Master's Thesis

Title

Autonomous Localization System in Wireless Sensor Networks

Supervisor

Prof. Masayuki Murata

Author

Yoshikazu Ohta

February 15th, 2005

Department of Information Networking

Graduate School of Information Science and Technology

Osaka University

Master's Thesis

Autonomous Localization System in Wireless Sensor Networks

Yoshikazu Ohta

Abstract

In wireless sensor networks, it is major that localization systems use data from sensors which receive signals from moving targets, measure Received Signal Strength Indicator (RSSI), and translate RSSI into the distance between the sensor and the target. It is expected that sensors monitor objects or an area for a long time. However, sensors have small battery capacity and it is important to save the energy of sensors and collect the necessary data. Furthermore, in an actual environment, it is difficult to arrange sensors uniformly like an grid, and the deployment density of the sensor has been biased. Therefore, it is needed to collect the applicable amount of data, regardless of sensors deployment.

In this thesis, we first focus localization systems in sensor networks and propose one model of localization systems. We also propose an efficient data collecting technique for localization system to get the accuracy required for the applications while saving energy. It is not needed many control messages and the system can collect data autonomously. In addition, we analyze data transfer in localization system and derive the number of data which system can collect and energy consumptions of a sensor, a target and a whole system, respectively. Finally, we verify that our proposal can efficiently collect necessary data to get accuracy. The results show that our proposed data collecting technique can work independently of the sensor density and topology. If localization system controls sensors appropriately, localization error gets about 1 m with actual measurement error model when sensors are deployed as deployment densities averaging $0.1/\mathrm{m}^2$.

Keywords

sensor networks, localization, Received Signal Strength Indicator (RSSI), data collection, Minimum Mean Square Error (MMSE)

Contents

List of Figures

List of Tables

1 Introduction

Recent advances in wireless communications and electronics have enabled the development of micro-sensors that can manage wireless communication and also have calculation power. By deploying a large number of sensors, wireless sensor networks can monitor large areas and be applied in a variety of fields, such as monitoring the environment; air, water, and soil. Also, wireless sensor networks can offer sensing data to context-aware applications that can adapt to user situations in ubiquitous computing. If properly conducted, sensors can work autonomously to measure temperature, humidity, luminosity and so on. Sensors send sensing data to a sink that has been deployed for data collection. Sensors can be easily deployed because they can communicate using wireless devices and so sensor networks need no other infrastructure [1].

In the future, sensors will be cheaper and deployed almost everywhere; therefore, services which depend on user location and localization of sensors will become more important. GPS [2] is a popular location estimation system, but since it needs signals from GPS satellites, it cannot work indoors. Using sensor networks instead of GPS makes indoor localization possible [3]. In the future, we expect that applications that satisfy location information requirements, such as navigation systems and target tracking systems in office buildings or in supermarkets will increase. The location of sensors is important too, because sensing data without knowing the location is usually meaningless in environmental sensing applications such as water quality monitoring, seismic intensity, and indoor air quality [4, 5].

For localization in wireless sensor networks is necessary to know the distance (or angle) between sensors and the target. The TDoA method using ultrasound or lasers realizes high accuracy; but, each device adds to size, costs, and energy requirements. Thus, such methods are not suitable for sensor networks. An inexpensive RF-based approach with low configuration requirements was researched. However, RSSI needs more data than other methods to achieve high accuracy, and collecting a large amounts of data causes an increase in traffic and sensors' energy consumption; thus, decreasing the lifetime of sensor networks. Furthermore, increasing delays in collecting data has a bad influence on real-time operations to obtain location information.

In this thesis, we propose a localization system that uses RSSI in wireless sensor networks to estimate the position of moving targets. Under a no control dense network, many sensors try to send data, and there are a long delay and a large energy consumption. To reduce data collected in a system and to extend the lifetime of sensor networks, we propose a data collecting technique in which sensors recognize the number of surrounding sensors. They autonomously decide whether to send sensing data, and can work with independence from the deployment of other sensors. This system does not need centralized control, complicated calculations, or the sending of a lot of packets. We evaluate performance of the data collecting technique by the amount of data collected and the energy consumption. We also evaluate the accuracy of our localization system by means of simulation experiments. The results show localization systems do not need a large amount of data and the localization errors depend on the density of sensors.

The remainder of this thesis is organized as follows. In Section 2, we explain background techniques for localization and our model of localization systems. In Section 3, we describe our proposed data collecting technique. In Section 4, we analyze data transfer in localization system and evaluate the number of data collected and energy consumption. In Section 5, we simulate the localization system and evaluate the accuracy. Finally, we conclude this thesis and mention future work in Section 6.

2 Localization system

2.1 Background

Localization methods are divided into range based localization and range free localization. Range based localization needs to know the distance (or angle) between sensors and the target. There are several ways to measure the distance (or angle) and estimate the position. The most popular methods to measure distances (or angles) for localization are:

- Received Signal Strength Indicator (RSSI)

  Receivers measure the power of the signal and translate it into distance, based on the known transmission power and effective propagation loss (Fig. 2(a)). Ordinarily, this method needs to measure the effective propagation loss beforehand and input this prior knowledge to receivers. The target position is calculated using data over three points. Maximum Likelihood (ML) estimation that estimates the position of a target by minimizing the differences between measured and estimated distances (Fig. 1). This method needs only RF and is suitable for wireless sensor networks, but measurement errors are larger than those with TDoA.

- Time of Arrival (ToA) and Time difference of Arrival (TDoA)

  The propagation time can be directly translated into distance, based on the known signal propagation speed. ToA measures the arrival time of ultrasound, but this method needs to synchronize senders (Fig. 2(b)). TDoA needs two types of signals and measures the arrival time difference of the two signals (Fig. 2(c)). These methods can be applied to many different signals, such as RF, acoustic, infrared and ultrasound. The methods can obtain high accuracy; however, each device adds to the size, cost, and energy requirements. Target positions are calculated by the same method as with RSSI measurement.

- Angle of Arrival (AoA)

  Systems estimate the angle at which signals are received and use simple geometric relationships to calculate node positions (Fig. 2(d)). This method needs a special antenna.

On the other hand, Range free localization does not need to measure the distance (or angle). Basic methods for Range free localization are:

- Centroid [6]

Beacons that know their positions are deployed and send signals. Nodes that have received beacon signals and calculate the centroid of the beacon points (Fig. 3(a)). However the centroid of the beacon points set the node position, and estimated accuracy depends on the density of the beacons. This method needs more nodes that know their positions than do other methods.

- DV-hop [7]

  DV-Hop consists of two phases:

  First phase; all landmarks that know their position exchange position and hop to reach each other and calculate 1hop distance c_i.

$$c_i = \frac{\sum \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2}}{\sum h_{i,j}} \quad \text{, for all landmarks } j \tag{1}$$

  Second phase; landmarks send X, Y, c and nodes get the distance from the landmarks. In Fig. 3(b), landmarks L_1, L_2 and L_3 compute 1hop distance, for example, c_1 is $\frac{100+40}{6+2} = 17.5$. Next, nodes get the distance from landmarks, which is a multiplication of the hops between nodes and landmarks i and c_i.

- RF-based [8, 9]

  This method needs RSSI measurements of the various points in a research area to create a signal strength map. Nodes can estimate their position by searching for the nearest RSSI points on a map. Location estimation accuracy depends on the map and how many points have been measured by RSSI.

However, many of these research works assume lots of senders and one receiver (target), if the target moves, all the senders need to be synchronized and this consumes a lot of energy. For this reason, we consider a localization system in which targets send packets for localization and the sensor receives packets and obtains the distance.

2.1.1 Research for localization

In sensor networks, localization system estimates target's location. Systems use only above measurement and do not estimate the position with accuracy. Correction for estimated position is proposed in Refs. [3, 10, 11]. Basic idea about correction are that targets exchange estimated

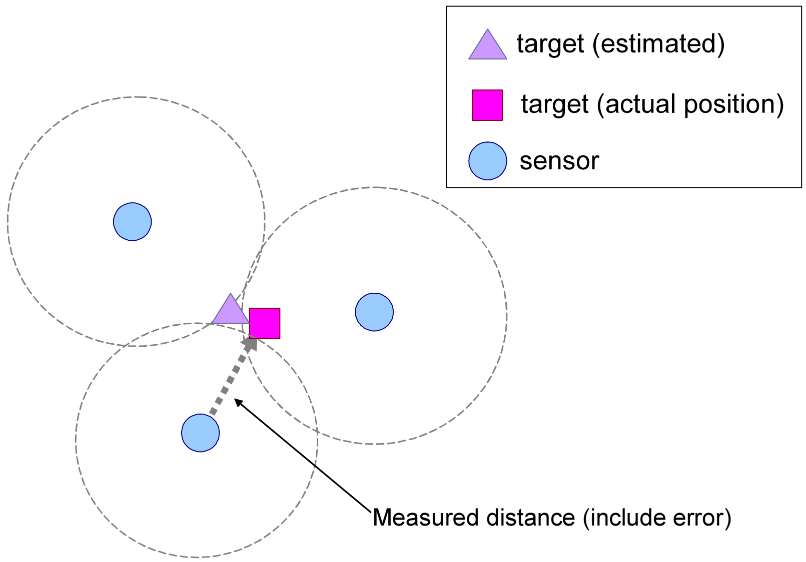
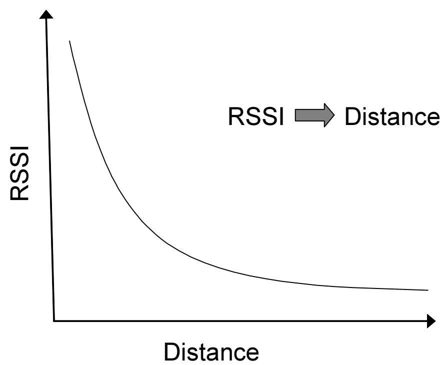


Figure 1: Localization algorithm: ML multilateration

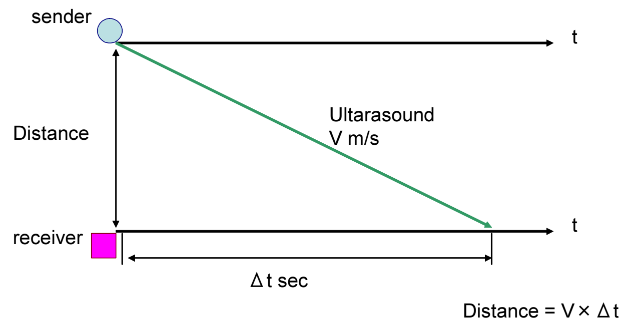
positions with neighbors or filter bad results of measurement. The former is not suitable to sensor networks and real-time localization, because it needs extra messages and long time to receive all messages and recalculate. The latter is suitable to sensor networks and real-time localization, because it does not need extra messages and long time for recalculation.

