Distributed wireless sensor node localization based on penguin search optimization

Md Al SHAYOKH¹*, Soo Young SHIN²

¹,²Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea
ORCID iD: https://orcid.org/0000-0002-6652-8487
ORCID iD: https://orcid.org/0000-0002-2526-2395

Abstract: Wireless sensor networks (WSNs) have become popular for sensing areas-of-interest and performing assigned
tasks based on information on the location of sensor devices. Localization in WSNs is aimed at designating distinct
geographical information to the inordinate nodes within a search area. Biologically inspired algorithms are being applied
extensively in WSN localization to determine inordinate nodes more precisely while consuming minimal computation
time. An optimization algorithm belonging to the meta-heuristic class and named penguin search optimization (PeSOA)
is presented in this paper. It utilizes the hunting approaches in a collaborative manner to determine the inordinate nodes
within an area of interest. Subsequently, the proposed algorithm is compared with four popular algorithms, namely
particle swarm optimization (PSO), binary particle swarm optimization (BPSO), bat algorithm (BA), and cuckoo search
algorithm (CS). The comparison is based on two performance metrics: localization accuracy and computation time to
determine inordinate nodes. The results obtained from the simulation illustrate that PeSOA outperforms the other
algorithms, achieving an accuracy higher than 30%. In terms of computation time to determine inordinate nodes, the
proposed algorithm requires 28% less time (on average) than the other algorithms do.

Key words: Wireless sensor networks, Localization, Penguin search algorithm, Optimization, Computation time

1. Introduction

The rapid advancement in the field of embedded devices and radio communication systems has enabled the
development of wireless sensor networks (WSNs). A WSN is a network of spatially distributed sensors that
function independently and sense the surrounding area-of-interest centrally. WSNs can be used in various smart
application scenarios, such as exploration of hazardous environments, tracking of specific objects for surveil-
lance systems, continuous monitoring of patients in hospitals, and tracking of assets. In these scenarios, where
network coverage or routing plays an essential role, determination of the appropriate positions of the sensor
nodes or localization of nodes is of paramount importance. Therefore, localization of WSN nodes is considered
a fundamental challenge in any WSN-based application [1].

Conventionally, a large number of WSN sensors are deployed in an altruistic manner to determine the location
information of inordinate nodes in an environment. Data collected by these sensors for use in diverse applica-
tions would be ineffective and labor-intensive without information on the appropriate location of these sensors
[2]. These challenges intensify the priority of localization in WSNs. In conventional methods, any device can
identify its position with the help of a global positioning system (GPS) and numerous localization techniques.
Due to high installation cost of GPS and relatively low performance in terms of accuracy in indoor environments, GPS is not considered to be a cost-effective solution to deploy WSNs on a large scale. This specific challenge inspires researchers to develop a reasonable solution in the area of WSN localization for real-life scenarios wherein non-GPS-equipped devices and sensors use positioning techniques in conjunction with few GPS-enabled devices [3].

Diverse techniques that apply nature-inspired and geometric algorithms have been implemented to solve localization issues. Herein, the primary objective of implementing bio-inspired techniques was to minimize the localization error or reduce computation time. Biologically inspired algorithms that are categorized as meta-heuristic algorithms are computational intelligence paradigms where problems are modeled as multimodal and multidimensional and solutions are provided by population-based stochastic techniques. However, in the area of WSN localization, the implementation of these bio-inspired algorithms to achieve higher accuracy with faster convergence is still under research [4]. Considering the challenges mentioned above, a newly developed meta-heuristic algorithm based on the collaborative hunting strategy of penguins, named the Penguin Search Optimization Algorithm (PeSOA), is proposed in this paper to solve the localization problems of distributed WSNs. Important metrics to measure the performance in the area of localization (such as accuracy and computation time) are analyzed subsequently with four nature-inspired algorithms: particle swarm optimization (PSO), binary particle swarm optimization (BPSO), bat algorithm (BA), and cuckoo search algorithm (CS). The remainder of this research article includes a detailed literature survey on WSN localization. This is followed by a description of the PeSOA algorithm and its implementation for WSN localization. Subsequently, the simulation results are illustrated and investigated, and feasible areas of future progress are discussed with concluding remarks.

