AI-WSN: Direction of Arrival Estimation Based on Bee Swarm Optimization for Wireless Sensor Networks

Devika E *

*Corresponding Author, Research Scholar, Department of Computer Science, Sree Saraswathi Thyagaraja College, Pollachi, India-642 107. E-mail: devikache6@gmail.com

Saravanan A

Associate Professor, Department of Computer Science, Sree Saraswathi Thyagaraja College, Pollachi, India-642 107. E-mail: a.saravanan21@gmail.com

Abstract

An Artificial Intelligence (AI) technique plays the most crucial factor to consider in energy utilization in a wireless sensor network (WSN). AI transforms industrial operations by optimizing the energy consumption in sensor nodes. As a result, it is crucial for improving sensor node location accuracy, particularly in unbalanced or Adhoc environments. Because of this, the purpose of this research is to improve the accuracy of the localization process in locations where sensor nodes encounter barriers or obstacles on a regular basis. The Bees Swarm Optimization (BSO) algorithm is used to segment sensor nodes in order to increase the accuracy of the Direction of Arrival (DoA) estimate between the anchor and unknown node pairs. Even in the presence of unbalanced conditions, the proposed DoA-BSO involving three separate bee colonies can identify plausible anchor nodes as well as segment nodes arranged in clusters. In order to obtain the intended result, the objective function is designed to take into consideration the hops, energy, and transmission distance of the anchor and unknown node pairs, among other factors. The studies are carried out in a large-scale WSN using sensor node pairs in order to determine the precision with which the DoA-BSO can be located. When comparing DoA-BSO to conventional approaches, the findings of the meta-heuristic algorithm show that it improves the accuracy and segmentation of nodes significantly.

Keywords: Wireless sensor network, Direction of arrival, Bees Swarm Optimization, Energy estimation.
Introduction

Increasingly important in a wide range of industries and applications, wireless sensor networks (WSNs) have gained traction due to their ability to revolutionize our surroundings (Wu et al., 2019). These networks include everything from healthcare and security sensors in the home and workplace to military applications in the field (Ullah et al., 2019). Nodes, which are the smallest components of a WSN, make up the vast bulk of the network hardware because they are so small (Datta and Dasgupta, 2021). Because of the limited resources available to a sensor node. For the network’s long-term viability, it is necessary to consider all aspects of energy conservation (e.g., leakage of power). RF transmissions consume approximately ten times the amount of energy that computation consumes (Hao et al., 2020).

When it comes to detecting events via nodes, WSNs are frequently deployed in a random fashion across large geography with the purpose of events detecting on sensor nodes. A short-range wireless transmission system is required for transferring information between nodes until they reach one or more sink nodes where the information can be reported. As has been demonstrated (Dai et al., 2020), routing systems play an important role in multi-hop networks. Using the shortest path between two sensor nodes can aid in energy conservation in the overall network. Data reporting, on the other hand, cannot be considered complete unless the source of any recognized events can be determined (Wang et al., 2019).

To put it another way, the time-location parameter is more important for a broadcast event than the actual sensor identification itself because it is more reliable (Wang et al., 2020). In order to successfully execute a WSN application, the positions of the sensor nodes must be known in advance (Zhang et al., 2021). As a result, the localization stage is always present in the majority of WSN applications. This stage necessitates the usage of RF communications with the surrounding sensors as well as specific algorithms in order to determine the geographic location (Vashistha and Law, 2019). As part of the DV-Hop scheme, unknown sensors are promoted to the position of anchor nodes (Diez et al., 2019). The localization is based on RSSI estimates of their distances from the real anchors.

As a result, sensor placements can be influenced by hardware defects, changing environmental conditions, network topologies, and node densities in distributed algorithms that deal with real-world multi-hop scenarios. For example, the issue of localization in severe multipath environments is exacerbated by the changing environment. It takes advantage of non-line-of-sight to optimize a simple probability for guessing the locations of unknown sensors in order to estimate their locations.