It is important problem to reduce message when sensors send it to the sink for multi-hop and large scale networks. Tree-based approaches are proposed to reduce message [12, 13]. Sensors configure tree, which the sink is root, and send data to the sink with tree. Sensors which receive data forward it to next when it is new, otherwise break off.

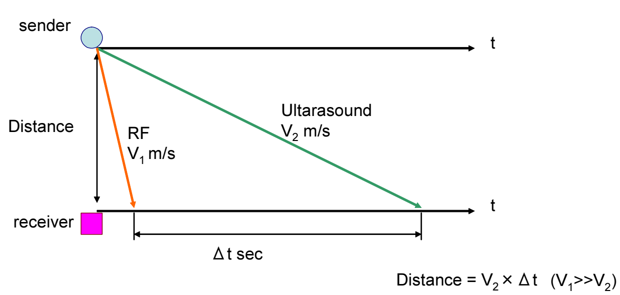
Placement of sensors has also been researched. Accuracy of the localization systems are differs according to the placement plan, grid and randomness. Results in Ref. [14] show that a grid is a good placement plan, however suggest that it depends on the localization algorithm and so sensors always cannot be placed as grid. Virtual Forces is approached to maximize coverage area [15]. Virtual Forces calculate optimal position using sensors' position and coverage area, and then sensors are moved by results of calculation to maximize coverage area.



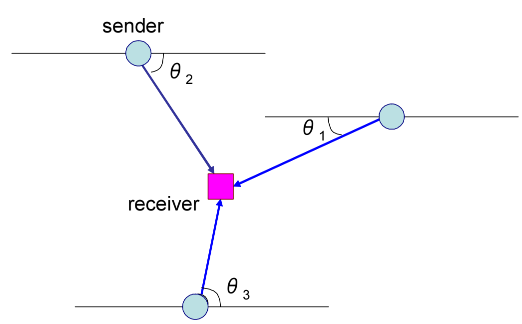
(a) Received Signal Strength Indicator



(b) Time of Arrival

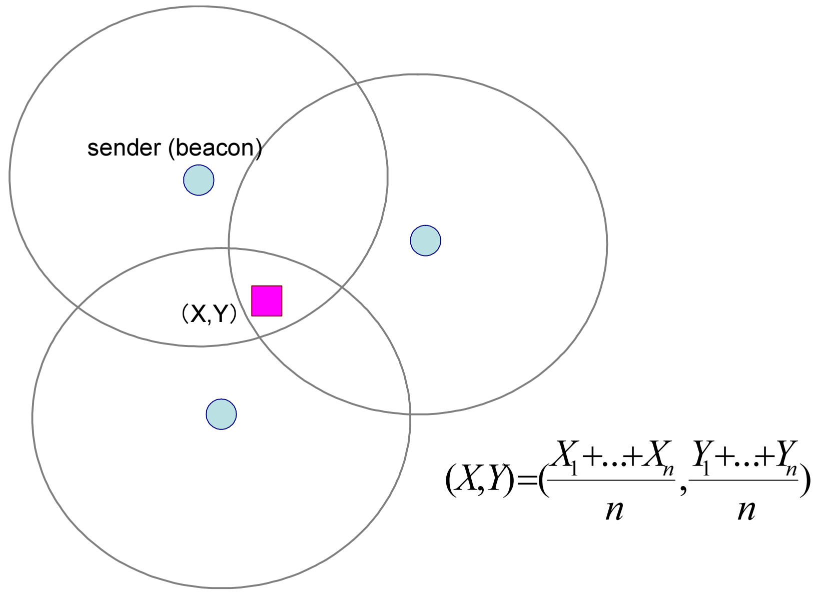


(c) Time Difference of Arrival

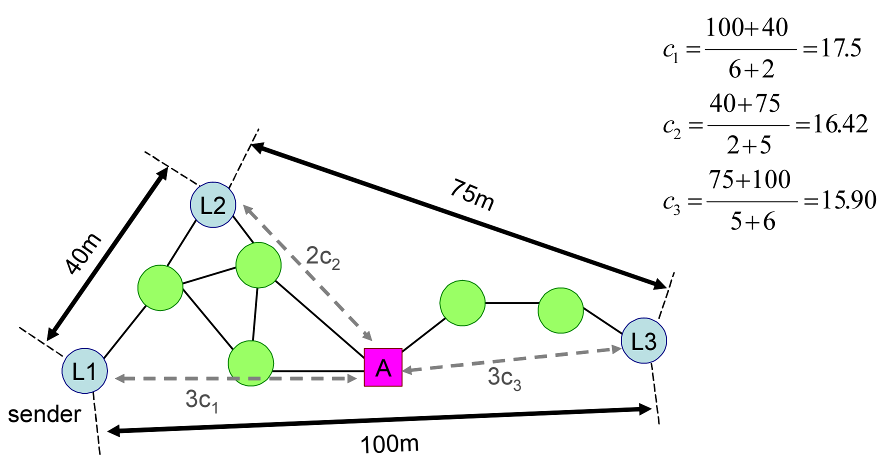


(d) Angle of Arrival

Figure 2: Range based localization



(a) Centroid



(b) DV-hop

Figure 3: Range free localization

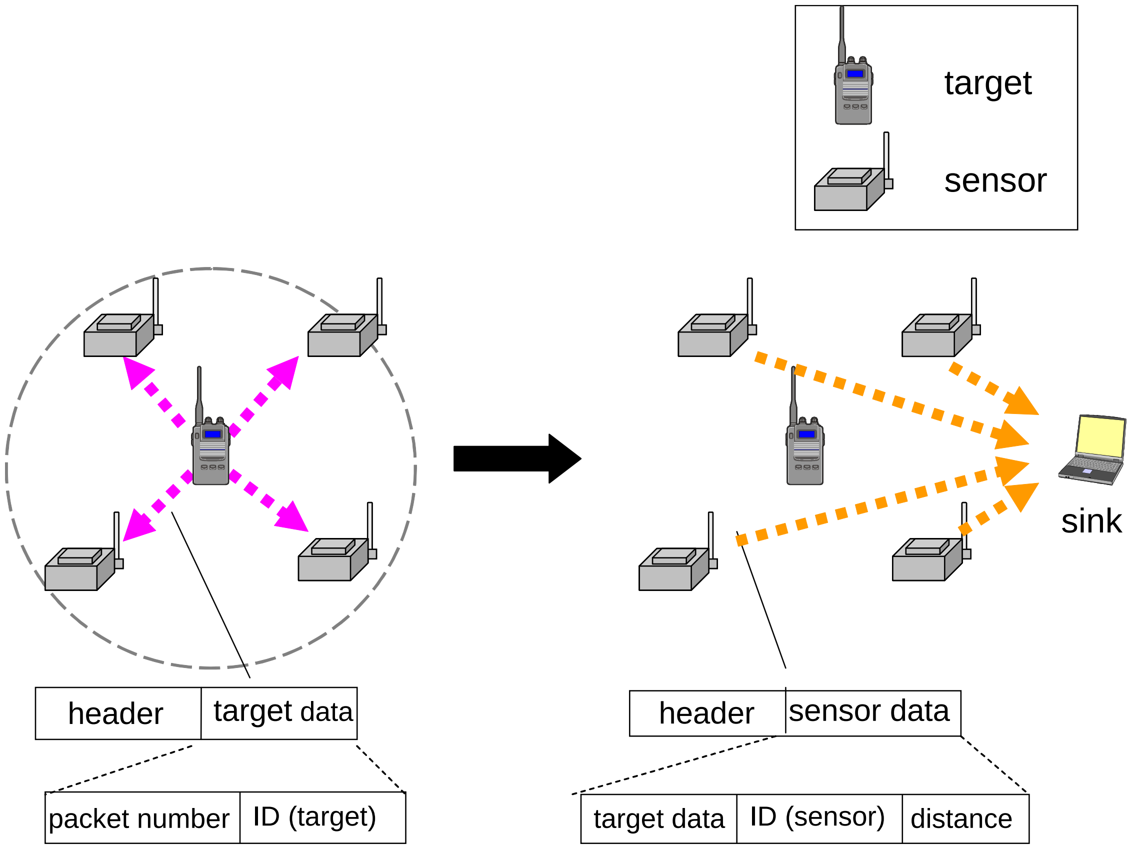


Figure 4: Localization system behavior

2.2 Framework of intended localization system

We consider a system that estimates the position of targets using sensors in an observation area, and the positions of the targets are stored in the sink's database. Targets have a wireless device and send a packet for position estimation. For multiple targets, a packet includes a target ID. After receiving a packet, sensors measure RSSI and transform it into distance using prior knowledge. Sensors send sensing data to the sink that calculates the targets' positions from the sensing data. We also consider the following details concerning localization systems.

2.2.1 Sensor placement

We assume that all sensors have already been deployed and that they do not move. Positions of sensors are needed to estimate targets' positions. There are two ways to learn sensors' positions.

First, a manager registers a sensors' position to the sink's database. If sensors need to know their position, the sink sends the position of the sensors. It resolves positions when sensors are placed on a grid or if only a few sensors are placed randomly. But it cannot resolve the problem when a lot of sensors are placed randomly. Second, a manager places a few beacon nodes that know their positions, and sensors estimate their positions by using information from a beacon node [5]. Beacons can handle a lot of sensors placed randomly, but sensors' positions include bigger errors than the first method.

