Multimedia Tools and Applications
https://doi.org/10.1007/s11042-023-14741-2

Real-time TDOA-based stationary sound source direction finding

Zahra Heydari[1,2] · **Aminollah Mahabadi**[1,2,3]

Received: 25 August 2021 / Revised: 28 April 2022 / Accepted: 4 February 2023

Abstract
Performance improvement of sound source direction finding is a critical acoustic task in target localization, tracking, and navigation as an NP problem that suffers from sound reflection and noise in passive or active methods. The accuracy of prediction increases by integrating more information about the signal specification, source position, sensor attributes, microphone array topology, characteristics of hardware architecture, dimension of distance, the status of environmental sounds, climatic conditions, effects of signal propagation, properties of barriers, and setting of initial estimation which is time-consuming. This paper presents a scalable method of sound source direction finding based on the time-difference-of-arrival approach to improve the accuracy of predictions of three-dimensional space in an outdoor environment. The real-time passive process is robust to sound reflection and ambient noise that decreases the running-time and increases the accuracy significantly through complexity reduction by proposing a novel Krill Herd algorithm based on successive unconstrained minimization technique. Experimental results of different actual and simulated datasets show the angle error amounts of Azimuth and Elevation are 0.7164° and 1.5054° in the near-field, and are 1.0260° and 0.2071° in the far-field, respectively. Performance evaluation of the passive proposed process shows that higher accuracy can be reached by using more parallel distributed sensor arrays.

Keywords Acoustic · Sound source direction finding · Time difference of arrival · Krill Herd algorithm · Successive unconstrained minimization · Real-time cyber-physical system · Three-dimensional space

1 Introduction

The human ear is capable of hearing and isolating sounds for positioning sources to *direction-finding*, *localization*, *tracking*, and *navigation*. Sound direction finding (SDF) is

✉ Aminollah Mahabadi
mahabadi@shahed.ac.ir; am@ipm.ir

[1] Computer Engineering Department, Shahed University, Tehran, Iran

[2] Acoustic Research Center, Shahed University, Tehran, Iran

[3] School of Computer Science, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

one of the essential characteristics found in humans and other creatures equipped with this tool through the sense of hearing [13]. Like the human ear, real-time embedded system designers try to build an automatic *machine hearing*, have the same ability to *hear* and *process* input sounds for robust *recognition* of sound sources direction to *localize* and *track* the positions [2]. The trend of *machine hearing* equipment designed in a minimum of two microphones includes critical tasks such as sound signals *sensing*, data *acquisition*, data *validation*, signal *processing*, accurate *direction-finding*, speed *localizing*, and safe *tracking* the positions [14]. Considering the wide variety of input sounds, the analysis and complete inference of the system hearing, is very complex and suffers from high computational *complexity*, sound *reflection*, presence of ambient *noise*, and artificially generated noise to produce *accurate speed* results [23]. The signal recognition process involves receiving, converting, and pre-processing the signals, denoising them, applying the suitable algorithm, using appropriate architecture, and other needed operations. The primary issue for performing automatic positioning is finding the *direction* of the leading sound from interfering sounds such as *noise* and *environmental* sounds [28]. The basic *direction-finding* application of sources is for the automatic localization and saving historian directions, to distinguish signals for tracking the sound sources, in any *indoor* or *outdoor* environment [15].

1.1 Source direction finding

The qualified speed positioning of the source is a critical task in an automatic localization system because it will not function correctly in the absence of a suitable signal, the presence of interfering sounds, created sound reflection, generated artificial noise, and combined ambient noise. Some problems of two ears existence in human hearing are the limitations in hearing range, low-speed positioning, and the accurate estimation of the direction. There are other various limitations of human hearing such as signal barriers, moving sources, signal interference, and noise problems, as well as the complexity of recognizing the signals. Speed accurate source positioning by minimum sensors is one of the classical and new research topics in real-time cyber-physical system (RCPS) design [26]. Now, research in mathematical acoustic-data modeling is vital to speed up the accepted accurate results of RCPS designing in the Internet of Things (IoT), Internet of Everything (IoE), Internet of Multimedia Things (IoMT), and Internet of Vehicles (IoV) for both civilian and military environments [11]. Some applications of positioning are sniper direction estimation [1], tracking unmanned aerial vehicles (UAV) [17], smart video conferencing (SVC) [25], speaker localization [16], speaker tracking [22], robot motion planning (RMP) in an unknown environment [18], Drone Audition Systems (DAS) [21], and in Wireless Acoustic Sensor Networks (WASN) [9]. Direction-finding is the primary technique of object tracking and agent navigation. The fundamental challenges of tracking are speed determining the source position, rapid memorizing the source positions to navigation for quick finding the direction, and accurate localization by using different sensor architectures and arranging various topology. In the positioning process, after determining the source direction, if the distance of the source to the sound sensors is obtained, then the process of locating the source coordinates is completed [6]. The real-time tracking process of each source suffers from accurate *direction-finding* and speed localizing. The choice of the appropriate architecture, best approach, and simple algorithm help to increase the speed, reduce the complexity, resistance to ambient noise, and creates robustness in environmental reflection for applying to light and low-power design of sound-based real-time embedded systems.

1.2 Acoustic source direction finding

Acoustic direction finding means determining the direction of a sound source with a minimum of two microphones to find the distance and angle between the source and the sensors, which is determined in two-dimensional (2D) and three-dimensional (3D) spaces. Orientation in terms of distance and proximity of resources, an active or passive method of determination, the type of technique, the accuracy of the position, and the speed of source detection have different complexity. If the sound sources at a great distance from the origin of the coordinates are located, they are called *far-field* and other *near-field* [6]. Cartesian coordinates of sound sources at long distances cannot obtain directly, so it is sufficient to estimate the direction of the source. Direction-finding is divided into *active* (detectable) and *passive* (non-detectable) approaches [12, 33]. In the active approach, in contrast to the passive, directional signals are transmitted to the source to obtain information in which its position is detectable and can be deceived. Sound source direction-finding methods are divided into two categories: *narrowband* and *broadband* [10]; *Impulsive* broadband sources such as explosions are the basis of this passive approach. The importance of estimating the direction of the impulsive sounds to the speed decision-making on actuator operation, after the estimation, is like the rapid deployment of active forces to the scene. SDF is a method for estimating the direction over time with several challenges, including (*natural*) ambient noise, (*artificial*) computational noise, and reflection. The estimator, by encountering the method architecture, sensor arrangement, outdoor environment, far distance sources, 2D or 3D coordinates, and the complexity of the estimation, leads to an increase in the *complexity* and *execution time* of the estimator algorithm. The choice is critical in *risk-aware* decision-making of multiple sound source direction finding for active parallel processing methods.

A passive automatic direction-finding system, in addition to not being detected by the enemy, can help to increase the processing speed of calculations, reduce the complexity, eliminate pre-processing, decrease the steps of the SDF algorithm, increase accuracy, and speed the response [6]. Automation addresses its challenges such as scalability for changing the dimensions of work, expanding the system equipment, reducing computational complexity, eliminating serial operations, increasing processing power, using parallel processing, increasing decision-making speed, using local sensor information in decreasing computational complexity, speed accurate extraction of global information to improve the accuracy for creating real-time equipment. The accuracy of multidimensional direction finding is related to the number of microphones and the spatial arrangement and distribution of the sensors or the used *topology*. Also, the classically used topology of sensors is transiting from a traditional structure to a parallel distributed model in the form of wireless sensor networks (WSN) for speed processing in online RCPS. The use of direction-finding in the new automatic detector equipment such as locators and voice trackers requires simplicity of approach, suitable microphone array topology, speed real-time detection, and high accuracy in determining the source direction. The simplicity of the selected method, low computational complexity, and low execution time compared to other standard forms are the reasons for the tendency of the time-of-arrival (TOA) and time-difference-of-arrival (TDOA) approaches [33].

1.3 TDOA-based direction finding

After receiving the sound signal, its specifications are extracted using standard techniques such as TOA and TDOA. The good signals decompose into waves with differences in

amplitude, phase, and frequency. By measuring the differences in these characteristics such as signal power changes, phase changes, Doppler shift, frequency modulated continuous wave (FMCW), channel impulse response (CIR), arrival time, and arrival time differences is possible the designing of various applications for direction-finding, localization, and tracking. The presence of ambient noise and environmental sounds can be harmful to multiple diagnostic applications [19]. In these technologies, the signal arrival time is the same as measuring the propagation time of the signal from the transmitter to the receiver to determine the propagation distance. This measuring requires the high synchronization of transmitter and receiver, which due to a large-scale problem, prevents its development in commercial devices. To smooth the synchronization of the devices in the arrival time method, the measurement of the time difference of the received signals in a pair of microphones is used, which determines the target's position at the intersection of the hyperbolas. This method is used for applications such as dual recording or noise cancellation in smartphones with dual *microphones*. To remove the need for *sensors* synchronization and develop the use of low-cost sensors, the TDOA implementation expands to direction-finding, localization, and navigation devices such as IoT devices and mobile phones with at least two microphones.

The passive TDOA technique in near-field direction finding is shown in Fig. 1(a). In the time difference of signal arrival, one of the microphones is the reference sensor, and the difference of signal arrival time in other microphones with this sensor is measured. MIC_i and MIC_j microphones have a distance of $d_i = r_j - r_i$ by measuring the difference in latency of the time of arrival signal between these two microphones. By considering the r_i and r_j distances from the signal source and drawing two hyperbolic forms, the intersection of two found half hyperbolic forms determines the position of the signal source. From MIC1 and MIC2 in Fig. 1(a), the d_1 distance between them and the first half hyperbolic form is obtained. Also, from MIC2 and MIC3, the d_2 distance between them and the other half hyperbolic form is obtained. Finally, the intersection of the two hyperbolic shapes defines the position of the SS1 source. The Azimuth and Elevation angles of a 3D coordinate system are shown in Fig. 1(b), where $[\phi, \theta]$ defines the Azimuth and the Elevation angles, and P proposes the sound source 3D place; r presents the sound distance from the origin of the coordinate where the reference sensor is located to the sound source. Due to its simplicity

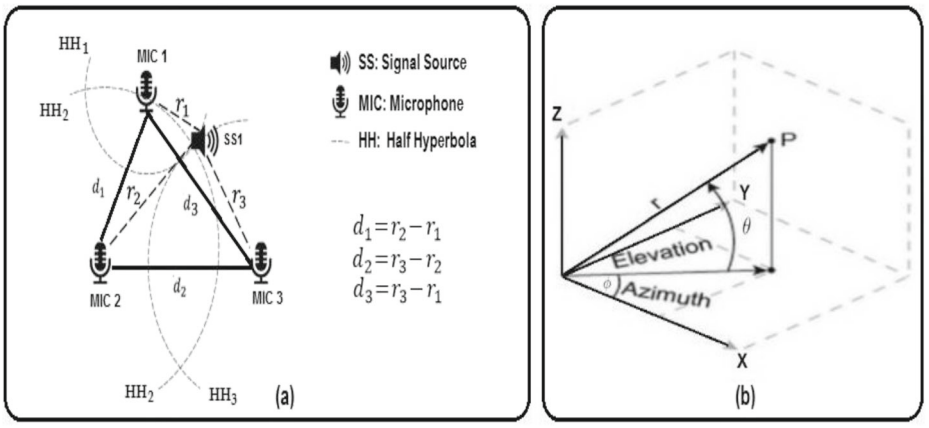

Fig. 1 (a) TDOA passive technique of sound source direction finding method in free-field and (b) Azimuth and Elevation angles in 3D coordinate system

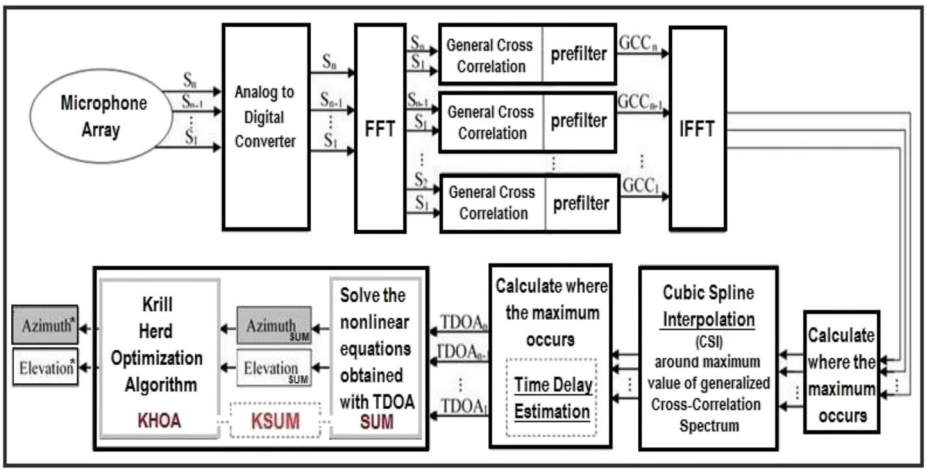

Fig. 2 Conceptual model of proposed TDOA-ASDF passive sound source direction finding method

and low computational complexity, the TDOA approach is a suitable method for solving broadband sound source problems. Additionally, by implementing the approach in combining the interpolation technique and adding the Krill Herd optimization algorithm (KHOA), the accuracy of estimation for the source direction increases. These capabilities are our reason and motivation for selecting the time delay estimation in the proposed TDOA-based acoustic source direction finding (TDOA-ASDF) method, as shown in Fig. 2.

