Master's Thesis

Title

Collecting and Analyzing Chorus of Japanese Tree Frogs and its Application to Energy-Efficient Control of Wireless Sensor Networks

Supervisor
Professor Masayuki Murata

Author
Yasuharu Hirano

February 6th, 2017

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

Master's Thesis

Collecting and Analyzing Chorus of Japanese Tree Frogs and its Application to Energy-Efficient Control of Wireless Sensor Networks

Yasuharu Hirano

Abstract

Nature gives us so much inspiration. Self-organization phenomenon in nature is ubiquitous and our research group have ever applied this phenomena to various network control methods. These methods aim at bringing robustness and adaptability features of biological systems against environmental changes into information networks. Our research started from the interesting characteristics of the chorus of Japanese tree frogs that are much well-known animals in Japan. For example, male Japanese tree frogs call alternatively among their neighbors in order to attract female frogs. This emerging property of their chorus is one of so-called *swarm intelligence*. Previous work applied the property to transmission scheduling in wireless communication. Further observation on them let us notice other properties.

This thesis presents our work for revealing previously unknown behaviors of Japanese tree frogs and applying them in a network control method, which has three parts: (1) identifying the location and the vocalizing timing of calling Japanese tree frogs, (2) creating mathematical models of Japanese tree frogs' calling behavior, and (3) applying the models into a network control method. The collaborator of this work takes charge of the second part, and this thesis mainly includes the first and third parts.

For identifying sound source locations we propose a novel localization method using direction of arrival (DOA) measurements obtained from microphone arrays. In order to identifying the locations of sound sources, at least three microphones have to record the same sound. Existing methods assumed that sound sources were in an area surrounded by microphone arrays, however, which required that the sound-observable range of the microphones had to exceed the maximum distance among microphones. In an outdoor field, microphone-array-deployable areas are limited, and therefore, these assumptions are not always satisfied. Our proposal estimates the positions of sound sources out of the area surrounded by microphone arrays. By locating at least three microphone arrays closer to each other, the area observed by the three microphone arrays becomes larger in our method. Especially, locating microphone arrays close to each other does not require much consideration of the difficulty of deploying microphone arrays. We show that our proposed method achieves an estimation accuracy with an error of less than 20 cm.

On the network control inspired by the Japanese tree frog, we focus on two calling behaviors of Japanese tree frogs. One is antiphase-synchronizing calling in a short time scale and the other is

synchronizing calling in a long time scale. These calling behaviors are described by mathematical models proposed by our research collaborator. We propose a self-organizing scheduling method inspired by the frogs' calling model for energy-efficient data transmission in wireless sensor networks. We show that our proposal reduces 78.5 % energy consumption.

Keywords

Acoustic localization, Japanese tree frog, direction of arrival (DOA), wireless sensor network

Contents

List of Figures

List of Tables

1 Introduction

Mathematical models inspired by biological mechanisms help us to develop robust and adaptive systems in the ICT field [2]. In the background of these interdisciplinary research progress, a lot of studies of the mathematical modeling of biological systems have been performed thanks to the development of experimental techniques and a computer performance. The cooperative behavior with sociality emerging from autonomous motion controls of individuals is called *swarm intelligence* and there are lots of research that apply swarm intelligence to the network control [3].

Figure 1: Japanese tree frog calling at the ridge of rice paddies

In this thesis we focus on the interesting behavior of calling Japanese tree frogs (Fig. 1). I think if the reader has ever been in Japan in summer, he/she might experience loud frog sounds, *ribbit*. . . *ribbit*. . . *ribbit* from rice paddies. They often disrupt our sleep, but once we listen deeply to the chorus, rhythmical vocalization of frogs lets us notice its cyclic nature. It is known that only male Japanese frogs vocalize, which is for attracting female frogs and for claiming their own territory to other males. For such purpose, their vocalization does not overlap with each other [4], which is called antiphase synchronization as shown in Fig. 2. Our research group previously focused on this frog's feature and applied it to data transmission control method in wireless communication [5].

Recently, the collaborator of our research discovered an another aspect of Japanese tree frog's calling [6]. In a longer time scale than the above mentioned antiphasing vocalization, male

Figure 2: Short-term dynamics of the Japanese tree frog

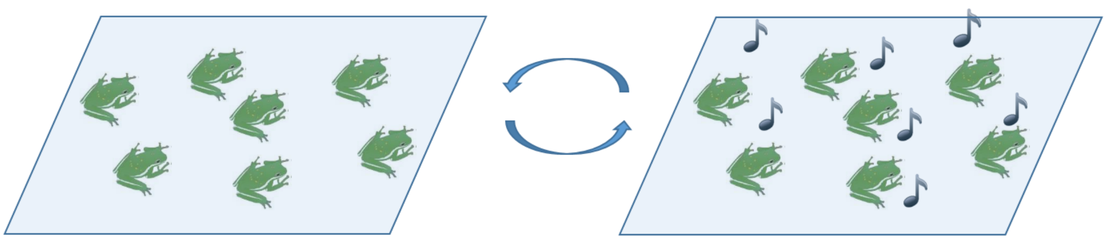


Figure 3: Long term dynamics of the Japanese tree frog

Japanese tree frogs vocalize advertisement sounds to inform females about their own position in a synchronized manner. That is, frogs transit in unison between two states: chorus state and resting state (Fig. 3). Once one of the individuals begins to call, the other individuals also begin to call. After chorusing for a few minutes, they stop calling and return to the rest state. In the chorus state, all frogs call while keeping antiphase synchronization. In the rest state, all of frogs do not call and stay silent. This calling behavior is thought to disperse the risks to be targeted by their enemies such as bats and snakes. Fatigue produced by vocalizing is huge [7] and therefore, by switching the chorus and the rest states, it is possible to reduce fatigue. However, the spatio and temporal structure of this calling behavior remains to be unclear.

In this thesis, we propose a method identifying the sound source location and the vocalizing timing of calling Japanese frogs, which is for building an elaborate model of the long-term dynamics of Japanese tree frog's calling. And also, we propose a energy-efficient network control method to which we applies the chorus models of Japanese tree frogs proposed in [6].

It is important to observe how individuals communicate with each other for investigating the mechanisms of their behavior. For modeling their communication, exploring *when* and *where* individuals interact with each other is necessary. Thus, the identification of individual positions is important. However, to find frogs in an outdoor environment is hard because they often are small and conceal themselves in the environment.

Many localization techniques have been proposed so far, but most of them are based on the assumption that a radio transmitter or receiver is directly mounted on the target animals [8, 9]. However, it is hard to put such a device on a target in advance in an outdoor environment. Then, we localize each animal based on the information that are detectable with some devices. One of such information obtained from their communication behavior is their calling. It is a natural idea to make a localization system that utilizes microphones to record their calling communication.

Outdoor environments make it difficult to deploy a localization system that consists of a large number of devices with wired connections. Therefore, we should conduct a localization method with a small number of wireless devices with a microphone for reducing the deployment cost. Most of existing sound-source localization methods are classified into two methods: TDOA-based (time difference of arrival) methods [10] and DOA-based (direction of arrival) methods. TDOA-based methods estimate sound-source positions using microphones' position and the time differences of a sound arrivals between all pairs of two microphone nodes. DOA-based localization methods estimate the sound-source position using sensor nodes' position and the angle of the signal arrival at each microphone array. In our previous research, we implemented a TDOA-based method into low-spec wireless sensor nodes and constructed a sound source localization system [11]. This system had various limitations such as an accuracy, observable range, and the number of sound source,

In this thesis we propose a novel DOA-based sound-source localization method using wireless sensor nodes with a microphone array. We implement it into micro computers, Raspberry Pi 2 [12] connecting with an 8-ch microphone array. Previously, we conducted some fieldwork for revealing spatio-temporal structures inherent in frogs' calling communication. Of course, it was difficult to detect the position of frogs visually because they call from inside of grass or underground. We found that deployable space for system equipment is very limited and that as the distance among nodes is longer, the measured positions of devices have larger errors. According to the feedback of the fieldwork, we design our localization system. We carry out experiments to clarify the accuracy of the estimated sound-source position taking into consideration of the outdoor environment.

