# Master's Thesis


Title


# Implementation of a spatial monitoring system
# for Japanese tree frogs
# and application of their satellite behavior
# to an LPWA network coverage method

Supervisor

Professor Masayuki Murata


Author

Yushi Hosokawa

February 5th, 2020

Department of Information Networking

Graduate School of Information Science and Technology

Osaka University

Master's Thesis


Implementation of a spatial monitoring system for Japanese tree frogs
and application of their satellite behavior to an LPWA network coverage method

Yushi Hosokawa

## Abstract

There have been many studies on biomimetics focusing on the superior functions and behaviors of living organisms. Various studies have applied swarm intelligence to network control. For that purpose, we are now focusing on the sound communication of Japanese tree frogs and applying it to network control methods. Sound communication is found in many living organisms and important strategies for increasing the survival rate of the swarm would be latent in it. It is known that Japanese tree frogs synchronize each timing of the start and stop of their calling with other frogs over a long time-scale, which forms a chorus. And over a short time-scale, they vocalize at the almost same interval and they also avoid call overlap. Moreover, in order to save energy, a male frog sometimes rests without calling near another calling male frog, which is called a satellite behavior. Satellite males begin calling when they have a chance to get females. Such a satellite behavior seems selfish for breeding, but it makes the time for the swarm to continue the chorus for making an appeal to females be extended.

It is thought that there is an important relationship between the behavior of Japanese tree frogs and their spatial distribution, but it has not been elucidated yet at this time. Thus, our research is proceeding in the following three steps: (1) measuring the calling timing and position of Japanese tree frogs, (2) constructing a mathematical model based on the measured data, and (3) applying the mathematical model to a network control. The second step is proceeded with collaborators, and this thesis deals mainly with the first and third steps.

For measuring when and where frogs interact with each other, we implemented a sound localization system with wireless devices connecting with a microphone. From an outdoor

experiment using the system, a sound position was estimated with an error of 57 cm. Regarding the application to network control, we focused on the satellite behavior of Japanese tree frog. By applying a mathematical model that can reproduce the satellite behavior to wireless sensor nodes, we propose a design method for the coverage in a low power wide area network (LPWAN) that can provide a long-term service while maintaining constant network performance. In this thesis, we evaluate the proposed method by computer simulation. Simulation results show that the proposed method can extend the time by 20.9% that at least one device can sense 60% of the observed field.

# 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 field of information communications technology [1]. As part of this interdisciplinary research progress, mathematical modeling research into biological systems has been facilitated by the development of experimental techniques and enhanced computer performance. Specifically, various studies have applied swarm intelligence; i.e., cooperative social behavior that emerges from the autonomous motion of individuals, to network control [2, 3].

For that purpose, we are now focusing on the calling behavior of Japanese tree frogs (Fig. 1), in which only the males produce successive calls to attract female frogs and advertise their territory to other males. In general, females tend to be attracted by males with large body lengths, and such males tend to make a low, large, and long call in chorus.

It is known that some kinds of frogs synchronize each timing of the start and stop of their calling with other frogs over a long time-scale, which forms a chorus. This is because if a frog does not call even when other males begin to call, the frog loses a chance to get a female. Also, a male stops calling, when the other males stop calling, to reduce its risk of being found by external enemies. As a result, frogs make a chorus synchronously. Over a short time-scale, frogs vocalize at the almost same interval and avoid call overlap by shifting the vocalization timing, which is called anti-phase synchronization. This is for appealing to the females in a chorus. Moreover, some males rest close to the calling males, which is called a satellite behavior. A satellite behavior can be seen when there are nearby more attractive individuals for females, such as length, voice quality, and high calling frequency [4]. Males with low attractiveness waits without calling near males with high attractiveness, and when a female was attracted by high attractiveness, satellite males try to get that female. Since a lot of energy is used for calling, males take a satellite behavior to court females effectively. These behaviors can be said to be selfish for breeding, but can be said to be energy saving and efficient behaviors as a swarm. From the viewpoint of a swarm, by performing synchronous chorus, it is possible to appeal to females with loud chorus, which gives females the impression of a high quality area for a breeding. In addition, while frogs perform a chorus, a certain rate of males are rested in the satellite

Figure 1: Photograph of a Japanese tree frog

behavior so that the time for the swarm to continue the chorus can be extended.

It is known that different species of frogs have a difference between the short time-scale and the long time-scale of chorus. There is a hypothesis that this difference is due to differences in the external environment in which they exist, and important strategies for increasing the survival rate of the swarm may be latent in them. Male Japanese tree frogs collectively switch between a calling state and a silent state and avoid call overlap, which is a quite interesting phenomenon. Some research results have been obtained on efforts to apply such features to information networks [5, 6].

We believe that there is an important relationship between the behavior of Japanese tree frogs and their spatial distribution, but it has not been elucidated yet at this time. Therefore, our research is proceeding in the following three steps: (1) measuring the calling timing and position of Japanese tree frogs, (2) constructing a mathematical model based on the measured data, and (3) applying the mathematical model to a network control method. The second step is proceeded with collaborators, and this thesis deals mainly with the first and third steps.

**Sound source localization**    We need to know *when* and *where* individuals interact with each other. The *when* can be obtained from recorded sounds using sound separation techniques such as independent component analysis (ICA); however, it is difficult to identify frogs in an outdoor environment because they are typically small and able to conceal themselves. For resolving the *where* problem, it is true that many sound-source localization methods have been proposed.

Sound-source localization methods use the time-difference-of-arrival (TDOA) of the sound from the sound source at multiple microphones, and use the direction-of-arrival (DOA)

of the sound from the sound source [7]. In the localization method using TDOA, a hyperbola with each microphone as a focal point is obtained as an estimated position from the coordinates of each microphone and the value of TDOA. There is a method of calculating the TDOA of a signal, such as GCC-PHAT [8], which calculates the cross-correlation of two signals. However, it is difficult to distinguish when sounds arrive from multiple sound sources simultaneously. Accurate time synchronization between microphones is required when calculating TDOA. In the localization method using DOA, a half-line from the microphone position in the direction from which the sound arrives is obtained as the estimated position. On a two-dimensional plane, an estimated position of a sound source is determined at one point by installing at least three microphones so as not to be arranged in the same straight line in localization method using TDOA and DOA. A method using a signal subspace, such as Multiple Signal Classification (MUSIC) [8], is widely used as a method for obtaining a DOA because it is robust to noise. In order to calculate the DOA, multiple time-synchronized microphones are required, and in many cases, time synchronization is realized by connecting multiple microphones in hardware. Such a device is called a microphone array, and a highly accurate DOA can be obtained by using the microphone array. Also, with the DOA estimation method represented by MUSIC, when the number of microphones in the microphone array is $M$, even if sounds arrival simultaneous from $M - 1$ sound sources, each DOA can be estimated.

There are many localization method in outdoor areas such as our field experiments [9–12]. Our typical case of experiments is shown in Figure 2. Frogs are distributed in the field, but the placement area of microphones is limited. On the other hand, existing methods assume that the sound sources are surrounded by microphones, which requires the sound-observable range of the microphones to exceed the maximum distance between the microphones. However, this assumption is typically unavailable in an outdoor setting. Moreover, because the deployable space for system equipment is very limited in the outdoor environment, it may be difficult to locate the devices in their optimal positions. In this thesis, we implement a DOA-based sound-source localization method proposed in [13]. In order to overcome the aforementioned limitations, our proposed method allocates microphones closer to each other than previous methods; e.g., on the four corners of a square with sides measuring 100 cm side length. The method then estimates locations

outside the area surrounded by the microphones. We implement a localization system with wireless devices connecting with a microphone to reduce the deployment cost. This brings advantages that time synchronization is easy and that the installation of devices can be flexibly changed. Besides, by locating the microphones near to each other, our method has three significant advantages: (1) it can accurately measure the positions of installed microphones, (2) it can capture the majority of generated sounds with all microphones, which is a key requirement of sound-source localization methods, and (3) the system equipment requires less space for deployment.



Figure 2: Photograph of a rice paddy where we did experiments

**Network coverage method inspired by Japanese tree frogs** We are currently focusing on the satellite behavior of Japanese tree frogs, which is a kind of sound communication of them, and applying a mathematical model that can reproduce the satellite behavior to wireless nodes in order to extend the lifetime of the entire wireless network. Our collaborators develop a mathematical model and in this thesis, we apply the mathematical model of Japanese tree frog to a design method of an LPWAN (low power wide area network) so that it can provide sustainable service where a gateway collects location-aware information from sensing nodes while maintaining a certain rate of coverage of the network.

It is assumed that LPWAN collects sensing data (temperature, humidity, etc.) that has a correlation with the position of the wireless node. Under this assumption, at least one node in a certain area has to transmits sensing data to the gateway and the remaining nodes in the area can sleep, which extends the service time of the network. When the accuracy of sensing data are required, it is necessary to maintain sensing coverage of an

observed area [14]. We show that by applying the satellite behavior of frogs, network lifetime can be extended while satisfying a certain degree of coverage.

The remainder of this thesis is organized as follows. In Section 2, we describe related work of sound-source localization methods. In Section 3, we describe the characteristics of the Japanese tree frog, the requirements of the localization system based on the their features, the explanation of the implementation method, and the evaluation results in simulations and outdoor experiments. In Section 4, a mathematical model expressing chorus and satellite behavior of the Japanese tree frog is explained, and we describe how our coverage design method of an LPWA network is realized, and we show the simulation results. Finally, we present our conclusions in Section 5.