1.4 Proposed method and evaluation overview

Most direction-finding methods, such as our proposed real-time TDOA-ASDF method, use a two-step processing model, in which the *positioning parameters* estimated in the *first step*. Then in the *second step*, the *accurate source direction* is determined according to the estimated parameters [6]. The proposed speed *scalable* method, resistant to ambient noise and sound reflection, for 3D estimation of the direction using the TDOA approach, consists of two steps the arrival time difference and the combining interception technique. The conceptual model of the TDOA-ASDF method in an outdoor environment is shown in Fig. 2. After receiving the signal through the sensors, the direction finder enters the frequency domain to increase the computing speed by digitizing the output. In the first step, Step 1 (Time Delay Estimation), the correlation of each signal with the reference sensor signal is calculated and entered into the time domain. Then cubic spline interpolation (CSI) is applied to the results to make a more accurate estimate; The delay rate of each sensor is obtained by the ratio of the reference sensor to the location with the maximum value. In the second step, Step 2 (Accurate Direction Estimation), the computational complexity reduces and the speed of finding direction increases by using the proposed novel Krill Herd with successive unconstrained minimization (KSUM) algorithm; this is estimated in the form of receiving time delay estimates, getting successive unconstrained minimization (SUM), applying Krill Herd optimization, having appropriate sensor architecture and using phase spectrum smoothing through CSI without pre-processing to reduce ambient noise. In experiments with actual and simulated data sets, the angle amounts of Azimuth and Elevation error in the near-field are 0.7164° (*degree*) and 1.5054°, and in the far-field are 1.0260° and 0.2071°, respectively.

The real-time direction determination speed is 1.935 *Sec* and the accuracy estimating of the center position of the source is high.

1.5 Innovation and organization

The challenges of automatic free-field direction-finding include the presence of ambient noise, produced artificial noise, combined sound reflection, high computational complexity, hard 3D direction finding, inadequate real-time SDF accuracy, low-speed direction-finding, relocation of sound sources, instability of sensor architecture, non scalability of the method, and the applicability of the method in far-field. Appropriate arrangement and distribution of acoustic sensors can not only simplify the calculation process and reduce the computational complexity, but also help to improve the positioning accuracy. This paper presents an automatic and scalable SDF method for estimating the 3D position of a stationary sound source that is resistant to noise and reflection in far- and near-field. The proposed passive method provides a two-step model with a minimal number of the sensor array and quick real-time estimation using different sensor topologies. Also, the proposed method is much faster and more accurate than the state-of-the-art methods. It has provided a significant increase in accuracy and speed by reducing computational complexity in expanding the number of sensors and using parallel architecture. In addition, the method is reliable enough to cope with the distributing sensors to large-scale topologies toward real-time decision-making for the free-field. With these observations, we are more interested in combined techniques to contribute to innovations in scalability, concurrency, accuracy, complexity, architecture, and algorithm.

Our essential *motivation* in our proposed sound source direction-finding is the low computational complexity and simplicity of the best TDOA-based method [29]. We designed a novel optimization (KSUM) algorithm as our *innovation* in the proposed TDOA-ASDF method, to increase the accuracy of direction estimation and reduce the computational complexity for resistance to noise and reflection. In summary, the *innovations* of this paper are presenting a reliable *method*, designing a novel optimal Krill Herd *algorithm*, and combining interpolation *techniques* to speed and accurately find the 3D direction of stationary sound sources for applying in *real-time* systems. Also, it presents a good *architecture* of SDF involving two arrays, offers a more accurate direction *estimator* by implementing the CSI technique, reduces the *dimension* of computational matrices by using the reference sensor assumption at the origin of the coordinate system, improves the *accuracy* of direction-finding by designing an optimal Krill Herd algorithm with using the result of SUM technique fast direction estimation, reduces the additional computational *complexity* by applying an interpolation technique, improves the *scalability* of the sensor arrays by implementing various topologies, increases the capability of *connecting* multiple robust arrays by planning a distributed architecture, and reduces the dimensions and *values* of computational matrices with the placement of the reference sensor of each array at the center of the coordinate system. In addition, the proposed TDOA-ASDF method is *resistant* to ambient noise and sound reflection to use in far-field applications for real-time localization and tracking.

The organization of this paper is as follows: Section 2 indicates required state-of-the-art TDOA-based related works produced about the passive direction-finding, Section 3 explains the problem modeling with details of two-step direction finding, which combined with the SUM technique and the used novel Krill Herd algorithm in the proposed TDOA-ASDF method, Section 4 shows the empirical experiments used to evaluate the results of our method with its used algorithms, and Section 5 summarizes the conclusions and future directions.

2 Preliminaries and related works

SDF algorithms evaluate for noise resistance, the degree of reflection sensitivity, dependence on environmental sounds, nonlinear equations solver, convergence speed, computational complexity volume, computational accuracy, dimensions of source, stationary and non-stationary source, sensor information, error rate direction determination, and direction-finding speed [20]. These algorithms for real-time response suffer from computational complexity, lack of resistance to ambient noise and sound reflection, and the presence of environmental sounds [24]. Therefore, we must try to divide the computational process into different steps, decrease the computational steps to reduce computational complexity, reduce directional dimensions to low computational volume, eliminate additional information to decrease computational operations, optimize the answer to determine the exact direction, insert pipeline structures for speed direction-finding, draw a reliable method with resistance to ambient noise, reduce computational artificial noise, and decrease the sensitivity to sound reflection for determining the exact direction [2]. Also, we must use the minimum number of cheap sensor arrays to increase scalability, reduce dependence on equipment infrastructure for rapid power changes in the sound source location, empower the work in near- and far-field to generalize the method, and apply the technique in free-field. Finally, we must try to increase the accuracy, reduce the error rate, improve the resolution, generate passive and indistinguishable speed methods, be fast in tracking, and be flexible in using different cheap sensor array topologies such as linear, circular, spherical, and distributed topologies.

Studies from 1976 onwards show that methods based on solving nonlinear hyperbolic equations resulting from time delays between sensors are among the most widely used SDF methods. A computationally efficient algorithm was presented for the direction-finding of a single source in far-field using a multi-sensor array. The time delay direction finding (TDDF) basic algorithm extracts the Azimuth and Elevation angles directly from the estimated time delays between the array elements [3]. Now, state-of-the-art methods focus on TOA and try to speed up the solving of hyperbolic equations of time delays between sensors. One of the advantages of the hyperbolic direction-finding method is that its results are resistant to noise-contaminated inputs. The solution to the *non-linear* problem is linearization and determining least-squares estimation. The *linearization* operations, contrary to algorithm simplification, make these algorithms more sensitive to noise [12]. Researchers have proposed various methods to solve stationary source direction finding based on TDOA, including the iterative process [10] and closed-form methods [27]. Non-linear equations can also be solved using linear Taylor expansion [10]. After solving the non-linear equations, the calculated source direction uses an iterative method and initial guessing with an error metric. The advantage of this method is the possibility of using additional delays to increase the accuracy of the calculation. However, this method is *sensitive* to the selection of the *initial guess* and the appropriate *convergence limit*. If the reasonable value is not selected, then it will not give an acceptable answer. Also, its disadvantages are that it created linearization *error* and generated a large volume of operations in different repetitions. All proposed methods could not integrate with the sensor features, source information, microphone array topology, indoor or outdoor environment properties, and a qualified linear approach.

First, Chan proposed a method to make it possible to use the additional sensor features [4]. Also, other researchers proposed speed-solving the Chan algorithm. The solving algorithms usually include iterative, analytical, and search methods [30]. The linear approach

mainly consists of two linear least squares (LLS) and weighted least squares (WLS) methods. The LLS method does not require prior knowledge of noise and is a simple and efficient solution in terms of the computational complexity volume. The linear weighted least squares (LWLS) method is a weighted version of the LLS method and has a higher estimation accuracy. Since using the constraint relationship between the target location and the auxiliary variables, the performance of the WLS improved, so a set of methods based on the WLS was proposed [19, 30]. The two-step linear WLS (2LWLS) is a direction-finding algorithm that uses two-step calculations to estimate the direction [5]. To solve the confinement effect that is created due to the fixed number of sampling points in the passive TDOA system, a time delay estimation algorithm based on cubic spline interpolation was proposed. This algorithm used CSI to interpolate the spectral peak curve of the correlation function. As a result, the accuracy of the peak value and the accurate estimation of the time delay of the signal source improved [31]. The proposed 2WLS method suffers from sign ambiguity resulting from a square root operation in the second WLS step, and the accuracy degraded. The position estimate in the first WLS step is not accurate enough. In a paper, an improved 2WLS method was proposed to circumvent this problem [5]. The tactic is mainly used to make a proper transformation of the constraint relation to keep only the linear items on the left side of the equation in the second WLS step. The experimental results show better performance, especially when the source is close to the reference sensor [8].

TDOA, due to low computational complexity and simplicity, is the primary choice for the two-step broadband direction finding of sound sources. This method attracted a lot of attention because, in a wide range of scenarios, the sound source emits only a transient signal without a predefined waveform pattern. Therefore, only the measurement of the time difference in finding the direction is used [7], which is the base method of our paper to estimate the direction of the broadband impulsive sound source. The main bottleneck in direction estimation is more related to the estimating time difference between the signals received from the sensors. In another research, the authors estimated the difference in arrival time using all permutations between microphones without knowledge of sound speed [32]. In our paper, we only estimate the time difference between the sensors and the reference sensor to reduce the time complexity of the process. We can even use the information to estimate the difference in arrival time and other measurements between the sensors to increase the accuracy of the final direction estimation. In an article, measuring the input phase difference between sensors was used along with the arrival time difference [5].

A summary of the base and state-of-the-art TDOA-based direction-finding methods and which will compare to our proposed method is given in Table 1 along with the characteristics of each technique. Also, the advantages and disadvantages of produced results are considered in the table columns. This table addresses critical direction-finding challenges such as *Method and Year* (reference and published year), *Approach* (time delay direction finding or time difference of arrival approach), *Environment* (outdoor or indoor environmental experiences), *Robust Noise-Reflection* (noise and reflection resistance), *Real Time* (online real-time or offline applicable ability of method), *Field* (Far (far-field) on near (near-field) distance of its sound source), *Scalable* (method scalability of hardware in using new and more sensors, adding more arrays, and applying different topology for near- or far-field), *Optimization Algorithm* (optimized or non-optimized for results accuracy), *Solving Method* (linear or non-linear solving method), *Array Architecture* (topology of the sensor array for symmetric or asymmetric used sensor and linear or other sensors type in the form of random, arbitrary, linear, etc.), *No. Sensor Type* (number of sensors and sensor type), *DIM* (direction finding dimension in 2D or 3D space), and *Goal* (complexity reduction, accuracy

Table 1 Summarized comparing the proposed method to the basic and state-of-the-art TDOA-based passive SDF methods

Method Year	Approach	Environment	Robust Noise-Reflection	Real Time	Field	Scalable	Optimization Algorithm	Solving Method	Array Architecture	No. Sensors Type	DIM	Goal
[3] 1999	TDDF	Indoor	Yes-Yes	No	Far	No	————	LLS	Random & Symmetric	7 *Hetrogen*	3D	Efficient Computationally
[7] 2018	TDOA	Outdoor	Yes-No	Yes	Far	Yes	Lagrange Multiplier Quadratic Constraints	CLS	Asymmetric	8 *Homogen*	3D	Low Complexity High Accuracy
[4] 1994	TDOA	Indoor	Yes-No	No	Near	No	————	CHAN	Arbitrary & Linear	$\geqslant 3$ *Homogen*	2D	Low Complexity High Accuracy
[30] 2019	TDOA	Indoor	Yes-No	Yes	Near	No	Firefly	WLS	Symmetric	4 *Homogen*	2D	Low Complexity High Accuracy
[5] 2018	TDOA	Outdoor	Yes-No	No	Far & Near	No	————	ITWLS	Symmetric	4 *Homogen*	2D	Good Accuracy Medium Complexity Good Performance
[29] 2018	TDOA	Outdoor	Yes-No	No	Far & Near	Yes	————	SUM	Asymmetric	7 *Homogen*	3D	Good Accuracy Low Complexity Good Performance
[TM] Proposed Method	TDOA	Outdoor	Yes-Yes	Yes	Far & Near	Yes	Krill Herd	KSUM	Symmetric	9*or*7 *Homogen*	3D	High Accuracy Low Complexity High Performance

improvement, and/or performance promotion). In the last two rows of the table, the primary motivating method of our problem-solving in the first row and our proposed method in the last row, are compared in terms of the above challenges.