Figure 4 shows an overview of our system, where wireless sensor nodes with a microphone array record sounds emitted from a sound source, estimate DOAs for the sound source, and transmit the estimated DOAs to a localization server. The lap-top computer finally estimates the sound-source position.

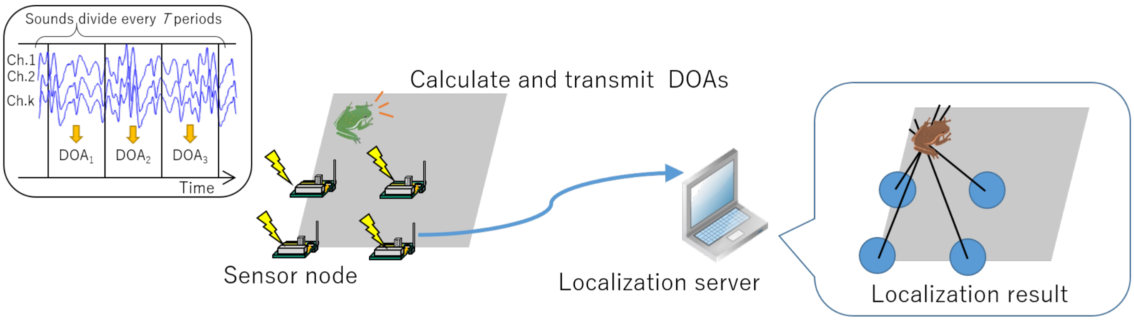


Figure 4: Summary of the implemented system

Most of existing methods assume that sound sources is within the area surrounded by sensor nodes [1]. However, considering the operation in an outdoor environment, it may be difficult to set devices on ideal positions. Therefore, our proposed method allocates wireless devices comparatively close to each other, e.g., on the 4 corners of a square 50 cm or 1 m on a side. Then, our method estimates out of the area surrounded by sensor nodes. It is noteworthy that in our method, since sensor nodes are located near each other, most of generated sounds are captured by all sensor nodes, which is an indispensable factor for sound source localization methods.

On the network control inspired by the Japanese tree frog, we focus on two calling behaviors of Japanese tree frogs that we mentioned above. We propose a self-organizing scheduling method inspired by the frogs' calling model for energy-efficient data transmission in wireless sensor networks. The short-term antiphase synchronization is used for contention-free data transmission among nodes. For applying the long-term synchronized state transition into a network control method, we interpret fatigue of frogs into the not used buffer size for data queueing in wireless sensor devices. From this interpretation, only when a device has a sufficiently much data to be transmitted, the device wakes up and starts data transmission. This triggers other devices wake-up and after the all data delivery to a data-collecting sink node completes, all devices return to the sleep mode. We show the proposal can reduce energy consumption by a computer simulation experiment.

The rest of this thesis is organized as follows. In the next section, we show some related work. In Section 3, we propose and evaluate a sound-source localization method for calling Japanese tree frogs. In Section 4, we apply the models of calling behavior of Japanese tree frogs to data transmission scheduling in wireless sensor networks. We give a conclusion of this thesis in Section 5.

2 Related work

In this thesis, we implement a sound-source localization system using wireless devices with a microphone aiming at biological research of the Japanese tree frog. Previously, we conducted some fieldwork for revealing spatiotemporal structures inherent in frogs' calling communication. Then, we found that it was difficult to detect the position of frogs because they call from inside of grass or underground, but found that their calling was loud and continued for a long time. According to the feedback of the fieldwork, we design the localization system using their chorus.

We present existing source localization methods in Section 2.1. We show chorus models of Japanese tree frogs that our research collaborator currently proposed in Section 2.2

2.1 Acoustic source localization methods

Most of existing sound-source localization methods are classified into two types of methods: TDOA-based (time difference of arrival) methods and DOA-based (direction of arrival) methods [10]. Both methods assume that three or more microphones or microphone arrays can record the sound from the same source.

TDOA-based methods estimate sound-source positions using microphones' position and the time differences of the sound arrivals between all pairs of two microphones. This type of methods requires comparatively highly accurate time synchronization among all microphones [13]. Comparing two sounds recorded by different two microphones, we can obtain one TDOA measurement from a phase-difference between them. The possible positions of a sound source are obtained as two hyperbolas whose foci are on the microphone locations from the TDOA between two microphones. Then, the intersection that all hyperbolas obtained from all sets of microphones meet is the estimated position of the sound source. Due to the errors of the timer in a sensor node as well as various environmental noises, all hyperbolic curves do not intersect on the same point. Many estimation techniques have been proposed to solve this problem [14–16].

DOA-based localization methods estimate the sound-source position using microphone arrays' position and the angle of the signal arrival. DOA means the direction from which a sound arises a microphone array. Nodes with a microphone array can estimate a DOA using some methods, e.g., multiple signal classification (MUSIC) [17], which is one of the best known methods. Since a microphone array consists of multiple microphones that connect with each other via a hardware circuit, it does not need to synchronize the clock of all microphones, which improves the measurement accuracy. We can localize the sound sources' position by finding the point where each DOA line originiated from each microphone array intersects. As well as TDOA, DOA-based methods suffer from various types of errors, resulting in an estimation error in the position of the source. Also a lot of estimation techniques have been proposed to solve the problem [1, 18, 19].

Existing multiple sound sources makes it more difficult to estimate them. Given that there are multiple sound sources and there is a localization server that collects TDOA/DOA measurements from all microphones, so-called data association problem happens. Namely, the localization server cannot know which sound source arises the TDOA/DOA measurements. The erroneous TDOA/DOA combinations often make ghost source, which is not exist actually. References [20, 21] try to deal with this problem.

In this thesis we propose a sound-source localization method focusing on identifying the location of calling Japanese tree frogs. This requires some challenges to overcome. To explain about them, we should describe the characteristics of the Japanese tree frog at first. Details are provided in Section 3.

2.2 Chorus model of Japanese tree frog

Our research group is working on modeling the chorus of the Japanese tree frog. We focus on the different behaviors of Japanese tree frogs, which are observed in the short term and the long term [4, 22]. In this section, we describe these models as dynamics.

2.2.1 Short-term dynamics

This section is not available now. We use the following equation as a short-term dynamics of frog chorus.

$$\frac{d\theta_n}{dt} = \omega_n + \sum_{r_{n,m}<R} \delta(\theta_m)\Gamma_{nm}(\theta_n - \theta_m) \tag{1}$$

2.2.2 Long-term dynamics

This section is also not available now.

(2)

(3)

(4)

(5)

(6)

(7)

(8)

(9)

(10)

(11)

(12)

(13)

3 Acoustic source localization of Japanese tree frogs using a wireless sensor network

3.1 System Requirements

3.1.1 Characteristics of Japanese tree frogs

Japanese tree frogs have unique and interesting characteristics. Male Japanese tree frogs vocalize *advertisement calls* early at night to inform their existence to a conspecific female (Fig. 1). When one frog begins to call, other frogs that hear it also begin to call following the frog (chorus). The chorus of a few frogs synchronizes in anti-phase so that their calling does not overlap. Their chorus behavior is considered to be for letting female frogs distinguish them individually.

The body length of Japanese tree frogs is 2.0–4.5 cm. They inhabit rice paddies or forests and their positions are sparsely-distributed. They do not call under the water and do not move while it calls. Once the Japanese tree frog begins a chorus, it often continues more than several minutes. The chorus of Japanese tree frogs can be observed in rice paddies in the spring of Japan, around which there are not the tall trees, but are growing thick grass.

3.1.2 Requirements for the Japanese tree frog localization

In this research, we identify the location of Japanese tree frogs calling in a rice paddy. Most of rice paddies are flat and they are filled with water, which makes it difficult to put microphones in the paddies. Therefore, we put them on the ridge of a rice paddies, however, which often is muddy and most part of it is not flat. For the operation of localization in an outdoor environment, it is desirable to use a light and small device for reducing deployment cost.