## 2    Related work

Sound localization techniques enable many applications, such as robot audition, automatic meeting processing, and sound-source tracking. Most existing sound-source localization methods can be classified as time difference of arrival (TDOA)-based or DOA-based approaches [7]. Both methods assume that three or more microphones or microphone arrays can record the sound from the same source.

TDOA-based methods estimate the sound-source positions using the locations of the microphones and the time differences of the sound arrivals between all pairs of microphones [15]. This type of method requires comparatively accurate time synchronization among all microphones [16]. Comparing two sounds recorded by different microphones, we can obtain one TDOA measurement from the phase-difference between them. The possible positions of a sound source are obtained from the TDOA as two hyperbolas whose foci are on the microphone locations. The intersection of all hyperbolas obtained from all sets of microphones is the estimated position of the sound source. Errors in the timer of a sensor node as well as environmental noise mean that not all hyperbolic curves intersect at the same point. Many estimation techniques have been proposed to solve this problem [17–20].

DOA-based localization methods estimate the sound-source position using the location of microphone arrays and the angle of the signal arrival. DOA refers to the direction from which a sound travels to a microphone array. Nodes with a microphone array can estimate the DOA using methods such as the well-known multiple signal classification (MUSIC) [21]. As a microphone array consists of multiple microphones connected with one another via a hardware circuit, there is no need to synchronize the clock among all microphones, which improves the measurement accuracy. We can localize the sound-source position by finding the point where each DOA line from each microphone array intersects. As for TDOA, DOA-based methods suffer from various types of errors, resulting in an estimation error in the source position. A number of estimation techniques have been proposed to solve this problem [22–24].

The existence of multiple sound sources makes it more difficult to identify them. Given that there are multiple sound sources and a localization server that collects TDOA/DOA measurements from all microphones, so-called data association problems occur. Namely,

the localization server does not know which sound source generates which TDOA/DOA measurements. The erroneous TDOA/DOA combinations often give rise to a "ghost source" that does not actually exist. References [25, 26] discuss this problem.

## 2.1 MUSIC

This section describes MUSIC which is a DOA estimation method. MUSIC [21] is proposed as a method for frequency estimation, and has been focused on estimating the direction of arrival of a signal as one of its applications. MUSIC converts a steering vector $\boldsymbol{A}(\omega, \theta)$ representing a sound transfer characteristic of a microphone array to be used into a space spanned by eigenvectors of an observation signal called signal subspace. Here, $\omega$ represents a frequency, $\theta$ represents an angle, and it is assumed that the steering vector is in the form of a transfer function. In an environment where there is no sound source other than the observation sound source and there is no noise, the eigenvalue of the signal correlation matrix of the observation signal takes a positive value by the number of sound sources, and otherwise becomes 0. However, it is assumed that all signals are uncorrelated with each other. After the above conversion, MUSIC uses that the signal subspace corresponding to the 0 eigenvalue is orthogonal to the steering vector whose sound arrival direction matches.

Let the Fourier transform of the time domain audio signal $\boldsymbol{x}(\tau)$ observed by the microphone array be $\boldsymbol{X}(\omega)$. The number of each element is the number of microphones $M$ included in the array. Also, let $\boldsymbol{S}(\omega)$ be a complex vector representing the signal of the sound source. The number of elements of $\boldsymbol{S}$ is the number of sound sources $L$. The signal $\boldsymbol{X}$ observed by the microphone array changes in amplitude and phase from the original $\boldsymbol{S}$ when transmitting in space, and a sound source signal arriving from the direction of $\theta$ can be expressed as follows using a steering vector as follows:

$$\boldsymbol{X}(\omega) = \boldsymbol{A}(\omega, \theta)\boldsymbol{S}(\omega). \tag{1}$$

At this time, the signal correlation matrix $\boldsymbol{R}(\omega)$ is a statistic indicating the spatial property of the observed signal. This is given by

$$\boldsymbol{R}(\omega) = \boldsymbol{X}(\omega)\boldsymbol{X}^H(\omega) \tag{2}$$

where $\boldsymbol{X}^H$ represents the Hermite transpose of $\boldsymbol{X}$. Eigenvalue decomposition of $\boldsymbol{R}(\omega)$ is given by

$$\boldsymbol{R}(\omega) = \boldsymbol{E}(\omega)\boldsymbol{\Lambda}(\omega)\boldsymbol{E}^{-1}(\omega) \tag{3}$$

where $\lambda_i(\omega)$, $\boldsymbol{e}_i(\omega)(1 \leq i \leq M)$ are the eigenvalues and eigenvectors of $\boldsymbol{R}(\omega)$, respectively. $\boldsymbol{E}(\omega) = \{\boldsymbol{e}_1(\omega), \boldsymbol{e}_2(\omega), \cdots, \boldsymbol{e}_M(\omega)\}$, and $\boldsymbol{\Lambda}(\omega)$ is a diagonal matrix with a diagonal component $(i,i)$ as $\lambda_i(\omega)$. $\lambda(\omega)$ represents the energy of each sound, and if the number of sound sources is $L$, $\lambda_i(\omega)$, $\boldsymbol{e}_i(\omega)$ is , $1 \leq i \leq L$, the eigenvalue and eigenvector corresponding to the sound source $i$.

From the above, if the direction $\theta$ of the steering vector $\boldsymbol{A}(\omega,\theta)$ matches the direction of arrival from the sound source, in the case of $L+1 \leq i \leq M$, Since the eigenvalue is 0, the expression (4) is obtained.

$$|\boldsymbol{e}_i(\omega)\boldsymbol{A}(\omega,\theta)|^2 = 0 \tag{4}$$

We obtain a spatial spectrum (which called MUSIC spectrum) by summing the expression(4) in the range of $L+1 \leq i \leq M$ and multiplying by the normalization term as follows:

$$P(\omega,\theta) = \frac{|\boldsymbol{A}^H(\omega,\theta)\boldsymbol{A}(\omega,\theta)|}{\sum_{m=L+1}^{M} |\boldsymbol{A}^H(\omega,\theta)\boldsymbol{e}_m|^2}. \tag{5}$$

As described above, if the direction $\theta$ of the steering vector $\boldsymbol{A}(\omega,\theta)$ matches the direction of arrival from the sound source, the denominator is 0 due to the orthogonality of the subspace (Actually, it is not completely 0 due to noise) and $P$ is a peak. By searching the observed signal for $\theta$ of the steering vector such that $P$ has a peak, the arrival direction of the sound, the DOA can be estimated.

## 2.2 Grid-based localization method

This section describes Grid-based localization method [24] on which the implemented method is based. This method divides an area into $N$ equal-sized cells. The cell whose direction from the microphones most closely matches the estimated DOA is then identified. The localization algorithm is as follows:

1. Discretize the area of interest into $N$ cells and calculate the coordinates of the center of each cell.

2. Calculate the $(M \times N)$ matrix $\mathbf{\Psi}$ whose elements $\psi_{m,n}$ give the angle from the $m$th microphone array to the $n$th cell center ($M$ is the number of microphone arrays).

3. Define a cost function $Cost$ that represents the degree of coincidence between the true DOA and the calculated angle in (6)

$$Cost(n) = \sum_{m=1}^{M} \left[ A\left( \hat{\theta}_m, \psi_{m,n} \right) \right]^2 \tag{6}$$

where $\hat{\theta}_m$ is the DOA obtained from the $m$th sensor node.

4. Find the cell that minimizes the cost function, that is, $n^* = \arg \min Cost(n)$.

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

$$A(X,Y) = 2 \sin^{-1} \frac{|\exp(jX) - \exp(jY)|}{2}. \tag{7}$$

In this method, the resolution of the grid, which depends on the number of cells, $N$, affects the estimation accuracy. Increasing $N$ will decrease the estimation error but increase the computational cost. Therefore, the authors of [24] proposed a recursive search method for the cell that has minimum $Cost$.

This localization method can deal with multiple sources given the correct number of sound sources. To determine the positions of multiple sources, the authors of [24] used a two-step procedure. First, the set $Q$ containing all possible combinations of DOAs is calculated. Second, for each cell, $Cost$ is calculated using a combination of DOAs, denoted by $q$ ($q \in Q$). The $S$ cells that have the $s$th minimum $Cost$ are selected as the source locations ($s = 1, 2, \cdots, S$), where $S$ is the highest number of DOAs detected by all microphones.

# 3 Real-time sound source localization system using wireless sensor network in the outdoor environment

## 3.1 Characteristics of Japanese tree frog

Japanese tree frogs are about $2.0 \sim 4.5$ cm in size and inhabit paddy fields and forests. The breeding season of Japanese tree frog is the spring, and during the breeding season, from the evening to around midnight, the male makes a sound called the advertising sound, which informs the female of its presence. When one or more Japanese tree frogs start to call, they start to call (which called chorus). At that time, it is known that a small number of Japanese tree frogs, about two to four, avoid call overlap so that they do not sound simultaneously [27]. At this time, they often away 0.5 m more than each other. Japanese tree frogs do not call very much in the water or on the water, do not move while they are calling, and once they start calling, they often continue chorus for more than five minutes. Each advertisement sound of a Japanese tree frog is about 0.1 seconds to 0.2 seconds long, and its fundamental frequency is about $2,000$ Hz.