2.2.2 Data collection

Sensors receive packets from targets, measure the power of the packet, and transform RSSI into distances to use theoretical and empirical models. For this purpose, we need prior knowledge about a target's transmission power and the deleterious effect of a fading channel in an observation area. We need to measure the effect of a fading channel, because it depends on its environment. The packet includes a target ID (1 byte) and a packet number (1 byte). After reading the packet, a sensor gets a target ID, a packet number, and a distance between the sensor and the target. Then the sensor sends the following data: its ID (1 byte), the target ID (1 byte), packet number (1 byte), and the distance between the sensor and the target to the sink (1 byte). System behaviors are shown in Fig.4

2.2.3 Calculation of target's position at the sink

ML estimation of a target's position can be obtained by Minimum Mean Square Error (MMSE) [5]. can resolve the position from data including errors for calculating a target's position. We explain the calculations for a two-dimensional case. It needs more than three sensors to resolve a target's position. First, the sink searches for the same data in terms of a target ID and a packet number from the data collected from sensors. The difference between measured and estimated distances is defined as Eq. (2) below.

$$f_i(x_0, y_0) = d_i - \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \quad (2)$$

(x_0, y_0) is the unknown position of the target, (x_i, y_i) for $i = 1, 2, 3...N$ is the sensor position, and N is the total number of data that the sink has collected. d_i is a distance between sensor i and a target. The target's position x_0 and y_0 can obtained by . Eq. (3) is obtained by setting $f_i = 0$, squaring and rearranging.

$$-{x_i}^2 - {y_i}^2 + {d_i}^2 = ({x_0}^2 + {y_0}^2) + x_0(-2x_i) + y_0(-2y_i) \tag{3}$$

After getting Eq. (3), we can eliminate the $(x_0^2 + y_0^2)$ terms by subtracting the kth eqnarray from the rest.

$$-{x_i}^2 - {y_i}^2 + {d_i}^2 - (-{x_k}^2 - {y_k}^2 + {d_k}^2) = 2x_0(x_k - x_i) + 2y_0(y_k - y_i) \tag{4}$$

Eq. (4) transforms Eq. (5), which can be solved using matrix solution given by Eq. (6). Position (x_0, y_0) can be obtained by calculating Eq. (6).

$$y = Xb \tag{5}$$

$$b = (X^T X)^{-1} X^T y \tag{6}$$

where

$$X = \begin{bmatrix} 2(x_k - x_1) & 2(y_k - y_1) \\ \vdots & \vdots \\ 2(x_k - x_{k-1}) & 2(y_k - x_{y-1}) \end{bmatrix} \tag{7}$$

$$y = \begin{bmatrix} -{x_1}^2 - {y_1}^2 + {d_1}^2 - (-{x_k}^2 - {y_k}^2 + {d_k}^2) \\ \vdots \\ -{x_{k-1}}^2 - {y_{k-1}}^2 + {d_{k-1}}^2 - (-{x_k}^2 - {y_k}^2 + {d_k}^2) \end{bmatrix} \tag{8}$$

$$b = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \tag{9}$$

3 Data collecting technique

In sensor networks, energy consumption of sensors and network capacities are an important problem. If there is a dense network, many sensors try to send data but there are long delays and failures of transmission in wireless sensor networks that have a small capacity. Long delays have a bad influence on real-time operations to get location information. Due to failures of transmission, a system may not be able to localize target positions. For these reasons, we propose a technique to control how many sensors send data. In our proposal, whether sensors send data depends on the density around the sensor and the distance between the sensor and the target. Sensors send data if the distance between the sensor and the target is shorter than a certain distance that is calculated by each sensor.

We design a data collecting technique that does not need complicated communications with each sensor and central control at the sink, and add just a small energy consumption and the network load. It need only sensors' measurement of density at intervals which takes short time and consumes a little energy. System can autonomously adapt to change of density in case of sensors' addition and failure, and collect in order amount of data. Also, this technique does not depend on MMSE and can be adapted to other localization algorithms or data collection except localization systems.

3.1 Measurement of sensors' deployment density

Each sensor measures the deployment density of sensors in surroundings of itself, by receiving packets sent for information of their existence at each period of time and for measuring the communication range. ρ_i which is the density around sensor i is approximately determined by Eq. (10). R is communication range and M_i is the number of sensors within R from sensor i.

$$\rho_i = \frac{M_i}{\pi R^2} \tag{10}$$

ρ_i remains unchanged for a long time, because sensors rarely move or stop on account of failure or energy loss. In a dense sensor network, if one sensor stops, the density changes just a little. Our proposal therefore does not need to measure density frequently. It needs to be done every hour or at an even longer interval. If more sensors were to be added or removed, then density would be greatly changed, but we think that sensors are only added or removed on rare occasions.

Measurement of the density does not affect the energy consumption of the sensors.

3.2 Expression to control sensor

We define the amount of data required by the system by Z. Sensor i sends data if the measured distance is shorter than distance D_i to collect Z. The number of sensors within D_i is proportional to density and D_i is defined in Eq. (11).

$$\frac{Z}{\pi {D_i}^2} = \rho_i \tag{11}$$

Arranging Eq. (11), Eq. (12) is obtained.

$$D_i = \sqrt{\frac{Z}{\pi \rho_i}} \tag{12}$$

D_i depends on ρ_i. The sink can collect the same amount of data independent of sensors' deployment density, because if ρ_i is high, D_i is small and if ρ_i is low, D_i is high.

3.3 Data fusion at target node

If sensors send data directly to the sink, we consider that there are many backoffs and long delays. Even if a system controls the number of sensors under our proposal, too much sensors send data for many targets and the system cannot collect enough data for localization. Using data fusion can solve this problem because sensor's data is only 4 bytes and the header is too long; for example, a 802.11 header length is 48 bytes and a 802.15.4 header length is between 15 and 31 bytes. However, data needs to be collected to one node like clustering [16], the election of a cluster head demands an additional energy cost to the sensors. In a localization system, it is easy for the target to collect sensors' data. Sensors send data to the target instead of the sink and the target puts the sensors data together and sends it (see Fig. 5). We call the *direct method* by which sensors send data directly to the sink and the *fusion method* by which sensors send data to the target and the target sends the data to the sink. We assume that the target has enough power as much as the sensor and the amount of data collected is free from the target power.

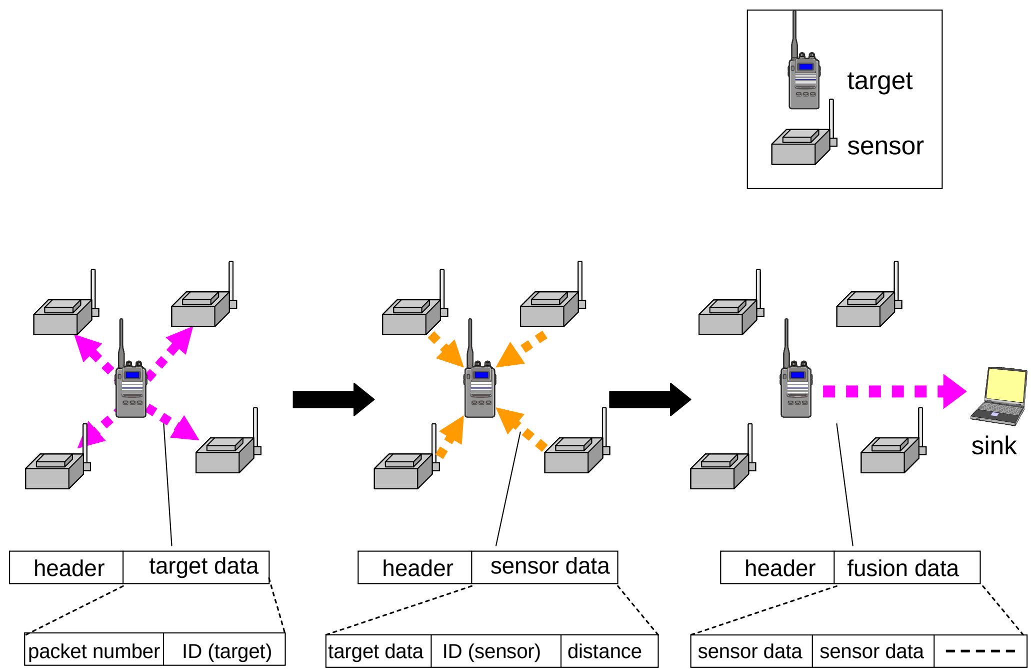


Figure 5: Data fusion in the target

4 Analysis of data transfer and energy consumption

In this section, we consider an effect of data transfer for localization system in this thesis. We assume that the MAC protocol is IEEE 802.15.4 for a low power device defined in Ref. [17]. IEEE 802.15.4 uses Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). First, we set out IEEE 802.15.4 MAC in Section 4.1. Second, we analyze the transmission success rate and the energy consumption in a localization system and demonstrate the performance of the localization system.

4.1 IEEE 802.15.4 MAC protocol

In wireless sensor networks, sensors, which are very small, have little memory and battery capacity cannot store data for a long time and also consume a lot of energy. In wireless communications, it is important to keep energy consumption low. IEEE 802.11 [18] for wireless LAN, which has been designed for high power devices such as PCs, is not suitable for application to wireless sensor networks, Many Protocols have been proposed to cut off wireless devices [19, 20, 21], but a standard has not been defined and so sensors are not subject to standardization, and a protocol has not been disseminated. IEEE 802.15.4 [17] for Low-Rate Wireless Personal Area Networks has been appeared recently and is expected to be suitable for wireless sensor networks. This standard defines Medium Access Control (MAC) and Physical Layer (PHY) for low power devices.

4.1.1 SuperFrame structure

IEEE 802.15.4 uses a SuperFrame structure for beacons as in Fig. 6. The SuperFrame consists of 16 slots, the first slot in the SuperFrame is always used for beacon transmission and the sensors send data for the other 15 slots. We assume that all 15 slots are Contention Access Period (CAP) for which all sensors can send data, if the slot is free. Sensors send data using CSMA/CA in CAP. The length of the SuperFrame is the same as the Beacon Interval (BI), which is defined as Eq. (13).

$$BI = aBaseSuperFrameDuration \cdot 2^{BO} \tag{13}$$

In the 2.4 GHz band, IEEE 802.15.4 define that $aBaseSuperFrameDuration$ is 15.36 ms and Beacon Order (BO) is the control parameter for BI from 0 through 14. Length of 1 slot is $\frac{BI}{16}$.

