



A Hybrid Approach for Localisation of Sensor Nodes in Remote Locations

RUPENDRA PRATAP SINGH HADA, Computer Science and Engineering, Indian Institute of Technology Indore, Indore, India

ABHISHEK SRIVASTAVA, Computer Science and Engineering, Indian Institute of Technology Indore, Indore, India

A Wireless Sensor Network (WSN) is a network of sensor nodes using low-power wireless technology to collect data in a region of interest (ROI). Due to their low energy, locating sensor nodes in large outdoor areas is challenging, which precludes GPS integration. WSNs typically comprise a small number of beacon nodes (BN) whose locations are known in advance, with most nodes deployed at unknown coordinates within the ROI. Endeavours to determine the locations of such unknown WSN nodes are largely based on the impractical assumption that every unknown node (UN) is within the communication range of BNs. Subsequently, these approaches utilise at least two BNs to determine the position of one UN. The Received Signal Strength Indicator (RSSI) or Angle of Arrival (AoA) values of the signals from the BNs form the basis for such localisation. This article suggests an iterative hybrid approach incorporating AoA and RSSI techniques, achieving accurate localisation with just one BN. The iterative method gradually covers the region, avoiding the unrealistic assumption of having all UNs within range. It also presents an innovative use of a unipolar stepper motor for AoA measurements. Experiments in a simulated environment and a real-world prototype validate the approach's effectiveness.

CCS Concepts: • Networks → Network types; Ad hoc networks; Mobile ad hoc networks; Network performance evaluation; Network experimentation;

Additional Key Words and Phrases: Localisation, outdoor localisation, ad-hoc networks, sensor network

ACM Reference Format:

Rupendra Pratap Singh Hada and Abhishek Srivastava. 2025. A Hybrid Approach for Localisation of Sensor Nodes in Remote Locations. *ACM Trans. Sensor Netw.* 21, 2, Article 23 (March 2025), 33 pages. <https://doi.org/10.1145/3715914>

1 Introduction

A **wireless sensor network (WSN)** comprises a number of nodes, each laden with sensors of various kinds and capable of communicating with other nodes through low-power signals. The nodes gather ambient information from their surroundings and transmit the same from one node to another, forming an ad-hoc network until a base station is reached [26]. To monitor a **region of interest (ROI)**, which is usually not easily accessible, a WSN is deployed with appropriate sensors on each node. Any deviation from the normal in the ROI is expected to be picked up by sensors of one or more nodes, and the base station is alerted. Environmental monitoring,

Authors' Contact Information: Rupendra Pratap Singh Hada, Computer Science and Engineering, Indian Institute of Technology Indore, Indore, India; e-mail: phd2101101004@iiti.ac.in; Abhishek Srivastava, Computer Science and Engineering, Indian Institute of Technology Indore, Indore, Madhya Pradesh, India; e-mail: asrivastava@iiti.ac.in.



This work is licensed under a Creative Commons Attribution 4.0 International License.

© 2025 Copyright held by the owner/author(s).