We assume that a spring-summer rice field in which Japanese tree frogs inhabit and are actively calling. One side of the rice field is about 20 m, and there are often no tall trees around the rice field, but there are many grasses. The rice fields are covered with water, the shores are muddy and there are few flat parts. In this thesis, a single rice field is targeted, and estimation is not performed over multiple rice fields, such as terraced rice fields. That is, vertical estimation is not performed, and the location where the equipment can be installed must be on the same plane as the rice field.

## 3.2 System performance requirements

The implemented system can be applied not only to observation of frogs but also to observation of various organisms performing voice communication. Here, the system is designed in consideration of the characteristics of Japanese tree frogs and their use in the habitat. As described above, the localization system is used in rice fields. At this time, there are various obstacles, weather, terrain, and other factors, and it is not always possible to guarantee that the devices are arranged as expected. When laying a large

number of devices, the preparation cost is also a problem. From the above, it is required that a system be configured with a small number of devices, that the device be installed with a certain degree of arbitrariness, and that the devices be easily adjusted.

In this thesis, the signal generated from the sound source is acquired by multiple microphones, and the position is estimated by collecting necessary information for the computer by using a wireless network. Localization in a two-dimensional plane can be realized by collecting DOAs from three devices that are not aligned on the same straight line. In addition, by using a wireless network, if devices are within wireless communication range of each other, they can communicate with each other, so that the cost of installation is reduced.

As for performance requirements, the following numerical values are targeted, taking into account the characteristics of the Japanese tree frog described in section 3.1.

- Estimation accuracy : error 50 cm or less

- Estimation time : 5 minutes or less

- Estimable area : $20 \times 20$ m$^2$

## 3.3 Implementation of real-time sound source localization system

### 3.3.1 Source number estimation in MUSIC for outdoor environments

Since the MUSIC described in Section 2.1 estimates a DOA at each frequency, it is necessary to consider the frequency characteristics of the sound source. In a practical implementation, using the fact that the eigenvalue $\lambda(\omega)$ represents the energy level of the signal at each frequency $\omega$, as shown in Eq. (8), DOA estimation is performed by summing up the MUSIC spectra weighted by the eigenvalue $(\lambda(\omega))$ obtained at each frequency.

$$P(\theta) = \sum_{\omega=\omega_{min}}^{\omega_{max}} \sum_{i=1}^{L} \lambda_i(\omega) P(\omega, \theta) \tag{8}$$

Here, $\omega_{min}$ and $\omega_{max}$ represent the lower and upper limits of the frequency for DOA estimation, respectively.

As described in Section 2.1, the MUSIC can estimate multiple DOA from a recorded sound data. To do so, the number of sound sources is required. However, it is a difficult

challenge to estimate the number of sound sources from recorded sound data. Some methods have been proposed for estimating the number of sound sources from the eigenvalue distribution of a signal correlation matrix calculated by MUSIC using Akaike's Information Criterion (AIC). AIC is used to balance model complexity with data fit. For using an AIC-based source number estimation, it is necessary to set the number of sound sources as the number of model parameters and define a likelihood function for estimating the number of sound sources from the distribution of observed eigenvalues. Here, since the likelihood generally increases when the number of model parameters (the number of sound source) becomes larger, an appropriate number of sound sources is estimated by using AIC. A general AIC formula is as follows, and the number of sound sources that minimizes the AIC is obtained.

$$AIC(M) = -2 \ln \mathcal{L}(M) + 2\eta \tag{9}$$

Here, $\mathcal{L}$ is a likelihood function, $M$ is a model, and $\eta$ is the number of parameters. In [28], the likelihood function $\mathcal{L}$ when the number of sound sources is $n$ is defined as in Eq. (10). $\mathcal{L}(n)$ becomes larger and closer to 1 as the $(n+1)$ th and subsequent eigenvalues are equal. That is, it has a high value when $n$ is equal to or greater than the true number of sound sources.

$$\mathcal{L}(n) = \frac{(\lambda_{n+1}\lambda_{n+2}\cdots\lambda_m)^{\frac{1}{m-n}}}{\dfrac{1}{m-1}(\lambda_{n+1} + \lambda_{n+2} + \cdots + \lambda_m)} \tag{10}$$

The number of sound sources is estimated by using this likelihood function. Since the number of sound sources is estimated for each frequency, the estimated value $L^{est}(\omega)$ of the number of sound sources at the frequency $\omega$ is given by the follows:

$$L^{est}(\omega) = \arg \min_{n} \left\{ -2 \ln \left( \frac{(\lambda_{n+1}\lambda_{n+2}\cdots\lambda_m)^{\frac{1}{m-n}}}{\dfrac{1}{m-1}(\lambda_{n+1} + \lambda_{n+2} + \cdots + \lambda_m)} \right) + 2n \right\}. \tag{11}$$

In the existing sound-source number estimation method, there is a possibility that the influence of noise appears strongly. With the weighting method described in Eq. (8), when

17

$L^{est} > L$ due to the influence of noise such as wind at a specific frequency, the MUSIC spectrum peaks in the direction in which no sound source exists. It is considered that the noise such as wind causes a smaller eigenvalue indicating power as compared with the advertisement sound of the Japanese tree frog. Therefore, we propose a method to estimate the number of sound source where the estimation result of the number of sound sources in each frequency is weighted by the first eigenvalue obtained in the frequency. In the current implementation, the number of sound sources $\hat{L}$ is represented by Eq. (12). We use $\hat{L}$ instead of $L$ in Eq. (8).

$$\hat{L} = \left\lceil \frac{2 \sum_{\omega=\omega_{min}}^{\omega_{max}} L^{est}(\omega)\lambda_1(\omega) + \sum_{\omega=\omega_{min}}^{\omega_{max}} \lambda_1(\omega)}{2 \sum_{\omega=\omega_{min}}^{\omega_{max}} \lambda_1(\omega)} \right\rceil \tag{12}$$

### 3.3.2 Localization method used for implementation

Here, we describe the extension of the grid-based method proposed in [24]. When we use the original grid-based localization method to estimate the positions of sound sources outside the area surrounded by the microphone arrays, a much higher grid resolution is required to avoid the estimation error. However, this involves a greater computational cost, which results in a longer calculation time. With the recursive method proposed in [24], although the calculation time can be reduced, the estimation accuracy might decrease. In our method, we first calculate a directional cost, $Cost_d$, for each direction from the center of the microphone arrays. Then, the direction that has the minimum $Cost_d$ can be obtained. The localization server calculates the $Cost$ defined by Eq. (6) for each cell whose center is close to the line running from the center of the microphone arrays according to this direction. In the following, we describe the proposed method for an example with only one sound source.

Our proposed method is divided into two steps. First, we estimate the direction in which the sound source exists and then we perform grid-based sound-source localization.

Let $\hat{\boldsymbol{\theta}}$ be an $M \times 1$ vector in which each element $\hat{\theta}_m$ is the DOA estimated by the microphone array $m$. Here, without loss of generality, we can assume that the coordinates of the center of the microphone arrays describe the origin. First, we estimate the direction

18

from the origin to the sound source. To estimate the sound-source direction, we use the sum of the angular distances between a vector from the origin to the direction $\theta_s$ and the estimated DOAs of each sensor node. This is because incorrect cells in the grid-based localization for an outside area surrounded by a microphone array often have the same direction from the origin as the true cell.

The cost of the sound-source direction for each $\theta_s$ ($0 \leq \theta_s < 2\pi$) is

$$Cost_d(\theta_s) = \sum_{m=1}^{M} \left[ A\left(\hat{\theta}_m, \theta_s\right) \right]^2 \tag{13}$$

where $A(X, Y)$ is the angular distance defined in (7).

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

$$\theta^* = \arg\min_{\theta_s}(Cost_d(\theta_s)). \tag{14}$$

For the second step, the grid-based sound-source estimation is conducted. We start by dividing the area of interest into cells with side lengths of $x$; sets of cells are denoted as $\boldsymbol{P}$. The value of $x$ affects the accuracy and computational cost of our method and is adjusted to meet the required estimation accuracy. Next, we determine the cell set $\boldsymbol{P}'$ that intersects with a vector whose starting point is the origin and whose direction is $\theta^*$ (Fig. 3). Accordingly, the computation cost of our proposed algorithm is $\mathcal{O}(\sqrt{N})$, while that of the original grid-based method is $\mathcal{O}(N)$ (or $\mathcal{O}(\log(N))$ if the above mentioned recursive approach is used). Reducing the computational cost is important for localizing multiple sound sources because, in most techniques, this requires repeating the calculation of single sound source localization.