2. Literature Survey

Numerous localization schemes have been analyzed and discussed in [5]. Considering the requirements of localization schemes, algorithms that are recommended in the literature can be categorized into range-based and range-free localization schemes. The balance between localized and anchor nodes is crucial for range-free schemes. In [6], [7] and [8], categories of range-free localization is detailed. In contrast to range-free localization schemes, both radio signal strength (RSS) and estimated distance between the anchor and localized nodes are essential for performing range-based localization schemes [5]. Most of the approaches in the literature considered localization as a multidimensional problem and addressed numerous techniques whose precision rates are relatively high. In [9], highly effective localization schemes were introduced to WSN through an accurate positioning system (APS). Here, the capability of identifying unknown nodes with the help of GPS is enhanced to non-GPS equipped nodes. A system that attained high accuracy by flooding the location of the anchor or GPS-equipped nodes to the area of interest and utilizing a triangulation approach was proposed in [10]. Refinement techniques to enhance this further were proposed in [11].

As localization accuracy remains prime metrics in WSN localization scenario, a Kalman filter was developed based on the least squares method and implemented successfully to achieve higher accuracy in the area of WSN localization [12]. Through consideration of the localization problem in WSNs in the field of multidimensional problems, this approach enabled the authors to apply gradient search techniques to analyze unknown data and calculate minimal path distances. Subsequently, the information related to the shortest path distances was fed to the algorithm to identify the unknown nodes more accurately. In [13], authors introduced a convex
optimization algorithm for WSN localization. The dependency on the distance calculation to achieve higher localization accuracy is omitted here. Convex programming assists the localization algorithm to achieve higher accuracy with a trade-off between the computational time and power consumption of the nodes. Apart from that, in [14] and [15], authors illustrated the convex programming in WSN localization having energy consumption as a trade-off factor. The GF (Genetic Fuzzy) algorithm in WSN localization is mentioned in [16]. It mimics the features of the DV-HOP algorithm and utilizes hop counting techniques for calculating the distance between anchor nodes and unknown nodes. The algorithm relies on faultless measurements and dense deployment of WSNs to achieve high precision. Authors of [17] proposed a micro-genetic algorithm to perform localization tasks for an APS as a post optimizer. In recent years, approaches based on the use of biological models and generic population-based techniques, commonly known as bio-inspired algorithms, have been applied in the area of localization to solve localization challenges. This is due to the high accuracy and faster convergence of such approaches [17].

In [18], unknown nodes were localized using a simulated annealing algorithm and densely deployed anchor nodes within the area of interest. A genetic algorithm (GA) was used in [19] to determine the ideal position of unknown nodes. A multi-phased localization approach that combines GA and simulated annealing was proposed in [20] to resolve the flip ambiguity issues between neighboring nodes. PSO- and iterative-PSO-based schemes were proposed in [21] and [22], respectively, to localize inordinate nodes and achieve higher accuracy with the help of three or more surrounding nodes. An objective function was developed and implemented in [23], however, the fitness value of the particles remains a significant concern. In [24], a system to achieve high precision by calculating fuzzy logic system (FLS) weights was introduced. It utilizes RSS from nodes equipped with GPS information and integrates it with the edge weights of these nodes. Nonetheless, it is a noteworthy research area for gaining advantages in terms of computational time and localization accuracy simultaneously. Accordingly, in this study, a new bio-inspired algorithm, called penguin search optimization algorithm (PeSOA) [25], was implemented to address the localization issue in WSNs. PeSOA is based on the collaborative hunting strategy of penguins. The proposed scheme performs better than other related approaches in terms of computation time and accuracy of locating inordinate nodes.