Existing frog localization system: "Firefly" The authors of [24] implement a sound-source localization system for the tree frogs using a special device with a light-emitting diode (LED) that turns on response to a nearby sound. This system can acquire the position of frogs with a precision of about 30 cm since in this system, respective devices are deployed in the observation area at 30 cm intervals, which takes comparatively much time to prepare and manage the system. The goal of our localization system is to achieve the precision comparable to that of [24] by a wireless sensor network composed of less number of wireless sensor devices with a microphone array.

System requirements Localization systems has various requirements for their application. Here, we discuss the requirements for our localization system that is for identifying the position of Japanese tree frog in an outdoor environment.

- Accuracy:
  Japanese tree frogs have a habit that each individuals does not make an advertisement call

while another frog is calling close to each other. Since this distance is about half a meter, therefore, our goal is to achieve an accuracy of 50 cm so as to distinguish any two frogs. This is not that different to that of "Firefly."

- Difficulty in deployment:
  In order to identify the position of the sound source in a 2-dimensional flat field, at least three microphones have to obtain the sound from the source. However, for the operation of localization in an outdoor environment, it is not always possible to place devices at arbitrary points. Therefore, it is desired that the localization method does not require inflexible positions of devices.

- Number of sound sources:
  Since Japanese tree frogs call in antiphase synchronization, it is rare to observe vocalization from multiple frogs simultaneously. However, multiple frogs can sometimes vocalize at the same time transiently, so we need to consider multiple sound-source localization problem.

- Real-time property:
  Frogs do not move while calling. It is known that once it begins to call it will keep calling for about a few minutes. Therefore, considering the case where it is necessary to find the position of the frog in the real environment, the time required for the localization is within a few minutes.

For the accuracy of the system, we use a DOA-based method due to its accuracy of measurement. In the following section, we show an existing localization method using DOA measurements and its problem to use for the frog observation.

3.2 Problems of existing methods

Many localization techniques using DOA measurements have been proposed to localize sound sources. A very simple idea to identify the position of a sound source is using intersection points of DOAs. Since this method is simple, the computational cost of it is very small. However, it suffers from an estimation error due to many irrelevant intersections when there are multiple sources and microphone arrays. In this thesis, we focus on the algorithm proposed in [1] since it has a comparatively high accuracy and its computational complexity is comparatively low.

3.2.1 Grid-based localization method

The method in [1] divides an area into equal sized N grids. Then finding the grid whose directions from microphones most closely match the estimated DOA. The localization algorithm is as follows:

1. Discretize the area of interest into N grids, and calculate the coordinates of the center of each grid.
2. Calculate a $(M \times N)$ matrix $\mathbf{\Psi}$ whose elements $\psi_{m,n}$ is the angle from the mth microphone array to the nth grid center (M is the number of microphone arrays).
3. Define a cost function $Cost_a$ that represents the degree of coincidence between the true DOA and an estimated DOA.

$$Cost_a = \sum_{m=1}^{M} \left[A\left(\hat{\theta}_m, \psi_{m,n}\right)\right]^2 \tag{14}$$

   where $\hat{\theta}_m$ is DOA from the mth sensor node.
4. Get the grid that minimises the cost function, that is, $n^* = \arg\min Cost_a$.

$A(X, Y)$ is angular distance between X and Y. This is given by

$$A(X, Y) = 2\sin^{-1} \frac{|\exp(jX) - \exp(jY)|}{2}. \tag{15}$$

In this method, the resolution of the grid, which depends on the number of grid N, determines an estimation accuracy. When increasing N, the estimation error will decrease while increasing computation cost.

This localization method can deal with multiple sources give the correct number of sound sources. To determine the positions of multiple sources, the authors of [1] uses a two-step procedure. At first, it calculates the set (denoted by J) that contains all possible combinations of DOAs. Second, for each grid, it calculate $Cost_a$ using a combination of DOAs, denoted by j $(j \in J)$. S grids that has sth minimum $Cost_a$ are selected as the sources' location ($s = 1, 2, \cdots, S$), where S is the highest number among the number of DOAs all microphone arrays detect.

When we execute this procedure naively, the amount of calculation becomes very large. Therefore, a recursive approach is also proposed. At first coarse grids are used for dividing the area, and the above-mentioned method are applied. Then, obtained grids are recursively divided into finer grids.

As the numbers of the sound sources, microphone arrays, or grids increase, the computational cost also increases. In particular, increasing the number of sound sources and grids significantly affect the amount of computation.

3.2.2 Problems of the grid-based localization method

As we discussed above, it is not always possible to place devices at arbitrary points in an outdoor environment. In addition, the condition where at least two microphone arrays observe the same sound sources may not be satisfied. This is because the range that a microphone array can estimate

DOA in an outdoor environment becomes smaller than that in an ideal environment due to various noises caused by wind and insects' calling.

These problems can be solved by a very simple idea, that is, we only place at least two microphone arrays close to each other. However, most of existing acoustic localization methods using multiple microphone arrays considers that sound sources are within the area surrounded by microphone arrays. The grid-based method is one of these methods. The grid based localization method assumes that the sound sources are in the area surrounded by microphone arrays and all microphone arrays can hear the sound generated from within the area. Once this assumption is broken, there happens an unexpected estimation error.

Figure 5 shows the situation which an incorrect grid has smaller cost than the correct grid. In this way, the existing method may choose an erroneous grid, especially from more outward grids. Such a problem can be suppressed by taking a finer grid division. However, if grids are divided smaller, the computational cost for the method becomes significantly higher, which results in not completing the localization within a prescribed time limit.

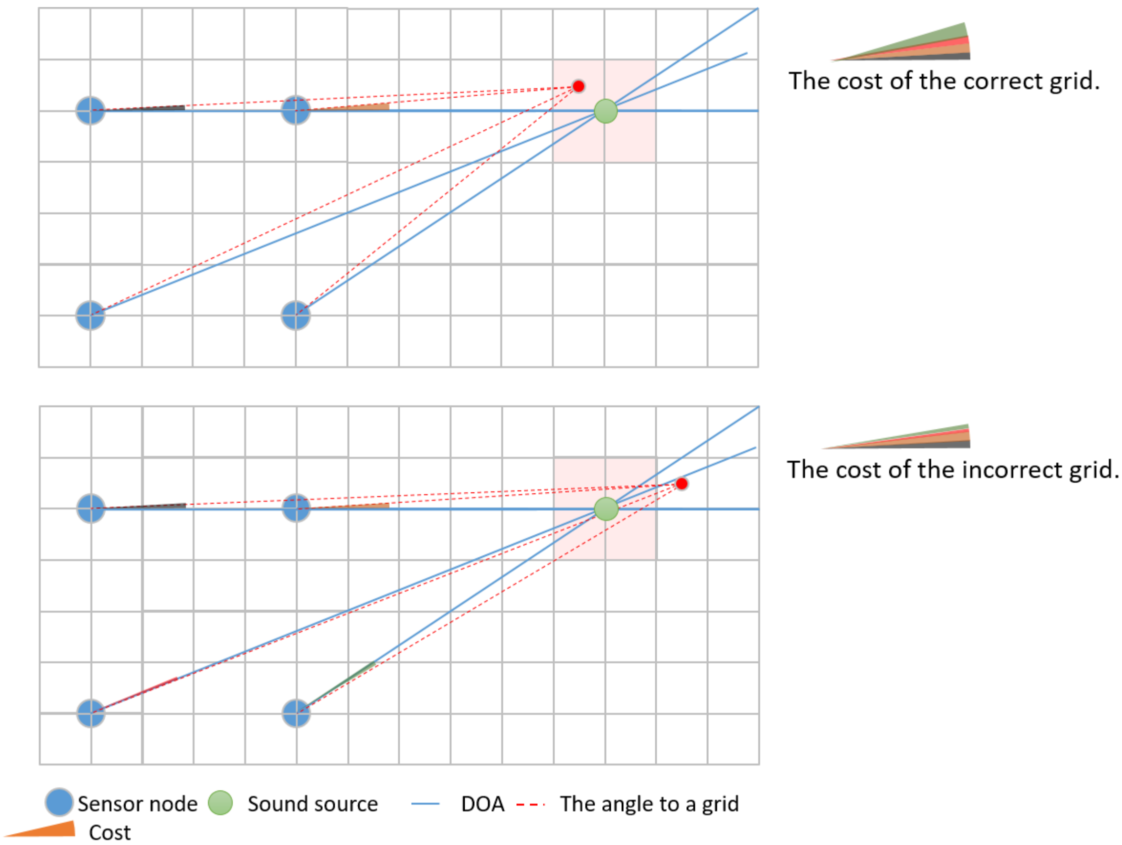


Figure 5: Incorrect grid has a smaller cost

3.3 Light-weighted multiple sound source localization method using DOA measurements

3.3.1 Overview

In this section, we propose a multiple sound source localization method. Our method is based on the grid-based method proposed in [1]. When we use the original grid-based localization method to estimate positions of sound sources in the outside of the area which surrounded by microphone arrays, it requires much higher resolution of grids than inside of the area in order to obtain an accurate result.