ACM 1550-4859/2025/03-ART23

<https://doi.org/10.1145/3715914>

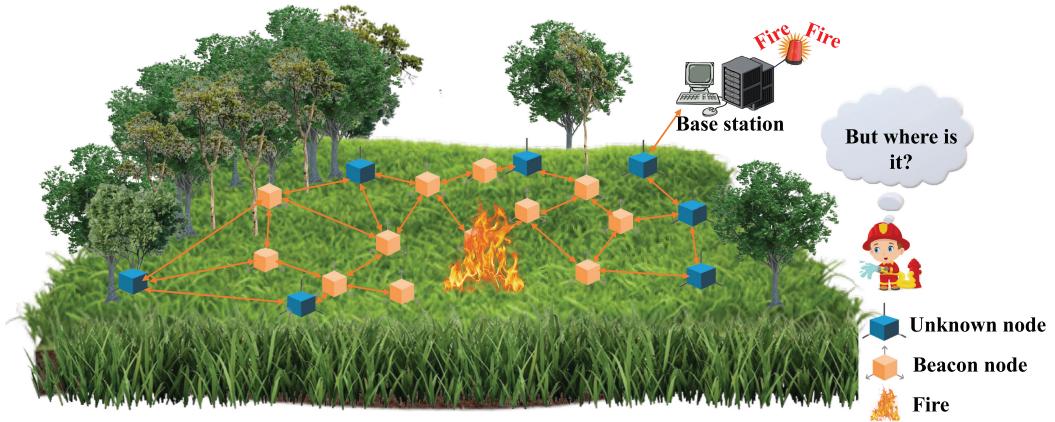


Fig. 1. Application of outdoor localisation in effective forest fire detection.

industrial automation, healthcare, agriculture, disaster management, and military surveillance are a few applications of WSN [3]. Effective surveillance using WSN has several impeding factors like limited energy, limited communication range, scalability [38], heterogeneity, and network security [15]. Of these issues, the limited energy of nodes is crucial and has far-reaching consequences. Energy limitations arise because the nodes in a WSN are usually deployed in inaccessible regions and are most commonly powered by non-rechargeable batteries. It is impractical to access the nodes and change the batteries. Thus, the utility of a node is limited by the lifetime of the batteries. It is imperative, therefore, to prolong the life of the batteries as much as possible. One way of doing this is through the use of exceedingly lightweight applications on the node. WSN nodes, therefore, cannot make use of several common applications, a crucial one of which is the **Global Positioning System (GPS)** used for accurate localisation. GPS sensors are typically heavyweight and energy inefficient and are thus not appropriate for energy-starved WSN nodes. Localisation of nodes in a WSN, therefore, becomes non-trivial [35]. Localisation of nodes is an important exercise, especially when the WSN is deployed over a large area. An example of this is in one of our projects Figure 1 wherein we utilise a WSN for early detection of forest fires in the Melghat Tiger Reserve (a vital forest area located in Central India). In such scenarios, the WSN nodes containing various sensors (like heat and temperature) are randomly deployed throughout the area sometimes manually but mostly using **unmanned aerial vehicles (UAV)**. This is because most parts of the reserve are inaccessible to humans and UAVs are preferable. Although the placed nodes are stable and static, the deployment through UAVs makes it impracticable to tag the nodes *a-priori*. The deployed WSN nodes can detect a fire (using the sensors) and send an alert signal to the base station using intermediate nodes and forming an ad-hoc network.

The base station learns that there is a fire but is unable to establish the precise location of the fire. This is a serious problem as the tiger reserve spreads over an area of 3,000 square kilometres, and pinpointing the fire location is very difficult. Satellite images and drones are sometimes employed for the localisation of fires, but these are effective only after the fire has become large and can be seen above the trees, by which time significant damage is done. They are not helpful for early fire detection systems.

There are several techniques used for localising nodes in a WSN. The approach taken by these techniques is to utilise nodes whose locations are known called *beacon nodes* as a reference and compute the locations of unknown nodes relative to beacon nodes. The localisation techniques fall into one of two categories: range-based [14] and range-free [20]. Range-based techniques

[18] utilise hardware devices to determine the relative positions of unknown nodes from beacon nodes by measuring the distance between them. Factors like the **Received Signal Strength Indicator (RSSI)** [4], **Angle of Arrival (AoA)** [43], and **Time of Arrival (ToA)** [44] are employed to compute this distance. Range-free [21] techniques use information such as node IDs, connection patterns, or topological data [48] to estimate the node's location. Existing localisation methods have several limitations: (1) They perform well in small regions where it is common for the unknown node to be within the communication range of two or more beacon nodes. However, in more extensive regions, especially the outdoors, it is sometimes unrealistic for two or more beacon nodes to be within the communication range of the unknown node. (2) multiple (two or more) beacon nodes are required to localise an unknown node; and (3) the accuracy of localisation is limited and the time required to localise the entire region is significant.