3. Penguin Search Optimization Algorithm

A beneficial search activity of an animal is defined as a search activity where the energy gain is relatively substantial compared with the energy consumed. Penguins, a group of aquatic flightless birds, utilize this specific characteristic to extract information such as search times, cost of food, and energy content in the prey. Resource availability and distance between nesting areas become prime factors in hunting behavior within the area-of-interest of penguins. Penguins, commonly known as sea birds, adapted their activities for swimming by utilizing wings. For penguins, the breathing capacity remains a base factor while diving because the dive is dependent on the reserve oxygen. The more speed and depth they gain, the more they consume oxygen, and the trip time starts to reduce [25]. Food required by large numbers of groups varies by species, age, and availability of food within the area of interest. An optimization has been developed straightforwardly by utilizing the hunting technique of penguins. This is outlined below:

- The total number of penguins is divided into several groups based on the food availability within the area of interest. Furthermore, their reserve oxygen before diving into the water is assessed.
- The entire area of interest is divided into several depths to search for the necessary food. In any specific
group, each penguin searches for food in an arbitrary manner within the specified boundary of the group. Penguins return to the surface after performing several dives and share their information regarding food availability and the location of food sources, through unique vocalizations.

- Based on the availability and food quantity, penguins reorder their groups to attain equilibrium. Finally, the group that locates the largest food source shares the location information among the penguins.

To develop the algorithm, the position of penguins within the area of interest is denoted as "i". The group distribution is completed based on the presence of food sources within the area of interest. Elementary steps are developed by applying these rules and summarized as a pseudo-code, which is illustrated in algorithm 1 [25].

**Algorithm 1 Penguin Search Optimization Algorithm (PeSOA)**

1. Generate a random population of P penguins in groups;
2. Initialize the probability of presence of fish in the holes and levels;
3. For i=1 to number of generations;
4. For each individual i ∈ P, do
5. While oxygen reserves are not depleted
   6. Take a random step
   7. Improve the penguin positions using the position update equation
   8. Update the quantities of fish eaten by this penguin
9. EndWhile
10. End For
11. Update the quantities of eaten fish in the holes, levels, and best groups
12. Redistribute the probabilities of penguins in the holes and levels
13. Update best solution
14. End For

By considering the search area as a multi-dimensional search space, the best solution is developed using the food dispersal probability to achieve an optimal value and obtain the maximum amount of food. All the members of a group utilize an identical solution within the search space. In general, each group performs several dives based on the amount of reserve oxygen and the probability of food expediency within the search area. Penguins update their search positions after each cycle using the following solution:

$$D_{new} = D_{LastLast} + \text{rand}() |X_{LocalBest} - X_{LocalLast}|$$  \hspace{1cm} (1)

where $D_{new}$ represents the updated position of a penguin and rand() is a random number for the distribution. $D_{LastLast}$, $X_{LocalBest}$ and $X_{LocalLast}$ represent the final and local best solutions that are used to update the positions of the penguins within a group. Penguins tend to swap necessary information to determine an optimal solution and rearrange the group after performing several dives within the area of interest. The attainment of the global optima without falling into the local optima after several iterations enables the PeSOA to perform better than conventional population-based approaches.

**4. Proposed Localization Technique Using PeSOA**

The main objective of performing WSN localization is to determine the maximum inordinate nodes with the help of anchor nodes. In the conventional approach, the total number of inordinate nodes (N) is deployed randomly within the search. In the proposed approach, a minimal number (M) of nodes are deployed with a GPS facility (anchor nodes). With M anchor nodes and a transmission capability over the range R, the number
of localizable nodes within the area of interest is $N - M$. The location information of GPS-equipped devices needs to be transmitted in each iteration during the localization scheme. In each iteration, localized nodes are considered as nodes of reference for the subsequent iteration. If a non-GPS-equipped node has at least three nodes within its surroundings, the node will be considered as a localizable node. The environment plays an important role in the calculation and affects the accuracy in any WSN scenario. This is because all the nodes are part of the wireless system. This effect is considered to be a Gaussian noise and is denoted as $n_{\text{noise}}$. It is used in 2, where the localizable nodes are considered for calculating the distance $\hat{d}_i$:

$$\hat{d}_i = d_i + n_{\text{noise}}$$

(2)

