then a cluster head fuses received data and data of itself into one and sends the aggregated data to the sink node by multi-hop communication. On the contrary, when data fusion is not applicable, the single-multi topology is the best. First, cluster members directly transmit their sensor data to a cluster head. Then, cluster heads transmit the aggregated data, which amounts in proportional to the number of cluster members plus one to a sink by multi-hop communication among cluster heads. In the case that data fusion is not applicable, the number of hops should be kept small to avoid the increased



(a) With data fusion

(b) Without data fusion

Figure 11: Point density of cluster heads with a sink node which is arranged at center



(a) With data fusion

(b) Without data fusion

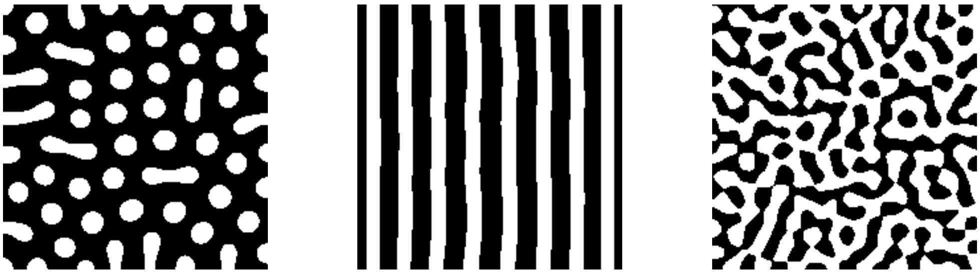
Figure 12: Point density of cluster heads with a sink node which is arranged at remote area

energy consumption in transmitting and receiving the large size of sensor data accumulated hop by hop. Consequently, the distance between cluster heads becomes large. Therefore, to save energy consumption by limiting the transmission power, multi-hop communication is used for sending sensor data from cluster heads to the sink node. However, multi-hop communication also leads to the higher energy consumption for increased data size. It means that there is the trade-off. In Fig. 10(b), although cluster heads can form a tree rooted at the sink node, they prefer to transmit sensor data directly to the sink considering the trade-off. As the size of the monitored region becomes larger, cluster heads begin to organize a tree.

Finally, we consider the point density of cluster heads to investigate the distribution of cluster heads in the best topologies. The point density ρ_{CH} of cluster heads can be calculated as,

$$\rho_{CH} = \frac{n}{A}, \tag{6}$$

where is A the size of region and n is the number of cluster heads in the region. In Fig. 11, the point density of cluster heads, averaged over 100 simulation runs, is shown for the topologies in Fig. 10. The 200×200 m^2 monitored region is divided into one hundred of 20×20 m^2 regions. The total numbers of cluster heads are 33 and 7 in Fig. 10(a) and Fig. 10(b), respectively. From the figure, it can be seen that the point density is high around the center, i.e. the location of a sink. In Fig. 11(a), the lowest density of 0.0004 is at the corner (9,0) and the highest density of 0.0012 is near the center (5,5). In Fig. 11(b), the lowest density of 0.0001 is also at the corner (0,0) and the highest density of 0.0004 is also near the center (5,5). It means that there are more clusters near a sink node at the center. It is because that the closer to the sink node the cluster head are, the more data the cluster heads must transfer to the sink nodes, since the number of nodes near the sink node is smaller than that of nodes father form the sink node. A lot of data brings high delay, because the cluster heads near the sink become a bottleneck of data flow. Therefore, a lot of node become cluster heads to assuage the bottleneck. Figure 12 also shows the point density of cluster heads with a distant sink node located in the region (15, 5). These figure also show that there are more clusters near a sink node. Therefore, we need to distribute cluster heads to have the heterogeneous distribution in a self-organizing way.



(a) spots (b) stripes (c) maze

Figure 13: Example of generated patterns by reaction-diffusion model

4 Reaction-Diffusion based Topology Control for Data Gathering

In this thesis, we adopt a reaction-diffusion model to organize the best topology revealed in the previous section in an autonomous and self-organizing manner. We focus on the similarity among the best topology for data gathering and the pattern generated by a reaction-diffusion model. The best topology has the periodic distribution of cluster heads, whereas the point density of cluster heads depends on the distance to a sink node. On the other hand, a reaction-diffusion model also organizes a periodic pattern such as spots, stripes, and maze as illustrated in Fig. 13. In addition, there is the mathematical form to derive the wavelength of a generated pattern from parameters. It makes it possible to adjust the point density of cluster heads in accordance with the distance from the sink by regulating the wavelength of a reaction-diffution pattern.