In this article, we propose a localisation approach that is effective over vast regions of interest without the need for all nodes to be within communication range with each other. In fact, in our technique, very few nodes are within the communication range of the initial beacon nodes. The proposed approach takes an iterative approach to localise unknown nodes and is able to localise nodes in the entire region in just a few iterations. A few unknown nodes close to the beacon nodes are first localised. These newly localised nodes now become beacon nodes themselves and, in the next round, localise other nodes close to them, and this process continues until the nodes in the entire region are localised. This is made possible by the requirement in our approach for there to be only one beacon node within the communication range of the unknown node for the latter to be localised. This is a marked improvement over existing techniques that require at least two and, ideally, more beacon nodes to be within the communication range of the unknown node. Our approach employs a combination of the angle between the unknown node and the beacon node (computed through the Angle of Arrival of signals) and the distance between the two (computed using the RSSI values) to achieve localisation. The key contributions of this article are summarised as

- (1) Localisation of an unknown node using just one other node through the use of two factors: angle and distance from a known node.
- (2) An iterative approach to localisation is proposed that enables the coverage of a very large area.
- (3) An innovative and inexpensive approach to compute angle of arrival is presented in this article and used for validation wherein a stepper motor is customised to move 360 steps per revolution and is used for measuring the angle with an accuracy of 1° .

The remainder of the article is organised as follows: Section 2 discusses existing techniques for outdoor localisation in a WSN. This is followed by a detailed description of the proposed approach in Section 3. Section 4 validates the efficacy of the approach through simulations and real-world prototypical implementation, and finally, Section 5 concludes the article with pointers to future work.

2 Previous Work

There is significant work done on localisation approaches for WSNs that are based on RSSI and AoA. In this section, we discuss a few prominent endeavours.

The work described in [41] is closely related to ours. An RSSI and AoA-based method is proposed in the article for localising multiple unknown targets (unknown nodes). Different from existing algorithms that rely on complex mathematical tools that do not always yield feasible solutions, a more straightforward approach was taken here: the article utilises available AoA measurements and transitions from Cartesian to Spherical coordinates which allows the approximation of initially non-convex measurement models to linear ones. The iterative methods first localise some

unknown nodes with a localisation error; then, these newly localised (with some localisation error) nodes help localise other unknown nodes. The localisation error propagated from one localised to another in subsequent rounds is known as accumulative error. Our proposed method works iteratively and considers the accumulative errors from the previous rounds (just like feasible solutions), representing the clear idea of waste areas. Also, our proposed approach is able to localise multiple nodes per iteration, leading to much fewer iterations. In addition to this, the work in [41] uses signals for the calculation of AoA, which is susceptible to disturbance and interference adversely affecting accuracy. To overcome this, we propose the use of a simple stepper motor-based mechanism, which is the most effective in free space models [9] (without any obstructions) for 2-dimensional spaces. We also study the effectiveness of our approach in different types of terrains that include: sandy, long grassy, and sparse tree terrains; which is again something that the earlier approach does not look at.