Here, $d_i$ denotes the actual distance and is obtained as follows [23]:

$$d_i = \sqrt{(x - x_1)^2 + (y - y_1)^2}$$

(3)

In the above equation, the positions of the destination node and $i^{th}$ anchor node are denoted as $(x, y)$ and $(x_i, y_i)$, respectively. Owing to the considerable impact on the RSS measurement in the calculation of distance, the log-normal shadowing effect is considered as noise. The central objective of an optimization problem is to minimize the objective function. The determination of the minimal localization error is regarded as the objective function of the optimization algorithm in WSN localization schemes. Herein, this issue is formulated as a multimodal problem. Considering this, each node that needs to be localized would apply the PeSOA algorithm to determine the desired location $(x, y)$. Nodes that utilize the optimization algorithm use the following objective function:

$$f(x, y) = \frac{1}{M} \sum_{i=1}^{M} (d_i - \hat{d}_i)^2$$

(4)

In this equation, $M$ denotes the number of GPS-equipped nodes positioned in the transmission area of the localizable nodes. If a node that is likely to be localized is located in between the transmission range of three or more GPS-equipped nodes (anchor nodes), that specific node would be identified first. It then performs localization by functioning as a reference node for others [23]. Because this algorithm is based on iterative behavior, the above-mentioned process is repeated until all the inordinate nodes within the area of interest are located. The most favorable location of the nodes $(x, y)$ that would be localized is provided by the optimization algorithm by minimizing the localization error. After the performance of successive iterations and determination of each inordinate node, the error in the determination of the inordinate nodes needs to be calculated using the mean of the squares of distances between the actual node position $(X_i, Y_i)$ and computed node position $(x_i, y_i)$. Here, $i$ ranges from 1 to $N_L$. $N_L$ denotes the number of nodes that must be localized. The following equation is preferred for calculating the localization error of the inordinate nodes:

$$E_L = \frac{\sum_{i=M+1}^{N} \sqrt{(x_i - X_i)^2 + (y_i - Y_i)^2}}{N_L}$$

(5)

where $E_L$ is the localization error of the PeSOA algorithm. The aforementioned process is repeated until each node within the area of interest is identified. Thus, the number of identified nodes increases. In addition, the amount of calculation for determining the error progressively increases as the number of iterations increase. The number of reference nodes in each iteration may vary from three to a number larger than that in the previous iteration. By applying PeSOA algorithm, parameters such as localization accuracy and computational time are improved compared to other algorithms.
5. Results and Discussion
A simulated environment was developed to explore the effect of the PeSOA algorithm within the area of interest. The experiment was carried out using MATLAB, and the results are illustrated using a graph. The parameters used to develop the simulated environment are depicted in table 1. Ten GPS-equipped nodes (anchor nodes) and 40 inordinate or unknown nodes are positioned randomly within the area of interest (100 × 100 m). The nodes that remain constant in all the iterations of the experiment are assigned a transmission range of 25 m. To conduct this experiment, the total number of iterations in this experiment is fixed at 150 [26] to its highest level. Considering the above parameters, the intended positions of the inordinate nodes are calculated using PeSOA, as illustrated in figure 1. The localization error calculated using equation 5 is presented in figure 2. The simulated environment is considered to be 2D. Furthermore, the geometric dilution of precision (GDOP) that is used in satellite navigation is not calculated, because the nodes are stationary.