In our method, microphone arrays estimate DOA using the MUSIC method. Then, all DOAs are collected in one localization server that estimates the positions of sound sources. The localization server divides the observation area into fine grids and selects small numbers of grids whose $Cost_a$ should be calculated. To do so, the localization server first calculates a directional cost $Cost_d$ for each direction from the center of microphone arrays. Then, the direction that has the minimum $Cost_d$ can be obtained. The localization server calculates $Cost_a$ for each grid whose center is close to the line running from the center of microphone arrays according to the direction.

3.3.2 Multiple signal classification method

The multiple signal classification (MUSIC) method [17] is one of methods to estimate DOA. The MUSIC method estimates the DOA of a signal using an eigenspace method.

Definition of a sound signal model The model of a signal measured by mth microphone in a microphone-array is defined as following.

$$x_m(t) = \sum_{l=1}^{L} (a_{m,l}(t, \theta_l) s_l(t)) + n_m(t), \tag{16}$$

where θ_l is direction of the lth sound source to microphone m, $s_l(t)$ is a signal of the lth sound source in time domain, $a_{m,l}(t)$ is a transfer function between lth sound source and mth microphone, and $n_m(t)$ means noise. The Fourier transform of $\boldsymbol{x}$ is

$$\boldsymbol{X}(\omega) = \sum_{l=1}^{L} (\boldsymbol{A}_l(\omega, \theta_l) S_l(\omega)) + \boldsymbol{N}(\omega) \tag{17}$$

where $S_l(\omega)$ is a sound signal in frequency domain in the lth sound source, $\boldsymbol{A}_l = [A_{1,l}(\omega), \ldots, A_{m,l}(\omega)]$ is a transfer function in frequency domain between the lth sound source and the mth microphone.

MUSIC method The MUSIC method estimates the angle of sound arrival by using a steering vector $\boldsymbol{A}(\omega, \theta)$. A steering vector is derived from actual measurements of impulse signals.

Let $s_l(t,\theta_l)$ be an impulse signal. The Fourier transform of $x_m(t) = a_{m,1}(t,\theta_1)s_1(t,\theta_1)$ is represented as

$$\boldsymbol{X}_m(\omega) = \boldsymbol{A}_1(\omega,\theta_1)S_1(\omega,\theta_1) = \boldsymbol{A}_1(\omega,\theta_1). \tag{18}$$

In this way, steering vector $\boldsymbol{X}(\omega)$ at arbitrary angle θ_1 is obtained in advance.

Let an observed signal in time domain be $\boldsymbol{x}(\tau)$, and its Fourier transform is $\boldsymbol{X}(\omega,\tau)$. The correlation matrices of $\boldsymbol{x}(\tau)$ are defined as

$$\boldsymbol{R}(\omega) = \boldsymbol{X}(\omega,\tau)\boldsymbol{X}^*(\omega,\tau) \tag{19}$$

where $()^*$ represent the complex conjugate. The eigenvalue decomposition of $\boldsymbol{R}$ is obtained as

$$\boldsymbol{R}(\omega) = \boldsymbol{E}(\omega)\Lambda\boldsymbol{E}^{-1}(\omega) \tag{20}$$

Eigenvalue λ_m represents the power of each sound. Therefore, λ_i and $\boldsymbol{e}_i$ when $1 \le i \le L$ represent eigenvalues and the vectors of sound sources. And others are those of noises. The spatial spectrum is defined as

$$P(\omega,\theta) = \frac{|\boldsymbol{A}^*(\omega,\theta)\boldsymbol{A}(\omega,\theta)|}{\sum_{m=1}^{M}|\boldsymbol{A}^*(\omega,\theta)\boldsymbol{e}_m|}. \tag{21}$$

When the direction of steering vector $\boldsymbol{A}(\omega,\theta)$ is equal to the direction of sound sources, the $P(\omega,\theta)$ becomes infinity. Then the sound source exists in the direction to θ.

3.3.3 Single source localization

At first, we describe our method in the case where the number of sound source is one. Our proposed method is divided into two steps. At the first step, we estimate the direciton which sound source exists. At the second step, the grid-based sound source localization is conducted.

Let $\hat{\boldsymbol{\theta}}$ be a $M \times 1$ vector and each element $\hat{\theta}_m$ is DOA that microphone array m estimated. Here we can assume that the coordinate of the center of microphone arrays is the origin without loss of generality. First, we estimate the direction from the origin to a sound source. For the estimation of the sound source direction, we use the sum of the angular distances between a vector from the origin to the direction θ_s and DOAs each sensor node estimated (Fig. 6). This is because incorrect grids in the grid-based localization for an outside area that is surrounded by microphone arrays often have the same direction from the origin as the true grid.

The cost of sound source direction for each θ_s ($0 \le \theta_s \le 2\pi$) is

$$Cost_d(\theta_s) = \sum_{m=1}^{M}\left[A\left(\hat{\theta}_m,\theta_s\right)\right]^2 \tag{22}$$

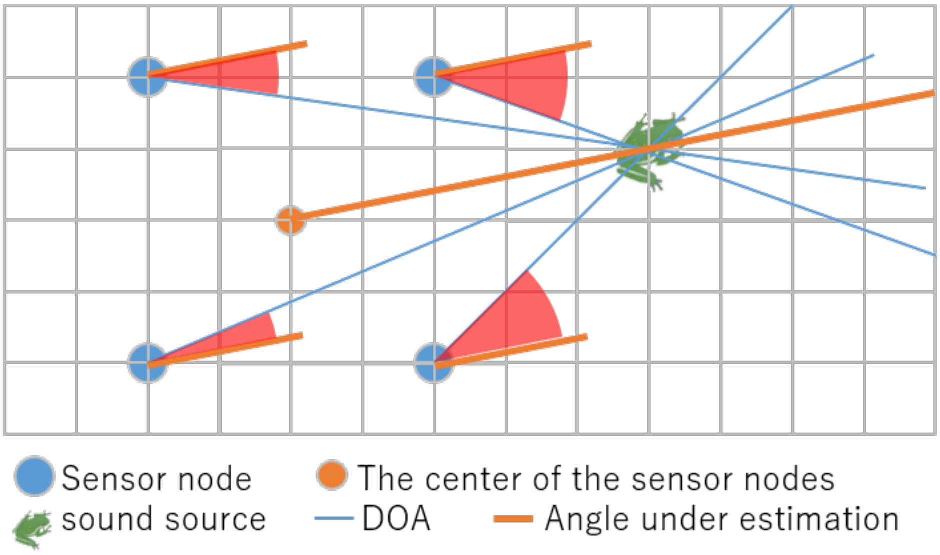


Figure 6: Estimation of direction

where $A(X, Y)$ is angular distance defined in (15).

Then, we can estimate the direction of a sound source as follows.

$$\theta^* = \arg\min_{\theta_s}(Cost_d). \tag{23}$$

Secondly, the grid-based sound source estimation is conducted. In advance, we divide the area of interest into grids with a side length of x meters, and sets of grids are denoted as $\boldsymbol{P}$. x affects the accuracy and the computational cost of our method, and is adjusted to the required estimation accuracy. Next, we find the grids $\boldsymbol{P}'$ that intersects with a vector whose starting point is the origin and direction is θ^* (Figure 7).