4.1 Reaction-Diffusion Model

A reaction-diffusion model is a mathematical model for pattern generation on the surface of body of fishes and mammals. Generally, pattern generation in a reaction-diffusion model is based on interactions among two virtual morphogens, i.e. activator and inhibitor. Reaction and diffusion of the two morphogens make spatially heterogeneous distribution of their concentrations, i.e. a pattern. Depending on the form of reaction-diffusion equations and their parameters, a variety of patterns can be generated such as illustrated in Fig. 13. A reaction-diffusion equation of two

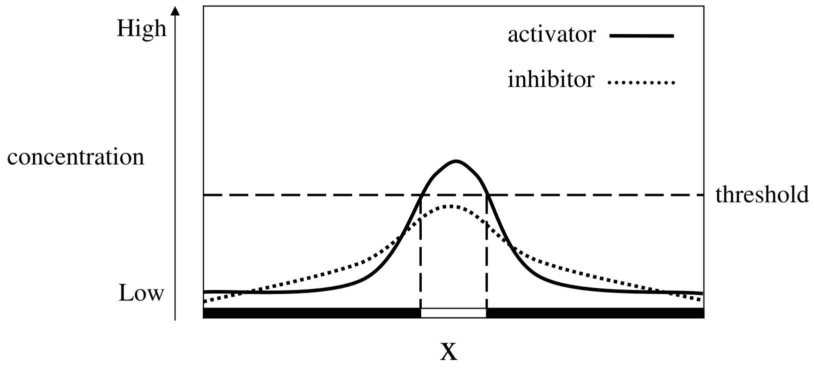


Figure 14: Pattern generation in reaction-diffusion model

morphogens can be written as,

$$\begin{cases} \dfrac{\partial u}{\partial t} = F(u, v) + D_u \nabla^2 u, \\ \dfrac{\partial v}{\partial t} = G(u, v) + D_v \nabla^2 v, \end{cases} \tag{7}$$

where u and v are the concentrations of activator and inhibitor, respectively. The first term of the right-hand side is a reaction team and the second term is a diffusion term. F and G are nonlinear functions for chemical reactions. D_u and D_v are the diffusion rate of activator and inhibitor, respectively. ∇^2 is the Laplacian operator.

In a reaction-diffusion model, the following conditions must be satisfied to generate patterns,

(1) The activator activates itself and the inhibitor, whereas the inhibitor restraints the activator.

(2) The inhibitor diffuses faster than the activator ($D_v > D_u$).

A mechanism of pattern generation can be explained as follows. In Fig. 14, those hypothetical chemicals are arranged in a line on the x-axis. The y-axis corresponds to the concentrations of activator and inhibitor. Now, consider that the concentration of activator has a peak at the center by a slight perturbation. The concentrations of activator and inhibitor are increased around the peak by self-activation. The generated inhibitor diffuses faster than the activator and restrains generation of activator at further regions. On the other hand at the peak, the concentration of activator is kept higher than that of inhibitor for different rates of diffusion. Consequently, the diversity in the concentration of activator emerges and a pattern appears. For example, when we color a point where the concentration of activator exceeds a certain threshold with white and others with black, we can see a black-white-black pattern shown at the bottom of Fig. 14.

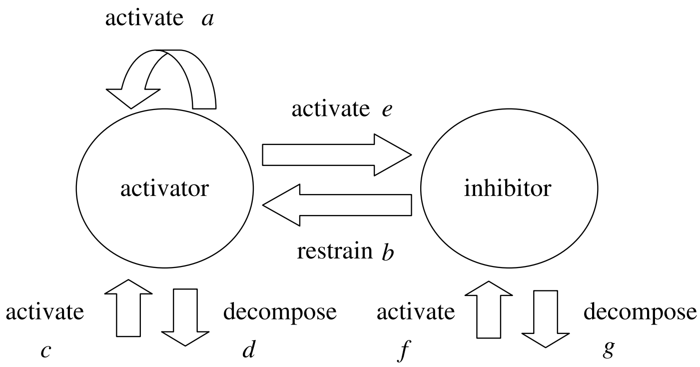


Figure 15: Dynamics of morphogens

In this thesis, we use the equations below for F and G, which model pattern generation on an emperor angelfish *pomacanthus imperator* [37].

$$\begin{cases} F(u,v) = \max\{0, \min\{au - bv + c, M\}\} - du, \\ G(u,v) = \max\{0, \min\{eu + f, N\}\} - gv, \end{cases} \tag{8}$$

where a and e correspond to the rate of activation and b is for that of inhibition. c and f are parameters for synthesis or increase of morphogens per unit time. d and g are for decomposition or decrease of morphogens per unit time. M and N are constants of limit. Figure 15 illustrates chemical reactions of the morphogens following the above reaction-diffusion equation. In order to generate patterns, the parameters must satisfy Turing conditions shown below.

$$a - d - g < 0, \tag{9}$$

$$eb - (a - d)g > 0, \tag{10}$$

$$D_v(a - d) - D_u g > 0, \tag{11}$$

$$(D_v(a - d) - D_u g)^2 - 4D_u D_v(eb - (a - d)g) > 0. \tag{12}$$

As far as these conditions are satisfied, the space will have the spatial heterogeneity in terms of the concentration of morphogens and a variety of patterns such as spots, stripes, and maze can be generated [38, 39].