Unique extensions for ToA measurements are used to handle complex radio propagation signals in order the identification of unknown sensor nodes, and these extensions are only available from this vendor. This approach provides excellent precision at a low cost in terms
of hardware and computing, making it an attractive option. In order to remove iteratively dubious weak range values, a robust least squares algorithm (RLS) is used to recalculate node placements. In our research, we are pursuing a different technique in which the outlier degree of each measurement is mainly controlled via weights, which has the potential to reduce their influence on the solution to the issue. For example, the application of a robust statistical model during the estimation procedure reduces the number of outliers.

Therefore, range-based iterative approaches are extremely sensitive to network topologies, necessitating the employment of robust methodologies to reduce location errors to an absolute minimum. (Li et al., 2020) described, it is frequently necessary to deal with nonlinear and non-convex optimization issues when developing a mathematical localization method.

In this paper, we develop a DoA-BSO-based robust localization technique that is based on BSO. Our experiments reveal that our technique functions effectively in multi-hop networks when sensor node distributions are not uniform. Compared to similar systems, this one outperforms them both in terms of accuracy of location estimate and the number of repetitions necessary.

**Literature Review**

In order to obtain preliminary estimates of the DOA, the authors used a modified evolutionary population dynamics-based distributed Grey Wolf Optimization (GWO) approach established by (Maruthi et al. 2020) to obtain the first estimations of the DOA. For applications where accurate DOA calculations are essential, NM faster convergence and lower communication costs are a significant advantage. By utilizing a distributed hybrid strategy, it is possible to develop a method for estimating DOA that is both energy-efficient and accurate.

(Kumar, S., & Zaveri, M. 2018) described an approach for identifying an event that makes use of a parametric estimation-based direction of arrival technique to locate the event. Aiming to predict the location of events and evaluate their effectiveness in an Internet of Things context, this project will deploy devices in a quasi-random manner in order to do this.

A new collaborative beamforming approach has been developed by (Ben Smida, O., et al. 2019) to be used over dual-hop transmissions via the K-node WSN that is robust to substantial channel estimation flaws. Initially, the signal is received by the WSN directly from the source. Then, before passing the signal to the next node, each node multiplies the received signal by the beamforming weight that it has calculated. Its purpose is to limit the amount of noise that is reflected back to the receiver while maintaining the appropriate power level of the signal sent. In some cases, CSI considerations have an impact on how these weights are calculated. In order to eliminate channel estimate errors that could have a detrimental effect on CB performance, they must be evaluated locally at each node. The fact that our unique RCB
approaches do not require any information transmission between nodes implies that they are scattered, which means that they enhance WSN spectral and power efficiency by orders of magnitude greater than earlier systems.

The distributed ML technique employed by (Prasad, M. S., & Panigrahi, T. 2019) to improve the accuracy of DOA estimation has improved the accuracy of DOA estimation. Improved DOA estimates from surrounding nodes can assist in obtaining a more accurate estimate of its own DOA estimate. As a result, it was discovered that the disseminated version of the CRLB performed significantly better than the original CRLB. There are two types of cluster resource locators (CRLBs): local CRLBs for each node subarray and global CRLBs for the WSN array in its entirety. According to the findings, even for coupled signals with the highest CRLB, simulation results suggest that the estimate accuracy improves at a node that employs a distributed technique.

(Sharma, A., & Chauhan, S. 2020) successfully solved the passive intruder detection in mobile vehicles. The presence of a mobile intruder can be deduced through the use of a hierarchical structure with three levels of hierarchy. K-mean clustering is utilized to organize the sensor nodes that have been deployed from the very beginning. When determining detection probability, the best threshold value is determined by performing a numerical analysis of the signals that have been received. Because of the new fusion rule, the possibility of detection is improved while the number of false alarms is maintained at a bare minimum. The correctness of the suggested fusion rule will be decided by the number of actual detections recorded.

(Shenadeh, et al., 2018) the suggested approach, Multi-Objective Optimization Algorithm Based on Sperm Fertilization Procedure (MOSFP), as well as three additional state-of-the-art multi-objective optimization algorithms, OMOPSO, NSGA-II, and SPEA2. The algorithms are given different packet payload sizes to work with, and the best value is determined. (Sampath Kumar et al., 2020) has developed the energy consumption of sensor nodes, this LBR-GSO uses a pseudo-random route-finding method and an upgraded pheromone trail-based updating strategy. To optimize route establishment, it employs an effective heuristic update method based on a cost-effective energy metric.