We use the angular distance function (7) to obtain the cost in each direction. To obtain $\theta^*$ using an algorithm, we discretize $\theta_s$ by equally dividing the angle of $2\pi$ by $N_\theta$. To achieve high accuracy, we must divide $\theta_s$ finite, which increases the computational cost. Therefore, we use a recursive algorithm. First, we start with a coarse angle then obtain $\theta_1$ and $\theta_2$, which are the minimum and next-lowest $Cost_d$ values, respectively. Once $\theta_1$ and $\theta_2$ have been determined, we repeat this step in the range $\theta_1 \leq \theta_s \leq \theta_2$ (here, we assume that $\theta_1 < \theta_2$). This results in the desired direction and also reduces the search cost.
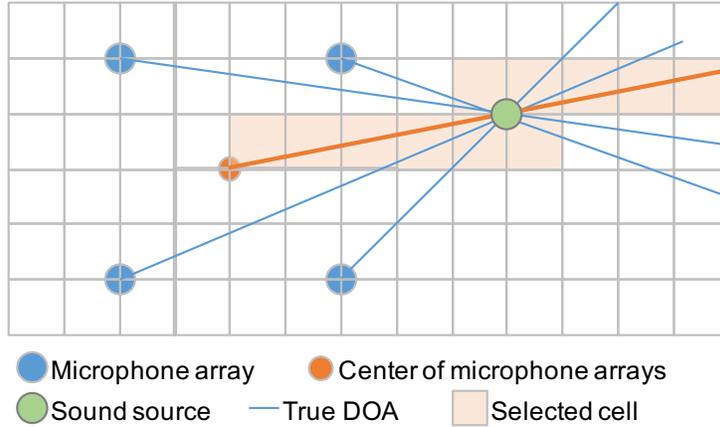
Figure 3: Schematic showing the method for reducing the number of cells to be searched

Table 1: Specification of devices

|  | Raspberry Pi 3 | Laptop PC |
|---|---|---|
| Clock frequency | 1.2 GHz, 4 core | 1.9 GHz, 2 core |
| RAM | 1 GB | 8 GB |
| OS | Raspbian stretch | Windows 7 |

### 3.3.3 Devices and implementation

First, we describe the devices used in our experiments. To obtain the DOA of the sound source, we use an 8-channel microphone array with a height of 12 cm (TAMAGO-03, System in Frontier Inc. [29]). Each TAMAGO-03 is connected to a Raspberry Pi 3 Model B with a USB cable (Fig. 4), on which we implemented the MUSIC [21] method to calculate the DOA. The TAMAGO-03 digitally converts an analog sound signal as 24-bit amplitude information at a sampling frequency of 16 kHz. The Raspberry Pi is equipped with a wireless LAN adapter (IEEE 802.11b/g/n) as standard. In the experiment, all Raspberry Pis are wirelessly connected with each other, constituting an IEEE 802.11 ad-hoc network. Sound-source localization is conducted on a laptop computer that collects DOAs from all Raspberry Pis; thus, the lap-top also belongs to the ad-hoc network. Table 1 summarizes the specifications of these devices.

By connecting the microphone arrays with each other by wireless communication, it is

Figure 4: Raspberry Pi 3 Model B with an 8-channel microphone array (TAMAGO-03)

easy to place and carry the devices. Localization is conducted according to the following steps.

1. Time synchronization of Raspberry Pis and the laptop PC is performed by using the network time protocol (ntp) via wireless communication.

2. Each Raspberry Pi records 8-ch sound data received from a connected microphone array for $T$ s.

3. Each Raspberry Pi divides the sound data into $\Delta$ s and estimates a DOA for $\Delta$-second sound data.

4. Each Raspberry Pi transmits the estimated DOAs to the laptop PC.

5. The laptop PC conducts the proposed grid-based localization method utilizing the received DOAs.

All programs for estimating the DOA and sound-source position are written in C++ language. As mentioned above, DOAs are estimated using the MUSIC method. We set $\Delta$ to 0.5 so that it is long enough to record the bout length of a Japanese tree frog of about 0.2 s. For avoiding the influence of a temporal noise, we set $T$ to 30 and each Raspberry Pi calculates the mode of generated DOA estimates. Note that if $\Delta$-s sound data has a

Table 2: Simulation parameters

| Parameters | Value | Description |
|---|---|---|
| $N$ | $100 \times 100$<br>$1000 \times 1000$ | Number of cells |
| $M$ | 4 | Number of microphone arrays |
| $N_\theta$ | $360/0.05$ | $\theta_s$ resolution |

very low sound pressure level, the Raspberry Pi ignores the data and does not conduct DOA estimation. We used the squared amplitude of the recorded sound as the threshold.

To estimate a DOA by the MUSIC method, an array manifold matrix is required, which most closely fits the signal subspace of a microphone array. The array manifold matrix of TAMAGO-03 is provided by *HARK open source robot audition software* [30]. The DOA estimation resolution is 5° when using the original array manifold matrix obtained from [30]. We use an interpolation method proposed in [31], which can interpolate the array manifold matrix to any degree in the time domain and frequency domain. According to this interpolation method, the Raspberry Pi estimates the DOA to an accuracy of 1°. Note that a higher interpolation resolution increases the size of the array manifold matrix file.

## 3.4 Numerical analysis of localization method used for implementation

### 3.4.1 Simulation settings

In this section, we evaluate the estimation accuracy of the proposed method by comparing it with the original grid-based method with and without a recursive approach using a computer simulation for clarifying the characteristics of our method. The estimation accuracy is defined as the localization error that reflects the distance between the true and estimated positions of a sound source. In the simulation, we also consider the case where a DOA error occurs. We summarize the simulation parameters of the evaluation in Table 2.

In the simulation, the observation area is an $A \times A$ square and the corners of the area are assigned the coordinates $(0, 0)$, $(0, A)$, $(A, A)$, and $(A, 0)$. Here, we set $A$ as 10 m.

This area is divided into $N$ square cells; that is, each cell is a square with sides measuring $A/\sqrt{N}$. A sound source is randomly placed in the observation area according to a uniform distribution. Microphone arrays are placed at $(-1, -1)$, $(-1, 0)$, $(0, 0)$, and $(0, -1)$. In the grid-based method with a recursive approach, the observation area was divided into $2 \times 2$ square cells, and the search was performed recursively until the side length of the cell became less than $A/\sqrt{N}$.

We assume that the DOA error follows the same uniform distribution regardless of the distance between the microphone array and the sound source when the microphone array can obtain a sufficient SNR. This assumption is based on our actual outdoor measurements. Note that, in the DOA estimation method, an estimated DOA is chosen from predefined discrete angles [21]. Therefore, we assume that a DOA error of $m$, denoted by $e_m^{DOA}$, follows the discrete uniform distribution whose probability density function $P(e_m^{DOA} = k)$ is $1/(e_{DOA} + 1)$, where $k = 0, 1, \cdots, e_{DOA}$.

### 3.4.2 Simulation Results

First, we present the estimation accuracy of our proposed method without DOA errors in Fig. 5 and Table 3. In the figure, we show the cumulative distribution function (CDF) of the estimation error in the proposed method (red line), the original grid-based method (blue line) and the grid-based method with a recursive approach (green line) when $N = 100 \times 100$. Since there is no significant difference between the results of $N = 100 \times 100$ and $N = 1000 \times 1000$, we only show the figure of the former result.

The localization accuracy of the proposed method was the same as that of the original grid-based method for both average and maximum error, and it was shown that the localization can be performed with a higher accuracy than the grid-based method with a recursive approach. In the grid-based method with a recursive approach, when the size of grid division is rough, the sound source does not necessarily belong to the cell with the minimum $Cost$. Thus, although the computation cost is smaller than that of the original one, the localization accuracy becomes lower.

Next, we evaluate the localization error by considering DOA errors. We set $e_{DOA}$ to 1 and 2.

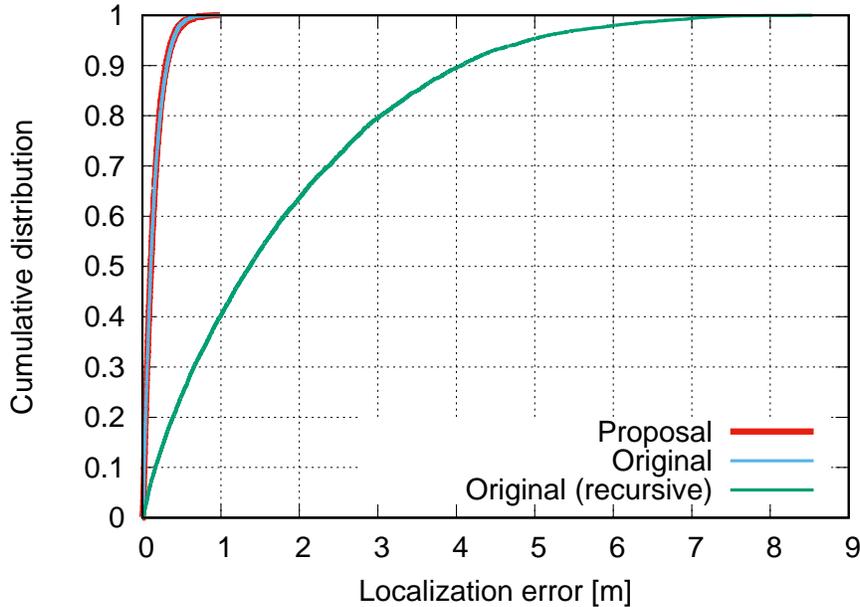Figure 7 and Table 4 show the results of simulating DOA errors. Our proposal showed

Figure 5: Localization-error distribution without DOA errors

almost the same accuracy as that of the original grid-based method. And both of the methods are superior to the recursive method in terms of the localization accuracy. Note that in all methods, the localization error increases when the DOA error was given, but the proposed method showed slightly better performance than that of the original grid-based method. As shown in Fig. 6, the estimation error increases near the edge of the observation area, which suggests that a localization system that utilizes our proposed method should be installed near the target of localization.