In Reference [16], the authors address the node localisability problem in a network where a few nodes are localisable. The authors propose the concept of partially localisable networks by presenting theoretical conditions for a node to be considered uniquely localisable. Using this, localisable nodes are identified by dividing the network into redundantly rigid and reconnected components. The method uses 10% of the nodes as **beacon nodes (BN)**, whereas our proposed approach uses only 1% of nodes as BNs. In spite of this, we are able to localise nodes over an extensive area and with greater accuracy. In [37], the authors propose an iterative multilateration method that localises the **unknown node (UN)** using trilateration (using three BNs) repetitively. The ToA used for distance measures have several disadvantages and obstacles in crowded areas when compared with RSSI-based methods [19]. The proposed method takes advantage of iterative multilateration to a great extent through its ability to localise a UN using a single BN. In the work cited, a UN node needs to be within the communication range of at least three BNs. The method, therefore, requires a very large number of beacon nodes when the density of unknown nodes is large, which is quite cost-ineffective when deployed in the real-world. In [22], the authors propose the sequential Monte Carlo localisation method and exploit the node's mobility to improve localisation accuracy and precision. The localisation accuracy is established for three scenarios: when UNs are static and BNs are movable; when UNs are moving and BNs are stable; and when both UNs and BNs are moveable. The approach utilises range-free methods for distance estimation. Our proposed work uses a range-based approach for calculating the parameters, which is much more accurate than range-free methods. Furthermore, the above method evaluates the results based on simulations only and, even in these, uses a high BN density, which is hardly possible in realistic deployments. In [47], the authors address an important issue in trilateration wherein the latter marks a localisable graph as non-localisable. The approach can appropriately localise UNs that are one hop away from participating in the trilateration exercise. The method uses trilateration and three BNs to determine the location. The method uses trilateration and requires at least three BNs to get the location of every unknown node. [27] proposes a novel localisability algorithm (i.e., Patch and Stitching) wherein the whole network is initially divided into small localisable networks, and subsequently joined together to form a global localisable network. Using an algebraic approach, the work proposes a subset of localisable network merging conditions for 2D/3D networks. The method uses trilateration for node localisability and can localise 90% of the network with 5% of BNs. Our proposed method, on the other hand, takes a hybrid approach for localisation, and in simulations is able to localise the whole network with only 1% of nodes acting as BNs. In [46], the authors work on the localisability of the network to answer the following two questions: whether it possible to say that a node is localisable or not while using a network graph; how many nodes of the network are localisable and whether these can be identified? The localisability method uses more nodes to localise unknown nodes in the network, where the proposed method takes a hybrid

approach and is able to efficiently localise most nodes using a very small number of known nodes thus producing significantly superior localisability results. [45] deals with localisation accuracies affected by outliers in range-based localisation methods. Earlier solutions, in general, use triangle inequalities to deal with noisy data with outliers. The method proposed, on the other hand, utilises a theoretical model based on graph embeddability and rigidity theory for the same. The article designs a bilateration generic cycles-based outlier detection algorithm and tries to evaluate its effectiveness through simulations and practical deployments. The method uses trilateration and cycles-based outlier detection. The detection method is ineffective for more extensive settings and likely fails to produce good results for larger outdoor areas. The proposed method discards the outliers if they are out of the predefined region of the network and is thus more robust.

In Reference [12], a machine learning-based localisation approach for outdoor settings is proposed. The approach employs virtual nodes to widen the dataset to address the issue of limited training data. The method collects information from multiple access points and uses this information to appropriately localise other unknown nodes. The approach is tested using simulations that vary parameters like the number of sensors and anchor nodes, radio transmission power, and wireless signal quality. In [30], the ARBL algorithm for outdoor localisation is proposed. This approach uses trilateration and reference node selection to determine the locations of sensor nodes. A reference triangle through triangulation is formed in this approach, and using this, the ranging inaccuracies are evaluated. The sensor node positions are based on these inaccuracies. Notably, the innovation in this work lies in selecting the optimal anchor node for the localisation process. The approaches in [12] and [30] are relatively impracticable. The former is because it is pretty challenging to implement, and the latter is because it requires at least three beacon nodes to localise an unknown node, which is hard to get in outdoor environments. In [23], the AoA and RSSI differences-based localisation method (ALRD) is proposed. In this method, the AoA is estimated by comparing the RSSI values of beacon signals received by two perpendicularly oriented directional antennas installed at the same place. Subsequently, two methods (maximum point minimum diameter and maximum point minimum rectangle) are proposed to minimise the ALRD localisation error. In [25], the authors propose a simplified combination of the AoA and RSSI methods. The calculations of the AoA require complex antenna arrays. The proposed approach (1AoA/nRSSI) uses the AoA values from only one anchor node in combination with n RSSI values to estimate the location of an unknown node. The method is suitable for shadowing environments and when more precise RSSI values are available.