(Álvarez et al., 2019) proposed a study of the applicability of these new systems based on a Cramér-Rao Lower Bound (CRLB) evaluation was conducted for various sensor sites in distinct 3D real situations. (Aubry et al., 2020) analysis was carried out CRLB evaluation for 2-D localization. (R. Rajakumar et al. 2017) was developed task is to use anchor nodes in the WSN to find the geographical location of unknown nodes. In this paper, the GWO algorithm is used to detect the right position of unknown nodes, hence resolving the node localization problem. (Yankui et al., 2019) provided a multiple screening-based K-means clustering
algorithm that can efficiently localize numerous sources based on DOA (direction of arrival) parameters.

Based on the above survey the DoA-BSO is developed in our research to overcome some of the issues in identifying the accurate path of routing and placing the sensor nodes inside the clusters has been discussed in the proposed method.

Methodology

In this paper, we develop a DoA-Bees Swarm Optimization (BSO) for a robust localization technique that is based on BSO. This model when applied in a distributed fashion, allows each unknown sensor to employ the reduction of localization errors in position predictions.

Energy Model

The converter, power amplifier, and radio receiver are the components of the node that consume the majority of its resources. Depending on the distance between the transmitter and receiver, the design makes use of free space and fading flow to communicate.

The quantity of energy consumed by sensor nodes is proportional to the square of the distance between them \((d^2)\). This condition is present if the propagation distance \(d < d_0\) (threshold distance) at any point in time. In reality, the distance between two points is inversely proportional to \(d^4\).

Calculating the greatest amount of energy required to deliver an \(l\)-bit packet over a certain distance between transmitter and receiver can be done using Equation (1) as follows:

\[
E_T(l,d) = \begin{cases} 
    lE_{elec} + l\varepsilon_{fs}d^2 & d < d_0 \\
    lE_{elec} + l\varepsilon_{mp}d^4 & d \geq d_0
\end{cases}
\]

(1)

Where

\(E_{elec}\) - energy consumed by transmitting node or receiving node over a \(l\)-bit data, and

\(\varepsilon_{fs}\) - amplifier coefficients of free-space model and

\(\varepsilon_{mp}\) - amplifier coefficient of multi-path fading model.

With the use of Equation (2), the shortest possible distance \(d_0\) between two sensor nodes can be found.
Here, the free-space model is used with the amplifier parameter $\varepsilon_{fs}$, which results in the consumption of energy by the sensor node in $d<d_0$ situation.

When $d \geq d_0$, the energy consumption of the sensor nodes is calculated using the multi-path fading model, with $\varepsilon_{mp}$ serving as the amplifier value.

It is possible for nodes within a cluster to communicate data to and from each other as part of the uneven dynamic clustering process based on the BSO given here because there is no constraint placed on the total length of node transmissions.

**DoA Estimation**

The Bees Swarm Optimization (BSO) meta-heuristic is proposed for continuous functions that are multivariable and multimodal in nature. The BSO meta-heuristic has gained widespread attention in the research community due to its low computational complexity and the use of a small number of control parameters. In addition, it is important to note that the optimization performance of ABC is comparable to that of existing metaheuristics.

In ABC, there are three types of bees: hired bees, observers, and scouts. The bee population is organized in the following ways: The employed bee does a waggle dance in order to communicate information about the location and quality of a food source with the spectator bee who is watching. The spectator bee then selects a food source based on a probability function connected to the quality of the food source, while the scout bee forages at random throughout the colony.

A food source will be completely depleted when all hired bees have completed their work and have returned to their previous responsibilities, which will result in a complete reversal of the roles. In the ABC meta-heuristic, which assigns each food source to a plausible solution to the optimization problem, one bee can be used for each conceivable solution to the optimization problem.