4.2 Application of Reaction-Diffusion Model to WSNs

We consider a WSN deployed in the region of $L \times L$ m^2. Sensor nodes are distributed at random or in the grid layout and nodes is able to control the transmission power at the necessary and sufficient level for data gathering. In the case of pattern formation, i.e. clustering, sensor nodes fix the transmission range at R m to exchange information about morphogen concentrations with close neighbors. Then, the maximum distance between the nodes in the cluster is R m.

In WSNs, since nodes are discretely arranged in the monitored region and messages are exchanged at regular intervals, we spatially and temporally discretize Eqs. (7) and (8) as follows [38, 40].

$$\begin{cases} u_t = u_{t-1} + \Delta t \Big\{ F(u_{t-1}, v_{t-1}) + D_u \sum_{j \neq i} \kappa_{i,j}(u_j - u_i) \Big\}, \\ v_t = v_{t-1} + \Delta t \Big\{ G(u_{t-1}, v_{t-1}) + D_v \sum_{j \neq i} \kappa_{i,j}(v_j - v_i) \Big\} \end{cases} \tag{13}$$

$$\begin{cases} F(u_{t-1}, v_{t-1}) = \max\{0, \min\{au_{t-1} - bv_{t-1} + c, M\}\} - du_{t-1}, \\ G(u_{t-1}, v_{t-1}) = \max\{0, \min\{eu_{t-1} + f, N\}\} - gv_{t-1}. \end{cases} \tag{14}$$

$$\kappa_{i,j} = \begin{cases} |x_j - x_i|^{-2}, & |x_j - x_i| \leq R, \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

We call the duration between the t-th control timing and the $t + 1$-th control timing as the t-th control interval. Each node calculates this equation at the timing t by using the information on morphogen concentrations received from the neighboring nodes during $t - 1$-th control interval. Consequently, a global spatial pattern of the concentration of activator emerges.

The wavelength λ of the generated pattern can be derived as [38, 41],

$$\lambda = \sqrt{\pi^3 \rho_{node} R^2} \sqrt[4]{\frac{D_u D_v}{eb - g(a - d)}}, \tag{16}$$

where, ρ_{node} is the point density of nodes, which is defined as the number of nodes per area size.

4.3 Self-Organizing Topology Control based on Reaction-Diffusion Model

We want to form the topology where cluster heads are close to each other near the sink node and sparsely distributed at the edge, as mentioned in Section 3.5. The basic behavior of our novel topology control mechanism based on a biological self-organizing mechanism in organizing the best topology is as follows. Our mechanism has two phases, i.e. pattern generation phase and topology organization phase. First, nodes generate a spatial pattern by exchanging information of

Table 6: Parameter setting for reaction-diffusion equation

parameter	value	parameter	value
a	0.08	g	0.06
b	0.08	M	0.2
c	0.02	N	0.5
d	0.03	D_u	0.02
e	0.1	D_v	0.5
f	0.15	Δt	0.01

morphogen concentration and calculating of the reaction-diffusion equation. By repeating message exchanges and reaction-diffusion calculation, a spatial pattern of morphogen concentrations emerges. Then, a topology for data gathering is organized based on the patten in the topology organization phase. Our mechanism organizes a topology for energy-efficient and low-delay data gathering, but the lifetime of a WSN also depends on the way how sensor data are gathered on the topology. In addition, a mechanism to rotate the role of cluster head among sensor nodes is also needed to prolong the lifetime of a WSN by balancing the energy consumption among nodes. However, they remain our future research topics.

4.3.1 Pattern Generation Phase

Each node maintains the concentration of activator and inhibitor, and an estimated distance to the sink which are used for heterogeneous distribution of cluster heads, for its neighboring nodes and itself. At the beginning of this phase, the sink node broadcasts the message, which is only a signal without any information, to announce the beginning of pattern generation phase and to synchronize all nodes. However, each node starts this phase after waiting for random milli-seconds. At the t-th control timing, a node calculates Eqs. (13) and (14) with parameters summarized in Table. 6 [42], by using the information on morphogen concentrations of neighbors, which the node has received during the $t-1$-th control interval. If a node did not receive concentration information from a neighbor in this interval, it uses the latest information it received instead. Then, it broadcasts information about its new morphogen concentrations to the neighbors. Nodes behave in an asynchronous manner. It means that timing of message emission and reaction-diffusion calculation are different among nodes. As nodes exchange information and calculate the reaction-diffusion equation, the spatial pattern eventually emerges. In this mechanism, we assume that morphogen concentrations converge at 3000 times of calculation and after this a node moves to the topology organization phase. However, it is able to achieve the same result with less calculations by using the acceleration methods in [23]. Since all nodes are synchronizing at the beginning of this phase, all nodes can finish the calculation at the almost same time.