Table 1: Parameters that used in the experiment

<table>
<thead>
<tr>
<th>Parameters of the system</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>Intel(R) Core (TM)i3-4005U CPU @ 1.70 GHz</td>
</tr>
<tr>
<td>Type of System</td>
<td>64-bit operating system</td>
</tr>
<tr>
<td>Memory</td>
<td>4 GB</td>
</tr>
<tr>
<td>Matlab Version</td>
<td>R2018a</td>
</tr>
</tbody>
</table>

Figure 1: Localization of WSN using PeSOA

In PeSOA, penguins form several groups and search for food. An increment in the number results in more profitable searches within the area of interest. This is because penguins subsequently exchange information among the groups and redistribute the group to search for more food sources. The more the groups of penguins, the higher is the accuracy achieved by the collaborative searching technique. The effect of the number of groups on the localization of the inordinate nodes is illustrated in figure 3.

The range of transmission is considered to be one of the important parameters (apart from localization accuracy and computation time) in WSN localization. As a trade-off, the computation time increases when the error in
Figure 2: Calculated localization Error of WSN nodes using PeSOA

Figure 3: Effect of group numbers on WSN localization using PeSOA

determining the locations of nodes decreases. Simulations are performed using PeSOA to analyze the effect of the range of transmission over the area of interest. Here, the percentage of noise (generally denoted as $P_n$) is considered to be 2 [27], and 150 iterations are performed. The effect of the range of transmission is illustrated in figure 4. Furthermore, an analysis is presented in table 2 in terms of minimization of the localization error and computation time. Figure 4 clearly indicates that the error decreases when the range increases. In addition, table 2 shows the increase in the computation time to localize all the inordinate nodes as a trade-off factor. If calculations cannot be performed by measuring distances from the anchor nodes then the chance of localizing any inordinate node within the area of interest becomes shallow. The chance of localizing inordinate nodes could
be soared through the increment of anchor nodes within the search area. Based on mentioned considerations, for analyzing the effect of the density of anchor nodes within the area of interest, simulations are performed and results are illustrated in Figure 5. In addition, by considering time of computation and error in localization as performance metrics, an analysis is presented in 3 to demonstrate the effect of the number of anchor nodes in search area. Figure 5 clearly depicts that the accuracy and to find out the inordinate increases when the number of anchor node increases. On other hand, table 3 illustrates the increase of total time to find out the inordinate nodes within the area of interest as a trade-off factor.

The total percentage of noise is considered as an important metric in WSN localization because it affects the calculation of the RSS value that is used in distance calculation. Simulations were performed to assess the
Anchor node analysis for PeSOA in WSN localization

Figure 5: Effect of anchor nodes number on WSN localization using PeSOA

Table 3: Analysis of the density of anchor nodes deployment within the area of interest

<table>
<thead>
<tr>
<th>Number of Anchor Nodes</th>
<th>Mean Localization Error (m)</th>
<th>Time of Computation (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.151</td>
<td>382.0775</td>
</tr>
<tr>
<td>11</td>
<td>0.123</td>
<td>393.0432</td>
</tr>
<tr>
<td>12</td>
<td>0.104</td>
<td>403.0197</td>
</tr>
<tr>
<td>13</td>
<td>0.070</td>
<td>408.9433</td>
</tr>
</tbody>
</table>

Noise effect on WSN localization using PeSOA

Figure 6: Effect of noise amount on WSN localization using PeSOA

impact, and the results are presented in figure 6. The figure shows that the error in determining the inordinate nodes increases with the increase in the amount of noise within the area of interest. This is owing to the
A comprehensive analysis of the computation time to identify all the nodes based on the mentioned algorithms is presented in Table 4. The figures clearly depict better performance for PeSOA over the other algorithms in each case owing to the collaborative search strategy. The ability of PeSOA to locate an object through division into several groups enables it to outperform the other algorithms. Because penguins search in different groups,