In Reference [49], a fingerprint-based WSN localisation technique is proposed for indoor and outdoor use. The objective of this work is to enhance fault tolerance and system efficiency. A precision of 5 m and 10 m was attained, respectively, in this technique for pedestrian and driving tests using phones for evaluation. In [36], multiple indoor and outdoor localisation approaches for WSN are analysed. The authors illustrate the applicability of **Multidimensional Scaling (MDS)** approaches in modern technologies like WSN-IoT [29], cognitive radios, and 5G networks. Centralised and distributed MDS techniques for indoor and outdoor localisation are discussed in the article. In [40], a localisation technique for mobile nodes in wireless sensor networks is proposed. The strategy targets short beaconing intervals and localisation deviations stemming from radio propagation. It solves these issues through geometric least square curve fitting. In [8], a technique is proposed that employs curve fitting for wireless sensor networks and a range-based approach for outdoor localisation is proposed. The techniques discussed in [36, 40, 49], and [8] are applicable in both indoor and outdoor environments. However, their localisation processes were tested on limited confines, potentially constraining their effectiveness in large areas.

For indoor localisation, [32] suggests a hybrid approach using **particle swarm optimisation (PSO)** and the global best local neighbourhood approach. This approach employs three anchor

nodes to locate an unknown node. It achieves a localisation inaccuracy of 0.44 m in a simulated environment. RSSI values are used for distance estimation between nodes, and the mean error varies with network size changes. In [11], a range-free localisation technique is proposed using a fusion of Harris Hawks optimisation and area minimisation. The technique categorises neighbours into incoming and outgoing for heterogeneous wireless sensor networks. It employs area minimisation to reduce the node's predicted region. The technique is simulated for small areas in 2D and 3D environments. In [39], a heuristic approach is introduced to tackle anisotropy-related challenges in WSN, which can lead to localisation errors. The proposed range-free approach incorporates geometric constraints and hop-based strategies to mitigate this problem. The techniques in [32], [11], and [39] are indoor localisation strategies that pose challenges when applied to larger areas. Implementing these for expansive spaces is arduous and mostly not feasible.

3 Proposed Approach

The proposed approach for localisation of nodes in a WSN is simple and is the first localisation endeavour, to the best of our knowledge, that is effective in very large areas. Localisation approaches until now have been limited by the assumption that the region of interest is small enough for all nodes to be within communication range of each other. This is quite unrealistic for large outdoor areas like forests or glaciers. Here, regions of interest are often spread over thousands of square kilometres, and the communication range of a WSN node is within 100 meters. In the proposed approach, we overcome this vital limitation and initiate with the modest requirement that the co-ordinates of a minimal number of the nodes in the region of interest be known. This is a realistic expectation as parts of the regions of interest, especially along the periphery, are usually accessible. The locations of such nodes are computed using external GPS devices. The nodes whose locations cannot be determined and need to be localised are deep within the regions of interest that are mostly hostile and inaccessible. Examples include forest regions whose peripheries are accessible, but the inner core is hostile and inaccessible; similarly, it is easy to access the periphery of a glacier but very difficult to move deep into it.

The first contribution of our approach is the capability of accurately localising an unknown node using the coordinates and signals from just one known beacon node. This is a significant improvement over existing techniques that require at least two beacon nodes to localise an unknown node. This contributes in substantial measure to the capability of localising a very large number of unknown nodes using a very small number of beacon nodes. Subsequently, in Section 4 (Experimental Evaluation), we try to localise most unknown nodes with just 1% of nodes being beacon nodes.

A predefined beacon node (mostly at the periphery of the region of interest) sends a signal to an unknown node that is within its range of communication. The unknown node registers the RSSI and the AoA of the signal from the beacon node. The RSSI value is used to determine the distance of the beacon node from the unknown node, and the AoA is used to determine the relative angle between the two. Combining these two factors leads to an accurate localisation of the unknown node.