3 Proposed TDOA-ASDF method

In this section, we describe the problem definition (Section 3.1), explain the TDOA-based two-step direction-finding method (Section 3.2), and present the successive unconstrained minimization technique (Section 3.3). Also, we propose the Krill Herd optimization algorithm (Section 3.4), explain the onlooker search mechanism (Section 3.5), illustrate the parameter improvement (Section 3.6), and clarify the fitness function and control variable constraint (Section 3.7).

3.1 Problem definition

We propose an advanced Krill Herd with Successive Unconstrained Minimization (KSUM) algorithm to combine the interpolation and the Krill Herd algorithm with the results of the SUM technique. In KSUM, the onlooker search mechanism is designed to reduce the falling probability into the local optimum. The parameter values of the proposed algorithm, including inertia weight and step-length scale factor, are varied according to the iteration of the evolutionary process, which improves the exploration and exploitation capabilities. Moreover, the SUM handling technique is proposed to guide the individual to the feasible space and ensure that the optimal solution satisfies the protection constraints. Then, KSUM is constructed with the result of the SUM technique to solve the multi-constrained optimization problem for improving its performance by including appropriate objective functions. Our method has a strong ability of diversity and a few adjusting parameters, which is suitable for optimized direction-finding. However, the search precision and convergence speed of the Krill Herd algorithm cannot be guaranteed to fall into the local optimum. So far, the proposed advanced Krill Herd algorithm is designed to overcome the shortcomings.

Non-linear hyperbolic equations were formed to calculate the source direction using estimated delays. An important issue is to reduce the volume of computational matrices to accelerate direction-finding. Assuming the coordinates of the reference sensor at the origin of the coordinate system, the volume of the matrix calculations decreases by reducing the dimensions and the numerical value of the matrices. To solve the matrix equations, five proposed methods of linear least squares (LLS), constrained least squares (CLS), Chan (CHAN), improved two-step weighted least square (ITWLS), and successive unconstrained minimization (SUM) that have low computational complexity are examined in this paper. The proposed KSUM is designed and evaluated with experiments on actual and simulated data using two sensor arrays. This method has more accuracy and speed than the selected state-of-the-art methods. We have modeled the problem in two-step: the time difference between the correlation and the direction estimation. Then we present the primary SUM method and our proposed KSUM algorithm.

3.2 Two-Step direction finding

A two-step direction-finding model has been done in two steps, the first step (Time Delay Estimation) and the second step (Accurate Direction Estimation), which will be described below.

Step 1: Time Delay Estimation ► TDOA, due to its low complexity and simplicity, is suitable for broadband sound sources direction-finding. One of the sensors is assumed to be the reference sensor to differentiate the signal reception time of other sensors and is calculated with it. Delay time calculation algorithms solve the problem by assuming the existence of a source and a one-way propagation pattern. Finally, the results were generalized based on other cases. The most common method to calculate the delay between two sensors is to determine the maximum result value of the cross-correlation function between all received data from the source. This method is called cross-correlation. The pattern of the signal received from the source without considering the reflection is shown in Eqs. (1) and (2), where $s(t)$ is the source sign, and d_{i1} is the time delay of the received sign i relative to the base sign. M is the number of sensors in one array and ξ_1 is the accumulating noise.

$$x_1(t) = s(t) + \xi_1(t) \tag{1}$$

$$x_i(t) = s(t - d_{i1}) + \xi_1(t), i = 2, 3, \ldots, M \tag{2}$$

We obtain the amount of similarity between two signs by calculating the amount of delay between them and applying the correlation function to them. In practice, to increase the computational speed, the operations are performed in the frequency domain to calculate the delay between each pair of sensors. The use of the cross-correlation method to calculate the delay between two signs is not ideal. It has low accuracy in the presence of noise or environmental reflections. The smoothness of the phase spectrum is considered to increase the efficiency of this method and reduce the effect of noise and reflection. The value of the cross-correlation of the two symbols is obtained using their Fourier transform, which is shown in Eq. (3). This value by applying the filter $\Psi_g(\omega)$ becomes Eq. (4). Equation (5) shows the phase spectrum smoothness, where ω represents the frequency, X_1 and $\acute{X}_i$ are the Fourier transform of the reference sensor signal and the Fourier transform of the signals of the other sensors, respectively. Continuation of Eq. (5) is transmitted from frequency to time using the inverse fast Fourier transform (IFFT). In the next step, 100 data containing the maximum value of the function $R_{x_1x_i}(\tau)$ is added by interpolating a cubic spline between both sets of data. The area that includes the maximum value of the previous step is the amount of delay between the received signals of each pair of sensors.

$$G_{x_1x_i}(\omega) = X_1(\omega)\acute{X}_i(\omega), i = 1, 2, \ldots, M \tag{3}$$

$$R_{x_1x_i}(\omega) = \Psi(\omega)G_{x_1x_i}(\omega) \tag{4}$$

$$\Psi_g(\omega) = \frac{1}{|G_{x_1x_i}(\omega)|} \tag{5}$$

Interpolation is an estimation technique to build new data points within a discrete set of known data points. Spline interpolation is a method in which an interpolator is a particular fragment polynomial called a spline. Cube spline interpolation adds new specimens to the IFFT spectrum correlation at near-maximum points. In that case, the new maximum value of the IFFT of the fast correlation spectrum may change, leading to a more accurate estimate of the time delay of the sensors. Considering that the amount of delay between the sensors should be low value; if according to the results, the delay is more than one second, then the cross-correlation operation with the phase spectrum is performed in reverse. Because the sound reaches the secondary sensor earlier than the reference sensor, the obtained delay between the two signals is multiplied by the negative. Finally, the calculated delays are compared.

Step 2: Accurate Direction Estimation ► The direction of the sound source is determined using estimation delays with non-linear hyperbolic equations. By considering the coordinates of the reference sensor at the origin of the coordinate system, the volume of the matrix calculations in direction determination is reduced. The result decreases the dimensions and the numerical value of the matrices. Cartesian coordinates of the sound source and sensors are displayed in three dimensions with $s_i = [x_i, y_i, z_i]$ and $P_s = [x_s, y_s, z_s]$, so that $i = 1, 2, \ldots, M$ and M is the number of array sensors. The $[\phi, \theta]$ angles of the sound source in 3D are relative to the reference sensor. k_{3D} is the direction vector in the 3D coordinates and is given in Eq. (6). The distance between each sensor and the reference sensor in 3D is according to Eqs. (7) and (8), where c is the sound speed in meters per second and τ_{i1} is the delay between the two pairs of sensors i and 1. Also, r_i and r_1 are the distance of sensor i from the sound source, and the distance of the reference sensor (sensor 1) from the sound source, respectively. ϕ and θ are the Azimuth and the Elevation of the sound source relative to the reference sensor, respectively. Figure 1(b) is shown the Azimuth and Elevation angles and a 3D coordinate system; P and r are the sound source and the distance from the origin of the coordinates where the reference sensor is located to the sound source.

$$K_{3D} = \begin{bmatrix} k_x \\ k_y \\ k_z \end{bmatrix} = \begin{bmatrix} \cos(\theta)\cos(\phi) \\ \cos(\theta)\sin(\phi) \\ \sin(\theta) \end{bmatrix} \tag{6}$$

$$c_{\tau_i} = r_{i1} = r_i - r_1 = \sqrt{(x_i - x_s)^2 + (y_i - y_s)^2 + (z_i - z_s)^2} - \sqrt{(x_1 - x_s)^2 + (y_1 - y_s)^2 + (z_1 - z_s)^2} \tag{7}$$

$$c_{\tau_i} = x_i cos(\theta)cos(\phi) + y_i cos(\theta)sin(\phi) + z_i sin(\theta) \tag{8}$$

The positioning problem of estimating direction, can be modeled with the relationship matrix in Eq. (9). In which $S = [s_2, s_3, ..., s_{M-1}]^T$ is the matrix contains the coordinates of the sound sensors of the array, except for the reference sensor (sensor 1), where the coordinates of the reference sensor in the coordinate system in 3D is equal to $s_1 = [0, 0, 0]$. $\hat{\tau} = [\hat{\tau}_2, \hat{\tau}_3, ..., \hat{\tau}_M]$ is the difference between the estimated arrival times. The vector $\hat{k}$ for estimation, regardless of its spatial dimension, contains an error of Δk, where n is the measurement error vector. Measurement error takes into account the difference between the time of arrival and the speed of sound error. In this paper, the speed of sound is a constant value. The noise vector $\Delta\tau$ is a measure of the difference in arrival time.

$$\hat{c}\hat{\tau} = S\hat{k} + n, \hat{\tau} = \tau + \Delta\tau, \hat{k} = k + \Delta k \tag{9}$$

$$h_1 - G_1\psi_1^\circ = \epsilon_1, G_1 = \begin{bmatrix} S_2^T & \frac{1}{2}(r_{21}^2 - \|S_2\|^2) \\ S_3^T & \frac{1}{2}(r_{31}^2 - \|S_3\|^2) \\ \vdots & \vdots \\ S_M^T & \frac{1}{2}(r_{M1}^2 - \|S_M\|^2) \end{bmatrix}, h_1 = -r \tag{10}$$

$$\min(h_1 - G_1\psi_1)^T W_1(h_1 - G_1\psi_1) \tag{11}$$

$$\psi_1^T \Sigma \psi_1 = 1, \Sigma = diag\{[1_N^T, 0]^T\} \tag{12}$$

$$\psi_1 = (G_1^T W_1 G_1)^{-1} G_1^T W_1 h_1, W_1 = (B_1 Q B_1)^{-1}, B_1 = -diag(r_{i1}),$$

$$Q = \frac{1}{2}(I_{M-1} + 1_{M-1}) \tag{13}$$

$$cov(\psi_1) \approx (G_1^T W_1 G_1)^{-1} \tag{14}$$

$$\hat{\phi} = \tan^{-1}\left(\frac{\psi_1(2)}{\psi_1(1)}\right) \tag{15}$$

$$\hat{\theta} = \tan^{-1}\left(\frac{\psi_1(3)}{\sqrt{\psi_1(1)^2 + \psi_1(2)^2}}\right) \tag{16}$$

3.3 Successive unconstrained minimization

The optimization estimate of the weighted least squares decreases with the quadratic range, where the quadratic constraint is only related to the direction of signal entry and does not depend on the inverse distance or range. One of the optimization techniques is the SUM, in which the first minimizes the optimization limit and the second minimizes the relationship between the unknowns (the angles and range of the sound source relative to the reference sensor) to improve the previous estimate to generate the final solution [18]. The error matrix is defined by the difference between the arrival time of the sound source signal between each of the sensors and the reference sensor with Eq. (10), where the ϵ_1 is the error vector of the initial source direction estimation relative to the reference sensor; ψ_1° is the estimated vector in the first stage of the minimization SUM method without consecutive constraints $\psi_1^\circ = [\phi^\circ, \theta^\circ, \frac{1}{r^\circ}]$; ϕ° and θ° are the Azimuth and the Elevation angles; the third component of ψ_1° is the inverses of the estimated distance of the reference sensor from the source. $\|S_i\|$ is the Euclidean norm of the coordinate of microphone i. In the first stage (first minimization), the constraint of Eq. (12) is removed, and by assuming the independence of the ψ_1 variables, then Eq. (11) is easily defined with Eq. (13). Assuming that the error in G_1 is small enough, the related covariance matrix can be represented as Eq. (14). The Azimuth and Elevation angles are obtained by the initial estimation of ψ_1 from Eqs. (15) and (16), respectively.