To generate a pattern where cluster heads are close to each other near the sink node and sparsely distributed at the edge as in Fig. 11, we dynamically change the wavelength λ of a pattern depending on the distance to the sink node. Among parameters which dominates the wavelength in Eq. (16), we regulate the diffusion rates D_u and D_v, because D_u and D_v directly change the wavelength and the other parameters need more careful setting to satisfy the Turing condition. Since the ratio of D_v to D_u determines the type of pattern, we keep the fraction as $D_v = 25D_u$. Then, Eq. (16) can be solved for D_u as follows,

$$D_u(\rho_{CH}(d_i)) = \frac{\lambda(\rho_{CH}(d_i))^2}{5\pi^3\rho_{node}R^2}\sqrt{eb-(a-d)g} \tag{17}$$

$$\lambda(\rho_{CH}(d_i)) = \frac{2}{\sqrt{\pi\rho_{CH}(d_i)}}, \tag{18}$$

where, d_i is the distance of node i to the sink node, $\rho_{CH}(d_i)$ is the point density of cluster heads at the distance d_i, $\lambda(\rho_{CH}(d_i))$ is the wavelength at the distance d_i, and ρ_{node} is the point density of nodes, which are calculated by dividing the number of nodes by the area of the monitored region. Here, we assume that we know $\rho_{CH}(d_i)$ by simulating the best topology on the monitored region with the same density of nodes. The same $\lambda(\rho_{CH}(d_i))$ is applicable as far as the node density ρ_{node} and the location of sink node are the same, independently of the exact location of sensor nodes. For the stability of reaction-diffusion pattern generation, a node calculate $D_{u(i,j)}$ and $D_{v(i,j)}$, which are the diffusion rates at a middle point of the node i and the node j , by using Eq. (17) with the distance $d_{i,j}$ of the middle point to the sink node. A node calculates $D_{u(i,j)}$ and $D_{v(i,j)}$ with every neighboring nodes, and Eq. (13) is extended as,

$$\begin{cases} u_t = u_{t-1} + \Delta t\Big\{F(u_{t-1}, v_{t-1}) + \sum_{j\neq i} D_{u(i,j)}\kappa_{i,j}(u_j - u_i)\Big\}, \\ v_t = v_{t-1} + \Delta t\Big\{G(u_{t-1}, v_{t-1}) + \sum_{j\neq i} D_{v(i,j)}\kappa_{i,j}(v_j - v_i)\Big\}. \end{cases} \tag{19}$$

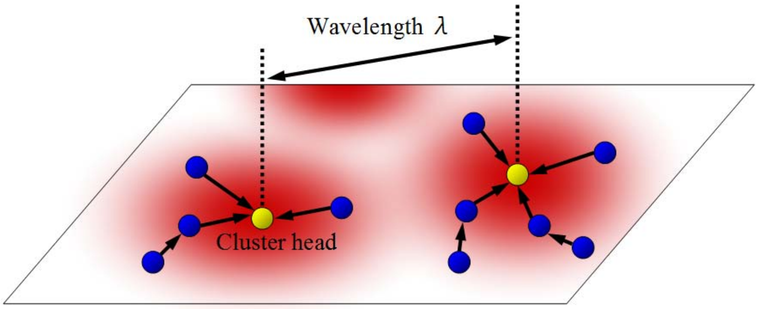


Figure 16: Topology control by reaction-diffusion spatial patten

In WSNs, since it is difficult for a node to know the exact distance d_{sink} from the sink node, a node estimates the distance. In our proposal, each node estimates the distance to the sink node by exchanging the estimated distance to the sink in the pattern generation phase. We assume that nodes can estimate distance to a neighbor node from received signal strength indicator (RSSI). Nodes near the sink node can estimate the distance to the sink node and inform neighbor nodes of the distance together with the information about the concentrations of morphogen. Then, neighbors receiving this information calculate the sum of the distance from a sender to the sink node and estimated distance to the sender. A node estimates its distance based on each of messages it received and chooses the smallest one as its distance d_{sink}. As the distance information propagates to the edge of a WSN through message exchanges, all nodes eventually obtain the estimate of distance.

4.3.2 Topology Organization Phase

Once a pattern is generated in the pattern generation phase, sensor nodes begin to organize a topology for data gathering, i.e. select cluster heads. A spatial pattern generated by the reaction-diffusion model has peaks of the activator concentration. The node i first identifies the minimum λ_{min} of $\lambda(\rho_{CH}(d_{i,j}))$ with the neighbor node j. The node i becomes a cluster head if it has the highest concentration of activator among nodes in the region of radius of $\lambda_{min}/2$. The other nodes can send sensor data to the closest cluster head by forwarding sensor data following the gradient of the activator concentration as shown in Fig. 16. Cluster heads need to organize an inter-cluster tree, but it is out of scope of our mechanism.