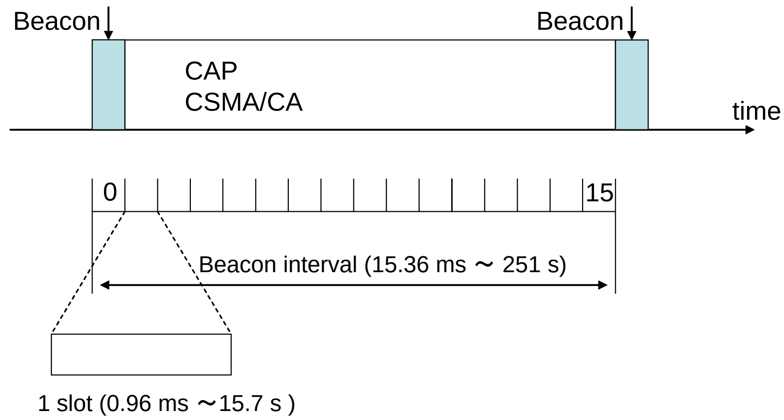


Figure 6: SuperFrame structure

4.2 CSMA/CA protocol

Sensors perform carrier sense and send data, if the channel is free. If the channel is not free, sensors wait for a backoff duration of a random variable between 0 and W_k, at k-th backoff, W_k is defined as Eq. (14).

$$W_k = 2^{\min(BE_{min}+k, BE_{max})} \quad 0 \leq k \leq M \tag{14}$$

BE_{min} is the minimum backoff period, BE_{max} is the maximum backoff period and M is the maximum number of backoff. These parameters are defined in IEEE 802.15.4 as Table 1. T_{max} is the maximum time slot, all sensors finish transmission by T_{max} regardless of whether successful.

$$T_{max} = \sum_{k=0}^{M} W_k \tag{15}$$

Table 1: Backoff parameters

	Value	Default
BE_{min}	0-3	3
BE_{max}	5	5
M	0-5	4

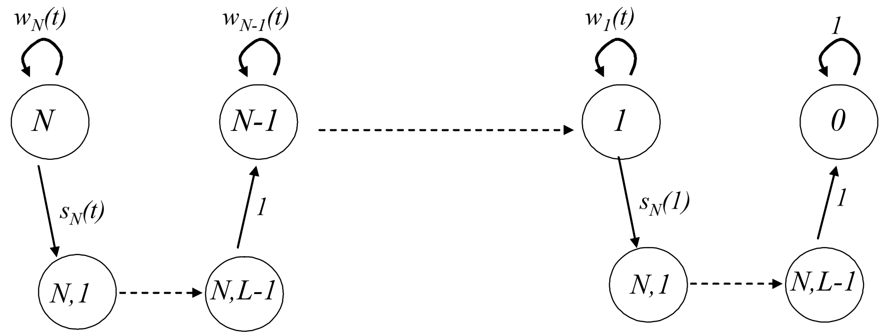


Figure 7: State transition

4.3 How many sensors can send data

In IEEE 802.15.4, sensors give up data transmission, because the maximum number of backoff is defined. Transmission success rate depends on the number of sensors trying to send data and the backoff parameters.

4.3.1 Analytical model

The analysis assumes that all sensors can perform carrier sense perfectly and sensors cannot send data to the same destination at the same time, and do not consider beacon transmissions. It allows that sensors can send data to different destinations. Communication speed is 250 kbps, duration of 1 slot is set to 0.96 ms and header length is set to 15 byte. In our localization system, we assume that all targets send packets synchronously and all sensors which receive packets simply start to send at the same time, and sensors do not add any on the way. We use an analytical model for state transition in CSMA/CA [22]. As parameter settings of original model are difficult, we change state transition simply; our model is shown as in Fig. 7. $s_j(t)$ is success probability and $w_j(t)$ is wait probability.

$$s_i(t) = j\ a_i(t)(1 - a_i(t))^{i-1} \quad (16)$$

$$w_i(t) = 1 - s_i(t) \quad (17)$$

$a_j(t)$ is the attempt probability that each sensor sends a packet following $a_j(t)$. L is the length of slot which transmission of the packet occupies. State probability vector $\mathbf{x}_{t+1}$ at time $t + 1$ is defined as

$$x_i(t+1) = w_i(t)x_i(t) + s_i(t)x_{i+1}(t) \qquad 0 \leq i \leq N, \;\; L = 1 \tag{18}$$

$$\left.\begin{aligned} x_i(t+1) &= w_i(t)x_i(t) + x_{i+1,L-1}(t) \\ x_{i,j}(t+1) &= x_{i,j-1}(t) \\ x_{i,1}(t+1) &= s_i(t)x_i(t) \end{aligned}\right\} \quad \begin{aligned} &0 \leq i \leq N, \\ &2 \leq j \leq L-1 \end{aligned} \tag{19}$$

At time T_{max}, all states are stable. This state transition gets the number of sensors with successful transmissions as Eq. (20).

$$\sum_{i=0}^{N} \{(N-i)x_i(T_{max}) + \sum_{j=1}^{L-1}(N-i)x_{i,j}(T_{max})\} \tag{20}$$

4.3.2 Evaluation

Figure 8 shows the results that the maximum number of backoff is set to 4 or 5 in analysis and simulation, respectively. BE_{min} and BE_{max} are the set default values and simulation runs of 1000 times at each point. Analysis fits to the simulation for a large number of sensors. If the number of sensor is small, the result of the analysis is a smaller success rate than in the simulation. Probability $s_i(t)$ should be larger when the number of sensors is small; an issue that needs further consideration. Figure 9 shows how many data systems can collect, In the case of 4 backoffs, the peak is about 35 sensors trying to send, which this means that the number of sensors trying to send should be set under 35 (peak value).

Next, we show the performance of data fusion at the target. We compare the two methods of data collection, direct and fusion in Figs. 10 and 11. However the direct method cannot collect data in the case of multiple targets, whereas the fusion method can collect data and is almost free from the effects of an increase of targets. The fusion method shows the good result, because the target fuses many packets into one. Slots needing to send data are defined as Eq. (21).

$$\begin{aligned} slots &= \left\lceil \frac{\frac{header+data}{250kbps}}{\text{slot length}} \right\rceil \\ &= \left\lceil \frac{120 + data}{240} \right\rceil \end{aligned} \tag{21}$$

$$data = 32 \cdot (\text{the number of received packet}) \tag{22}$$

The relationship between slots what targets have send the data for and the number of packets

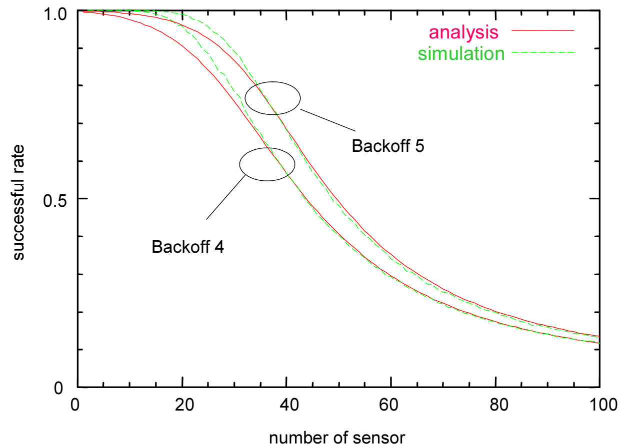


Figure 8: Comparison analysis and simulation

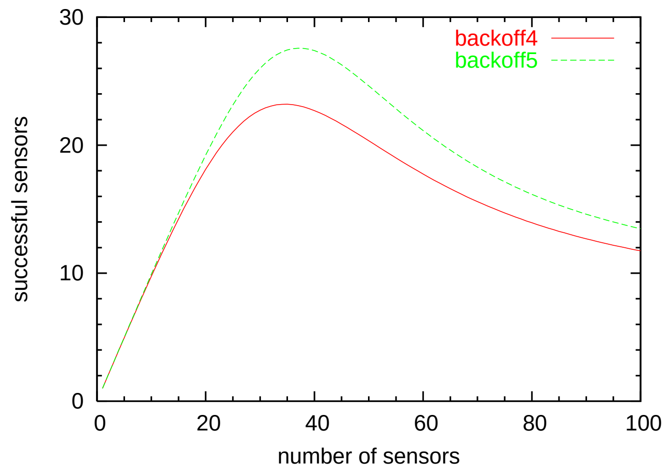


Figure 9: The number of sensors which succeeded in transmission

received by one target is shown in Fig. 12. These results show that the effect of packet length is small.

4.3.3 Modification of CSMA/CA parameter and evaluation

The fusion method is superior to the direct method in term of the success rate. But direct is needed when the target cannot store and send a lot of data. We modify BE_{max} and M and evaluate directly by using a modified CSMA/CA. We remove BE_{max} from Eq. (14) and M is the given parameter. The backoff duration of modified CSMA/CA is a random variable between 0 and W'_k

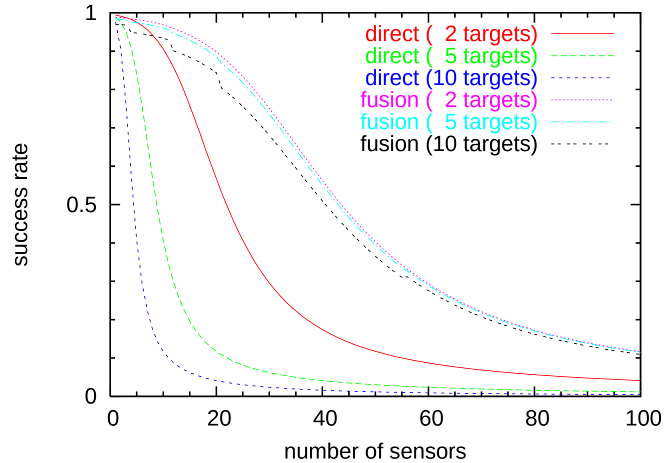


Figure 10: Transmission successful rate of data fusion: maximum backoff is set to 4

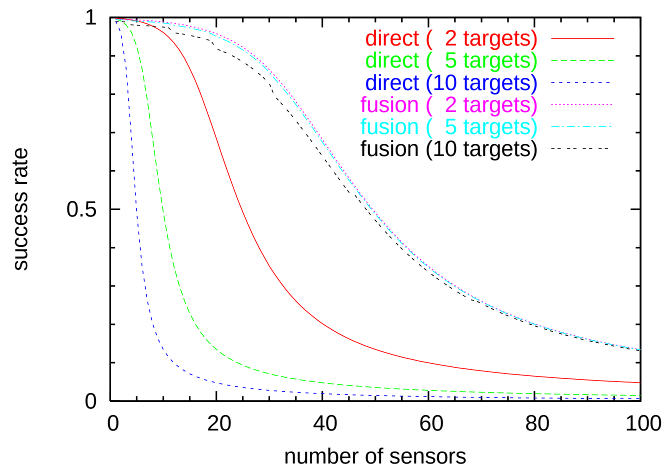


Figure 11: Transmission successful rate of data fusion: maximum backoff is set to 5

at the k-th backoff, W'_k is defined as Eq. (23).

$$W'_k = 2^{(BE_{min}+k)} \quad 0 \leq k \leq M \tag{23}$$

$$T_{max} = \sum_{k=0}^{M} W'_k \tag{24}$$

We compare normal and modified CSMA/CA at T_{max} and show result in Table 2.