As shown in Section 3, the computation cost of our proposed algorithm, the original grid-based method, and the grid-based method with a recursive approach are $\mathcal{O}(\sqrt{N})$, $\mathcal{O}(N)$, and $\mathcal{O}(\log(N))$, respectively. Here, the calculation time on the Laptop PC used for the localization of one sound source is evaluated. Table 5 shows the average calculation time of these methods when changing the value of $N$. As shown in Table 5, when $N = 10,000$, there is almost no difference in the calculation time among the three methods, but when $N = 1,000,000$, the original method takes about 0.4 s for localization. In the case of multiple sound-source localization, as the number of DOA combinations increases, the calculation time for localizing them increases. Then, it can be said that our proposed

Table 3: Localization error without a DOA error

| N | Error (m) | |
| | Average | Max |
|---|---|---|
| Original | $100 \times 100$ | 0.142 | 1.007 |
| | $1000 \times 1000$ | 0.092 | 0.652 |
| Original (recursive) | $100 \times 100$ | 1.790 | 8.538 |
| | $1000 \times 1000$ | 1.763 | 8.330 |
| Proposed | $100 \times 100$ | 0.141 | 0.987 |
| | $1000 \times 1000$ | 0.093 | 0.636 |

method is more advantageous than the original from the viewpoint of localization accuracy and calculation time.

## 3.5 Outdoor experiment and results

### 3.5.1 DOA estimation experiment for multiple sound sources

We conducted outdoor experiments to verify whether the correct DOA value could be estimated by the MUSIC method when there were multiple sound sources. In our experiments, there is no obstacles near the sound source or devices. If two sound sources are extremely close to each other, the two sound sources can not be able to distinguished by the MUSIC. Therefore, we placed a microphone array and two loud speakers (for sound sources) as shown in Fig. 8. The two loud speakers played back recorded sounds of different frogs. Spectrograms of the frogs' calling are as shown in Fig. 9. It can be seen that the frequency bands with high sound energy are common, but have different characteristics.

Figure 10 shows the MUSIC spectrum calculated by the MUSIC described in Section 3.3.1. The MUSIC spectrum have two peaks only in the direction where the sound source exists.

### 3.5.2 Localization experiment in outdoor experiment

In order to evaluate the accuracy of our localization system, we conducted localization experiments in an outdoor area with no obstacles near the sound source or devices. The
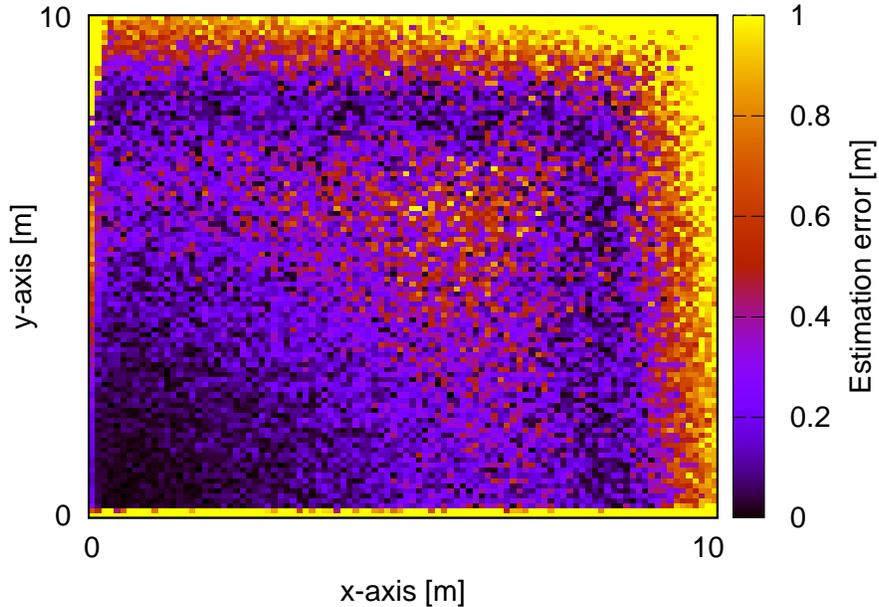
Figure 6: Heat map showing the spatial error distribution of the proposed method

true positions of the sound source had to be obtained in advance; however, it is difficult to measure their accurate positions in an outdoor environment. Therefore, we used a laser distance meter with an error of approximately 1 mm (Leica DISTO D210 [32]) and calculated the positions of the sound source by triangulation. It was also necessary to calibrate the direction of the microphone arrays in advance, using several localization results. According to the results from section 3.4, localization accuracy is higher when the sound source is located near to the microphone arrays; therefore, for the calibration, we installed the sound source close to the microphone arrays.

We installed microphone arrays as shown in Fig. 11. For the sound source, we used a loud speaker that replayed the advertisement calls of a Japanese tree frog. The maximum sound pressure level of the replay was approximately about 80 dB. The localization parameters were the same as those in the simulation shown in Table 2.

We show the results in Table 6. The average, maximum, and minimum values of the localization error are 0.57 m, 1.28 m, and 0.16 m, respectively. Note that when we can obtain the true DOA estimates, these values are 0.13 m, 0.38 m, and 0.06 m, respectively. Regarding computational time, DOA estimation takes approximately 0.07 s for a 0.5-s
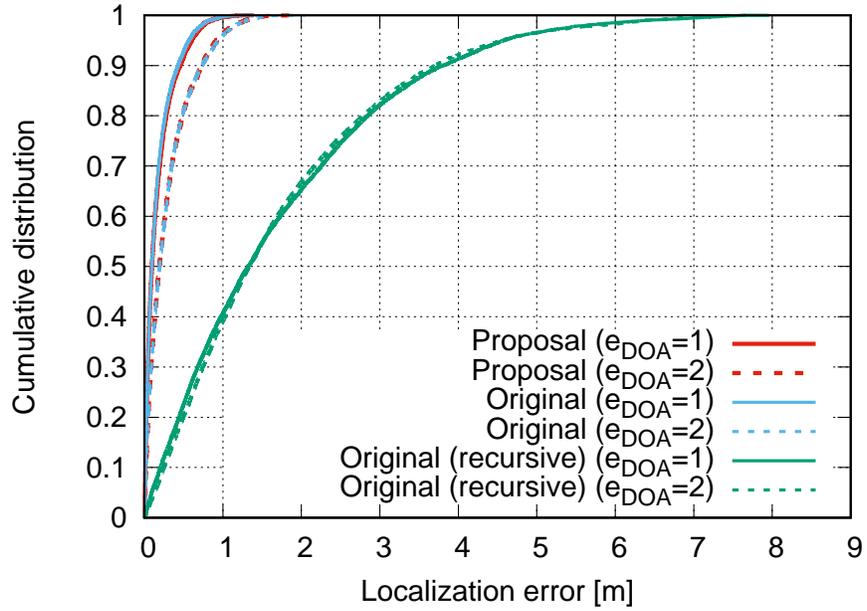
Figure 7: Localization-error distribution including DOA errors

8-ch sound data and location estimation takes approximately 0.1 s.

In the experiment, the estimated DOAs include an average error of 1.8° and a maximum error of 4°. These errors are caused by various factors, such as sound reverberation, the position error of the microphone arrays, and the sound source. Specifically, the Raspberry Pi connected to the microphone array has a strong influence on the DOA estimation error, likely due to sound reflection. For more accurate localization, increasing the number of microphone arrays is a simple and robust solution. This is easily achieved because they are connected by wireless communication.

Table 4: Localization error including a DOA error

| | $N$ | $e_{DOA}$ | Average error (m) | RMSE |
|---|---|---|---|---|
| Original | $100 \times 100$ | 1 | 0.163 | 0.404 |
| | | 2 | 0.293 | 0.541 |
| | $1000 \times 1000$ | 1 | 0.252 | 0.71 |
| | | 2 | 0.399 | 0.632 |
| Original (recursive) | $100 \times 100$ | 1 | 1.716 | 1.310 |
| | | 2 | 1.715 | 1.309 |
| | $1000 \times 1000$ | 1 | 1.687 | 1.299 |
| | | 2 | 1.670 | 1.292 |
| Proposal | $100 \times 100$ | 1 | 0.168 | 0.410 |
| | | 2 | 0.291 | 0.539 |
| | $1000 \times 1000$ | 1 | 0.157 | 0.397 |
| | | 2 | 0.288 | 0.537 |

Table 5: Calculation time

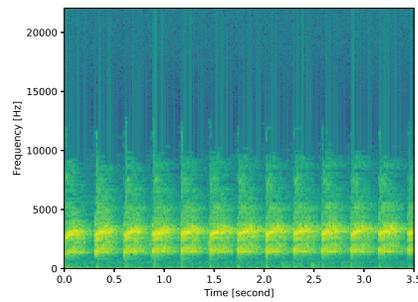| | $N = 100^2$ | $N = 500^2$ | $N = 1000^2$ |
|---|---|---|---|
| Original | 0.039 s | 0.124 s | 0.379 s |
| Original (recursive) | 0.038 s | 0.036 s | 0.038 s |
| Proposal | 0.038 s | 0.037 s | 0.038 s |
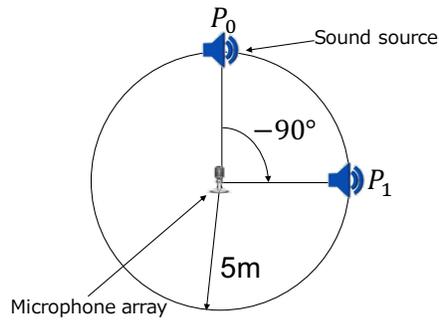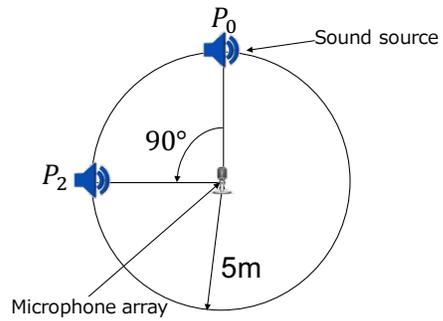
(a) Experiment 1        (b) Experiment 2