In the second stage (second minimization), the constraint is reduced to decrease the generated solution error of the first stage to guess the final estimation better. The solution of the first stage is expressed as Eq. (17), where $\Delta\psi_1$ is the estimation error of this stage. The error matrix of the second stage is estimated as Eq. (18). The square unitary matrix 1_{N-1} is a $matrix_{(N-1)\times(N-1)}$ and I_{N-1} is the identity matrix as a $matrix_{(N-1)\times(N-1)}$, where N is the dimension of the coordinate system. The weight matrix W_2 is expressed by Eq. (19). The final estimate ψ_2 is obtained by using Eq. (20), where the Azimuth and Elevation angles are estimated in 3D by Eqs. (21) and (22).

$$\psi_1 = \psi_1^\circ + \Delta\psi_1, h_2 - G_2\psi_2^\circ = \epsilon_2, h_2 = \begin{bmatrix} \psi_1(1:N-1) \bigodot \psi_1(1:N-1) \\ \psi_1(N)^2 - 1 \\ \psi_1(N+1) \end{bmatrix} \tag{17}$$

$$G_2 = \begin{bmatrix} I_{N-1} & 0_{N-1} \\ -1_{N-1}^T & 0 \\ 0_{N-1}^T & 1 \end{bmatrix} \tag{18}$$

$$W_2 = B_2^{-1}(G_1^T W_1 G_1)B_2^{-1}, B_2 = [2\psi_1(1:N)^T, 1]^T \quad (19)$$

$$\psi_2 = (G_2^T W_2 G_2)^{-1} G_2^T W_2 h_2 \quad (20)$$

$$\hat{\phi} = \tan^{-1}\left(\frac{sgn(\psi_1(2))\sqrt{\psi_2(2)}}{sgn(\psi_1(1))\sqrt{\psi_2(1)}}\right) \quad (21)$$

$$\hat{\theta} = \tan^{-1}\left(\frac{sgn(\psi_1(3))\sqrt{1-\psi_2(1)-\psi_2(2)}}{\sqrt{\psi_2(1)+\psi_2(2)}}\right) \quad (22)$$

Input: NumberOfRuns, NumberOfKrills, NumberOfNeighbors, SpeedOfSound, Max-iter, V^f, D_{max}, N^{max}, Azimuth, Elevation; {*Number of runs; Number of Krills; Number of neighbors; Speed of sound; Maximum iterations; Foraging speed; Maximum diffusion speed; Maximum induced speed; Azimuth and Elevation are extracted from SUM Algorithm*}
Output: $Azimuth^*, Elevation^*$ using (52) and (53); {*Azimuth Angle, Elevation Angle*}
1 **Initialize:** $G = 1$; $p \leftarrow rand(NP)$; {*Initialize counter, and population form Krill variables*}
2 **Default Value:** NumberOfRuns=10, NumberOfKrills=NP=50, NumberOfNeighbors=25, Max-iter=100, $N^{max} = 0.1$, $V^f = 0.03$, SpeedOfSound=349.02, $D_{max} = 0.008$ {*Default value of our simulation parameters*}
3 Calculate UB and LB using (51); {*Calculate upper band and lower band of search space*}
4 Generate initial population using (45); {*Generate the initial population, lists & variables*}
5 Evaluate each Krill individual fitness and memorize best solution according to its initial position; {*Evaluate each Krill individual according to its position and generate the best solution*}
6 **while** *(G < Max-iter)* **do**
7 Update inertia weight w and step-length scale factor using (48) and (49);
8 Sort the population according to (50);
9 **for** $i \leftarrow 1$ **to** NP **do**
10 Motion calculation using (24) to (40); {*Movement, Foraging motion, Random diffusion*}
11 Perform the genetic operators using (41) to (44);
12 Update and modify the population by using (39) and (40);
13 Evaluate each Krill individual fitness using (46) to (47);
14 Employ the onlooker search mechanism by using (48);
15 Modify best solution;
16 **end**
17 G++;
18 **end**
19 **return** $Azimuth^*$, $Elevation^*$ {Azimuth angle, Elevation angle}

Algorithm 1 The proposed Krill Herd algorithm for direction finding problem.

3.4 Krill Herd algorithm

The proposed Krill Herd algorithm according to Algorithm 1 is as follows: initially, in Step 1 (*Initialization and Bands Calculation*), the input parameters, the direction data, and the limits of variables are initialized (Lines 1 to 3 of Algorithm 1); in Step 2 (*Population Generation*), the Random values are assigned to each D-dimensional as an optimization phase (Line 4 of Algorithm 1); then, in Step 3 (*Fitness Evaluation*), each Krill individual is evaluated according to their position and memorized, solving the best solution (Line 5 of Algorithm 1); in Step 4 (*Motion Calculation*), movement is designed by other Krill individuals, foraging motion, and random diffusion (Line 10 of Algorithm 1). Also, in Step 5 (*Genetic Operations*), the genetic operators, including crossover and mutation, are performed (Line 11 of Algorithm 1). Eventually, in Step 6 (*Population Updates*), the population is updated, and the loop is repeated (Line 12 of Algorithm 1); finally, in Step 7 (*Output Generation*), if the stopping metric is matched, then the direction is extracted (Line 19 of Algorithm 1), else goto Step 4 (Line 6 of Algorithm 1). The main inputs of the Krill Herd

algorithm from the results of the SUM technique are formed to solve the multi-constrained optimization problem for improving its performance.

3.4.1 The Krill cluster

One of the marine studies is Cluster Krill Swarming. When crowds encounter natural enemies, some of them are killed and decreasing the population density. Then, Krill clusters gather to increase population density and find food. Based on these Krill clustering behaviors, some researchers have proposed a new heuristic algorithm to solve the Krill Herd algorithm for the global optimization problem. Each Krill individual represents a feasible solution to the optimization problem. If increasing population density and finding food are considered as the direction of advancing the optimization problem, then the process of individual aggregation of Krill is the process of the search algorithm for the optimal solution. The location of each Krill changes over time, and its changes are mainly influenced by the following three factors: movement induced by other Krill individuals, foraging motion, and random diffusion. For an n dimensional decision problem, the Lagrangian model is used in Eq. (23), where N_i is the induced motion of other Krill individuals; F_i is the foraging activity, and D_i is the physical diffusion.

$$\frac{dX_i}{dt} = N_i + F_i + D_i \tag{23}$$

3.4.2 Movement induced by other Krill individuals

To achieve the overall migration of the population, each Krill member will interact with each other and make the population very crowded. The direction α_i of air movement is influenced by the local effect of neighbors, the target effect of the desired person, and the repulsion effect of population elimination. For Krill, the N_i movement caused by other Krill can be defined as Eqs. (24) and (25), where N^{max} and N_i^{old} are the maxima induced speed and the last motion-induced, respectively; where $\omega_n \epsilon [0, 1]$ is the inertia weight of the motion; a_i^{local} and a_i^{target} represent the local effect and target effect, respectively. The local effect is induced by neighbors can be assumed as an attractive/repulsive tendency and is determined as Eq. (26), where NN is the number of neighbors, X and K represent the related position and fitness, respectively; $Kworst$ and $Kbest$ are the worst and the best value of the Krill Herds so far; ϵ is a small positive value to avoid singularities. Furthermore, the sensing distance of Krill individuals should be proposed to define the local effect, which determines the neighbors of a Krill individual using Eqs. (27) to (29), where NP represents the population size. The optimal global solution as the target direction will affect the motion of each Krill as Eqs. (30) and (31), where random variable $rand_1 \epsilon [0, 1]$ is uniformly distributed; g and g_{max} represent the actual iteration and the maximum iteration numbers, respectively.

$$N_i = N^{max}\alpha_i + \omega_n N_i^{old} \tag{24}$$

$$\alpha_i = \alpha_i^{local} + \alpha_i^{target} \tag{25}$$

$$\alpha_i^{local} = \sum_{j=1}^{NN} \hat{K}_{ij}\hat{X}_{ij} \tag{26}$$

$$\hat{X}_{ij} = \frac{X_j - X_i}{\| X_j - X_i \| + \epsilon} \tag{27}$$

$$\hat{K}_{ij} = \frac{K_i - K_j}{K_{worst} - K_{best}} \tag{28}$$

$$d_i = \frac{1}{Number\,Of\,Neighbor \times NP} \sum_{j=1}^{NP} \parallel X_i - X_j \parallel \tag{29}$$

$$\alpha_i^{target} = C^{best} \hat{K}_{i,best} \hat{X}_{i,best} \tag{30}$$

$$C^{best} = 2\left(rand_1 + \frac{g}{g_{max}}\right) \tag{31}$$

3.4.3 Foraging motion

In the foraging motion, the *food* searched by the population is estimated according to the fitness distribution of the Krill individuals. Its position is determined by the physical definition of the *center of mass* as Eq. (32). The foraging activity of the Krill is affected by two main factors: the current and the previous location of the food source, which can be expressed as Eqs. (33) and (34), where V_f is the foraging speed, w_f is the inertia weight between 0 and 1, F_i^{old} is the previous foraging food motion, b_i^{food} and b_i^{ibest} represent the food attraction and the effect of the best fitness of the i^{th} Krill so far, which are defined as Eqs. (35) and (36). C^{food} is the food coefficient that is varied by the iteration as Eq. (37), where the random variable $rand_2 \epsilon [0, 1]$ is uniformly distributed.

$$X^{food} = \frac{\sum_{j=1}^{NP} \frac{1}{k_i} X_i}{\sum_{j=1}^{NP} \frac{1}{k_i}} \tag{32}$$

$$F_i = V_f \beta_i \omega_f F_i^{old} \tag{33}$$

$$\beta_i = \beta_i^{food} \beta_i^{ibest} \tag{34}$$

$$\beta_i^{food} = C^{food} \hat{K}_{i,food} \hat{X}_{i,food} \tag{35}$$

$$\beta_i^{ibest} = \hat{K}_{i,ibest} \hat{X}_{i,ibest} \tag{36}$$

$$C^{food} = 2\left(1 - \frac{g}{g_{max}}\right) \tag{37}$$

3.4.4 Random diffusion and position update

The physical diffusion of Krill individuals is explained by the maximum diffusion speed and a random directional vector as Eq. (38), where D^{max} represents the maximum diffusion speed, and δ is the random vector whose arrays are generated randomly variables according to uniform distribution between -1 and 1. These three factors make each Krill individual change her position in the optimal direction. The position of the individual during the time interval $t + \Delta t$ is updated as Eq. (39), where Δt is the most important constant, which ultimately depends on the search space. The search is expressed by Eq. (40), where $C_t \epsilon [0, 2]$ is a constant number, NV is the total number of control variables, UB_j and LB_j express the upper and lower bounds of the j^{th} variable, respectively.

$$D_i = D^{max}\left(1 - \frac{g}{g_{max}}\right)\delta \tag{38}$$

$$X_i(t+\Delta t) = X_i(t) + \Delta t \frac{dX_i}{dt} \tag{39}$$

$$\Delta t = C_t \sum_{j=1}^{NV} (UB_j - LB_j) \tag{40}$$

3.4.5 Genetic operators

The crossover and mutation strategies of the Genetic algorithm are embedded into KSUM to improvement the performance of the algorithm. The crossover is modeled as Eq. (41), where D is the dimension of the optimal problem, random variable $rand_3 \epsilon [0, 1]$ is uniformly distributed, X_{r1} $(r1 \neq i)$ is randomly picked from the current population, C_R is the crossover probability which is equal to zero for the global best solution. The mutation is defined as Eqs. (42) and (43), where X_{best} is the global best position of the whole swarm, $\mu \epsilon$ [0, 1] is the mutant factor, random variable $rand_4 \epsilon [0, 1]$ is uniformly distributed, X_{r2} and X_{r3} $(r2 \neq r3 \neq i)$ are randomly chosen from the current population, Mu is the mutant probability as Eq. (44) which is also equal to zero for the best global solution. The population is generated as Eq. (45), where $x_{j,min}$ and $x_{j,max}$ express the upper and lower bounds of the j^{th} decision variable, NP is the population size, random variable $rand_5 \epsilon [0, 1]$ is uniformly distributed.