Figures 13, 14 and 15 show the success rate of transmissions at 1, 5 and 10 targets, respectively. Success rate increases with the maximum number of backoff. It is possible that many sensors try to send the data and success transmission, if we modify the backoff parameter. For T_{max} increases according to the exponential, we need to set a long period for the interval of localization. The

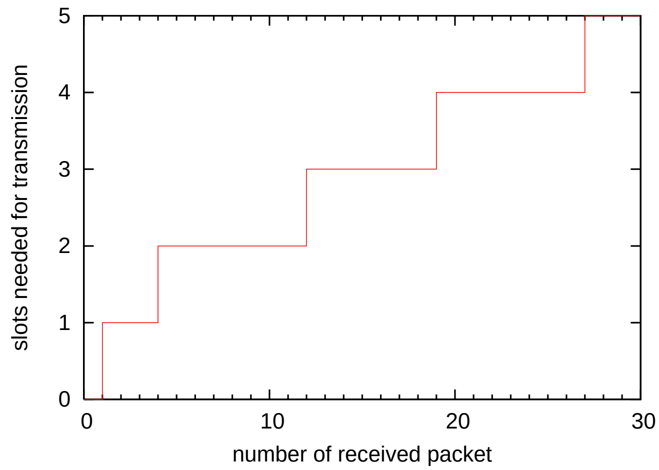


Figure 12: Relation between slots and the number of received packets

intervals of localization are set appropriately based on the number of targets, the maximum number of backoff and how to send data either by the direct or the fusion methods. Next we compare the direct and the fusion methods in the case of the same sensors trying to send the data. We consider that 20 items of data are sufficient for localization. Figures 16 and 17 show the relation of the number of targets and the success rate, respectively, in cases of 10 and 20 sensors trying to send data. By the direct method, when the modified CSMA/CA and the maximum number of backoff is set to 7, shows about same performance as by the fusion method, when the normal CSMA/CA and the maximum number of backoff is set to 4. It seems obvious that if there are more than 10 targets, the performance of the direct method become worse than the fusion method and it needs more backoffs to collect the same amount of data. In case of the direct method, to collect data needs a long time if many targets are in observation area, on the other hand, to collect the data needs just a short time if only a single target is in the observation area, because all sensors have successful transmissions and finish by T_{max}.

4.4 Performance evaluation of energy consumption

4.4.1 Energy consumption model

We analyze energy consumption with the energy model described in Ref. [16]. Energy consumption to transmit a k bit message a distance d is defined as Eq. (25) and to receive a k bit message

Table 2: Values of max delay T_{max}

CSMA/CA	M	T_{max}
normal	4	120
	5	152
modified	4	248
	5	504
	6	1016
	7	2040
	8	4088

is defined in Eq. (26). E_{elec} is to run circuitry in the transmitter and receiver, and ε_{amp} is needed for the amplifier to transmit.

$$E_T(k,d) = E_{elec} \cdot k + \varepsilon_{amp} \cdot k \cdot d^2 \tag{25}$$

$$E_R(k) = E_{elec} \cdot k \tag{26}$$

Table 3: Radio characteristics

Operation	Energy Dissipated
Transmitter and Receiver Electronics (E_{elec})	50 nJ/bit
Transmit Amplifier (ε_{amp})	100 pJ/bit/m^2

4.4.2 Analysis of energy consumption

We analyze energy consumption of the sensor and target in one data collection. Energy consumption at the sink is ignored because the sink commonly is assumed to have infinite energy. Assumptions in this analysis are that only nodes successfully transmitting consume energy for transmission, all targets are same distance from the sink in the case of multiple targets and energy consumption of carrier sense is never considered.

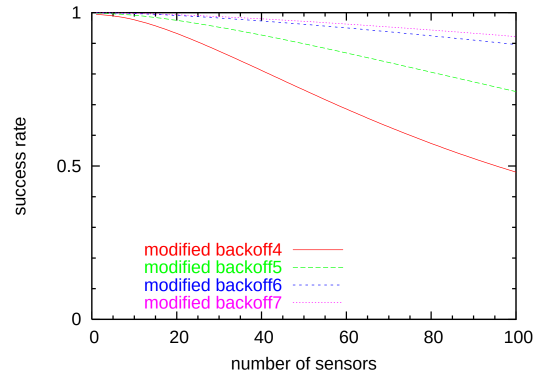


Figure 13: Transmission successful rate of modified CSMA/CA: sensors send data directly to the sink in the case of 1 target

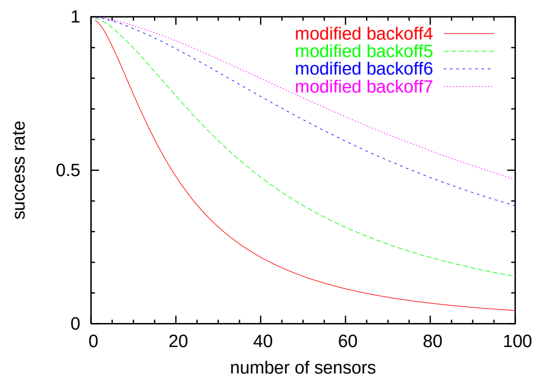


Figure 14: Transmission successful rate of modified CSMA/CA: sensors send data directly to the sink in the case of 5 targets

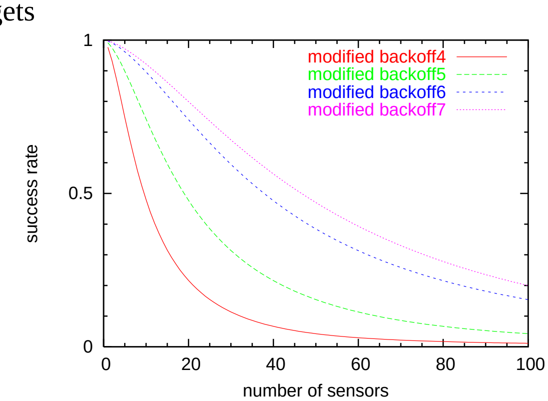


Figure 15: Transmission successful rate of modified CSMA/CA: sensors send data directly to the sink in the case of 10 targets

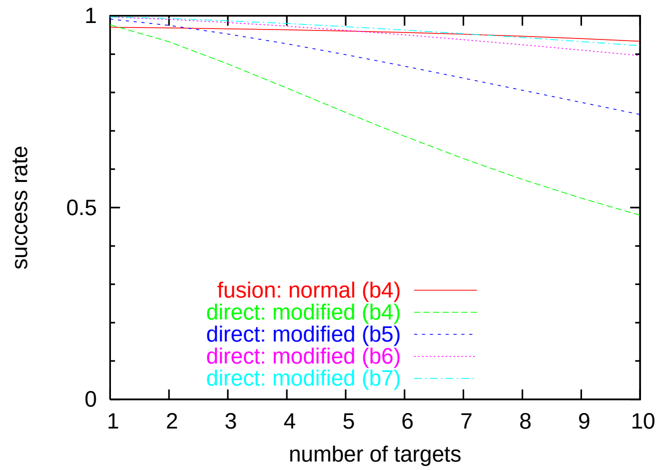


Figure 16: Comparison of the fusion method and the direct method: 10 sensors trying to send data

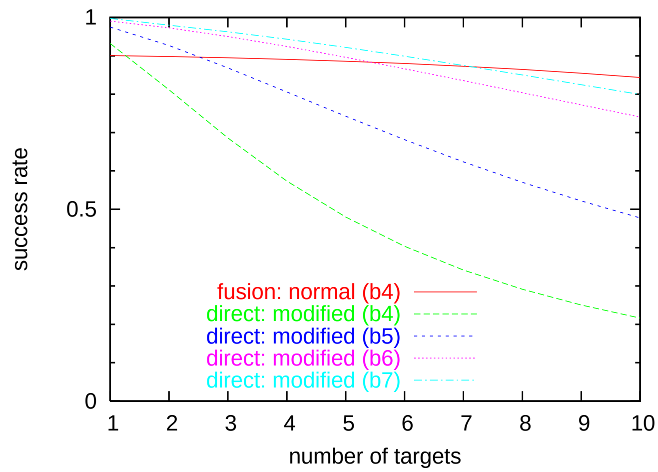


Figure 17: Comparison of the fusion method and the direct method: 20 sensors trying to send data

N_s is the number of sensors receiving a packet which a target sends for localization. N_s is defined as Eq. (27). $d_{t\to s}$ is the radius of the targets' packet transmission area and ρ is the sensors' density of the whole area.

$$N_s \quad = \quad \rho\pi {d_{t\to s}}^2 \tag{27}$$

The number of Targets is N_t. Sensors try to send is n $(\leq N_s)$. The radius of the area where n sensors exist is Eq. (29).

$$n \quad = \quad \rho\pi {d_n}^2 \tag{28}$$

$$d_n \quad = \quad \sqrt{\frac{\rho\pi}{n}} \tag{29}$$

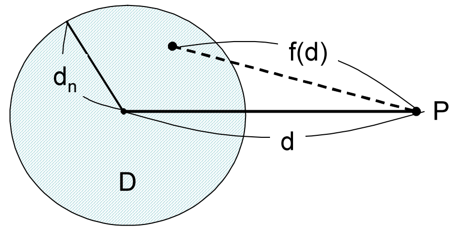


Figure 18: Distance between circle D and a point

The distance between any point of the circle D with d_n radius and a point P is defined as $f(d)$ (see Fig. 18), mean-square of f is defined as Eq. (30).

$$\begin{aligned} \overline{f(p,d_n)^2} &= \frac{\int\int_D f^2\,dxdy}{{d_n}^2\pi} \\ &= \frac{\int\int_D (x+p)^2 + y^2\,dxdy}{{d_n}^2\pi} \\ &= p^2 + \frac{1}{2}{d_n}^2 \end{aligned} \tag{30}$$

The direct method

Targets consume energy only to transmit a packet for localization, of which the length is $h + k_1$. h is the header length. Energy consumption of the target is

$$E_{target} = [E_{elec} + \varepsilon_{amp}\, {d_{t\to s}}^2](h + k_1) \tag{31}$$

Sensors consume energy to receive a packet from a target and then transmit sensing data to the sink. The length of the sensing data is k_2 and the distance between the target and the sink is d. Energy consumptions of receiving and transmission are defined in Eqs. (32) and (33), respectively.

$$E_{Rx\ sensor} = E_{elec}\,(h + k_1) \tag{32}$$

$$E_{Tx\ sensor}(n,d) = [E_{elec} + \varepsilon_{amp}\,\overline{f(d,d_n)^2}](k_2 + h) \tag{33}$$

The energy consumption of an average of N_s sensors is defined in Eq. (34).