Figure 8: Position of sound sources and a microphone array in DOA estimation experiments
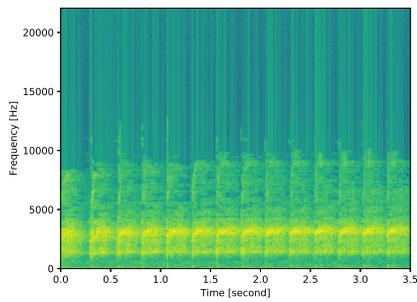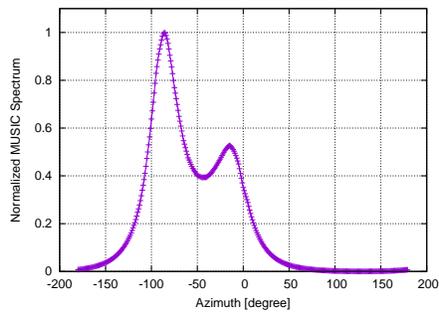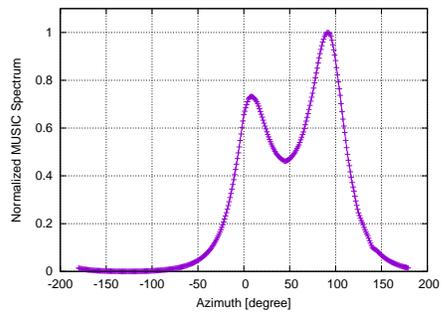


(a)        (b)

Figure 9: Spectrograms of sound sources



(a) When we placed sound sources as in experiment 1

(b) When we placed sound sources as in experiment 1
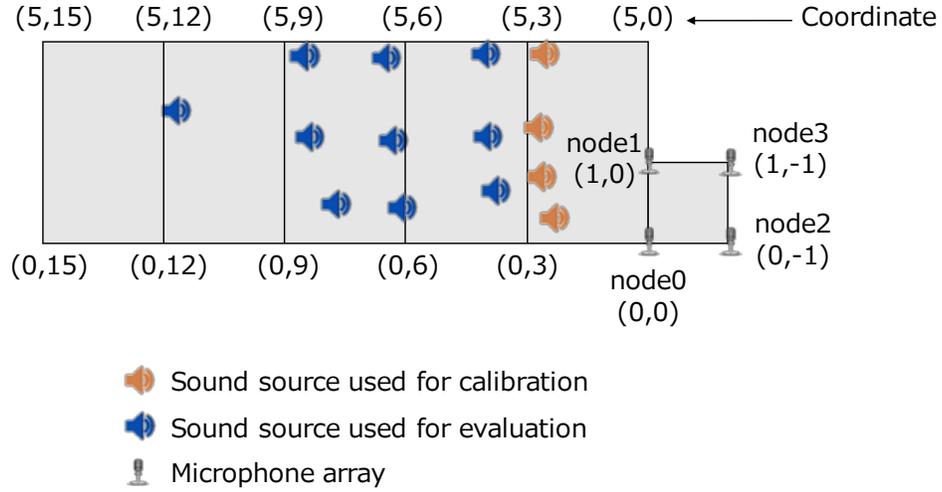
Figure 10: MUSIC spectrums

Figure 11: Position of sound sources and microphone arrays in localization experiment

Table 6: Result of a localization experiment

| Position of sound source | | Estimated result | | Localization error (m) |
|---|---|---|---|---|
| x | y | x | y | |
| 0.31 | 4.01 | 0.425 | 4.125 | 0.16 |
| 2.34 | 3.98 | 2.475 | 4.875 | 0.91 |
| 3.71 | 3.76 | 3.325 | 3.675 | 0.40 |
| 0.40 | 6.48 | 0.575 | 6.075 | 0.45 |
| 2.46 | 6.30 | 2.325 | 6.225 | 0.15 |
| 4.13 | 6.08 | 3.925 | 6.225 | 0.25 |
| 0.36 | 8.50 | 0.525 | 8.325 | 0.24 |
| 2.37 | 8.38 | 2.125 | 7.475 | 0.98 |
| 4.05 | 7.73 | 4.125 | 8.625 | 0.90 |
| 1.72 | 11.70 | 1.775 | 10.425 | 1.28 |

# 4 Efficient LPWA network coverage method

## 4.1 Mathematical model of chorus and satellite behavior of Japanese tree frogs

Temporally not available until the corresponding part will be published

$$\tag{15}$$

$$\tag{16}$$

$$\tag{17}$$

$$\tag{18}$$

$$\tag{19}$$

$$\tag{20}$$

$$\tag{21}$$

$$\tag{22}$$

$$\tag{23}$$

$$\tag{24}$$

$$\tag{25}$$

$$\tag{26}$$

$$\tag{27}$$

$$\tag{28}$$

$$\tag{29}$$

$$\tag{30}$$

Table 7: Parameter settings

$$(31)$$

$$(32)$$

$$(33)$$

Figure 12: Change in fatigue of three frogs (1,500–3,000 step)

Figure 13: Change in energy of three frogs (1,500–3,000 step)

## 4.2　Sleep control of LoRaWAN end-devices based on satellite behavior

In the following, we describe the sleep control method of LoRaWAN end device based on the mathematical model described in section 4.1. First, we describe how frogs chorus and satellite behavior are applied to end device control.

**State($s_n$)**　LoRaWAN end devices have three communication modes: ClassA, ClassB, and ClassC. Class A devices are only received when an ACK is received from the gateway. Class B devices can receive beacons that the gateway sends periodically, while class C devices can always receive and never sleep. End devices can consume less battery by properly switching between these three modes depending on whether reception is required. The state in which sensing data is being transmitted periodically, the state in which the sensor is in standby mode and senses periodically, and the state in which it is in sleep are called a *sending state*, a *standby state* and a *sleep state*, respectively. If the state of the $n$ th end device is represented as $s_n$, $s_n = 0$ corresponds to the sending state, $s_n = 1$ corresponds to the standby state, and $s_n = 2$ corresponds to the sleep state. $s_n = 0$ corresponds to ClassC. The mathematical model in Section 4.1 allows them to always know the calling of other individuals without consuming battery, but requires more battery to perform the same action on the end devices. Therefore, if $s_n = 1, 2$, the end devices will not receive initially in class A, but after a certain time the switch to class C will. Devices transit between $s_n = 0$ and $s_n = 1$ based on the Eqs. (21) and (21). In addition, the transition probability between $s_n = 1$ and $s_n = 2$ is calculated as in Eq. (34) to Eq. (38). Devices switch mode from Class A to Class C causes after the elapse of $\alpha \times \Delta T_{inter}$ time. Where $\Delta T_{inter}$ is the average duty cycle. In Class C, devices release the receiving slot for $\beta \times 2\pi$, and if no state transition occurs, switch to Class A.

Figure 14: Change in energy of three frogs

Figure 15: State transition of three frogs (1,500–3,000 step)

Figure 16: Ratio of frogs in each state

**Data transmission timing($\theta_n$)**   When $s_n = 0$, the end device transmits sensing data to the gateway. The data transmission timing is determined based on $\theta_n$ which is the phase. A end device transmits sensing data at the timing of $\theta_n = 0$, and when receives that another device is transmitting sensing data, the phase is updated based on the phase difference from itself. At the time of transmitting the sensing data, we add the remaining battery level to the header.

**Duty Cycle($T_n$)**   We can determine a duty cycle for determining an interval for collecting sensing data by adjusting $T_n$ according to a required value.

**Battery($E_n$)**   We express the battery capacity (mAh) of the end device as $E_n$. Here, we ignore the battery consumption in the sensing, and the power consumption only in the standby mode, the receive mode, and the transmit mode in Table 8 is considered. Table 8 is excerpted from the datasheet of Semtech's LoRa communication module SX1276 [33]. The change of $E_n$ for each class is expressed by the Eqs. (39) and (40). we determine the value of $\eta, \iota$ and, $\kappa$ with reference to Table 8.

$$P_n^{standby \to sleep} = I_3(E_n) \tag{34}$$

$$P_n^{sleep \to standby} = I_4(E_n) \tag{35}$$

In Section 4.1, we use the number of continuous calling to calculate the transition probability to the satellite state. The end device can directly compare the remaining battery power, so if the battery is lower than the surrounding device, the transition probability to the sleep state is increased, and if the battery is higher than the surrounding device, we determine transition probabilities as shown in Eqs. (37) and (38) so as to make a transition to the standby state.

Table 8: Power consumption in LoRa module SX1276 (DC=3.3V)

| Mode | Configuration | Value |
|---|---|---|
| sleep | | 0.2(uA) |
| standby | | 1.6(mA) |
| receive mode | bands 2&3 | 12.0(mA) |
| | band 1, LNA boost ON | 11.5(mA) |
| transmit mode | TX power=13dBm | 29(mA) |
| | TX power=7dBm | 20(mA) |

$$E_n^{sleep} = \tilde{\vec{E}} \tag{36}$$

$$I_3(E_n) = \frac{1}{\exp(-\epsilon(E_n^{sleep} - E_n - \delta)) + 1} \tag{37}$$

$$I_4(E_n) = \frac{1}{\exp(\epsilon(E_n^{sleep} - E_n + \delta)) + 1} \tag{38}$$

Here, $E_n^{sleep}$ indicates the remaining battery level that determines whether or not to transition to the sleep state, and $\vec{E}$ is the battery level of end device that was observed within $r_1$ from itself.