No.	R	Energy model	Fusion	Sink	Area
6	40 m	e-LEACH ($d_0 = 75$)	$\times$	Center (100,100)	200×200
12	40 m	e-LEACH ($d_0 = 75$)	$\bigcirc$	Center (100,100)	200×200

Table 7: Scenarios for simulation

5 Simulation

In this section, we verify our self-organizing topology control mechanism proposed in the previous section through simulation experiments. The feasibility of the proposal is evaluated in comparison with the optimal topology obtained in section 3.

5.1 Simulation Settings

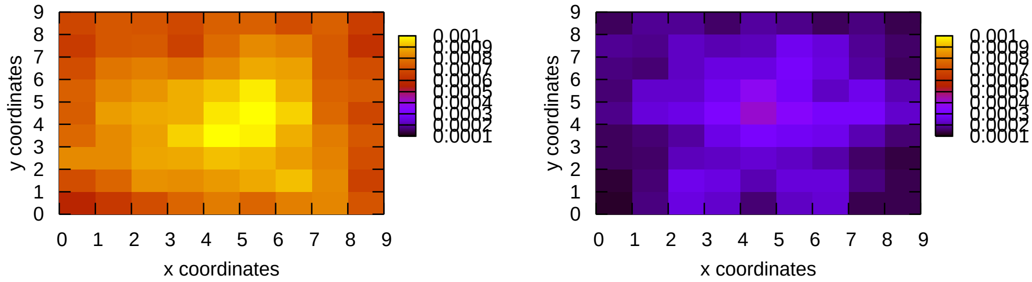
We conduct experiments based on the scenarios 6 and 12 (summarized in Table 7), which are used for generation of Figs. 10 and 11. The maximum transmission range is set at 40 m instead of infinity as R in Eq. (17). Within the range of 40 m, a node has 12–13 neighboring nodes, and it is the sufficient number for generating a pattern by the reaction-diffusion equation. In addition, to avoid the boundary condition effect, we simulate the case of 400 nodes distributed in the monitored region of 400×400 m^2, but results are obtained for the center region of 200×200 m^2 with 100 nodes.

We consider two cases where all nodes know the exact geographical distance to the sink node and where nodes estimate the distance $\tilde{d}_i$ by using the scheme explained in Section 4.3.

5.2 Simulation Results

We compare topologies generated by our self-organizing topology control mechanism with those generated by solving the optimization problem form viewpoints of energy consumption and delay. We should note here that energy consumed in topology generation and its delay are out of scope of evaluation here.

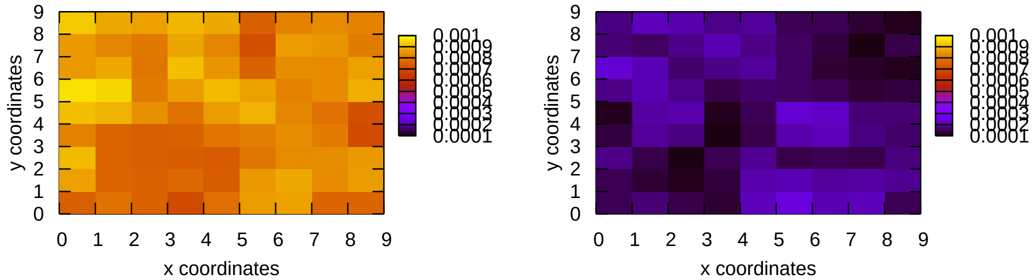
Figure 17 shows the point density of cluster heads averaged over 100 simulation runs for the case that the exact distance to the sink is known. We evaluate the similarity between Fig.17 and Fig. 11 by Peak Signal-to Noise Ratio (PSNR), which is widely used as a measure of quality of a



(a) Fusion

(b) No fusion

Figure 17: Point distribution of cluster heads with distance information



(a) Fusion

(b) No fusion

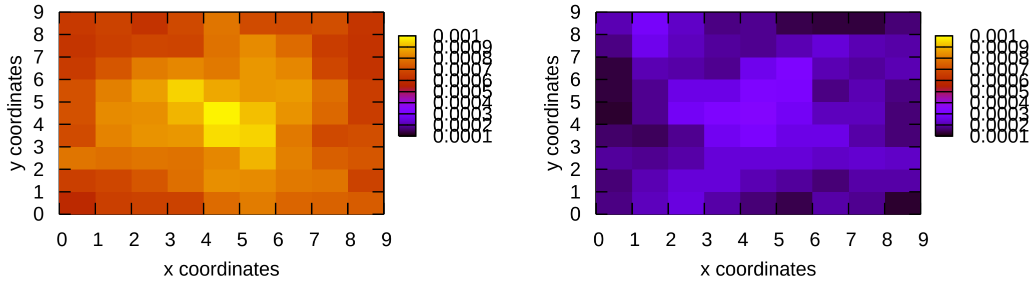
Figure 18: Point distribution of cluster heads without wavelength control

coded image to the original image. PSNR can be derived as,

$$PSNR = 10 \log_{10}(\frac{MAX_I^2}{MSE}), \tag{20}$$

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} ||I(i,j) - K(i,j)||^2, \tag{21}$$

where MAX_I is the maximum pixel value, for which we use the maximum density in deriving PSNR for Fig. 17, $I(x,y)$ and $K(x,y)$ gives the value of pixel at the coordinate (x,y) in an image of $m \times n$ pixels. PSNR are 18.4 dB in Fig. 17(a) and 11.7 dB in Fig. 17(b), respectively. For