$$E_{sensor}(n,d) = \frac{N_s\,E_{Rx\ sensor} + Succ(N_t\cdot n, M, 1)\,E_{Tx\ sensor}(n,d)}{N_s} \tag{34}$$

The total energy consumption in the system is

$$E_{total}(n,d) = N_t\,\{E_{target} + N_s\,E_{sensor}(n,d)\} \tag{35}$$

The fusion method

Targets consume energy to transmit a packet for localization, of which the length is $h + k_1$, and receive sensing data from sensors and send it to the sink. The length of the sensing data is k_2 and the distance between the target and the sink is d. h is the header length. Energy consumption of the target is

$$\begin{aligned}
E'_{target} &= \frac{N_t\ E_{Rx\ target} + Succ(N_t, M, L)\ E_{Tx\ target}}{N_t} && (36)\\
E'_{Rx\ target} &= Succ(n, M, 1)\ E_{elec}\ (h + k_2) && (37)\\
E'_{Tx\ target} &= (E_{elec} + \varepsilon_{amp}\ {d_{t\to s}}^2)\ (h + k_1) + \\
&\quad (E_{elec} + \varepsilon_{amp}\ d^2)\ (h + Succ(n, M, 1)\ k_2) && (38)\\
L &= \left\lceil \frac{h + Succ(n, M, 1)\ k_2}{240} \right\rceil && (39)
\end{aligned}$$

Sensors consume energy to receive a packet from a target and then transmit sensing data to the target. The distance between the target and the sink is $\overline{f(0, d_n)^2}$. The energy consumptions of receiving and transmission are defined in Eqs. (40) and (41), respectively.

$$\begin{aligned}
E'_{Rx\ sensor} &= E_{elec}\ (h + k_1) && (40)\\
E'_{Tx\ sensor}(n) &= (E_{elec} + \varepsilon_{amp}\ \overline{f(0, d_n)^2})\ (h + k_2) && (41)\\
& && (42)
\end{aligned}$$

The energy consumption of an average of N_s sensors is defined in Eq. (43).

$$E'_{sensor}(n) = \frac{N_s\ E'_{Rx\ sensor} + Succ(n, M, 1)\ E'_{Tx\ sensor}(n)}{N_s} \quad (43)$$

The total energy consumption in the system is

$$E'_{total}(n, d) = N_t\ \{E'_{target}(n, d) + N_s\ E'_{sensor}(n)\} \quad (44)$$

4.4.3 Results of single target

The parameter settings are shown in Table 4. M is set to 4 and we use a normal CSMA/CA. The number of sensors trying to send is set to 5, 10, 20, 30 and 40. We show the results of the energy consumption for various distance between the target and the sink.

In the case of which sensors send the data directly to the sink, the result of the mean energy consumption per sensor is shown in Fig. 19. $E_{sensor}(n, d)$ is according to a square of the

Table 4: Parameter settings

Parameter	Value
k_1	16 bit
k_2	32 bit
h	120 bit
$d_{t \to s}$	20 m
ρ	0.1

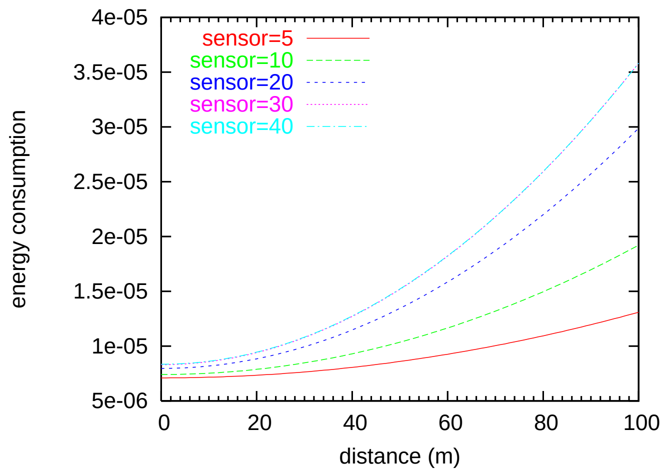


Figure 19: Mean energy consumption per sensor: the direct method and single target

distance from the sink because the power model in Section 4.4.1 accords to a square of the distance. $E_{sensor}(n, d)$ is usually in proportion to the number of sensors trying to send however, $E_{sensor}(40, d)$ shows about the same energy as $E_{sensor}(30, d)$. It means that $Succ(40, 4, 1)$ is equal to $Succ(30, 4, 1)$. If more sensors try to send, the energy consumption of the sensors becomes small because the system cannot collect data. Sensors deployed far from sink use a lot of energy for transmission and spend all their energy earlier than do other sensors. The energy consumption of the target is very small and do not depend on the number of sensors because the target consume energy only to send one packet. This value is 12.2400 nJ.

In the case of where a target fuses the data and sends it to a sink, the results are shown in Table 5 and Fig. 20. The tendency of $E'_{target}(n, d)$ is the same as $E_{sensor}(n, d)$. Target consumes a lot of energy, however the target can be supplied with energy easier than sensors. All sensors consume the same energy regardless of where they are deployed, for sensors send data to the target, not to sink. It allows for maintenances of sensors to become easy because all sensors consume the same energy and stop at the same time.

Figure 21 compares the direct method with the fusion method in terms of total energy consumption. When sensors are nearer than 30 m, the fusion method is worse than the direct method. When sensors are further than 30 m, the fusion method is better than the direct method. If the system must bring the total energy down, it is needed to control sensors send whether to the target or the sink, depending on the distance from the sink.

Table 5: Mean energy consumption per sensor: the fusion method and a single target

Number of sensors	Energy consumption
5	7.1041 [nJ]
10	7.4090 [nJ]
20	7.9659 [nJ]
30	8.3072 [nJ]
40	8.3487 [nJ]

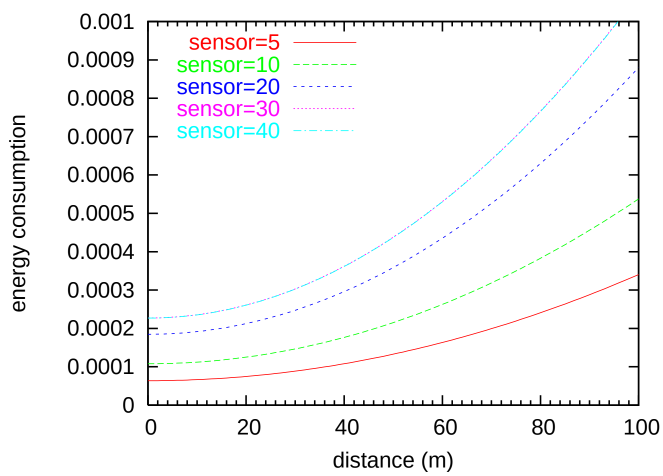


Figure 20: Energy consumption per target: the fusion method and a single target

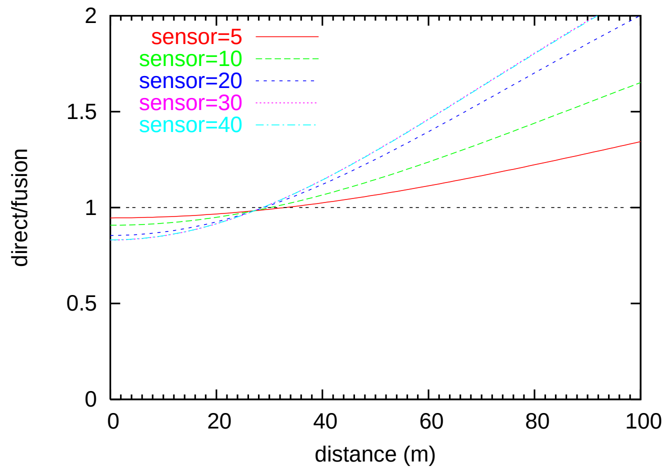


Figure 21: Comparison total energy consumption of the direct method with the fusion method: a single target

4.4.4 Results of multiple targets

We consider energy consumption in the case of multiple targets in an observation area. The direct method uses modified CSMA/CA which the maximum number of backoff is set to 7, because it has about the same success rate as the fusion method in Figs. 14 and 15. Results of multiple targets are shown in Figs. 22 through 24 and Table. 6, and the mean energy consumption per target in case of the direct method is 12.2400 nJ. These have the same tendency as the results of single targets. If the backoff parameters of the direct and fusion methods are adjusted to collect the same number of data, the results show the same tendency at any number of sensors and targets.

Table 6: Mean energy consumption per sensor: the fusion method and multiple targets

Number of targets	Number of sensors	Energy consumption
5	10	7.4090 [nJ]
5	20	7.9659 [nJ]
10	10	7.4090 [nJ]
10	20	7.9659 [nJ]

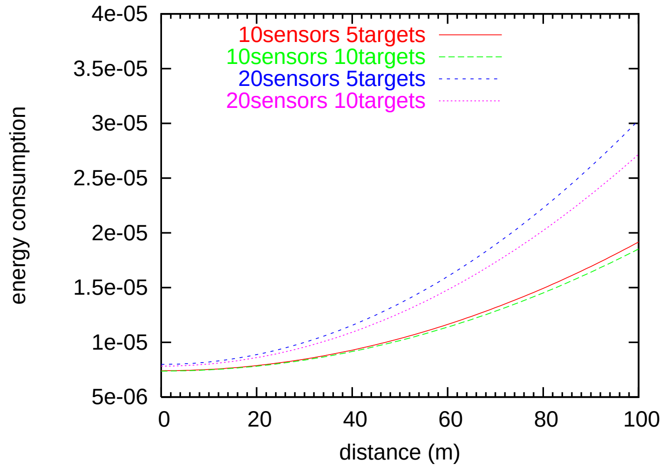


Figure 22: Mean energy consumption per sensor: the direct method and multiple targets

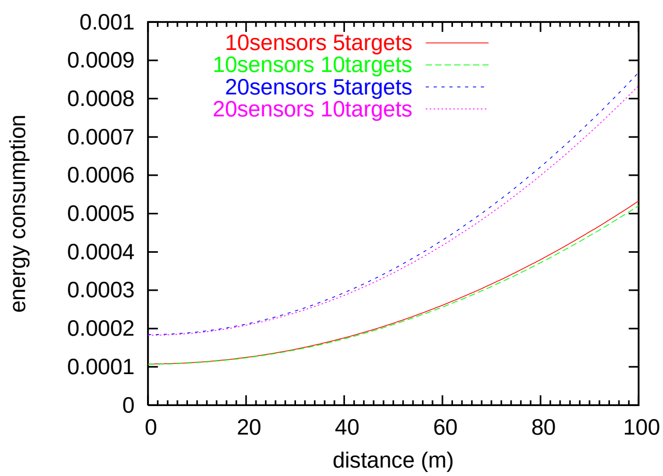


Figure 23: Mean energy consumption per target: the fusion method and multiple targets