$$X_{i,j} = \begin{cases} X_{r1,j} & rand_3 < C_R \\ X_{i,j} & else \end{cases} \quad j = 1, \ldots, D \& i = 1, \ldots, NP \tag{41}$$

$$C_R = 0.2/\hat{K}_{i,best} \tag{42}$$

$$X_{i,j} = \begin{cases} X_{best,j} + \mu(X_{r2,j} - X_{r3,j}) & rand_4 < Mu \\ X_{i,j} & else \end{cases} \quad j = 1, \ldots, D \& i = 1, \ldots, NP \tag{43}$$

$$Mu = 0.05/\hat{K}_{i,best} \tag{44}$$

$$X_{j,i|g=0} = x_{j,min} + rand_5 \times (x_{j,max} - x_{j,min}) \; j = 1, \ldots, D \& i = 1, \ldots, NP \tag{45}$$

3.5 Onlooker search mechanism

In the evolutionary process, the difference among individuals will gradually decrease as the number of iterations increases, which makes Krill Herd prone to fall into local optimum. Therefore, the onlooker search mechanism, base on the behavior of onlooker bees of the artificial bee colony algorithm, is introduced to overcome this shortcoming. All onlookers select individuals to be updated again according to the probability (Pi) value based on the roulette method as Eq. (46), where $i \epsilon 1, 2, \ldots, NP$ and NP represents the population size. fit_i is associated with the fitness value f_i of the i^{th} Krill calculated in Eq. (47). The individual with better fitness will attract more onlooker bees and be updated by Eq. (47), where $X_{r1,g}$ and $X_{r2,g}$ $(r1 \neq r2 \neq i)$ are randomly chosen from the current population, X_{best} is the best global position of the whole swarm, and $rand_1$ is a uniformly distributed random variable between 0 and 1. Then the selection combined with a novel constraint method, is utilized to choose a better vector between the trial vector X_i'' and the target

vector X_i. Moreover, the number of onlooker bees influences the convergence speed, and it is usually 1/3 of the population size.

$$fit_i = \begin{cases} \frac{1}{1+f_i} & f_i \geq 0 \\ 1+ \mid f_i \mid & f_i < 0 \end{cases} \tag{46}$$

$$X'' = X_i + rand_1 \times (X_{best} - X_i) + (1 - rand_1) \times (X_{r1,g} - X_{r2,g}) \tag{47}$$

3.6 Parameter improvement

Generally, ω_n and ω_f are defined as ω called inertia weight. When ω takes a more significant value, the algorithm can perform better globally and diversify the solution. Smaller ω is aimed to complete the local search. An excellent algorithm should have two capabilities: exploration and exploitation. The former can avoid the possibility of finding the optimal global solution quickly by falling into the local minimum. The latter refers to the ability to explore potential better solutions near the existing optimal value. C_t is the step-length scale factor. The larger the value of C_t is, the stronger the exploration capability will be, leading to slower convergence. On the contrary, the exploitation capability will be more vital, leading to premature convergence. Therefore, improving of the parameters should be implemented, which directly affects the search precision and convergence speed. In the proposed method, inertia weight ω is gradually decreased as seen in Eq. (48), where g and g_{max} represent the number of iterations and the maximum iteration number, respectively. The way of non-linear decline is to ensure that the motion can get a more excellent range of exploration in the former and accelerate the convergence of particles in the latter. The step-length scale factor C_t is changed according to the number of iterations of the algorithm, as seen in Eq. (49).

$$\omega = 0.1 + 0.8 \times \left(1 - \frac{g}{g_{max}}\right)^2 \tag{48}$$

$$C_t = \begin{cases} 0.7 \ if(g < o.4 \times g_{max}) \\ 0.4 \ otherwise \end{cases} \tag{49}$$

3.7 Fitness function and control variable constraint

The fitness function calculation is decided to generate the best solution for estimating the Azimuth and Elevation. The goal is to find the best answer for $\hat{k}$, the vector of the estimated directions. The calculation uses Eq. (50) to generate needed results, where c is the speed of sound; $\hat{\tau}_i$ is the estimated time difference between each sensor and the reference sensor of each array; s_i is the coordinate of sound sensor i without the reference sensor. The Krill Herd algorithm uses the estimated Azimuth and Elevation angles to limit the search space. Equation (51) shows the limitation of the search space of the upper and lower bands in three dimensions, respectively, based on the vector for the direction of Eq. (6). ρ_1 and ρ_2 are the angles in radians that create the search space around the vector for initial estimation. ρ_1 and ρ_2 are considered 2.5° and 0.5°, respectively, which form 5° of the search space for the estimated Azimuth angle and 1° of the search space for the estimated Elevation angle from the previous step. If the estimated Azimuth angle is close to 2.5° of the initial angles of each quadrilateral of the trigonometric circle (for example, 0°, 90°, 180°, and 270°), then ρ_1 is

equal to 0.5°. Finally, the best estimated $\hat{k}$ vector is selected from the Krill Herd algorithm and is estimated by Eqs. (52) and (53) under the Azimuth and Elevation angles.

$$Fitness\,Function = \sum_{i=1}^{N-1} (c\hat{\tau}_i - s_i\hat{k})^2 \tag{50}$$

$$UB = \begin{bmatrix} \cos(\theta+\rho_2)\cos(\phi+\rho_1) \\ \cos(\theta+\rho_2)\sin(\phi+\rho_1) \\ \sin(\theta+\rho_2) \end{bmatrix} LB = \begin{bmatrix} \cos(\theta-\rho_2)\cos(\phi-\rho_1) \\ \cos(\theta-\rho_2)\sin(\phi-\rho_1) \\ \sin(\theta-\rho_2) \end{bmatrix} \tag{51}$$

$$\hat{\phi} = \tan^{-1}\left(\frac{\hat{k_y}}{\hat{k_x}}\right) \tag{52}$$

$$\hat{\theta} = \tan^{-1}\left(\frac{\hat{k_z}}{\sqrt{\hat{k_x}^2 + \hat{k_y}^2}}\right) \tag{53}$$

4 Experimental results

In this section, we explain the experimental results produced for impulsive direction-finding. First, the experimental setup (Section 4.1), the state-of-the-art DFS comparing methods and dataset specifications, along with the data collection method to sound signal recording (Section 4.2), are presented. After validating the proposed TDOA-ASDF and other implemented methods (Section 4.3), we will explain the microphone array topology (Section 4.4) and finally present the empirical experiments to evaluate the results (Section 4.5).

4.1 Experimental setup

Initially, our implemented direction-finding techniques and methods have been implemented in the MATLAB 2019 software tool and executed on a machine with core(TM)i7-6700HQ 2.6 GHz and 12 GB RAM under the management of the Windows 10 operating system. Its performance was evaluated in different datasets. For a fair comparison, LLS, CLS, CHAN, ITWLS, SUM, and the proposed KSUM methods have been implemented and simulated different real and artificial datasets under the same conditions in a similar situation. Recording equipment for simultaneous sound signal sampling data includes microphones, processor cards, and *Motu* devices. The incoming sounds of microphones are analog and require the conversion of analog to digital for signal processing; to better record, the microphone array and digitally convert the sound signal, a multi-channel array microphone with *Motu* 24 I/O preamplifier and the corresponding processor card has been used for sampling and converting analog to digital data; this device also has the task of recording sound signals simultaneously of all microphones. All used capacitive microphones in the experiments have a sampling frequency of 44100 *samples per second* (Hz) and 1024 *samples per buffer*. We have implemented the TDOA-ASDF method in the form of KSUM (using the Krill Herd algorithm and the interpolation technique) and have obtained good results.

Table 2 Directions of sound sources relative to the reference sensor of each array

Source	Position			
	Array 1		Array 2	
Angle i	Azimuth $(degree)$	Elevation $(degree)$	Azimuth $(degree)$	Elevation $(degree)$
1	90	+0.0000	149.3865	0
2	30.6135	+0.0000	90	0
3	49.8022	+0.0000	130.1978	0
4	57.6355	+0.0000	122,3645	0
5	44.9757	-0.2920	90	-1.9473
6	74.3643	-0.1931	108.4349	-0.3080

4.2 Comparative methods and data sets

All conventional passive studied methods, and the proposed method, are based on TDOA. The state-of-the-art LLS, CLS, CHAN, ITWLS, and SUM methods are examined and compared with the proposed KSUM method. The results of the proposed method, which used the KSUM algorithm and was combined with the interpolation technique, and, in particular, the SUM method as the primary motivation of our process, are compared to an analysis by using experimental data sets. The used data sets in this paper are applied to a broadband stationary sound source organized into two categories of real and simulated data sets. The actual data from the signal recording is sampled in the natural environment. Synthetic data are created based on real data to confirm the various dimensions of the proposed method, reduced the limitations of the practical experiment, and allowed a broader investigation by changing the conditions of the problem. The presented angles in Table 2 are obtained by calculating the actual Azimuth and Elevation angles of the sound source coordinates of positions 1 to 6 of Table 3. In this article, we know the exact position of the source as the ground truth of data sets in advance to see the amount of error. Figure 3 shows a diagram of the actual signals received by the reference microphone of the sensor arrays 1 and 2 to present the simulated signs for the based microphone on the both sensor arrays.

Table 3 Sound sources coordinates

Source	Position		
Coordinate i	X $(meter)$	Y $(meter)$	Z $(meter)$
1	2	5	0
2	7.07	5	0
3	4.535	5	0
4	4.535	6	0
5	500	500	0
6	300	1000	0

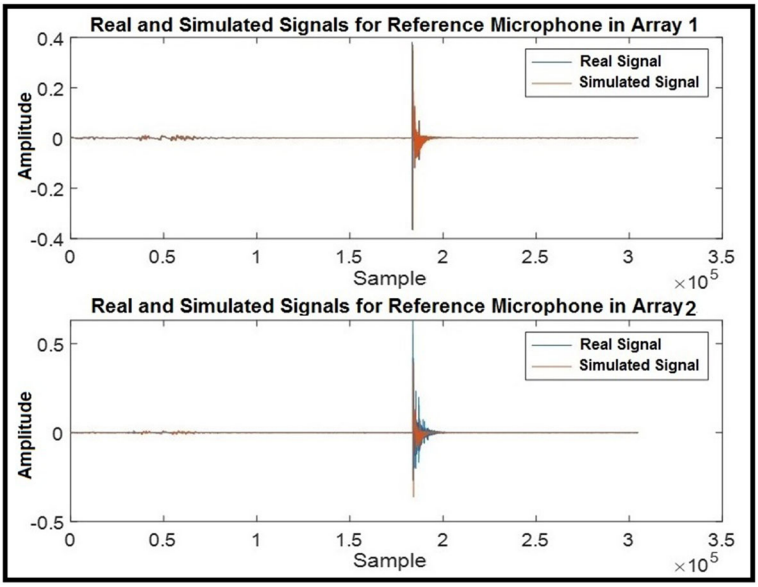

Fig. 3 The real and simulated signals for reference microphone in Array 1 and Array 2

Real accurate data set: Real data are sound signals recorded by data collection equipment. The location estimation error for the stationary source is evaluated in four different situations. The tested positions are considered at various points and distances. The sound source being tested is stationary and the cracker type, which causes an explosive shock signal. The site of the explosion is assumed to be a point in space, but we know that the site of the blast covers more area than that point. The speed of impulsive sound is 349.02 *meters* per second at a temperature of 30 degrees *Celsius*. The angle in the direction of the horizon and perpendicular to the four positions of the stationary source (location of the cracker explosion) tested in the coordinate system in terms of degrees relative to each sensor array, is shown in Table 2. This angle is derived from knowing the Cartesian coordinates of sound sources. In

Fig. 4 The microphone array topology of Array 1 and Array 2 for outdoor testing

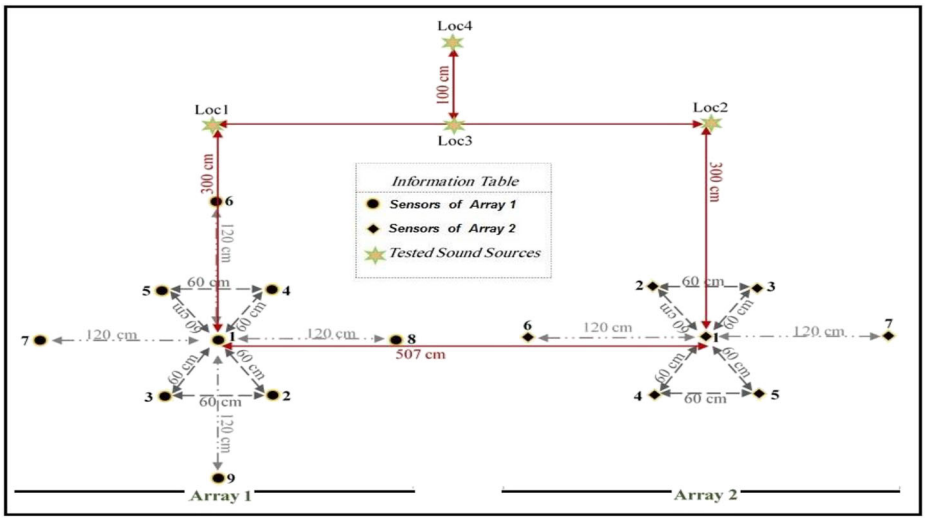

Fig. 5 The mesh topology and geometry of microphone array in Array 1 and Array 2

these experiments, two arrays with *seven* and *nine* identical microphone sensors were used. The structure of both microphone arrays is shown in Figs. 4 and 5. Array 1 contains *nine* microphones, and Array 2 consists of *seven* microphones that are arranged outdoors.