(a) Fusion

(b) No fusion

Figure 19: Point distribution of cluster heads without distance information

comparison, we show the case without the wavelength control of Eqs. (17) and (18) in Fig. 18. PSNR are 15.2 dB in Fig. 18(a) and 8.96 dB in Fig. 18(b), respectively. The increase of PSNR by 3.2 dB and 2.8 dB means that MSE (Mean Square Error) decreases by 52 % and 47 % in comparison with the case without the wavelength control, respectively. Therefore, the wavelength control contributes to organize a topology more similar to the optimal topology.

In Figs. 18 and 19, results for the case where a node estimates the distance $\tilde{d}_i$ from the sink node are shown. PSNR is increased from 15.2 dB to 16.7 dB when data fusion is possible, and from 8.96 dB to 11.3 dB when data fusion is not possible, by using the wavelength control, respectively.

Finally, we evaluate the energy$\times$delay when each node conjectures the distance $\tilde{d}_i$ to the sink. Since our proposal cannot organize a tree of cluster heads as in the multi-multi and single-multi topologies, we assume that SPT or MST is applied to inter-cluster tree formation in derivation of the amount of energy consumption and the gathering delay. Since the number of cluster heads is relatively small against the number of nodes, we can use centralized or fully-distributed methods [19, 28] for organizing the inter-cluster topology. Results are shown in Fig. 20, together with results of the optimal topology for the tree topology, the single-multi topology, and the multi-multi topology in the same scenarios. In Fig. 20, it can be seen that the product of the amount of energy consumption and the delay with the proposal is as small as that in the single-multi topology or the multi-multi topology. The difference is only about 8.4 % with data fusion, and 30 % without data fusion. Therefore, we can conclude that our proposal can generate a topology, which enables

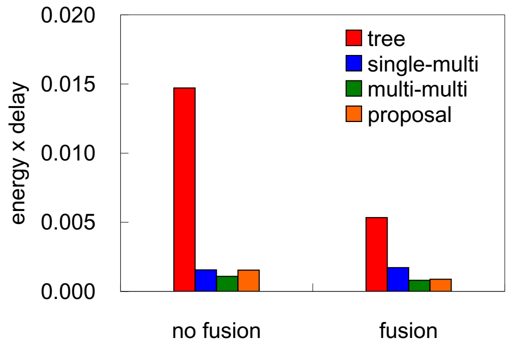


Figure 20: Comparison of our proposal with the best topologies

data gathering as energy efficient and low delay as the optimal topology does.

6 Conclusions

Energy-efficient and low-delay data gathering is the most important in WSNs. In this thesis, we first investigated the best topology for data gathering from the viewpoints of energy efficiency and delay. We considered six types of topologies, i.e. direct, tree, single-single, single-multi, multi-single and multi-multi. Through extensive comparison among them under various environments, we proved that adopting multi-hop communication in both of inter and intra clusters enabled energy-efficient and low-delay data gathering in a balanced manner for most of scenarios.

Following this result, we proposed a novel topology control mechanism based on a biological mechanism, that is, the reaction-diffusion model to organize the best topology in a self-organizing and autonomous way. Through simulation experiments, it was shown that a topology organized by our proposal accomplished as small energy consumption and delay as the optimal topology derived under an ideal condition does.

Although the effectiveness of our mechanism was verified through simulation experiments, we require further improvement. We plan to extend our proposal to have a mechanism of forming an inter-cluster tree. In addition, we also need a mechanism to rotate the role of cluster head among sensor nodes for balanced and homogeneous energy consumption leading to the longer lifetime of a WSN.

Acknowledgments

First, I would like to express my sincere gratitude to my supervisor, Professor Masayuki Murata of Osaka University, for his continuous support and expensive advice throughout my studies, and providing me this precious study opportunity in his laboratory.

I especially would like to express my deepest appreciation to Associate Professor Naoki Wakamiya of Osaka University, for his elaborated guidance and invaluable firsthand advice. All works of this thesis would not been possible without his support.

I am most grateful to Chief Researcher Takao Yamaguchi and Staff Researcher Astushi Yoshida of Intelligent Processing Technology Research Laboratory, Advanced Technology Research Laboratories, Matsushita Electric Industrial Co., Ltd., and Assistant Professor Etsushi Nakaguchi of Osaka University for their valuable cooperation and helpful advice.

I am very grateful to Professors Koso Murakami, Makoto Imase, Teruo Higashino, and Hirotaka Nakano of Osaka University, for their appropriate guidance.