Figure 7: Number of localized nodes after each iteration by using PSO, BPSO, BA, CS and PeSOA

Figure 8: Analysis of localization error using PSO, BPSO, BA, CS and PeSOA

shadowing effect that impacts the calculation of the distance in RSS-based localization schemes. An accurate determination of the number of localized nodes within the search area after each cycle is considered to be critical in real-time applications of the proposed scheme. To reveal the number of nodes that are localized after each iteration, an analysis was performed by applying two popular algorithms in conjunction with PeSOA in the area of WSN localization. The results are presented in Figure 7. As shown in the figure 7, PeSOA localizes all the inordinate nodes faster than the other four algorithms because of the collaborative searching strategy in an organized manner. To determine the overall performance, PeSOA is compared with the other four algorithms based on the accuracy of locating nodes, density of anchor nodes, range of transmission, and amount of noise present within the area of interest. The results are presented in Figures 8, 9, 10, and 11, respectively.
the velocities of the groups are continuously updated simultaneously with that of the group that enables faster search within the search area and the determination of the locations of the inordinate nodes. To achieve the desired solution, 150 iterations were performed in this experiment. Energy efficiency is considered to be one of the significant concerns in the WSN scenario owing to the constraints of sensor size. It is apparent that the energy utilized by sensors increases with an increase in computation time. The results illustrated in table 4 indicate that the minimal computation time obtained by PeSOA helps the nodes to maximize the network lifetime. All the results clearly indicate the higher performance of PeSOA compared with those of the other four bio-inspired algorithms (PSO, BPSO, BA, and CS) in terms of two important parameters of WSN localization.
Figure 11: Analysis of noise percentage for PSO, BPSO, BA, CS and PeSOA

Table 4: Analysis of mean localization error and time of computation to find out the inordinate nodes

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Mean Localization Error (m)</th>
<th>Time of Computation for 150 iterations(s)</th>
<th>Time Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSO</td>
<td>0.2891</td>
<td>846.6145</td>
<td>O (m*n)</td>
</tr>
<tr>
<td>BPSO</td>
<td>0.3755</td>
<td>653.4910</td>
<td>O (m*n)</td>
</tr>
<tr>
<td>BA</td>
<td>0.2533</td>
<td>569.6523</td>
<td>t* (O (P*N/2))</td>
</tr>
<tr>
<td>CS</td>
<td>0.1912</td>
<td>505.4653</td>
<td>O (n^2 Nq)</td>
</tr>
<tr>
<td>PeSOA</td>
<td>0.1134</td>
<td>407.5721</td>
<td>O (m)*n</td>
</tr>
</tbody>
</table>

6. Conclusion

Localization remains a significant challenge in WSNs. This study formulated the localization challenge as a multi-modal problem and applied the hunting strategy of penguins through the PeSOA algorithm to solve the challenge where nodes are deployed in a distributed manner. The simulated results reveal that PeSOA performs better than other algorithms in terms of both the important metrics of WSNs, i.e., localization accuracy and computation time for locating inordinate nodes. PeSOA demonstrated an average of 30% higher localization accuracy and a 28% faster convergence time compared to the other mentioned algorithms such as PSO, BPSO, Bat Algorithm etc. PeSOA more effectively locates inordinate nodes with minimal error by utilizing a collaborative approach for determining the objectives. As energy consumption, network life time and mobility becoming major concern in the future wireless sensor networks scenario, future research studies will be carried on by considering these parameters. Also, migration strategies of penguins especially emperor penguin’s migration strategy can be applied to WSN localization to achieve better results in terms of accuracy with minimal convergence time where emperor penguins strategy will tend to find out the best possible food sources in harsh environment within the limited resources by synchronizing their dives and saving more energies.
7. Acknowledgement

This work was supported by Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2018R1A6A1A03024003) and the MSIT (Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program (IITP-2021-2020-0-01612) supervised by the IITP(Institute for Information communications Technology Planning Evaluation).

References