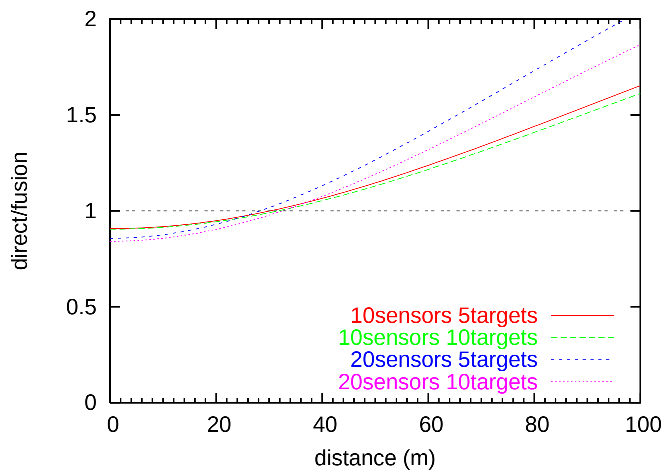


Figure 24: Comparison total energy consumption of the direct method with the fusion method: multiple targets

5 Performance evaluation of localization accuracy

5.1 Received power model of IEEE 802.15.4 radio module

When the proposed data collection method is evaluated, we try to apply the received signal characteristic of an actual sensor node. We adopt received power model for CC2420 produced by Chipcon, which is the commercially available RF transceiver in accordance with the IEEE 802.15.4 standard. We refer to propagation characteristics obtained by the experimental result in Ref. [23]. The relation between the distance and RSSI is denoted as Eq. (45), in case of transmission power is set to 1 mW. The distribution of RSSI is denoted as Eq. (46). It suggests that the distribution of RSSI is according to the exponential distribution.

$$\Lambda(r) = 0.00008r^{-2.56} \tag{45}$$

$$p(P_r) = \frac{1}{\Lambda(r)} \exp(\frac{-P_r}{\Lambda(r)}) \tag{46}$$

where P_r is receiving power, $\Lambda(r)$ is the mean and $\Lambda(r)^2$ is the variance of the distribution. Therefore, the variance becomes large with the distance.

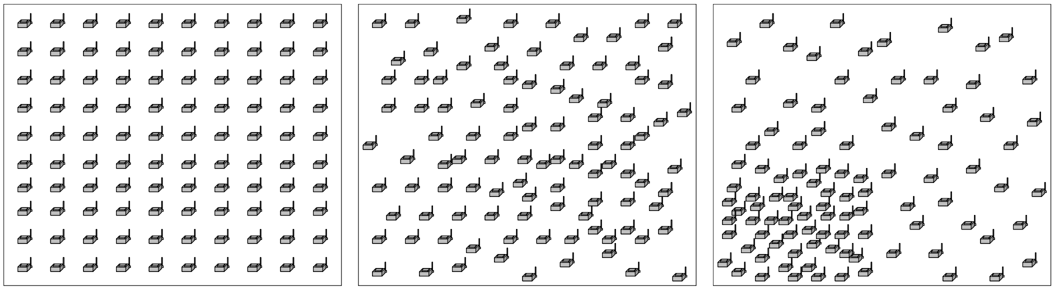
5.2 Simulation results

We adopt the distribution of RSSI for Eq. (46) and the maximum number of backoff for 4 in normal CSMA/CA algorithm following to IEEE 802.15.4. Sensors are deployed as grid in a simulation area as 100 m square. We simulate the localization error 100 times each number of sensors and the targets' position are generated randomly in the simulation area. Sensors get RSSI within 20 m from the target and determine whether send or not according to Eq. (12). Transmission success rate is according to the results in Section 4.3. If system cannot estimate the targets' location due to less than three data, the localization error is set to 100 m.

Figures 26 and 28 illustrate the mean estimated error at the direct and the fusion methods, respectively. Figures 27 and 29 illustrate the number of data collected at the direct and the fusion methods, respectively. In these cases, density of sensors are $0.1/\mathrm{m}^2$ and 1000 sensors are deployed as grid shown in Fig. 25(a) in the simulation area. The direct method gets worse when many targets exist in area, because sensors cannot send data. The result shows system cannot estimate positions of many targets when sensors send data in the direct method. In contrast, the fusion method stays constant and is scalable for increasing targets. There are great differences between two methods.

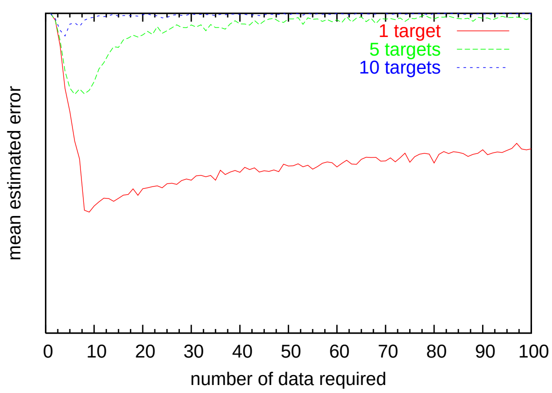
The mean estimated error is around 1 m when the number of sensors trying to send is set to 9. However, in both methods, the mean estimated error gets large when the number of sensors try to send has large. It has two reasons. First, sensors far from the target have large errors, and in a large number of sensors, the sink collects more data from far sensors than in a small number of sensors. Second, system cannot collect few data when a large number of sensors send data. Figure 30 and 31 show the mean estimated error and the number of data collected in grid topology, respectively, when single target exists in area and data is collected by the fusion method. In case of low density, $1/\mathrm{m}^2$, system cannot collect a large amount of data and the minimum mean estimated error gets large. On the other hand, in case of high density, $1/\mathrm{m}^2$, the minimum mean estimated error is good around 0.3 m. It can satisfy the requirement of almost all application for localization. The mean estimated error depends on the density of sensors and so the density is important factor to get applicable accuracy. If density is high, the system can get correct data from sensors which are located near the target. The number of sensors trying to send should be decided by the density and the capacity of network. The Results of the number of data collected show that our proposed data collecting technique can control the number of sensor trying to send, just as nearly we designed. It work out more accurate in high density because measurement of density around sensor become accuracy.

Next, we validate our proposed data collecting technique in random and biased topologies shown in Figs. 25(b) and 25(c), respectively, when single target exists in area and data is collected by the fusion method. We deploy half of the sensors in left bottom area and the other sensors in the other area for biased topology. Figure 32 through 35 show the mean estimated error and the number of data collected, respectively. Results of the number of data collected show that system can collect as same data as grid topology and proposed data collecting technique can work perfectly in any topology and any density. Results of the mean estimated error show that the minimum mean estimated error is around 0.5 m in random topology and around 0.6 m in biased topology. Results in this section show localization system get high accuracy with the small number of data and if many sensors send data, localization system cannot estimated position. It is very useful for localization to control the number of sensor trying to send.



(a) Grid (b) Random (c) Biased

Figure 25: Example of topologies in case of 100 sensors



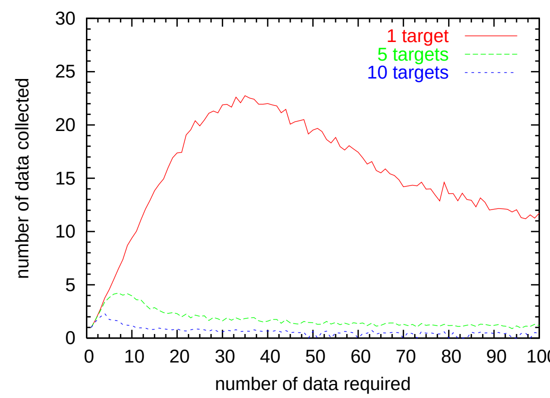
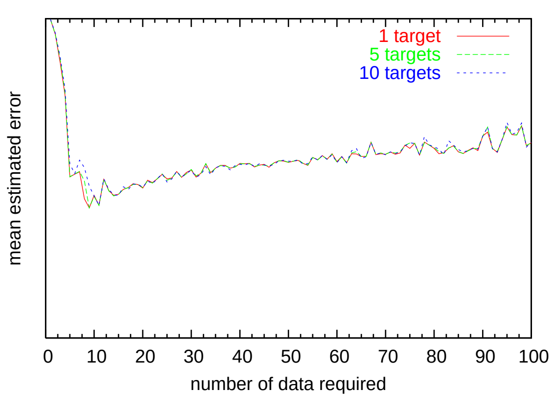


Figure 26: Mean estimated error of the direct method: density is set to 0.1 and grid topology

Figure 27: Mean estimated error of the direct method: density is set to 0.1 and grid topology



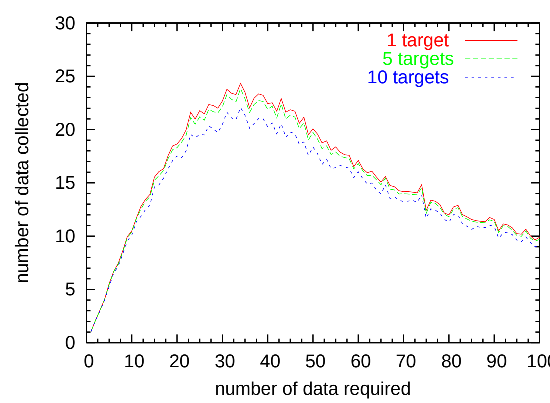


Figure 28: Mean estimated error of the fusion method: density is set to 0.1 and grid topology

Figure 29: Mean estimated error of the fusion method: density is set to 0.1 and grid topology