The now localised “unknown” node is no longer unknown and becomes a new beacon node. This new beacon node now sends a signal to another unknown node within its range of communication, and in a similar manner, this unknown node is localised and becomes yet another beacon node. This process continues till all the nodes are localised or a beacon node does not have any unknown node within its range of communication. Figure 2 is a high-level depiction of the localisation process, and Algorithm 1 describes the same in more detail. The following subsections describe in detail the process of computing the distance between nodes based on the RSSI value of the received signal, the computation of the angle between the nodes based on the Angle of Arrival of the signal, and

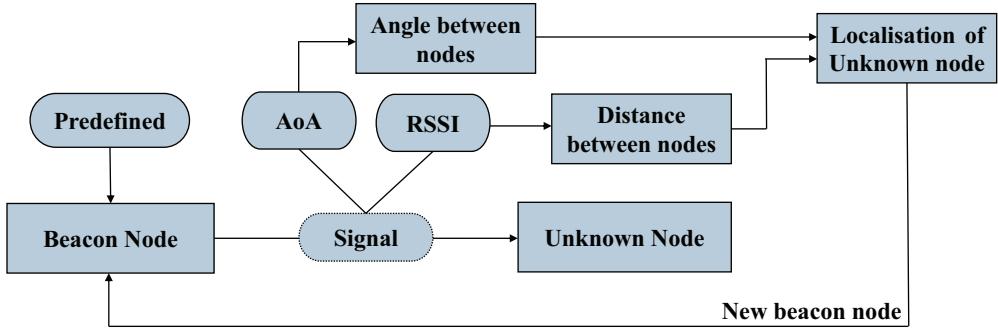


Fig. 2. Flow diagram of the proposed outdoor localisation method.

ALGORITHM 1: Hybrid localisation**Preconditions:**

```

1: Anchor nodes:  $A$ 
2: Unlocalised Sensor nodes:  $S$ 
3: Communication Range:  $C_R$ 
4: function RSSI( $A, S$ )
5:    $RSSI \leftarrow$  RSSI value between  $A$  and  $S$ 
6: end function
7: function AoA( $A, S$ )
8:    $AoA \leftarrow$  angle of arrival value between  $A$  and  $S$ 
9: end function
10: function LOCALISATION( $A, S$ )
11:    $RSSI \leftarrow$  RSSI( $A, S$ )
12:    $AoA \leftarrow$  AoA( $A, S$ )
13:   while  $num(S) \geq 0 \parallel l_{prev} = l_{now}$  do
14:     MULTI-ITERATION Localisation
15:      $l_{prev} \leftarrow$  # nodes localised in previous round
16:      $DIST \leftarrow$  Distance from  $RSSI$  and  $AoA$ 
17:     if  $DIST \leq C_R$  then
18:        $S^M \leftarrow$  node localised in this iteration
19:     else
20:       Ignore that node
21:     end if
22:      $l_{now} \leftarrow$  # nodes localised in current round
23:      $S \leftarrow S - S^M$ 
24:      $A \leftarrow A \cup S^M$ 
25:   end while
26: end function

```

the multiple iterations of the localisation process that ultimately enables coverage of most (if not all) of the region of interest.

3.1 Node Distance Using RSSI

The power level in a received radio transmission is measured by the RSSI. The strength of a wireless link is commonly measured in **decibel-milliwatts (dBm)**, with a large number indicating a strong signal and superior link quality. In addition to identifying issues with wireless interference, RSSI is used to calculate the distance between a wireless device and the access point to which it is

connected. The distance is calculated using the following relation:

$$\text{RSSI}(dBm) = 10 * \log_{10} (P_r/P_0), \quad (1)$$

where P_r is the power level received at the receiver, and P_0 is the reference power level (usually expressed as the received power at a distance of one metre from the transmitter). It is important to note that factors like distance, obstructions, and interference have a significant impact on the RSSI value. The precise value of RSSI, therefore, changes based on these. Also, the measurements and reference values used by various wireless technologies and manufacturers have a bearing on the computed RSSI value.