Using the ABC meta-heuristic approach, the following sections will break it down into its component parts: The converter, power amplifier, and radio receiver are the components of the node that consume the majority of its resources. Depending on the distance between the transmitter and receiver, the design makes use of free space and fading flow to communicate.
• Initialization Phase

A D-dimensional actual set of vectors is used as the starting point for the ABC meta-heuristic, which then generates a random Population Number (PN) from there. For example, suppose that \( x_{ij} = \{x_{i1}, x_{i2}, \ldots, x_{iD}\} \), where \( j = 1, 2, \ldots, D \) represents the \( i \)th food source.

\[
x_{ij} = x_{\text{min}j} + \text{rand}(0, 1)(x_{\text{max}j} - x_{\text{min}j})
\]

where

\( x_{\text{min}j} \) - lower limits

\( x_{\text{max}j} \) – upper limits

• Employed Bee Phase

With the expression: \( x_{ij} \), each bee is given a new solution, \( v_{ij} \), to work with.

\[
v_{ij} = x_{ij} + \Phi_{ij}(x_{ij} - x_{kj})
\]

where

\( k \) - Obtained randomly from \{1,2,\ldots, SN\} and

\( \Phi_{ij} \) - uniform random number \([-1, 1]\).

If the new answer is deemed to be more appropriate in terms of \( v_{ij} \) value than the prior one, Bee will forget the old one and remember \( v_{ij} \). The usage of \( x_{ij} \) as an old solution will continue as long as this issue is not solved.

• Onlooker Bee Phase

The waggle dance is performed by the bees in the hive to communicate to onlookers the nectar information about their food source. Following that, they chose a food source depending on the probability \( p_i \) in the following manner:

\[
p_i = \frac{f_i}{\sum f_i}
\]

where

\( f_i \) - fitness of \( x_{ij} \).

When \( x_{ij} \) is no longer available, a new solution with a higher fitness level will be introduced to the food sources that the observer bee chooses.
- **Scout Bee Phase**

In the event that a solution cannot be improved any further, the food source is abandoned, and the bee that was previously employed becomes a scout once more. After that, the scout will generate a random food supply for you to use in your quest.

**Algorithm for Cluster Head Selection**

**Step 1**

In this step, sensor nodes are placed throughout the area in order to begin the network configuration. In addition, information about the nodes, such as their distance from the base station and their energy level, is gathered. During the course of the network, each node delivered an advertisement message to the base station, which was received by the base station.

**Step 2**

The base station assigns one employee bee to each cluster head in order to evaluate the cluster head fitness. The fitness value is estimated for the cluster head by the worker bee.

\[
Fit_i = \eta e_i + \frac{\lambda}{n-1} \sum_{k=1}^{n} e_k \|d_{ik} - d_{ave}\| \\
\]

\[
d_{ave} = \frac{1}{n} \sum_{i=1}^{n} \sum_{k=1 & k \neq i}^{n} d_{ik} \\
\]

Where

- \(Fiti\) - fitness value of cluster head \(i\).
- \(\eta\) and \(\lambda\) - energy factors and it is expressed as \(\eta + \lambda = 1\), \(\eta, \lambda \in [0,1]\)
- \(i\) - cluster head,
- \(k\) - neighbour node,
- \(e_i\) - residual energy of the selected cluster head,
- \(e_k\) - residual energy of the kth neighbour node,
- \(d_{ik}\) - distance between the kth neighbour and ith CH and \(d_{ave}\) - distance between node and CH.
The residual energy of the node is calculated with the help of the first-order radio model. The spectator bees use a probability value that is defined in the third phase to determine which cluster heads are the most promising.

**Step 3**

Bees on cluster heads that have been randomly selected exchange physical fitness information with other bees in the surrounding region. In order to identify a cluster head with a high fitness value, an observer bee measures the fitness data of every bee in the colony with the help of a computer program.

\[
P_i = \frac{F(\theta_i)}{\sum_{k=1}^{s} F(\theta_k)}
\]

where

- \( F(\theta_i) \) - fitness of CH \( i \)
- \( F(\theta_k) \) - fitness of CH \( k \)
- \( i \) and \( k \) - randomly selected CH and
- \( s \) - total number of sensor nodes.

It is also determined what the lower and upper bounds of the fitness value

\[
Upper \ Range = \frac{\sum_{i=1}^{n} fit_i}{n}
\]

\[
Lower \ Range = \sqrt{\frac{\sum_{i=1}^{n} fit_i}{n}}
\]

Those cluster leaders will next be selected as the best candidates for the position of final cluster leaders. A technique known as cluster head selection was used to ensure that all sensor nodes in the network were able to submit their data to the sink, and it was successful.