Simulated data set: Simulated data were used to create signs similar to real data, one of the recorded signs was first selected and simulated. To do this, we use Eq. (7) and, having the coordinates of the sensors and the sound source, we delay the single sound signal; this is done by the number of sensors in each array. Then, by utilizing the Monte Carlo simulation method, we add a Gaussian noise with a mean of zero to the delays of each of the sensor symbols of Array 1 and Array 2. Monte Carlo is repeated 100 times per experiment. Two other arrays were used to test the far-field, and the results for testing positions 5 and 6 are given in Table 2. Positions 1 to 4 are close for field-testing.

4.3 Validation method

We validate the LLS, CLS, CHAN, ITWLS, SUM, and KSUM methods on each of the used real datasets and take the real signals and keep the ambient noise to experiment with the *error rate*, *running time*, and system *performance* in testing to *optimize* the finding direction. We apply the Krill Herd model to design the KSUM algorithm for creating the TDOA-ASDF method in this way and measure the intermediate results to find better results than those of the original SUM method. The development of our validation method, and evaluation metrics, show the correct modeling and implementation.

4.4 Microphone array topology

We experimented with two arrays of 7 and 9 similar microphones that simultaneously receive and record sound in different scenarios. The topology of these arrays is based on the optimal arrangement geometry of the arrays in Fig. 5 and uses for mapping the best parallel topology for WSN. The *black* circles represent *Array 1*, the rhombuses represent *Array 2*, and the *yellow* stars represent the test site of the sound source. The arrangement geometry

Table 4 Coordinate system of Array 1 and Array 2 in near-field scenario

Microphone	Position					
	Array 1 [9 microphones]			Array 2 [7 microphones]		
i	X $(meter)$	Y $(meter)$	Z $(meter)$	X $(meter)$	Y $(meter)$	Z $(meter)$
1	2.0	2	0.34	7.07	2	0.34
2	2.3	2-$\sqrt{27}$	0.34	6.77	2+$\sqrt{27}$	0.2
3	1.7	2-$\sqrt{27}$	0.2	7.37	2+$\sqrt{27}$	0.2
4	2.3	2+$\sqrt{27}$	0.2	6.77	2-$\sqrt{27}$	0.2
5	1.7	2+$\sqrt{27}$	0.2	7.37	2-$\sqrt{27}$	0.2
6	2.0	3.2	0.2	5.87	2	0.2
7	0.8	2.0	2.0	8.27	2	0.2
8	3.2	2.0	0.2			
9	2.0	0.8	0.2			

of the microphones in the two arrays is separate, and the arrays are $507cm$ apart from the center microphone in the direction of the horizon ($x = 5.07m$). These two arrays can be considered as two parallel distributed arrays because they are both connected to a microphone array device for using in analog to digital conversion, sampling, and synchronization. The height of the *microphone 1* (*reference sensor*) of each array from the ground is $34cm$ ($z = 34cm$), and the height of the other microphones from the ground is $20cm (z = 20cm)$.

The arrangement coordinates of the arrays are presented in terms of *meters* in Table 4. Table 5 shows the coordinates of the array for far-field application. For simulated data, the sensor arrays shown in Table 4 are expressed for the first four positions of Tables 2 and 3, and the two sensor arrays of Table 5 for the effect of scattering the sensors in the arrays

Table 5 Coordinate system of Array 1 and Array 2 in far-field scenario

Microphone	Position					
	Array 1 [7 microphones]			Array 2 [7 microphones]		
i	X $(meter)$	Y $(meter)$	Z $(meter)$	X $(meter)$	Y $(meter)$	Z $(meter)$
1	28.1	28.5	03.4	500.0	0004.0	0.034
2	00.0	27.0	03.5	471.9	3.985	0.035
3	06.2	09.9	01.8	478.1	3.814	0.018
4	38.1	16.0	04.0	5.230	3.807	0.010
5	23.1	00.0	00.8	5.302	4.035	0.012
6	51.1	09.2	01.0	5.222	4.188	0.000
7	58.3	32.0	01.2	4.871	4.229	0.017

and their distances from each other for the positions of five and six sound sources. The position of the tested sound source for the simulated data of the sensor arrays is extracted from Table 5, positions 5 and 6 are shown in Tables 2 and 3. In the proposed method, sensor 1, or the reference, is placed at the origin of the coordinates, and the rest of the sensors and sources are obtained relative to it.

4.5 Empirical experiments

In this section, we present our empirical experiments of the LLS, CLS, CHAN, ITWLS, SUM, and KSUM methods in *Sample Output* for visualizing the used *SUM* and KSUM algorithms in the proposed TDOA-ASDF, *Quality of positioning* for believing the quality of angles output signals, and *Performance* for admission of the improved signal finding run-time for showing the significant computational complexity reduction and accuracy increasing of the TDOA-ASDF method.

4.5.1 Simulation approach

Experiments have been performed on real and simulated data in 3D space. The angle between the sensors is the basis of each array, and the source is estimated. The actual sound source is recorded once in each position, so there is no average or standard deviation. In the KSUM method, this algorithm is repeated 100 times in its second step, and the Krill Herd optimization has a standard deviation in Fig. 11. However, this standard deviation is so small and not valuable. In all graphs, the absolute errors are in degrees. The value of the

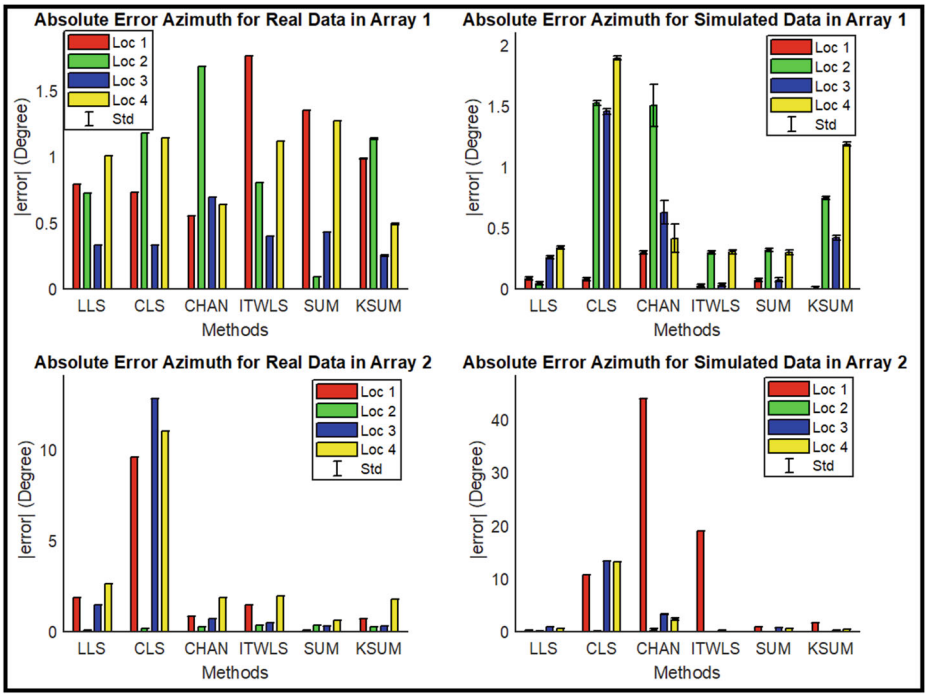

Fig. 6 Absolute error Azimuth angle in real and simulated data sets

absolute error of the estimated Azimuth angle in real data and the mean of absolute error and standard deviation of the Azimuth angle in simulated data relative to the two sensor arrays are shown in Fig. 6. Also, the value of the absolute error of the estimated Elevation angle in real data, and the average absolute error and standard deviation of the Elevation angle in simulated data relative to the two sensor arrays are shown in Fig. 7. Simulated data experiments in 3D space have been performed separately. The tables and graphs of the results are the mean and standard deviation in 100 repetitions of the Monte Carlo experiment.

In creating simulated data by inserting an error in the latency of each sensor, experiments have been performed on six different source positions. The amount of orientation error in the process of the proposed method is small. The four positions examined are the source for examining the correct operation of location algorithms. The fifth and sixth positions are designed to investigate the effect of source distance from sound arrays. Also, positions 5 and 6 are tested to examine the expansion of the algorithm from the near-field to the far-field. The average absolute error of the estimated angles of Azimuth (angle to the horizon) and Elevation (angle to height) in real and simulated data to eliminate the effect of the position of the source in the near-field and the far-field are shown in Figs. 8 and 10, respectively. Also, Fig. 9 shows the mean absolute error and the standard deviation of the Azimuth and Elevation angles in the simulated data set of positions 5 and 6 of the sound source in the far-field . Next, Figs. 12 and 13 show a comparison of the presence and absence of cubic spline interpolation in outdoor near-field data and simulated far-field data, respectively. In Fig. 11, the average of all four experimental positions of the sound source is calculated. In Fig. 12, the average of the two positions 5 and 6 of the sound source related to the sound source in the far-field is examined. Finally, Fig. 13 shows a comparison of the implementation times of the studied methods.