- ClassA

$$\begin{cases} \dfrac{dE_n}{dt} = 0 & \text{if } s_n = 2A, \\\\ \dfrac{dE_n}{dt} = -\eta & \text{if } s_n = 1A. \end{cases} \tag{39}$$

- ClassC

$$\begin{cases} \dfrac{dE_n}{dt} = -\iota & \text{if } \delta(\theta_n) = 1, \\\\ \dfrac{dE_n}{dt} = -\kappa & otherwise. \end{cases} \tag{40}$$

## 4.3 Simulation evaluation

In this section, we show that by applying satellite behavior to LoRaWAN end devices, it is possible to extend the service life by performing sleep while maintaining constant coverage by simulation. In order to show the effectiveness of satellite behavior, we compared the sleep control method described in Section 4.2 with the method that does not transition to the sleep state. LoRaWAN is one of the LPWA standards, and enables long-distance communication with the same transmission power as conventional wireless communication. According to the communication range of LoRaWAN, we install 100 end devices randomly in the observation area of 1 km × 1 km. The coverage is calculated on the intersections of the grid when the observation area is divided into squares of 20 m × 20 m. In this paper, coverage is defined as the ratio of intersections sensed by one or more sensors to the total number of intersections. Table 9 shows the parameter settings. End device can sense within a circle of radius $r_2$ around the end device.

We show changes in coverage and battery in Figs. 17 and 18. By performing the sleep control inspired by the satellite behavior, it was shown that the sleep was performed without decreasing the coverage and the sensing time could be extended. Also, the ratio of numbers of the end device in each state with respect to the total number of end devices when the sleep reference in the Eq. (36) is set to the maximum value or the median value of battery of which end devices at a distance of $r_1$ or less is shown in Fig. 19. We also show that by changing the criterion for entering the sleep state, it is possible to adjust according to the required life and coverage.

Table 9: Parameter settings on LoRaWAN end device

| Parameter | Value | Description |
|:---:|:---:|:---:|
| $r_0$ | 200 | Maximum distance activated by data transmission from another terminal |
| $r_1$ | 100 | Maximum distance to the individual that can be a reference in transition to sleep |
| $r_2$ | 150 | Maximum sensing distance |
| $\alpha$ | 0.39 | Recovery rate of fatigue |
| $T_{max}$ | 100.12 | Maximum value of fatigue |
| $\Delta T$ | $T_{max}/1.2$ | Parameter to determine inflection point in Eq. (25) and Eq. (27) |
| $\gamma$ | 0.5 | Steepness of logistic functions (Eq. (25) and Eq. (27)) |
| $E_{max}$ | 500 | Battery capacity (mAh) |
| $p_{high}$ | 0.8 | Parameter in Eq. (28) |
| $p_{low}$ | 0.01 | Parameter in Eq. (28) |
| $\epsilon$ | 0.5 | Steepness of logistic functions (Eq. (31) and Eq. (32)) |
| $\eta$ | 1.6 | Power consumption in standby mode |
| $\iota$ | 29.0 | Power consumption in transmit mode |
| $\kappa$ | 12.0 | Power consumption in transmit mode |
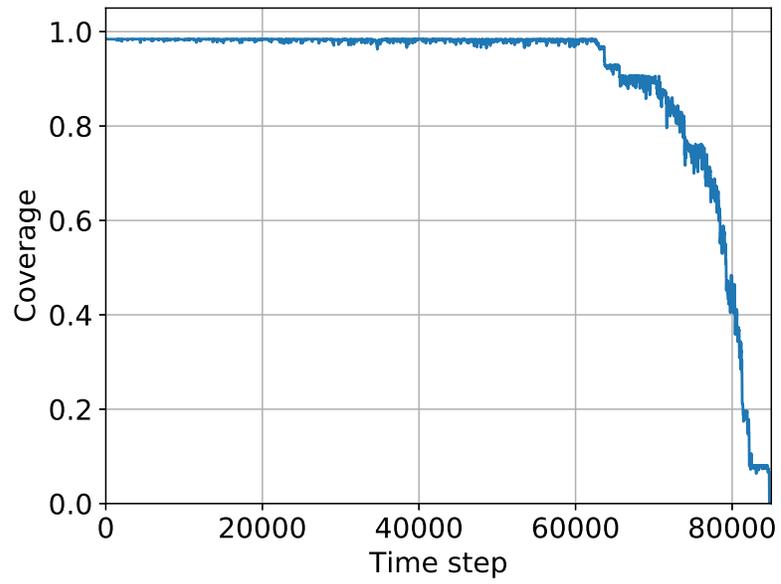
(a) Proposed method
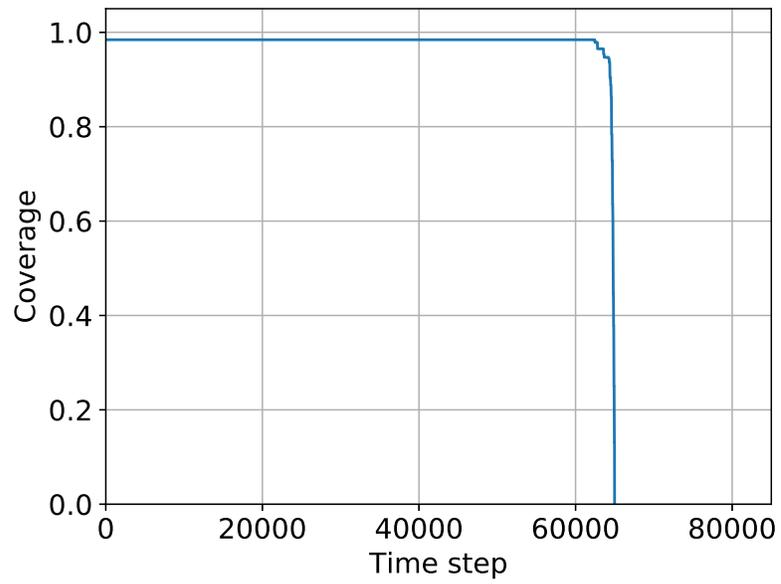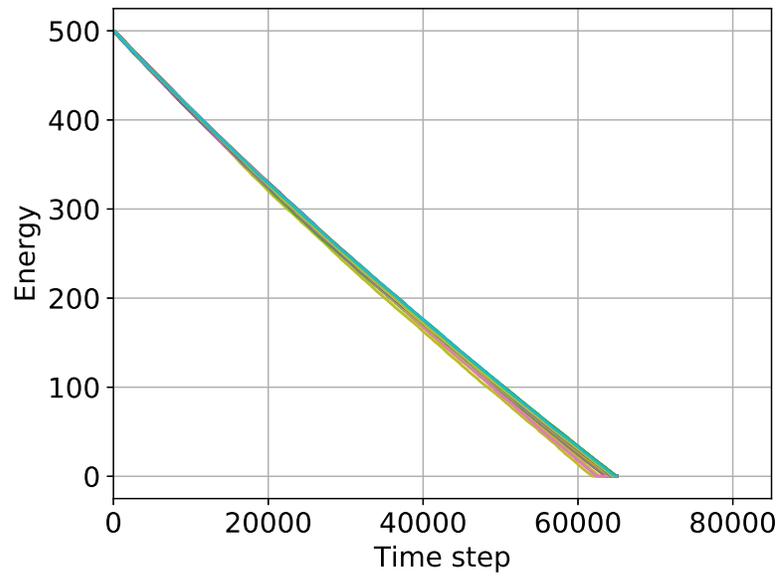