We use the angular distance function (15) to get the cost with each direction. In order to obtain θ_* using a computer program, we discretize θ_s. For obtaining a high accuracy, we have to divide $theta_s$ finely, which increase computational cost. Therefore, we use a recursive algorithm. First, we start an angle estimation with a coarse angle, and then we obtain θ_1 and θ_2, which are the minimum and the second minimum $Cost_d$ values, respectively. Once θ_1 and θ_2 is found, repeat the same step again in the range of $\theta_1 \leq \theta_s \leq \theta_2$ (here, we assume that $\theta_1 < \theta_2$). Then, the direction is found while reducing searching cost.

3.3.4 Multiple source localization

The idea of multiple source localization of our method is similar to the grid-based localization method. Let $\boldsymbol{J}$ denote the set of all possible unique combinations of DOAs. Then, we calculate

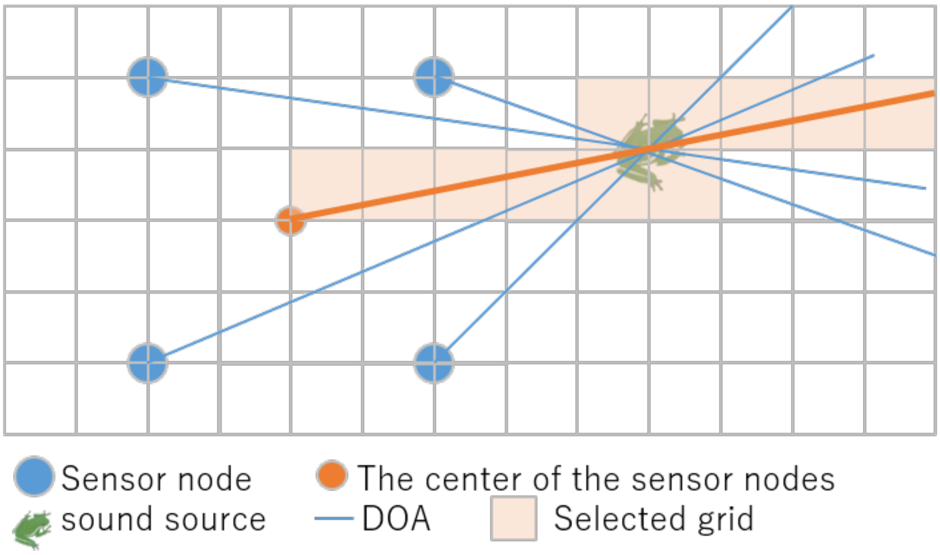


Figure 7: Reduction of searched grids

$Cost_d(\theta_s)$ for each $j \in \boldsymbol{J}$. Here, we assume that the number of sound sources S is the maximum number of DOAs that microphone arrays detected. Finally, we obtain S directions that has the Sth smallest cost. After that, the localization procedure is the same as the case in a single sound source.

3.4 Experiments and results

In this section, we evaluate the proposed localization method by computer simulation. It is noted that we implemented the MUSIC method into a wireless devices and evaluated the computational time and the accuracy to estimate DOA. The result of this preliminary experiments is used for a DOA error model in our simulation.

3.4.1 Single source localization

Simulation settings We assume that there are 4 sensor nodes with a microphone array in the observed area. Each sensor node is placed on the corner of a square region with a side (denoted by S_{dist}) of 0.5 m or 1.0 m. Microphone arrays can estimate DOA for the sound emitted within 5 m from them. Each sensor node estimates DOA at a short time interval. For that, it records a sound, divides the recorded sound into short parts, and estimates DOA for each part. After the estimation of DOA, it transmits DOA to a localization server. The overall system is shown in Fig.4.

- Sensor node settings
  We place a set of 4-sensor nodes with a microphone array on each corner of a square. We

assume that the center of the square is the origin. In the case of 0.5 m square, sensor nodes are placed at $(-0.25, -0.25)$, $(-0.25, 0.25)$, $(0.25, -0.25)$, and $(0.25, 0.25)$. Sensor Nodes can estimate DOA of the sound from a source if the distance between a sensor node and the source is less than 5 m.

- Sound source settings
  The observed area is a square region with a side of 9 m. The corners of the observed area is $(-4.5, -4.5)$, $(-4.5, 4.5)$, $(4.5, 4.5)$, and $(4.5, -4.5)$. Here, we divide the region into 18x18 grids, that is, each grid is a square with a side of 0.5. A sound source is arranged at an arbitrary place at a grid point, except for the same position as a sensor node. We evaluate the localization error when each grid point is chosen for the sound source position. Considering the symmetric property, we only evaluate the square region whose corners are $(-4.5, 0)$, $(0, 0)$, $(0, -4.5)$, and $(-4.5, -4.5)$.

- Localization settings
  The localization server conducts our localization method only when it receives DOAs from all of the sensor nodes, that is, if at least one of the sensor nodes can not get DOA, localization is not performed. In this simulation, we evaluate the estimation accuracy with and without errors in obtained DOA. An error model of DOA is as follows.

$$DOA = trueDOA + \epsilon(\epsilon = 5^\circ * \lfloor r + 0.5 \rfloor, r\ N(0, 1)) \tag{24}$$

  This simulates the actual result of DOA measurement by using the MUSIC method. We implement the MUSIC method into a Raspberry pi 2 B+ that connects with an 8-ch microphone array (TAMAGO-03, System in Frontier Inc.) as shown in Fig. 8. For obtaining one measurements of DOA, a Raspberry pi takes about 15 seconds in our preliminary experiment. For the localization method, we set the grid size to 0.10 m, the resolution of direction estimation for θ_s is set to $2\pi/36$. Simulation trial is performed 30 times per each sound source position when we consider the above error model.

- Evaluation metrics
  We evaluate the localization error by distance between the estimated point and the true point, and also by the root mean square error (RMSE) between the estimated-positions and the true source positions. RMSE shows the variation of the estimation result from the true position.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} ((x_i - \hat{x})^2 + (y_i - \hat{y})^2)} \tag{25}$$

  where x_i, y_i is the coordinate of result, $\hat{x}, \hat{y}$ is the coordinate of true sound source position and N is the number of the estimation.

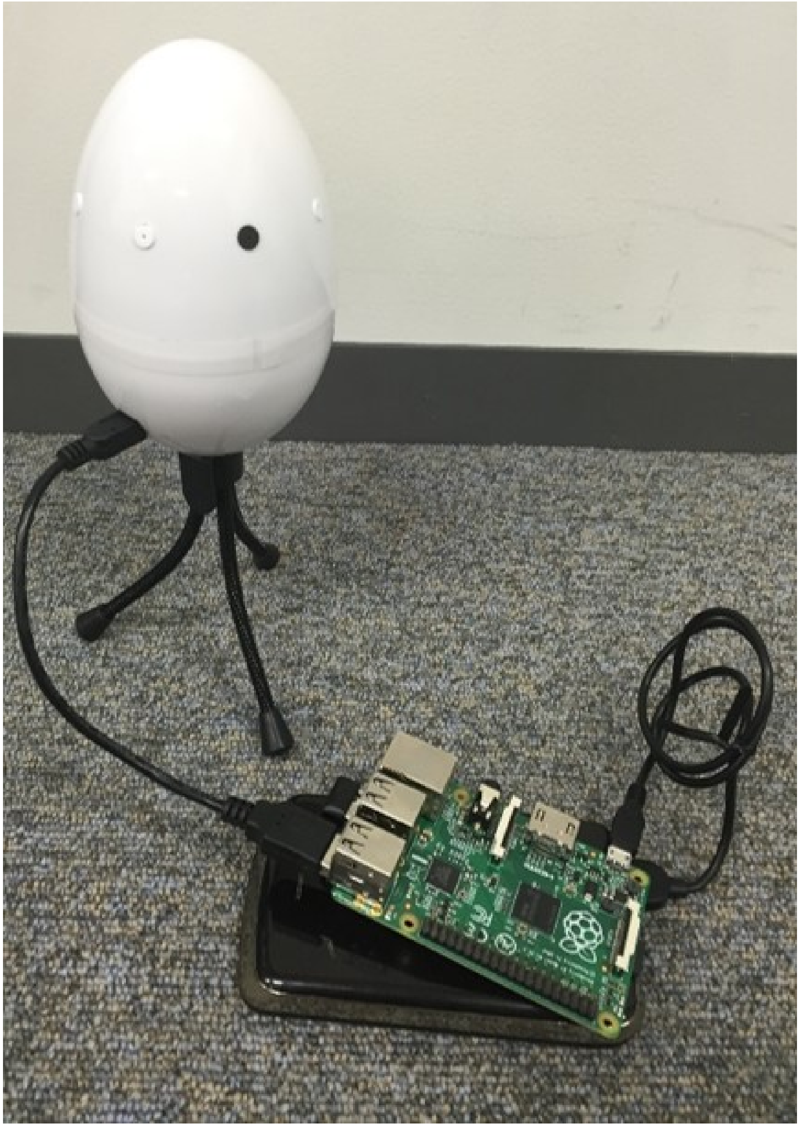


Figure 8: Raspberry pi 2 B+ with an 8-th microphone array (TAMAGO-03).