The approach proposed in this work also incorporates the path loss values incurred owing to the nature of the environment while computing the RSSI values. The environments include: the free space model (with no obstructions); sandy terrain [2]; long grassy terrain [1]; and terrains with sparse trees [2]. We endeavour to establish a mapping between these environments and the RSSI values by employing specific equations to calculate their respective path loss values:

$$L_p = PL_0 + 10 * \alpha * \log_{10} (P_r/P_0) + X_\sigma, \quad (2)$$

$$E_1 = 60.97 + 10 * 3.42 * \log_{10} (P_r/P_0) + 3.02, \quad (3)$$

$$E_2 = 59.42 + 10 * 2.56 * \log_{10} (P_r/P_0) + 3.84, \quad (4)$$

$$E_3 = 60.98 + 10 * 3.33 * \log_{10} (P_r/P_0) + 7.30. \quad (5)$$

In this context, PL_0 denotes the path loss value at a reference distance of d_0 measured in dB, while α signifies the path loss exponent, indicating that the rate of path loss increases as the logarithm of the distance. Additionally, X_σ represents a normally distributed random variable with a mean of zero and a standard deviation of σ . The path loss values in sandy terrains, long grassy terrains, and sparse tree terrains are denoted as E_1 , E_2 , and E_3 , respectively.

Our algorithm proposes the calculation of the RSSI value using Equation (1) to determine the signal strength between the unknown node and the beacon nodes.

3.1.1 RSSI to Distance. The computation of the RSSI value over various terrains makes it possible to determine the distance separating the unknown node from the beacon node. This distance, combined with the angle measurements, serves to determine the position of the unknown node. The distance, d , is calculated as follows using the RSSI value:

$$d = 10^{\left(\frac{P_m - \text{RSSI}}{10 \cdot N}\right)} \quad (6)$$

Where RSSI indicates the strength of the signal from the beacon node as received at the unknown node, measured power P_m indicates the RSSI value of the signal at 1 meter from the beacon node, and N is a constant (with a value between 2 and 4) that depends on the environmental factors.

3.2 Angle of Arrival (AoA)

AoA is the angle between a wave's propagation and a reference direction, also referred to as orientation. The orientation is expressed in degrees clockwise from the North. The AoA is absolute when the orientation is 0 or pointed to the North; otherwise, it is relative. The Angle of Arrival can be determined using radio waves, Bluetooth, RFID, or WiFi signals. Earlier, the AoA was determined via radio waves by arranging four antennas squarely on each anchor node. This facilitates angle-of-arrival computations [24] by adjusting the phases of signals across the antennas. A rotating beacon is generated by directing the maximum radiation from the antenna array. This process employs the scanning phased array technique, known as beamforming. The technique guarantees that four radio waves constructively interfere in a specified direction [7]. More recently, the demand for

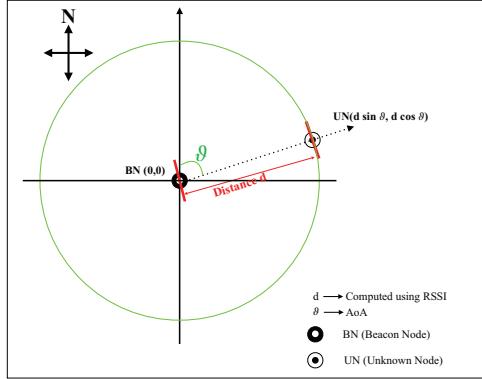


Fig. 3. Position calculation of any unknown node using the beacon node.

cost-effective indoor localisation systems has led to the adoption of more innovative approaches like **Bluetooth Low Energy (BLE)** and **Radio Frequency Identification (RFID)** technologies for computing the AoA. The AoA information is obtained using BLE devices in two ways: (1) through **Switched Beam Systems (SBS)** and (2) using **Adaptive Array Systems (AAS)**. The SBS approach scans the azimuth plane using a fixed number of beams to identify the direction with the highest received power or signal strength. The AAS, on the other hand, allows beam steering in any desired direction by altering weights across the antenna array elements [31]. Alternatively, RFID-based solutions use UHF RFID transponders and multiple directional antenna arrays. Given the normally narrow bandwidth of RFID communication channels, the AoA approach becomes essential, especially in environments tending to severe multi-path conditions. By analysing the phase difference between incoming signals, the AoA method relies on isolating a strong **line-of-sight (LOS)** path from other signal paths to estimate the transponder's position accurately [5].