(b) None satellite method

Figure 17: Change in coverage $(E_n^{sleep} = \max\vec{E})$
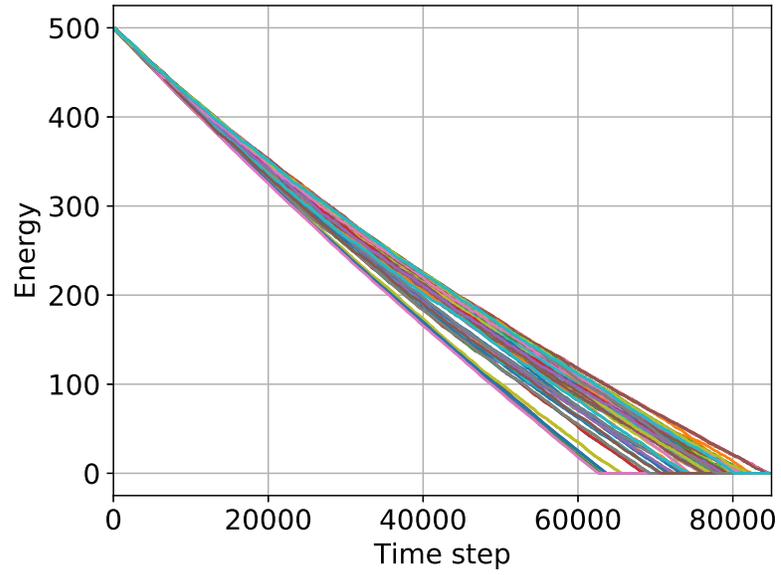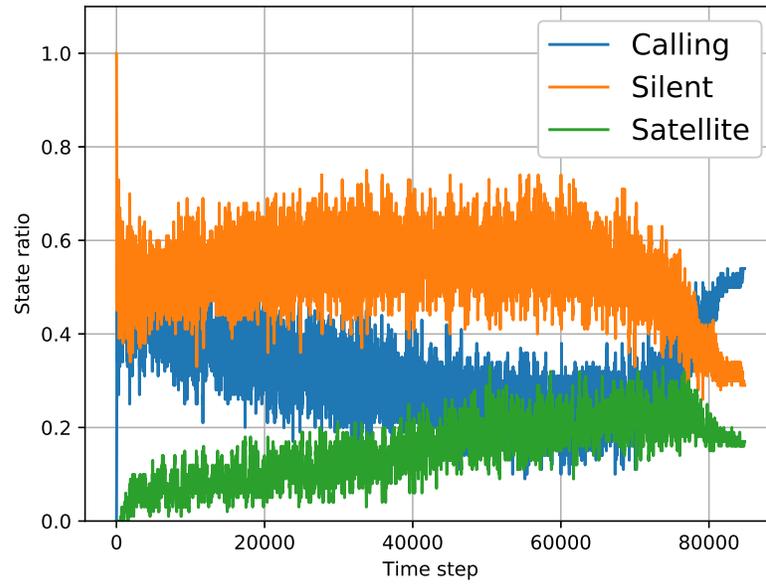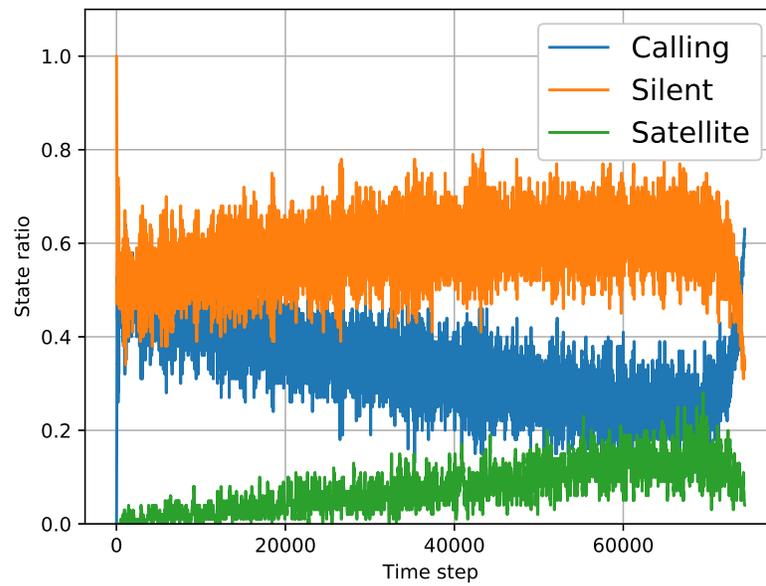
(a) Proposed method



(b) None satellite method

Figure 18: Change in battery ($E_n^{sleep} = \max\vec{E}$)

(a) $E_n^{sleep} = \max \vec{E}$)



(b) $E_n^{sleep} = \mathrm{median} \vec{E}$)

Figure 19: Ratio of end devices in each state

# 5 Conclusion

In this thesis, we implemented a sound-source localization method using a wireless microphone-array network for the outdoor environment. Simulation results showed that the proposed method can estimate the position of a sound source with an average error of 0.29 m for a 10 m×10 m area when errors related to the DOA estimation were considered. In the experiments, the average localization error of our proposed method was 0.57 m. We also show that the number of sound sources can be estimated by eigenvalue distribution and DOA estimation for multiple sound sources can be performed in outdoor environment.

We also applied a mathematical model representing Japanese tree frog chorus and its satellite behavior to an LPWA coverage method. Simulation results show that the lifetime of an LPWAN can be extended while maintaining a certain coverage.

Our future work is to propose and implement a localization method for the DOA combination problem that occurs with multiple sound sources, and to conduct experiments in a real environment. In addition, for the application of the mathematical model into an LPWAN, it is necessary to evaluate collision avoidance due to anti-phase synchronization of frogs over a short time scale, and to compare with existing methods.

# Acknowledgments

# References

[1] S. Barbarossa and G. Scutari, "Bio-inspired sensor network design," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 26–35, May 2007.

[2] F. Dressler and O. B. Akan, "A survey on bio-inspired networking," *Computer Networks*, vol. 54, no. 6, pp. 881 – 900, Oct. 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389128610000241

[3] Z. Zhang, K. Long, J. Wang, and F. Dressler, "On swarm intelligence inspired self-organized networking: Its bionic mechanisms, designing principles and optimization approaches," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 513–537, 1st Quarter 2014.

[4] J. R. Lucas, R. D. Howard, and J. G. Palmer, "Callers and satellites: Chorus behaviour in anurans as a stochastic dynamic game," *Animal Behaviour*, vol. 51, no. 3, pp. 501–518, Mar. 1996.

[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, no. 6, pp. 661–669, Mar. 2012.

[6] I. Aihara, D. Kominami, Y. Hirano, and M. Murata, "Mathematical modelling and application of frog choruses as an autonomous distributed communication system," *Royal Society Open Science*, vol. 6, no. 1, pp. 1–16, Jan. 2019.

[7] M. Cobos, F. Antonacci, A. Alexandridis, A. Mouchtaris, and B. Lee, "A survey of sound source localization methods in wireless acoustic sensor networks," *Wireless Communications and Mobile Computing*, vol. 2017, pp. 1–24, Aug. 2017.

[8] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 4, pp. 320–327, Aug. 1976.

[9] L. M. Kaplan, Q. Le, and N. Molnar, "Maximum likelihood methods for bearings-only target localization," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, May 2001, pp. 3001–3004.

[10] Y. Fu and Z. Tian, "Cramer–rao bounds for hybrid TOA/DOA-based location estimation in sensor networks," *IEEE signal processing letters*, vol. 16, no. 8, pp. 655–658, 2009.

[11] C.-E. Chen, A. M. Ali, and H. Wang, "Design and testing of robust acoustic arrays for localization and enhancement of several bird sources," in *Proceedings of the 5th international conference on Information processing in sensor networks*, 2006, pp. 268–275.

[12] T. C. Collier, A. N. Kirschel, and C. E. Taylor, "Acoustic localization of antbirds in a mexican rainforest using a wireless sensor network," *The Journal of the Acoustical Society of America*, vol. 128, no. 1, pp. 182–189, 2010.

[13] Y. Hosokawa, Y. Hirano, D. Kominami, I. Aihara, and M. Murata, "Implementation of a real-time sound source localization method for outdoor animal detection using wireless sensor networks," in *Proceedings of International Conference on Signal Processing and Communication Systems (ICSPCS)*, Dec. 2019, pp. 1–6.

[14] I. Khoufi, P. Minet, A. Laouiti, and S. Mahfoudh, "Survey of deployment algorithms in wireless sensor networks: coverage and connectivity issues and challenges," 2017.

[15] F. Gustafsson and F. Gunnarsson, "Positioning using time-difference of arrival measurements," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2003, pp. 553–556.

[16] 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 international conference on Embedded networked sensor systems*, Nov. 2004, pp. 1–12.

[17] 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, no. 2, pp. 439–443, Feb. 2013.

[18] 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, no. 2, pp. 337–359, Apr. 2014.

[19] T. Ajdler, I. Kozintsev, R. Lienhart, and M. Vetterli, "Acoustic source localization in distributed sensor networks," in *Proceedings of Asilomar Conference on Signals, Systems and Computers (ACSSC)*, vol. 2, Nov. 2004, pp. 1328–1332.

[20] F. Meyer, A. Tesei, and M. Z. Win, "Localization of multiple sources using time-difference of arrival measurements," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2017, pp. 3151–3155.

[21] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276 – 280, Mar. 1986.

[22] J. C. Chen, K. Yao, and R. E. Hudson, "Source localization and beamforming," *IEEE Signal Processing Magazine*, vol. 19, no. 2, pp. 30–39, Aug. 2002.

[23] 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 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 2, Dec. 2003, pp. 1228–1233.

[24] A. Griffin, A. Alexandridis, D. Pavlidi, and A. Mouchtaris, "Real-time localization of multiple audio sources in a wireless acoustic sensor network," in *Proceedings of European Signal Processing Conference (EUSIPCO)*. IEEE, Sep. 2014, pp. 306–310.

[25] 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, no. 1, pp. 44–53, Dec. 2005.

[26] 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, no. 8, pp. 1938–1949, Apr. 2014.

[27] I. Aihara, T. Mizumoto, T. Otsuka, H. Awano, K. Nagira, H. G. Okuno, and K. Aihara, "Spatio-temporal dynamics in collective frog choruses examined by mathematical modeling and field observations," *Scientific Reports*, vol. 4, pp. 1–8, Jan. 2014.

[28] H. C. Z. H. Wei Cheng, Zhousuo Zhang and G. Zhu, "A comparative study of information-based source number estimation methods and experimental validations on mechanical systems," *Sensors 2014*, 2014.

[29] System in Frontier inc., http://www.sifi.co.jp/en/, available at Jan. 2020.

[30] "HARK SaaS," https://api.hark.jp/web/login, available at Jan. 2020.

[31] K. Nakamura, K. Nakadai, and H. G. Okuno, "A real-time super-resolution robot audition system that improves the robustness of simultaneous speech recognition," *Advanced Robotics*, vol. 27, no. 12, pp. 933–945, May 2013.

[32] Leica Geosystems, https://lasers.leica-geosystems.com/, available at Jan. 2020.

[33] "SX1276-7-8-9 Datasheet," https://www.semtech.com/products/wireless-rf/lora-transceivers/sx1276, available at Jan. 2020.