**Cluster Formation**

Nodes in the network broadcast the results of each round of the selection process for cluster leaders to the other nodes in the network during the selection process for cluster leaders. In order to select the cluster to which each sensor node wishes to be added, the cluster head with the lowest energy consumption and the shortest distance between sensor nodes is picked as
the cluster to which each sensor node wishes to be added. In either case, data can be transported directly from a member node to the sink, or it can be routed to the cluster heads before being transmitted to the sink itself. As a result, there is a significant reduction in the volume of data traffic. When data is transferred from a member node to a cluster head, and subsequently to a sink, the amount of energy consumed during the transfer is significantly reduced.

The following are the stages that must be completed in order to create a cluster.

**Step 1:** Calculate the distance $d(n_i, CH_j)$ between each node $n_i$ and each cluster head $CH_j$

$$\text{Avg} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} d(n_i, CH_j)}{M}$$  \hspace{1cm} (11)

where

$N$ - network size, and

$M$ - number of selected cluster heads.

**Step 2:** The second step in the process involves taking an average of the distances between the sensor node and a specified CH.

**Step 3:** The sensor nodes choose the CH that is closer to them than the average distance. If more than one node fits this requirement, the CH with the largest quantity of remaining energy should be chosen. As a result of this, all of the nodes in the network are now connected to their respective CHs.

In order to maintain an equal number of members in each cluster, only a limited number of nodes can be used as members of a cluster. The node approaches a CH and accepts membership in the cluster as a member in order to form clusters. Any time a node’s stable factor is exceeded, the node begins searching for the next available CH. This is the method by which the clusters are created. Using this method of clustering, data is transported from a member node to a sink node while consuming the least amount of energy possible.

**Results and Discussion**

The study analyses the performance metrics of the proposed and present techniques in this section based on the network assumptions given in Table 1. The entire experiment is run in the NS2.34 environment, which has been simulated. These network assumptions are used as input parameters for simulation in order to construct the entire strategy using the current methodologies available at the time.
Table 1. Summary of Fall Detection Datasets

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network size</td>
<td>1000 ×1000m</td>
</tr>
<tr>
<td>Sensor nodes</td>
<td>200</td>
</tr>
<tr>
<td>Transmission Rate</td>
<td>50 to 250 Kbps</td>
</tr>
<tr>
<td>Number of Nodes</td>
<td>50 to 100</td>
</tr>
<tr>
<td>Data Flows</td>
<td>2 to 10</td>
</tr>
<tr>
<td>MAC Protocol</td>
<td>IEEE 802.11</td>
</tr>
<tr>
<td>Antenna Model</td>
<td>Omni Antenna</td>
</tr>
<tr>
<td>Initial Energy</td>
<td>14.0 Joules</td>
</tr>
<tr>
<td>Packet Size</td>
<td>512 bytes</td>
</tr>
<tr>
<td>Receiving Power</td>
<td>0.4 Watts</td>
</tr>
</tbody>
</table>

Packet Delivery Ratio (PDR), throughput, latency, average energy consumption, control overhead, normalized routing overhead, packet drop, and finally the packet drop ratio are some of the performance metrics against which the proposed technique is evaluated.

- The PDR is defined as the total number of packets received by sink nodes as a percentage of the total number of packets created by source nodes. It does not use the metric system at all.

- When a particular number of data packets are transferred to the sink nodes in a given amount of time, this is referred to as the throughput. The unit of measurement is kbps (kilobits per second).

- While traveling from a source node to a sink node, a data packet encounters a period of time known as a delay. It is measured in milliseconds, which is a unit of time.

- In computing average energy consumption, the total amount of energy consumed by sensor nodes across the network while packets are transmitted between the source and destination is taken into consideration. Milli-Joules is the unit of measurement used herein.

- Control overhead is defined as the ratio of control messages transmitted to the amount of original data received at each sensor node. It does not use the metric system at all.

- The transmission of data control packets is referred to as normalized routing overhead since they are transmitted with every data packet. Heuristics are computed by dividing the total number of routing packets transmitted by the total number of data packets that have been received.