I am also indebted to Associate Professor Go Hasegawa, Specially Appointed Associate Professor Kenji Leibnitz, Assistant Professors Shin'ichi Arakawa, Masahiro Sasabe, and Yuichi Ohsita of Osaka University for providing me helpful comments and feedback.

I heartily thank my seniors and colleagues in the Department of Information Networking, Graduate School of Information Science and Technology of Osaka University for their help. Our conversations and work together have greatly influenced this thesis.

This research was partly supported by "Special Coordination Funds for Promoting Science and Technology: *Yuragi Project*" of the Ministry of Education, Culture, Sports, Science and Technology, Japan.

Finally I would like to extend my indebtedness to my family for their understanding, support, encouragement and sacrifice throughout my study.

References

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless Sensor Networks: a Survey," *Computer Networks*, vol. 38, pp. 393–422, Mar. 2002.

[2] T. Arampatzis, J. Lygeros, and S. Manesis, "A Survey of Applications of Wireless Sensors and Wireless Sensor Networks," in *Proceedings of the 13th Mediterranean Conference on Control and Automation (MED 2005)*, pp. 719–724, June 2005.

[3] E. Bonabeau, M. Dorigo, G. Theraulaz, and I. NetLibrary, *Swarm Intelligence from Natural to Artificial Systems*. Oxford University Press, Aug. 1999.

[4] F. Dressler, *Self-Organization in Sensor and Actor Networks*. John Wiley & Sons, Nov. 2007.

[5] R. Huang, J. Zhu, X. Yu, and G. Xu, "The Ant-Based Algorithm for the Optimal Many-to-one Routing in Sensor Networks," in *Proceedings of International Conference on Communications, Circuits and Systems (ICCCAS 2006)*, vol. 3, pp. 1532–1536, June 2006.

[6] S. Selvakennedy, S. Sinnappann, and Y. Shang, "T-ANT: A Nature-Inspired Data Gathering Protocol for Wireless Sensor Networks," *Journal of Communication*, vol. 1, issue 2, pp. 22–29, May 2006.

[7] P. Boonma and J. Suzuki, "BiSNET: A Biologically-Inspired Middleware Architecture for Self-Managing Wireless Sensor Networks," *Computer Networks: The International Journal of Computer and Telecommunications Networking*, vol. 51, pp. 4599–4616, Nov. 2007.

[8] F. Dressler, "Bio-inspired Promoters and Inhibitors for Self-Organized Network Security Facilities," in *Proceedings of the 1st IEEE/ACM International Conference on Bio-Inspired Models of Network (BIONETICS 2006)*, pp. 1–7, Dec. 2006.

[9] M. Dorigo, *Ant Colony Optimization*. MIT Press, July 2004.

[10] K. Akkaya and M. Younis, "A Survey on Routing Protocols for Wireless Sensor Networks," *Ad Hoc Networks*, vol. 3, pp. 325–349, May 2005.

[11] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," in *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences (HICSS-33)*, pp. 1–10, Jan. 2000.

[12] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," *IEEE Transactions on Wireless Communications*, vol. 1, pp. 660–670, Oct. 2002.

[13] W. Choi, P. Shah, and S. K. Das, "A Framework for Energy-Saving Data Gathering using Two-Phase Clustering in Wireless Sensor Networks," in *Proceedings of the first Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MOBIQUITOUS 2004)*, pp. 203–212, Aug. 2004.

[14] O. Younis and S. Fahmy, "Distributed Clustering in Ad-hoc Sensor Networks: A Hybrid, Energy-Efficient Approach," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM 2004)*, vol. 1, pp. 629–640, Mar. 2004.

[15] H. Chan and A. Perrig, "ACE: An Emergent Algorithm for Highly Uniform Cluster Formation," in *Proceedings of the First European Workshop on Sensor Networks (EWSN 2004)*, pp. 154–171, Jan. 2004.

[16] B. J. Culpepper, L. Dung, and M. Moh, "Design and Analysis of Hybrid Indirect Transmissions (HIT) for Data Gathering in Wireless Micro Sensor Networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 8, pp. 61–83, Jan. 2004.

[17] S. Lindsey, C. Raghavendra, and K. M. Sivalingam, "Data Gathering Algorithms in Sensor Networks Using Energy Metrics," *IEEE Transactions on Parallel and Distributed Systems*, vol. 13, pp. 924–935, Sept. 2002.

[18] G. Lu, B. Krishnamachari, and C. S. Raghavendra, "An Adaptive Energy-Efficient and Low-Latency MAC for Tree-based Data Gathering in Sensor Networks," *Wireless Communications and Mobile Computing*, vol. 7, pp. 863–875, Sept. 2007.

[19] H. O. Tan and I. Korpeoglu, "Power Efficient Data Gathering and Aggregation in Wireless Sensor Networks," *ACM SIGMOD Record*, vol. 32, pp. 66–71, Dec. 2003.