Simulation results Figure 9 shows the result of localization. In the figure, x and y axes are the coordinate of a sound source and estimation errors are represented by the color map when the sound source is at the (x, y). Most of the position of a sound source is well estimated and computation takes about 30 seconds. As shown in Table 1, the mean error of our method achieves an error of less than 20 cm. When comparing Figs. 9a and 9b, the estimation error of the latter is smaller. On the other hand, an unobservable area of the latter colored in gray is larger.

The original grid-based localization method does not consider the sound source in the outside of the area covered by sensor nodes. Even if we use the method, it cannot estimate the position correctly as shown in Fig. 10).

Table 1: Estimation errors of the single sound source localization

S_{dist}	0.5 m	1.0 m	0.5 m (DOA error)	1.0 m (DOA error)
Mean error [m]	0.173	0.177	0.40	0.213
RMSE [m]	0.251	0.223	1.701	0.972

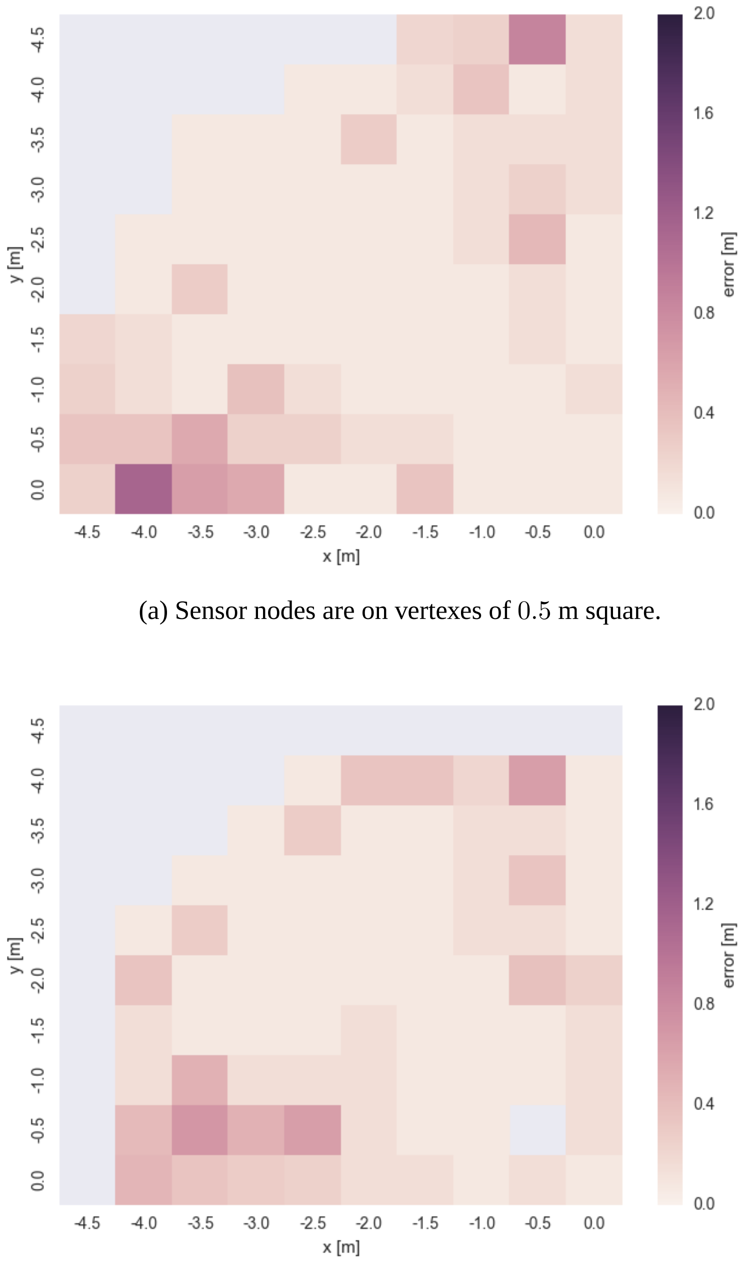
Figure 11 shows the result of localization considering DOA errors. In this case, we get the center of gravity of 30 estimation results per each sound source position. In this case, we simply take the center of gravity of all estimation results since there is only one sound source. DOA can be estimated with an error of about 50 cm.

3.4.2 Multiple source localization

We show the result of localization when considering multiple sound sources. Almost all of settings are the same as the prior section. In this section, we assume that 2 sound sources are arranged at the area. Since evaluating all possible positions of two sound sources takes large cost, we evaluate the estimation accuracy in 2 patterns where sound sources are located at $(2.5, 2.5)$ and $(-1.0, 2.0)$ or $(2.5, 2.5)$ and $(-2.0, -1.5)$.

Simulation results Figure 12 shows the result of localization without DOA errors. In this case we can correctly estimate the positions of sound sources. The estimation error is less than 10 cm.

Figure 13 and Table 2 shows the result of estimation considering multiple sound sources and DOA errors. Figure 13 presents the estimated points in all 30 trials. At this time, to obtain the gravity center of the estimated points, we should divide these points into two clusters. In most cases, the number of clusters k (k is the total number of the sound source in the field) is undetermined. Therefore, we have to use a clustering method that can deal with unknown k. In this thesis, we use the x-means method [25]. The accuracy for both of the sound sources are shown in Table.2. The distance between the estimated position and the true position is about 50 cm, which is our target value of accuracy. The calculation time for this simulation is about 90 seconds.



(a) Sensor nodes are on vertexes of 0.5 m square.

(b) Sensor nodes are on vertexes of 1 m square.

Figure 9: Estimation errors of the single source localization without DOA error

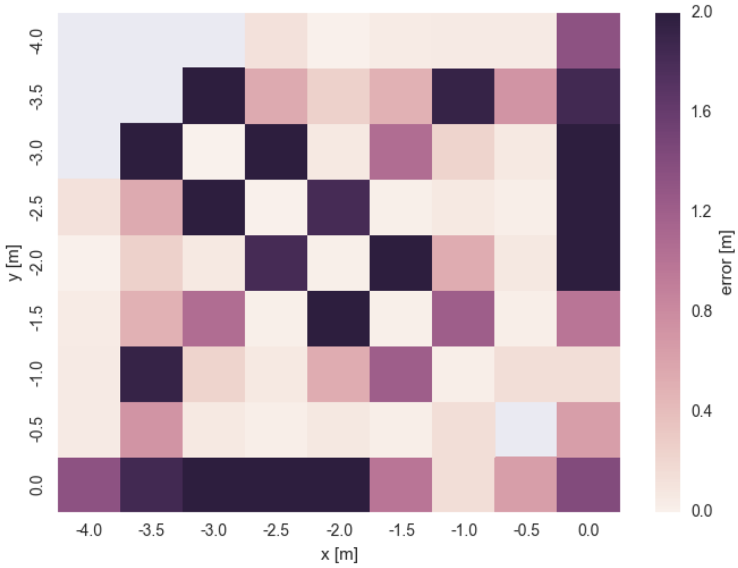
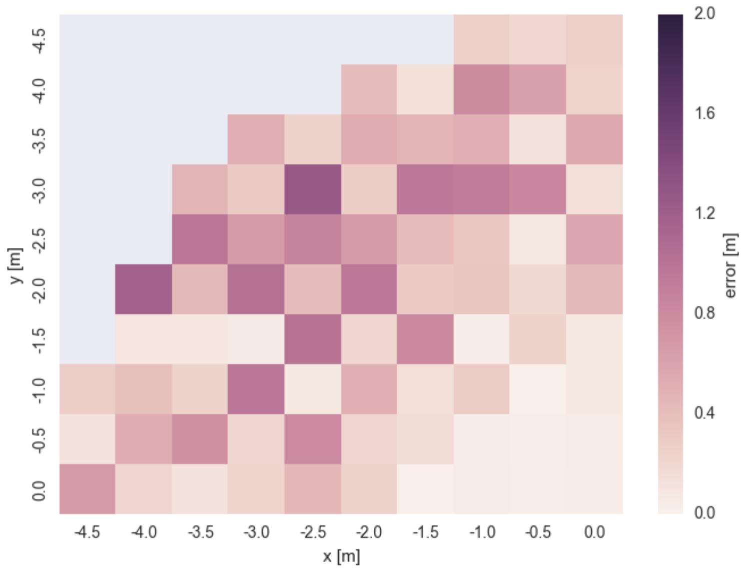