- Dropped packets, also known as packet losses, occur when a data packet fails to reach its intended destination after it has been transmitted from its point of origin. The most prevalent cause of packet loss is congestion in the network traffic. It does not use the metric system at all.

- Packet drop ratio, which is defined as the percentage of packets lost during transmission as a fraction of the total number of packets delivered, cannot be measured in any specific unit of measurement.
When it comes to packet delivery ratio, Figure 1 shows a comparison between the projected DoA-BSO and other state-of-the-art methods. As shown in the data, DoA-BSO achieves a higher packet delivery ratio for all node densities than the previous techniques, regardless of the node density.

![Figure 1. Packet Delivery Ratio between proposed DoA-BSO and existing methods](image1)

The proposed DoA-BSO approach exceeds the existing methods in terms of throughput at all node densities and outperforms them across the board in Figure 2.

![Figure 2. Throughput between proposed DoA-BSO and existing methods](image2)
When compared to existing techniques, the proposed DoA-BSO has a lower delay for all node densities in Figure 3.

![Figure 3. Delay between proposed DoA-BSO and existing methods](image)

A comparison of the DoA-BSO average energy use is depicted in Figure 4. As compared to existing methodologies, the DoA-BSO strategy consumes less energy across the full node density.

![Figure 4. Average energy consumption between proposed DoA-BSO and existing methods](image)
We compared the control overhead displayed in Figure 5 between the proposed DoA-BSO and the existing state of art methods. According to the data, DoA-BSO achieves lower control overhead across the board for all node densities than previously used techniques.

![Figure 5. Control overhead between proposed DoA-BSO and existing methods](image)

The routing overhead between a planned DoA-BSO and an existing state of art methods is presented in normalized form in Figure 6. According to the findings of this study, DoA-BSO achieves lower normalized routing overhead than existing systems for all node densities.

![Figure 6. Normalized routing overhead between proposed DoA-BSO and existing methods](image)
A packet drop occurs on a network as illustrated in Figure 7. The suggested DoA-BSO achieves lower packet drops than the current techniques across the board for all node densities. It is more reliable than other techniques in terms of packet delivery, with a lower drop rate.

![Figure 7. Packet drop between proposed DoA-BSO and existing methods](image1)

With regard to the number of packets dropped, Figure 8 compares the current DoA-BSO with the exiting methods. With respect to all node densities, the proposed DoA-BSO strategy has a lower packet drop ratio than the existing approaches.

![Figure 8. Packet drop ratio between proposed DoA-BSO and existing state of art methods](image2)
To put it another way, the proposed protocol overcomes the previous state-of-the-art direct-access techniques in terms of higher throughput and shorter delay, respectively. This reveals that the proposed DoA-BSO method surpasses earlier methods in terms of overall performance and efficiency.

**Conclusion**

To increase the accuracy of distance estimation between the anchor and unknown node pairs, sensor nodes are segmented in this work using the BSO Algorithm, which is used in this research. Even in the presence of unbalanced conditions, the proposed DoA-BSO involving three separate bee colonies can identify plausible anchor nodes as well as segment nodes arranged in clusters. The studies are carried out in a large-scale WSN using sensor node pairs in order to determine the precision with which the BSO can be located. When comparing DoA-BSO to conventional approaches, the findings of the meta-heuristic algorithm show that it improves the accuracy and segmentation of nodes significantly. In the absence of known sensors, weights are assigned to surrounding nodes based on their proximity to an anchor node and the error distances they have between themselves and the anchor node. The system has proved its capacity to tolerate outliers while still producing accurate distance estimations over a variety of radio ranges and levels of noise. Existing methods take more iterations to attain convergence with localization performance on isotropic networks is comparable with the DoA-BSO method. The proposed technique, which makes use of anisotropic networks, outperforms the competition in terms of positioning accuracy while requiring fewer calculations.

**Conflict of interest**

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

**Funding**

The author (s) received no financial support for the research, authorship, and /or publication of this article.
References


Shehadeh, Hisham A., Mohd Y. Idha Idris, Ismail Ahmedy, Roziana Ramli, and Noorzaily Mohamed Noor. 2018. The Multi-Objective Optimization Algorithm Based on Sperm Fertilization


Bibliographic information of this paper for citing:


Copyright © 2022, Devika E. and Saravanan A.