[20] J. Kamimura, N. Wakamiya, and M. Murata, "A Distributed Clustering Method for Energy-Efficient Data Gathering in Sensor Networks," *International Journal of Wireless and Mobile Computing*, vol. 1, pp. 113–120, Feb. 2006.

[21] Y. Taniguchi, N. Wakamiya, and M. Murata, "A Distributed and Selforganizing Data Gathering Scheme in Wireless Sensor Networks," in *Proceedings of the 6th Asia-Pacific Symposium on Information and Telecommunication Technologies (APSITT 2005)*, pp. 299–304, Nov. 2005.

[22] A. M. Turing, "The Chemical Basis of Morphogenesis," *Royal Society of London Philosophical Transactions Series B*, vol. 237, pp. 37–72, Aug. 1952.

[23] K. Hyodo, N. Wakamiya, E. Nakaguchi, M. Murata, Y. Kubo, and K. Yanagihara, "Experiments and Considerations on Reaction-Diffusion based Pattern Generation in a Wireless Sensor Network," in *Proceedings of the 1st IEEE International Workshop: From Theory to Practice in Wireless Sensor Networks (t2pWSN 2007)*, June 2007.

[24] Y. Chen and T. C. Henderson, "S-nets: Smart Sensor Networks," in *Proceedings of International Symposium on Experimental Robotics (ISER 2000)*, pp. 85–94, Dec. 2000.

[25] M. Durvy and P. Thiran, "Reaction-Diffusion Based Transmission Patterns for Ad Hoc Networks," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM 2005)*, vol. 3, pp. 2195–2205, Mar. 2005.

[26] A. Yoshida, K. Aoki, and S. Araki, "Cooperative Control Based on Reaction-Diffusion Equation for Surveillance System.," in *Proceedings of Ninth International Conference on Knowledge-Based Intelligent Information and Engineering Systems (KES 2002)*, pp. 533–539, Sept. 2005.

[27] K. Hyodo, N. Wakamiya, and M. Murata, "Reaction-Diffusion based Autonomous Control of Camera Sensor Networks," in *Proceedings of 2nd International Conference on Bio-Inspired Models of Network, Information, and Computing Systems (BIONETICS 2007)*, Dec. 2007.

[28] R. Cristescu and M. Vetterli, "Power Efficient Gathering of Correlated Data: Optimization, NP-Completeness and Heuristics," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 7, pp. 31–32, July 2003.

[29] E. Dijkstra, "A Note on Two Problems in Connexion with Graphs," *Numerische Mathematik*, vol. 1, pp. 269–271, Dec. 1959.

[30] R. Prim, "Shortest Connection Networks and Some Generalizations," *Bell System Technical Journal*, vol. 36, pp. 1389–1401, Nov. 1957.

[31] P. Agarwal and C. Procopiuc, "Exact and Approximation Algorithms for Clustering," *Algorithmica*, vol. 33, pp. 201–226, Dec. 2002.

[32] H. C. Jang, H. C. Lee, and J. X. Huang, "Optimal Energy Consumption for Wireless Sensor Networks," in *Proceedings of the 9th Joint Conference on Information Sciences (JCIS 2006)*, Oct. 2006.

[33] T. Murata and H. Ishibuchi, "Performance Evaluation of Genetic Algorithms for Flowshop Scheduling Problems," in *Proceedings of the 1st IEEE Conference on Evolutionary Computation (ICEC 2006)*, vol. 2, pp. 812–817, June 1994.

[34] S. Lindsey, C. Raghavendra, and K. Sivalingam, "Data Gathering in Sensor Networks using the Energy*Delay Metric," in *Proceedings of the 15th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2001)*, pp. 2001–2008, Apr. 2001.

[35] M. Beigl, T. Zimmer, A. Krohn, C. Decker, and P. Robinson, *Smart-its: Communication and Sensing Technology for UbiComp Environments*. University of Karlsruhe, Fak. fur Informatik, Feb. 2003.

[36] "MICAz Data Sheet." available at `http://www.xbow.com/Products/Product_pdf_files/Wireless_pdf/MICAz_Datasheet.pdf`.

[37] S. Kondo and R. Asai, "A Reaction-Diffusion Wave on the Kin of the Marine Angelfish Pomacanthus," *Nature*, vol. 376, pp. 765–768, Aug. 1995.

[38] J. D. Murray, *Mathematical Biology. I: An Introduction*. Springer, June 2002.

[39] J. D. Murray, *Mathematical Biology. II: Spatial Models and Biomedical Applications*. Springer, Jan. 2003.

[40] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill Companies, Feb. 1991.

[41] H. Shoji and Y. Iwasa, "Labyrinthine Versus Straight-Striped Patterns Generated by Two-Dimensional Turing System," *Jounal of Theoretical Biology*, vol. 237, pp. 104–116, Apr. 2005.

[42] "About Reaction-Diffusion Model." available at `http://www.bio.nagoya-u.ac.jp/~z3/research/rdsoft.htm`.