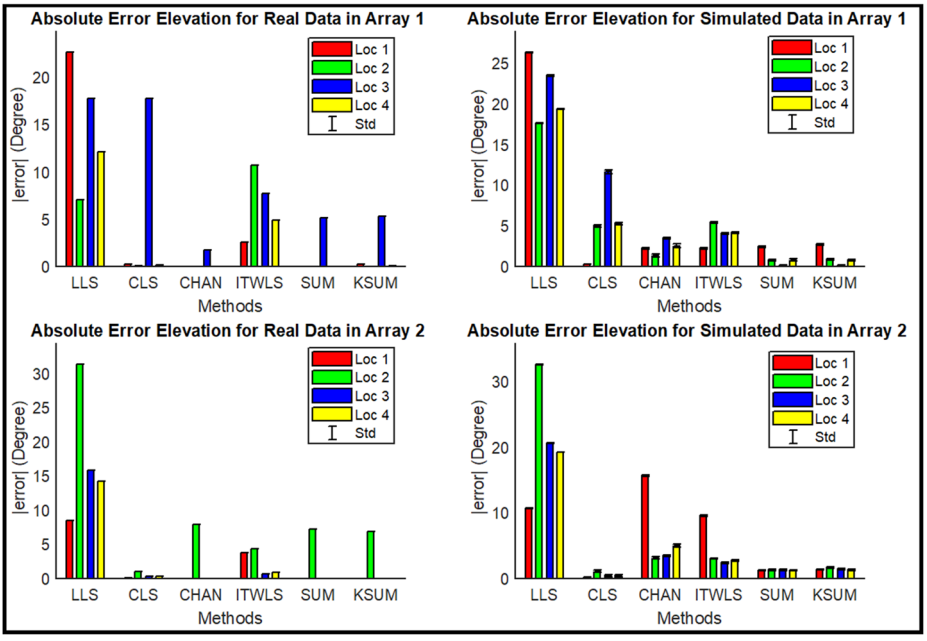

Fig. 7 Absolute error Elevation angle in real and simulated data sets

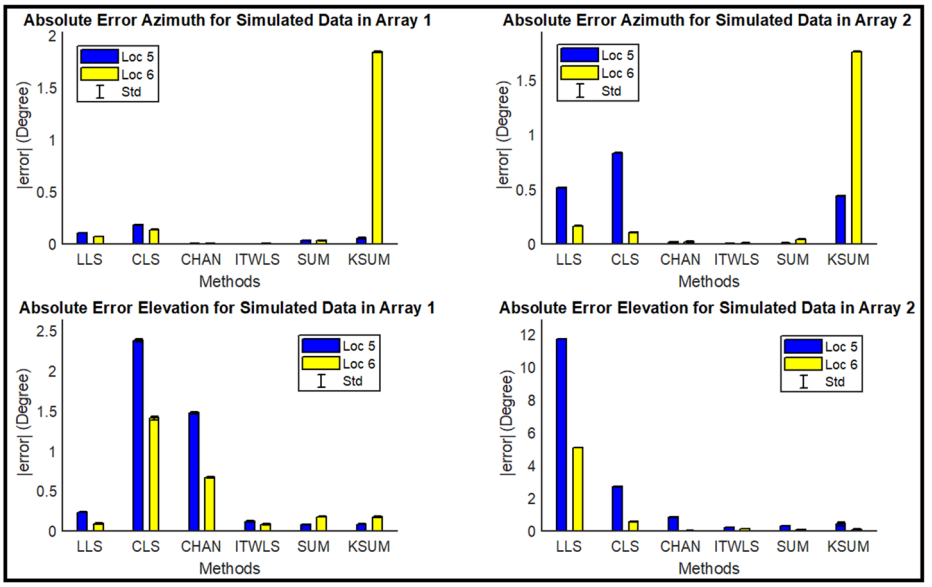

Fig. 8 Mean absolute error and standard deviation of Azimuth and Elevation angles in simulated data set in far-field

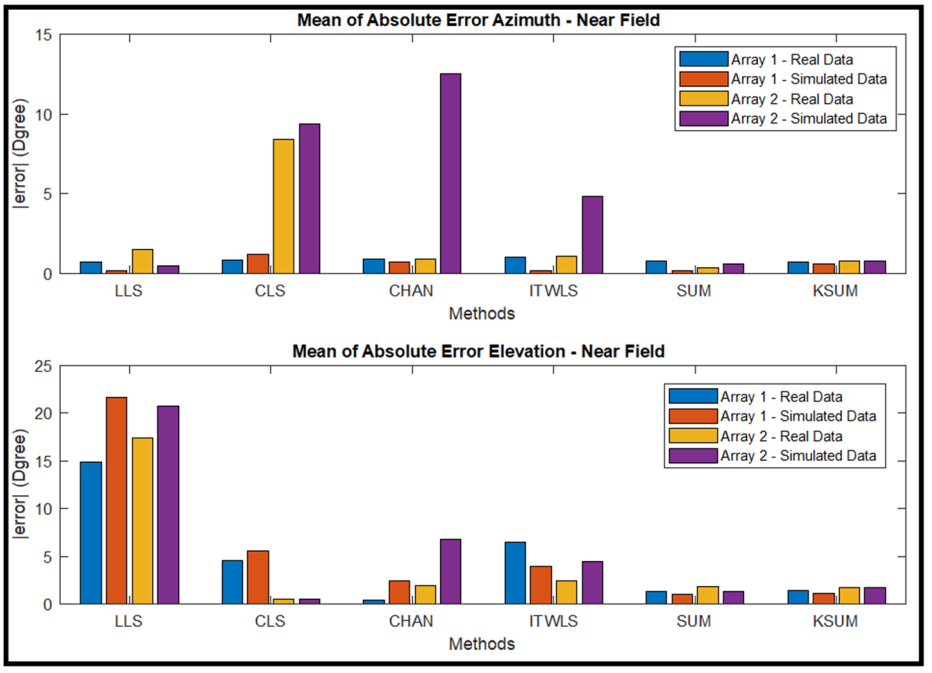

Fig. 9 Mean absolute error and standard deviation of Azimuth and Elevation angles in simulated data set in near-field

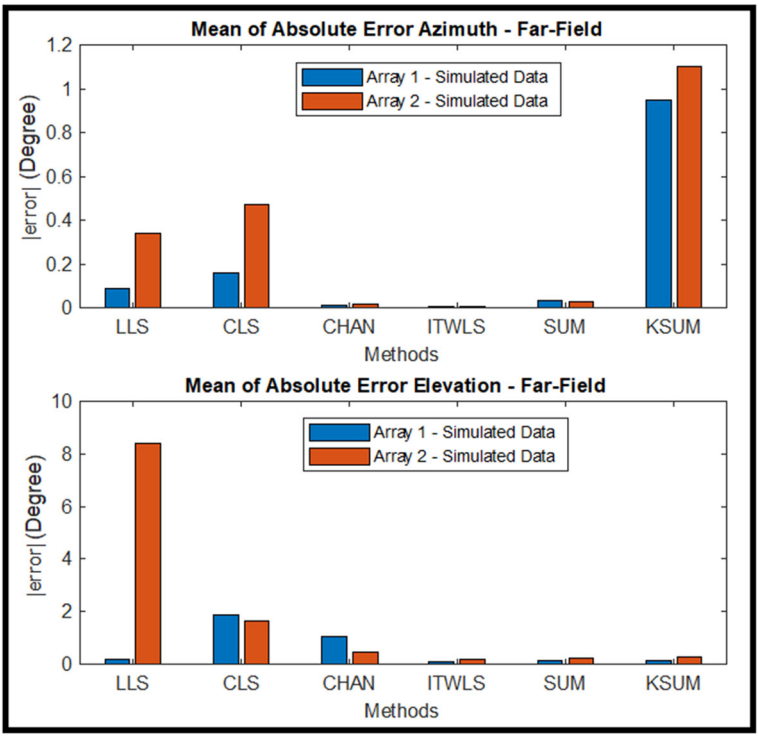

Fig. 10 Average absolute error of estimated angles of Azimuth and Elevation angles in real and simulated data to eliminate the effect of source position in far-field

4.5.2 Quality of direction finding

From the diagrams in Figs. 6 and 7 which are related to the 3D coordinate system and the positions 1 to 4 of the sound source in the near-field, by comparing the estimated Azimuth and Elevation angles, it is clear that the KSUM method, without consecutive constraint, is the most appropriate method with the least and the average variance among the studied methods. Also, according to Fig. 9, the KSUM, regardless of the position of the source, compared to the other methods in the near-field, has the lowest error rate. On the other hand, according to Figs. 8 and 10, which are related to simulated data in the far-field, the ITWLS method has a lower error rate; moreover, according to the scalability approach, the method must be responsive at close distance from the sound source to the sensors. The KSUM algorithm must not only operate at close distances to the sound source, but also work long distances away from the sound source, so the KSUM method is an excellent method to determine the direction of the sound source in the near- and far-field. In the far-field, the KSUM does not have very accurate Azimuth results, and the reason is that if we reduce the limitations of the Krill Herd algorithm, we will achieve better results. However, the angle error is about 1°, and this value is small and is ignored. The distance % error is the minimum error because our direction was measured without wind.

4.5.3 Interpolation effects

According to Figs. 11 and 12, the interpolation of the cubic spline, which has the most significant effect on the estimation of the difference in the arrival time difference, is effective in reducing the average absolute error in Azimuth and Elevation, especially in the sensor *Array 1*. In some cases, the Azimuth and Elevation angles without interpolation, are less error-prone. Still, overall the reduction in the Azimuth and the Elevation side-by-side is evident despite interpolation. But sometimes, this is not the case, and this is due to the lack of the pre-processing and elimination of noise, whose presence harms interpolation. The method of estimating the difference in arrival time, using cross-correlation with the smoothness of the phase spectrum, is one of the most common methods for determining the position of sound sources. On the other hand, by introducing the correlation spectrum in the time domain, a new and more efficient correlation peak can be achieved. Placing the reference sensor at the origin of the coordinates increases the algorithm's speed, and in other articles was not observed to reduce the complexity of the estimation.

The lack of smoothness at the beginning of the proposed algorithm, makes interpolation for the data around the cross-correlation peak, due to the presence of noise, difficult to select a new correlation peak and leads to the estimation of coordinates with more error than the absence of interpolation. However, the proposed method without any pre-processing to remove noise has good accuracy. The Azimuth angle typically has more minor errors than the Elevation angle. Therefore, the method is much more effective in 2D or 3D space and has excellent estimation accuracy. In general, in both near- and far-field, estimating the direction in 2D is more accurate than in 3D, and at the same time, the execution time of the

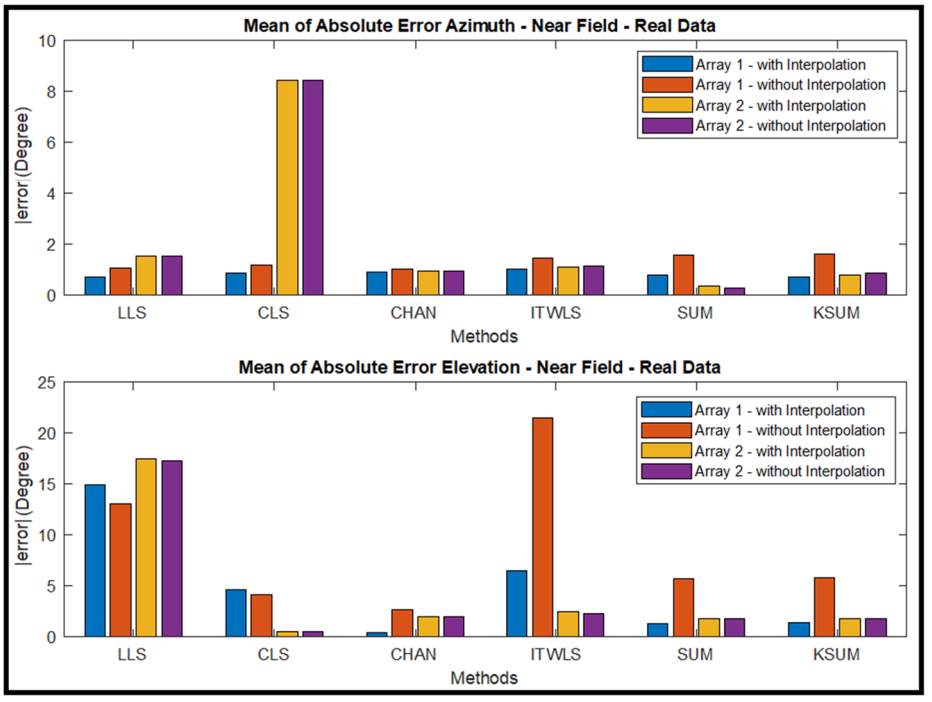

Fig. 11 Mean and standard deviation of absolute error of Azimuth and Elevation angles in real data set to investigate the effect of interpolation in near-field

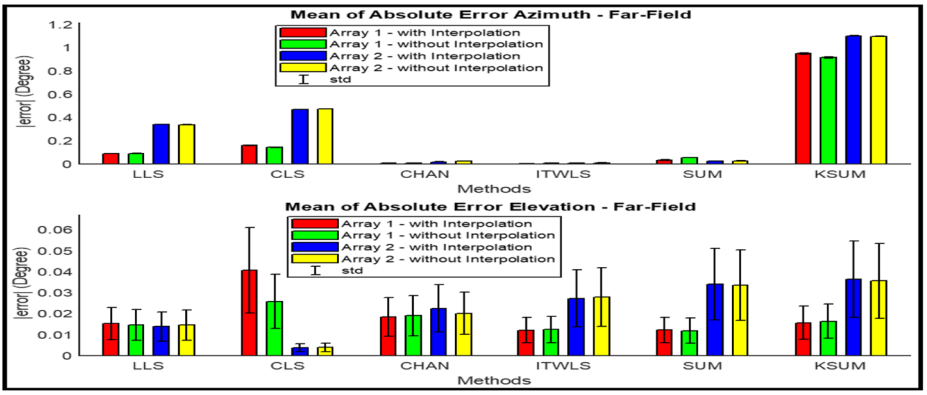

Fig. 12 Mean and standard deviation of absolute error of Azimuth and Elevation angles in simulated data set to investigate the effect of interpolation in far-field

algorithm, in 2D is less than the 3D. Connecting the execution time of using the method in Fig. 13, it is clear that the execution time of the proposed method with using interpolation (in Fig. 13(a)) and without using interpolation (in Fig. 13(b)), the TDOA estimation is almost the same and does not differ much. On the other hand, in the direction estimation using the proposed algorithm compared to different algorithms, KSUM has more time due to 100