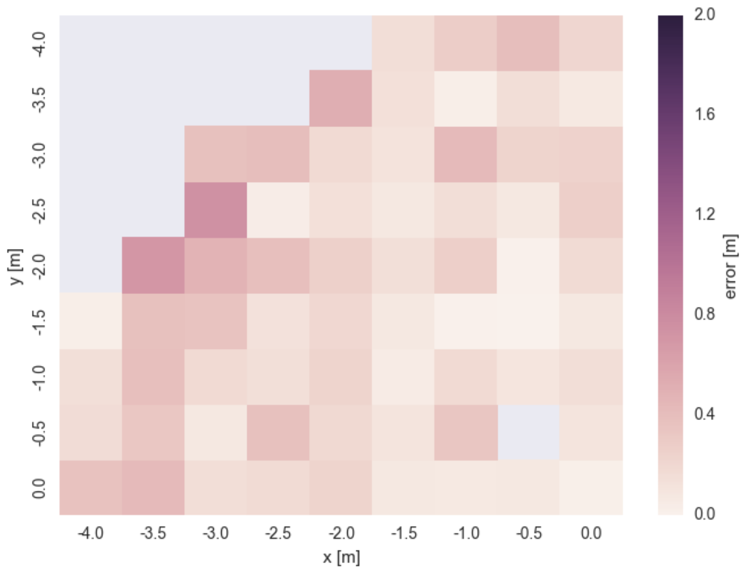
Figure 10: Grid-based estimation [1]

Table 2: Estimation errors of two sound sources

The sound source position [m,m]	$(2.5, 2.5)$	$(-1.0, 2.0)$
The center of gravity of the estimation [m,m]	$(2.87, 2.83)$	$(-1.15, 2.40)$
RMSE [m]	1.47	2.25

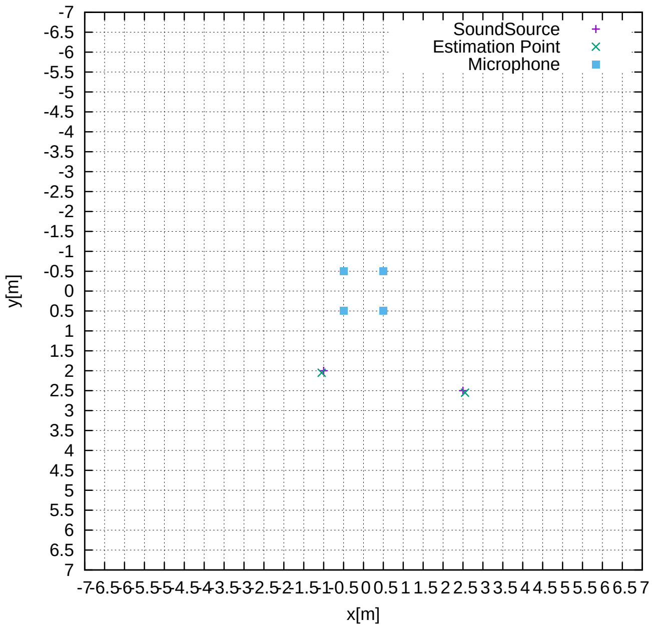


(a) Sensor nodes are on vertexes of 0.5 m square.

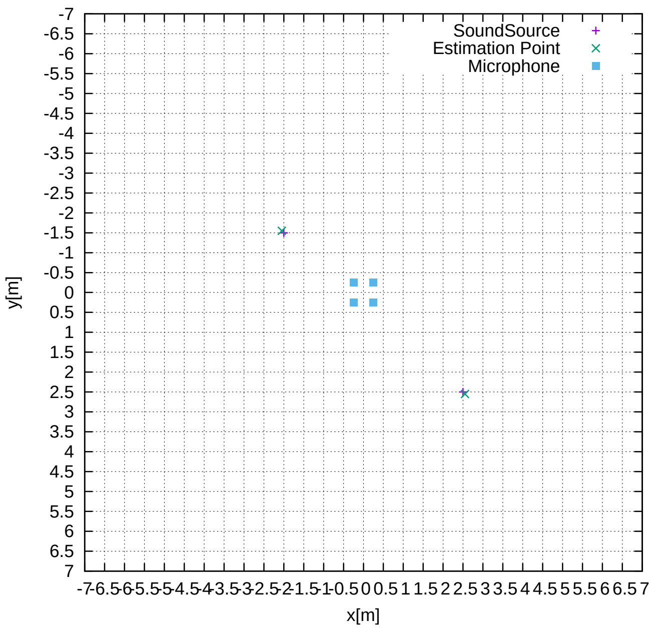


(b) Sensor nodes are on vertexes of 1 m square.

Figure 11: Estimation errors of the single source localization with DOA error



(a) The estimation result when sound sources are at $(2.5, 2.5)$ and $(-1.0, 2.0)$, sensor nodes are on vertexes of 0.5 m square (without DOA error).



(b) The estimation result when sound sources are at $(2.5, 2.5)$ and $(-2.0, -1.5)$, sensor nodes are on vertexes of 0.5 m square (without DOA error).

Figure 12: Estimation results of the two source localization with DOA error

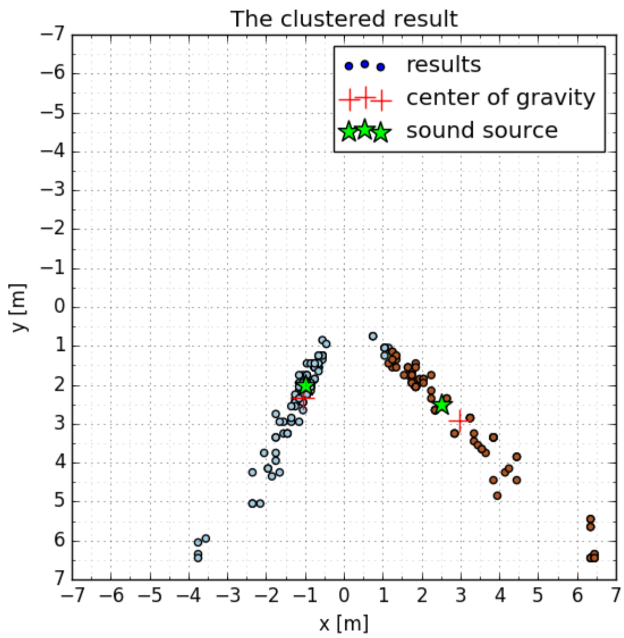


Figure 13: Estimation result when sound sources are at $(2.5, 2.5)$ and $(-1.0, 2.0)$, sensor nodes are on vertexes of 0.5 m square (with DOA error).