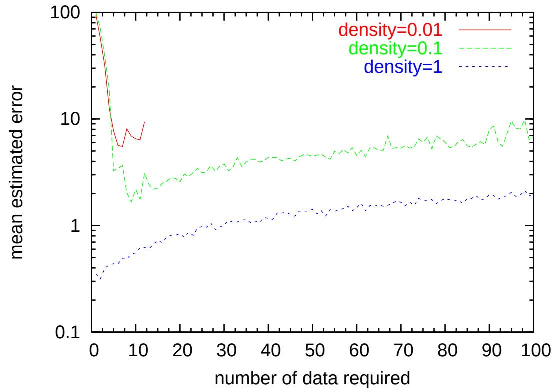


Figure 30: Mean estimated error of the fusion method in grid topology

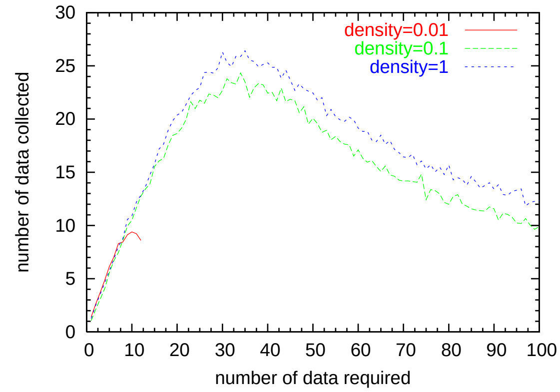


Figure 31: The number of data collected of the fusion method in grid topology

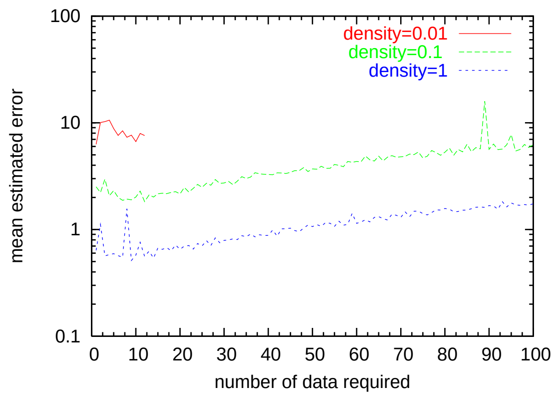


Figure 32: Mean estimated error of the fusion method in random topology

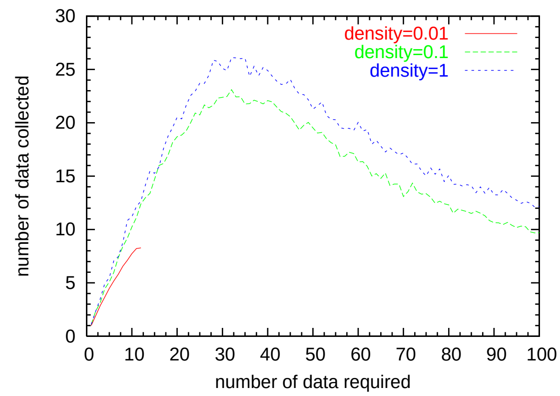


Figure 33: The number of data collected of the fusion method in random topology

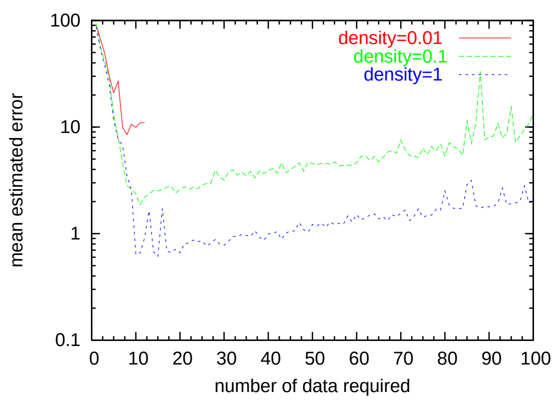


Figure 34: Mean estimated error of the fusion method in biased topology

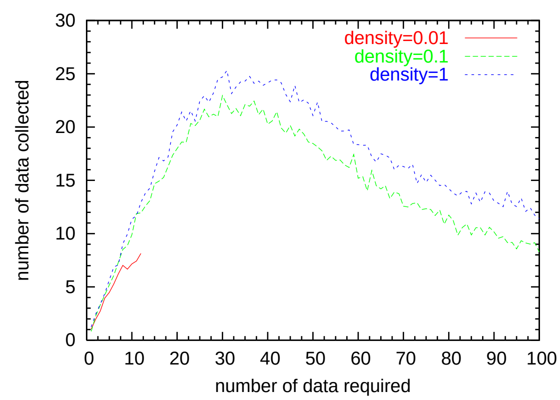


Figure 35: The number of data collected of the fusion method in biased topology

6 Conclusion

In this thesis, we have presented a localization system that uses RSSI to obtain the distance between sensors and targets for wireless sensor networks. It is important to save the energy of sensors and collect the necessary data in wireless sensor networks, because sensors have small battery capacity . We have also proposed data collecting technique and have shown availability of controlling the number of sensors trying to send and data fusion at the target. The results show that it can control the number of sensors trying to send even if sensors are deployed as non-regular and localization accuracy is enough to meet applications if the density is over $0.1/\mathrm{m}^2$. These results are led to use only MMSE using all data reached to the sink. Correction makes errors more small, system can estimate more accurate. Applications of localization, for example tracking user, are enough to realize in term of accuracy. As a matter of course, it needs massive sensors. Therefore, it is expected that sensors will be cheaper or measure distance more correct.

In this thesis, we assume all sensors and targets can communicate with the sink directly. However, this assumption is not always realized because of limitation of the radio range. Sensors must send data with multi-hop and routing algorithm is needed. It is needed to choice suitable routing algorithm and evaluate the performance of system about the number of data collected, the energy consumptions and the accuracy of localization. Furthermore, data collecting technique works out when the number of sensors are stationary and cannot adapt frequent change of the sensors' density due to addition or relocation of sensors. It is needed to consider framework to adapt dynamic change of the sensors' density. Next, we consider above to complete our research for localization system.

Acknowledgments

I would like to express my sincere appreciation to my supervisor, Prof. Masayuki Murata of Osaka University, who introduced me to the area of wireless sensor networks including the subjects in the thesis. He has also been a constant source of encouragement and advice throughout my studies and the preparation of this manuscript.

I would like to express my great gratitude to Associate Prof. Masashi Sugano of Osaka Prefecture College of Nursing. None of the work discussed in this thesis would have been possible without his advice and suggestions.

I am also indebted to Associate Professor Naoki Wakamiya, Associate Professor Go Hasegawa, Research Assistants Shin'ichi Arakawa and Research Associate Ichinoshin Maki of Osaka University for their helpful comments and feedbacks.

I deeply appreciate the helpful advice of Dr. Kenji Leibnitz. He has afforded result of his analysis and given helpful comments.

Finally, I want to thank my many friends and colleagues in the Department of Information Networking of the Graduate School of Information Science and Technology of Osaka University for their support.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, pp. 102–114, Aug. 2002.

[2] E. Kaplan, *Understanding GPS Principles and Applications*. Artech House, 1996.

[3] Y. Gwon, R. Jain, and T. Kawahara, "Robust indoor location estimation of stationary and mobile users," in *Proceedings of IEEE INFOCOM 2004*, Mar. 2004.

[4] N. Patwari and A. O. Hero III, "Using proximity and quantized RSS for sensor," in *Proceedings of the 2nd ACM International Conference on Wireless Sensor Networks and Applications*, Sept. 2003.

[5] A. Savvides, C.-C. Han, and M. B. Strivastava, "Dynamic fine-grained localization in ad-hoc networks of sensors," in *Proceedings of the 7th International Conference on Mobile Computing and Networking*, pp. 166–179, 2001.

[6] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less low cost outdoor localization for very small devices," *IEEE Personal Communications Magazine*, vol. 7, pp. 28–34, Oct. 2000.

[7] D. Niculescu and B. Nath, "Error characteristics of ad hoc positioning systems (APS)," in *Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing*, pp. 20–30, May 2004.

[8] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proceedings of IEEE INFOCOM 2000*, vol. 2, pp. 775–784, 2000.

[9] P. Krishnan, A. SKrishnakumar, W.-H. Ju, C. Mallow, and S. Ganu, "A system for LEASE: Location estimation assisted by stationary emitters for indoor rf wireless networks," in *Proceedings of IEEE INFOCOM 2004*, Mar. 2004.

[10] C. Savarese, J. M. Rabaey, and K. Langendoen, "Robust positioning algorithms for distributed ad-hoc wireless sensor networks," in *Proceedings of the General Track: 2002 USENIX Annual Technical Conference table of contents*, pp. 317–327, 2002.

[11] K. Langendoen and N. Reijers, "Distributed localization in wireless sensor networks: a quantitative comparison," *Computer Networks*, vol. 43, no. 4, pp. 499–518, 2003.

[12] H. Kung and D. Vlah, "Efficient location tracking using sensor networks," in *Proceedings of IEEE Wireless Communications and Networking Conference 2003 (WCNC 2003)*, pp. 1954–1962, Mar. 2003.

[13] W. Zhang and G. Cao, "Optimizing tree reconfiguration for mobile target tracking in sensor networks," in *Proceedings of IEEE INFOCOM 2004*, Mar. 2004.

[14] N. Bulusu, J. Heidemann, and D. Estrin, "Adaptive beacon placement," in *Proceedings of the Twenty-first International Conference on Distributed Computing Systems (ICDCS-21)*, Apr. 2001.

[15] Y. Zou and K. Chakrabarty, "Sensor deployment and target localization in distributed sensor networks," *Transaction on Embedded Computing Systems*, vol. 3, no. 1, pp. 61–91, 2004.

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

[17] I. C. Society, *Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs)*. The Institute of Electrical and Electronics Engineers, Inc., Aug. 2003.

[18] P.Brenner, *A Technical Tutorial on the IEEE 802.11 Protocol*. Breezecom Wireless Communications, 1997.

[19] W. Ye, J. Heidemann, and D. Estrin, "Medium access control with coordinated, adaptive sleeping for wireless sensor networks," *ACM/IEEE Transactions on Networking*, vol. 12, pp. 493–506, June 2004.

[20] T. van Dam and K. Langendoen, "An energy-efficient MAC protocol for wireless sensor networks," in *Proceedings of the 1st international conference on Embedded networked sensor systems*, pp. 171–180, June 2003.

[21] Y.-C. Tseng, C.-S. Hsu, and T.-Y. Hsieh, "Power-saving protocols for IEEE 802.11-based multi-hop ad hoc networks," in *Proceedings of IEEE INFOCOM*, pp. 200–209, June 2002.

[22] K. Leibnitz, N. Wakamiya, and M. Murata, “Analytical evaluation of the transmission range in clustered sensor networks,” *submitted to 19th International Teletraffic Congress (ITC-19)*, (Beijing, China), August 2005.

[23] D. Zhao, M. Takashima, K. Yanagihara, J. Taketsugu, K. Fukui, S. Fukunaga, S. Hara, and K. Kitayama, “A maximum likelihood estimation method of localization using RSSI in wireless sensor networks,” in *technical report of the Institute of Electronics, Information and Communication Engineers (IEICE)*, Mar. 2005.