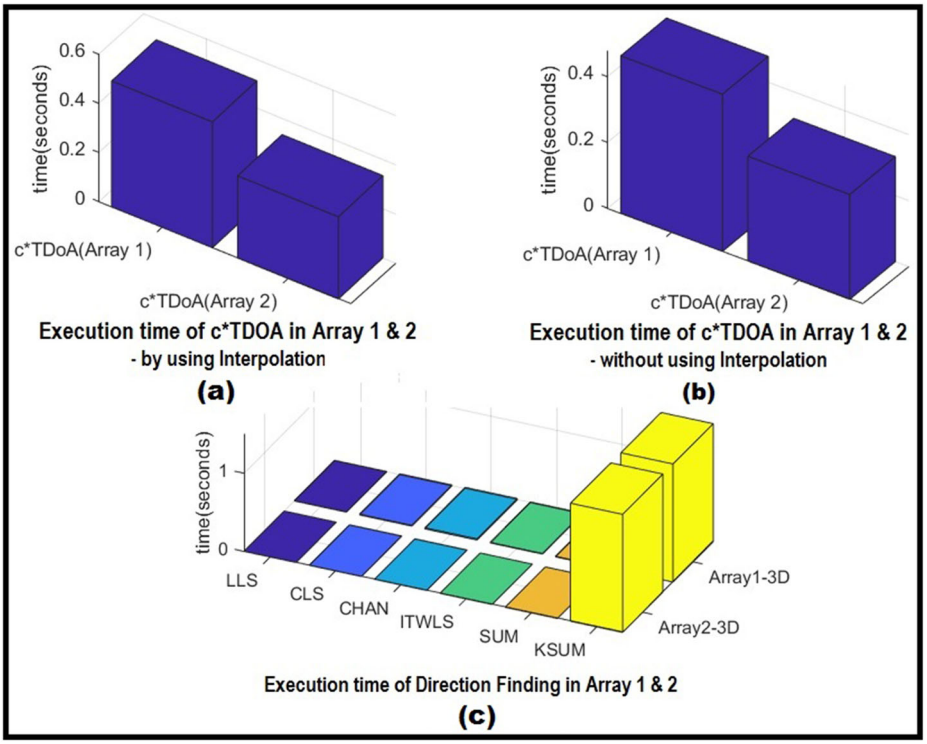

Fig. 13 Comparison of the execution time of studied direction finding methods

repetitions but relatively little time that is close to real-time and acceptable; therefore, its implementation does not have much overhead in real-time response.

4.5.4 Performance

We compared the performance of the proposed TDOA-ASDF (KSUM) with the primary and state-of-the-art TDOA-based methods and achieved the best performance, especially compared to the SUM method. This experiment has been repeated 100 consecutive times for each signal analyzed, and the average has been calculated. It can be seen that the real-time method has a minimum run-time and maximum accuracy for solving the problem. In Fig. 13(c), the execution time of the studied methods indicates that the maximum execution time is related to estimating the difference in arrival time. Moreover, as the number of microphones increases, the execution time increases. Also, in practice, increasing the number of sensors increases the estimation accuracy. In addition, in estimating the direction and solving hyperbolic equations, the minimization method, without consecutive constraints, has a short execution time. The KSUM has more execution time than the other algorithms, which is only due to 100 repetitions in the Krill Herd algorithm and its optimization. But this time is relatively short and prompt. *Array 1* has *nine*, and *Array 2* has *seven* sensors, so *Array 1* has a longer execution time than *Array 2*. The actual arrangement of the sensor arrays is shown in Fig. 4, which is not suitable for very long distances (far-field) due to the compactness of the topology. The layout of the sensor array in Table 5 is suitable for far-field sound sources.

Since KSUM is an iterative algorithm, a trade-off between the accuracy and the number of iterations was made, concluding that with a longer step of the analyzed signals from an unknown source position, we achieve even higher accuracy. The KSUM method, compares to the SUM base method, improving the accuracy of the direction-finding, has been able to execute run-time in the near-field and the far-field. Also, the method requires more calculations and has reduced the computational complexity in the matrix operations. So, we have been able to achieve suitable improvements in our used hardware equipment and method. Also, the measured running times of the implemented algorithms in the MATLAB tool show that our method can be a candidate for applying in RCPS and be better in light programming on IoT, IoE, and IoMT objects. The obtained experimental results indicate that the proposed method can be used to determine the direction of the stationary sound source with good precision in a free-field environment.

5 Conclusion

Accurate and fast passive sound source direction-finding, which is based on TDOA, suffers from the challenges of selecting the coordinate origin, source mobility, source distance, source size and height, microphone array topology, array architecture, number of the microphones in the array, computational complexity, and 3D positioning coordinates. Also, other challenges are there, such as direction accuracy, orientation speed, presence of ambient noise, generated artificial noises, sound reflections, environmental sounds and conditions, and the speed rate of the mobile source. The accuracy of the predictions can be improved by integrating microphone features and architecture, sound source information, indoor or outdoor environmental properties, optimizing the spacing between the microphones, and robust sensors in free-field. We need to provide a fast acquisition rate per channel, which is necessary for capturing sound signals and performing cross-correlation with high accuracy. The maximum latency of the physical hardware system is very important. The increase in

the number of microphones to estimate the direction leads to improved accuracy where noise exists. The multidimensional acoustic sensor array positioning is not only related to the number of microphones, but also has a critical relationship with the topology of the sensors. The performance of direction-finding increases by integrating more information, array architecture, microphone array topology, and sound source information. The proposed scalable method eliminates pre-processing, and is resistant to ambient noise and reflection problems, real-time, undetectable, fast, accurate, and is applicable in 2D and 3D positioning, which can significantly reduce the running time through complexity reduction by combining the interpolation technique and designing a novel algorithm.

Experimental results from different datasets showed that the error amounts of Azimuth and Elevation are 0.7164° and 1.5054° in the near-field, and are 1.0260° and 0.2071° in the far-field, respectively. The distance % error showed great dependency on the wind direction where the minimum error occurred when the wind direction was the same or opposite to the measurement direction. The real-time direction determination speed is 1.935 *Sec*, and the accurate estimation of the central position of the source is high. The maximum execution time is related to estimating the difference in arrival time. The obtained experimental results indicate that the proposed method can be used to determine the direction of the stationary sound source with good precision in a free-field environment and will be robust to sound reflection and ambient noise. For future works, due to the strong effect of estimating the time delay between each pair of microphones in the array in determining the source position, we try to reduce the error in estimating the time delay and the effect of signal window type and reduce the time delay estimate between each pair of microphones. Also, the following important tasks will be applying the method in the near-field, improving the performance of far-field direction finding, eliminating the generated computational noise, integrating more information, designing the best microphone array topology, and decomposing the generation model into a speed robust model. In addition, there is a need to limit the use of microphones in the setting to make the compaction of the system possible, complexity reduction of the method, and minimization of resource consumption possible in an operational system.

Declarations

Conflict of Interests This Article has been written by the stated authors, is original and has not been previously published. They have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Abiri A, Pourmohammad A (2020) The bullet Shockwave-Based Real-Time sniper sound source localization. IEEE Sensors J 20(13):7253–7264
2. Bai Y et al (2020) Acoustic-based sensing and applications: A survey. Comput Netw 181:107447
3. Berdugo B et al (1999) On direction finding of an emitting source from time delays. Acoust Soc Am 105(6):3355–3363
4. Chan YT, Ho KC (1994) A simple and efficient estimator for hyperbolic location. IEEE Trans Signal Process 42(8):1905–1915
5. Chen JY et al (2018) Improved two-step weighted least squares algorithm for TDOA-based source localization. 19th International Radar Symposium (IRS) IEEE: 1–6
6. Chen X et al (2019) Bias reduction for TDOA localization in the presence of receiver position errors and synchronization clock bias. EURASIP J Adv Signal Process 2019(1):1–26
7. Cui X et al (2018) Approximate closed-form TDOA-based estimator for acoustic direction finding via constrained optimization. IEEE Sensors J 18(8):3360–3371

8. Cui X et al (2019) Azimuth-only estimation for tdoa-based direction finding with 3-d acoustic array. IEEE Trans Instrum Meas 69(4):985–994
9. Dang X et al (2022) TDOA-based robust sound source localization with sparse regularization in wireless acoustic sensor networks. IEEE/ACM Trans Audio Speech Lang Process 30:1108–1123
10. Foy WH (1976) Position-location solutions by Taylor-series estimation. IEEE Trans Aerosp Electron Syst 2:187–194
11. Germán F. et al (2020) Design and implementation of acoustic source localization on a low-cost IoT edge platform. IEEE Trans Circuits Syst II: Express Br 67(12):3547–3551
12. Gillette MD, Silverman HF (2008) A linear closed-form algorithm for source localization from time-differences of arrival. IEEE Signal Process Lett 15:1–4
13. Gombots S et al (2021) Sound source localization state-of-the-art and new inverse scheme. Elektrotech Inftech 138:229–243
14. Grondin F, Michaud F (2019) Lightweight and optimized sound source localization and tracking methods for open and closed microphone array configurations. Robot Auton Syst 113:63–80
15. Heydari Z et al (2021) Speed impulsive sound source direction finding method based on TDOA. The CSI J Comput Sci Inf Technol (JCSE) 19(1):23–35
16. Heydari Z et al (2021) The effect of energy and interpolation combination on the accuracy of indoor localization based on the speech signal time difference of the arrival. J Acoust Eng Soc Iran 8(2):13–21
17. Invernizzi D et al (2020) Comparison of control methods for trajectory tracking in fully actuated unmanned aerial vehicles. IEEE Trans Control Syst Technol 29(3):1147–1160
18. Jahanshahi H et al (2019) Robot motion planning in an unknown environment with danger space. Electronics 8(2):201
19. Jin B et al (2018) Robust time-difference-of-arrival (TDOA) localization using weighted least squares with cone tangent plane constraint. Sensors 18(3):778
20. Liu M et al (2020) Indoor acoustic localization: A survey. HCIS 10(1):1–24
21. Manamperi WN et al (2022) Drone audition: Sound source localization using on-board microphones. IEEE/ACM Trans Audio Speech Lang Process 30:508–519
22. Mirbeygi M et al (2021) RPCA-based real-time speech and music separation method. Speech Comm 126:22–34
23. Mirbeygi M et al (2022) Speech and music separation approaches - a survey. J Multimed Tools Appl 1(2):1–43
24. Peng Y et al (2021) Research on positioning accuracy of passive acoustic positioning system based on feature matching in air. J Phys 1952(4):042052
25. Seidel A et al (2021) Scenario-based analysis of the carbon mitigation potential of 6G-enabled 3D video conferencing in 2030. Telematics Inform 23:101686
26. Shi Z et al (2020) An acoustic-based surveillance system for amateur drones detection and localization. IEEE Trans Veh Technol 69(3):2731–2739
27. Smith J, Abel J (1987) Closed-form least-squares source location estimation from range-difference measurements. IEEE Trans Acoust Speech Signal Process 35(12):1661–1669
28. Sudo Y et al (2021) Multi-channel Environmental Sound Segmentation utilizing Sound Source Localization and Separation U-Net. IEEE/SICE International Symposium on System Integration (SII): 382–387
29. Sun Y et al (2018) Solution and analysis of TDOA localization of a near or distant source in closed form. IEEE Trans Signal Process 67(2):320–335
30. Wu P et al (2019) Time difference of arrival (TDoa) localization combining weighted least squares and firefly algorithm. Sensors 19(11):2554
31. Yan T, Zhang Y (2018) TDOA Time delay estimation algorithm based on cubic spline interpolation
32. Zhou Z et al (2020) A closed-form method of acoustic emission source location for velocity-free system Using Complete TDOA Measurements. Sensors 20(12):3553
33. Zou Y, Liu H (2020) A simple and efficient iterative method for TOA localization. In: IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), pp 4881–4884