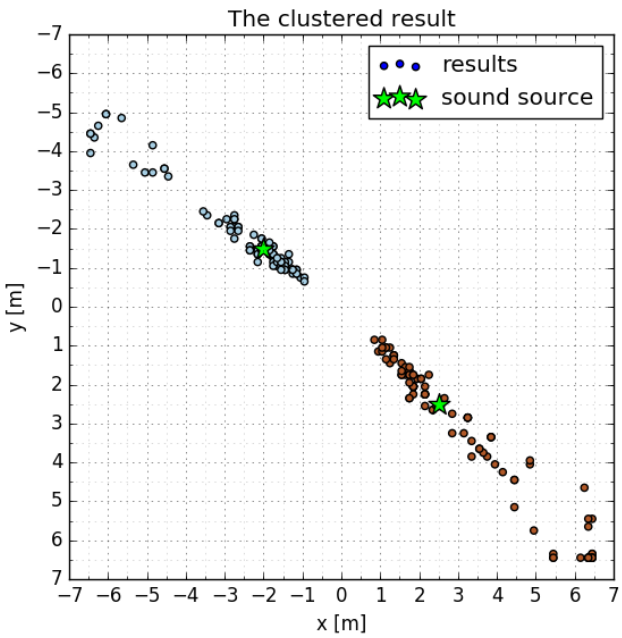


Figure 14: Estimation result when sound sources are at $(2.5, 2.5)$ and $(-2.0, -1.5)$, sensor nodes are on vertexes of 0.5 m square (with DOA error).

4 Energy-efficient transmission scheduling in wireless sensor networks

This section is currently not available.

Figure 15: Active state of sensor nodes.

Figure 16: Remaining energy.

5 Conclusion

In this thesis, we proposed a sound source localization method using a wireless sensor network. Even if there is a sound source in the area where estimation was difficult by the existing method, it was possible to estimate with an error of 20 cm with no DOA errors. We also showed the proposed method can estimate the positions of multiple sound sources under some DOA errors. Total calculation time to estimate the position of a sound source is a few minutes. These results can achieve the goal of our system. We also propose a data transmission scheduling method for wireless sensor networks inspired by the calling models of the Japanese tree frog. This can reduce energy consumption of such networks while it increases the total delay time.

Our future work is to conduct localization experiments in an actual environment to show the availability of our proposed system. We will also try to reduce the computational time for the localization to realize a real-time property.

Acknowledgements

There are so many people to thank for helping me during my master's degree studies at Osaka University. I would first like to express my appreciation to my thesis supervisor Professor Masayuki Murata of Osaka University. His continuous support and expensive advice throughout my studies steered me in the right direction. Engaging the tutelage of him is the valuable experience that I have ever had.

I would like to thank Assistant Professor Daichi Kominami of Osaka University, for his elaborated guidance and invaluable firsthand advice. I also would like to thank Assistant Professor Ikkyu Aihara of University of Tsukuba. I learned the fun and depth of studying Japanese tree frogs from him.

I am grateful to the members of Murata Laboratory, particularly to Associate Professor Shin'ichi Arakawa, Assistant Professor Yuichi Ohsita, and Assistant Professor Naomi Kuze of Osaka University for their valuable comments and suggestions on this study.

This thesis would not been possible without supports from the above people. And finally, I pay my heartfelt respects and gratitude to everyone offering me emotional suport.

References

[1] A. Griffin, A. Alexandridis, D. Pavlidi, and A. Mouchtaris, “Real-time localization of multiple audio sources in a wireless acoustic sensor network,” in *2014 22nd European Signal Processing Conference (EUSIPCO)*, pp. 306–310, IEEE, Sept. 2014.

[2] S. Barbarossa and G. Scutari, “Bio-inspired sensor network design,” *IEEE Signal Processing Magazine*, vol. 24, pp. 26–35, May 2007.

[3] F. Dressler and O. B. Akan, “A survey on bio-inspired networking,” *Computer Networks*, vol. 54, pp. 881 – 900, Oct. 2010.

[4] I. Aihara, R. Takeda, T. Mizumoto, T. Otsuka, T. Takahashi, H. G. Okuno, and K. Aihara, “Complex and transitive synchronization in a frustrated system of calling frogs,” *Physical Review E*, vol. 83, Mar. 2011.

[5] A. Mutazono, M. Sugano, and M. Murata, “Energy efficient self-organizing control for wireless sensor networks inspired by calling behavior of frogs,” *Computer Communications*, vol. 35, pp. 661–669, Mar. 2012.

[6] I. Aihara, D. Kominami, Y. Hirano, and M. Murata, “Autonomous distributed control methods for wireless sensor networks based on nonlinear dynamics of frog choruses,” *Technical report of IEICE IN*, Mar. 2017.

[7] H. C. Gerhardt and F. Huber, *Acoustic communication in insects and anurans: common problems and diverse solutions*. University of Chicago Press, July 2002.

[8] A. Wilson, M. Wikelski, R. Wilson, and S. Cooke, “Utility of biological sensor tags in animal conservation,” *Conservation Biology*, Mar. 2015.

[9] R. E. Floyd, “Rfid in animal-tracking applications,” *Potentials, IEEE*, vol. 34, pp. 32–33, Oct. 2015.

[10] Z. M. Saric, D. D. Kukolj, and N. D. Teslic, “Acoustic source localization in wireless sensor network,” *Circuits, Systems and Signal Processing*, vol. 29, pp. 837–856, Oct. 2010.

[11] Y. Hirano, T. Iwai, D. Kominami, I. Aihara, and M. Murata, “Implementation of a sound-source localization method for calling frog in an outdoor environment using a wireless sensor network,” in *Proceedings of International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pp. 2458–2462, Mar. 2016.

[12] T. R. P. Foundation, available at 10th Feb. 2017.

[13] G. Simon, M. Maróti, Á. Lédeczi, G. Balogh, B. Kusy, A. Nádas, G. Pap, J. Sallai, and K. Frampton, “Sensor network-based countersniper system,” in *Proceedings of 2nd international conference on Embedded networked sensor systems*, pp. 1 –12, Nov. 2004.

[14] A. Canclini, E. Antonacci, A. Sarti, and S. Tubaro, “Acoustic source localization with distributed asynchronous microphone networks,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 21, pp. 439–443, Feb. 2013.

[15] P. Bestagini, M. Compagnoni, F. Antonacci, A. Sarti, and S. Tubaro, “Tdoa-based acoustic source localization in the space–range reference frame,” *Multidimensional Systems and Signal Processing*, vol. 25, pp. 337–359, Apr. 2014.

[16] T. Ajdler, I. Kozintsev, R. Lienhart, and M. Vetterli, “Acoustic source localization in distributed sensor networks,” in *Proceedings of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1328–1332, Nov. 2004.

[17] R. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Transactions on Antennas and Propagation*, vol. 34, pp. 276 – 280, Mar. 1986.

[18] J. C. Chen, K. Yao, and R. E. Hudson, “Source localization and beamforming,” *IEEE Signal Processing Magazine*, vol. 19, pp. 30–39, Aug. 2002.

[19] J.-M. Valin, F. Michaud, J. Rouat, and D. Létourneau, “Robust sound source localization using a microphone array on a mobile robot,” in *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 2, pp. 1228–1233, Dec. 2003.

[20] X. Sheng and Y.-H. Hu, “Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks,” *IEEE Transactions on Signal Processing*, vol. 53, pp. 44–53, Dec. 2005.

[21] H. Shen, Z. Ding, S. Dasgupta, and C. Zhao, “Multiple source localization in wireless sensor networks based on time of arrival measurement,” *IEEE Transactions on Signal Processing*, vol. 62, pp. 1938–1949, Apr. 2014.

[22] I. Aihara, D. Kominami, Y. Hirano, and M. Murata, “Autonomous distributed control methods for wireless sensor networks based on nonlinear dynamics of frog choruses,” to appear in *Technical Report of IEICE, IN 2017-03*, Mar. 2017.

[23] Y. Kuramoto, *Chemical oscillations, waves, and turbulence*, vol. 19. Springer Science & Business Media, 2012.

[24] T. Mizumoto, I. Aihara, T. Otsuka, R. Takeda, K. Aihara, and H. G. Okuno, “Sound imaging of nocturnal animal calls in their natural habitat,” *Journal of Comparative Physiology A*, vol. 197, no. 9, pp. 915–921, 2011.

[25] D. Pelleg, A. W. Moore, *et al.*, “X-means: Extending k-means with efficient estimation of the number of clusters.,” in *ICML*, vol. 1, pp. 727–734, June 2000.