In the presented work, our primary emphasis is on localisation. Implementation of the AoA computation method is beyond the scope of this work, and we do not dwell on it. For our real-world prototypical implementation, we use a stepper motor mounted with a laser for AoA calculations (the detailed description is mentioned in Section 4.3). For simulation, the Equation (7) shows the calculation of AoA values between three points:

$$\text{Angle} = \text{degrees}(\text{atan2}(N_y - BN_y, N_x - BN_x) - \text{atan2}(UN_y - BN_y, UN_x - BN_x)). \quad (7)$$

Where UN , BN , and N are the unknown nodes, beacon node and the pointer towards the north axis, respectively. The two-argument arctangent or $\text{atan2}(y, x)$ calculates the angle in radians between the positive x -axis and the line from the origin to the point (x, y) in the Cartesian plane. The resulting angle falls within the range of $-\pi < \theta \leq \pi$. Equation (7) is the approach adopted to calculate the AoA in simulations.

The AoA of signals of at least two beacon nodes usually are required to localise an unknown node. However, the proposed approach overcomes the limitation of requiring at least two beacon nodes. It is able to effectively localise an unknown node using a single beacon node through the combined use of the RSSI value and the AoA measurement. Figure 3 demonstrates the combination of the distance computed using RSSI values and the AoA to determine the location of the unknown node. The X and Y axes of the unknown node are as computed in Equation (8) using simple trigonometry.

$$X_{UN} = d_{RSSI} \sin(AoA), \quad (8)$$

$$Y_{UN} = d_{RSSI} \cos(AoA). \quad (9)$$

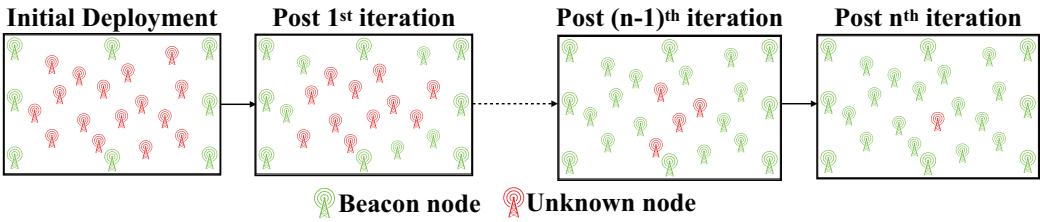


Fig. 4. Localisation using multi-iteration approach.

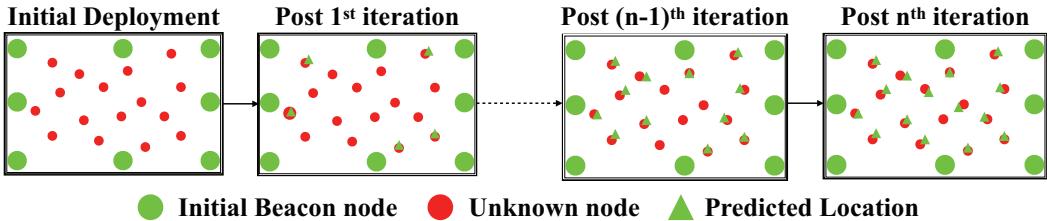


Fig. 5. The nodes with green and red circles represent the initial beacon nodes and unknown nodes, respectively. The green triangles represent the predicted location of a nearby unknown node. In multi-iteration, some unknown nodes are localised in every iteration with some error, and this will increase with the iteration round.

3.3 Multi-iteration

As explained in the last sections, the RSSI and AoA values of a signal from a beacon node enable the effective computation of the location of an unknown node that lies within the communication range of the beacon node. The use of RSSI and AoA together makes precise localisation possible with just one beacon node [17]. Subsequent to localisation, an unknown node becomes a “new” beacon node. The RSSI and AoA values from this new beacon node now become the basis for the localisation of subsequent nodes that fall within its communication range. This localisation process continues through several iterations until all the nodes in the region of interest are localised, or a scenario is arrived at with no unknown nodes within the communication range of beacon nodes. Figure 4 is a simple depiction of the progress of the localisation process.

In multi-iteration, initially, unknown nodes are localised (with some localisation error) using beacon nodes. The newly localised nodes subsequently play the role of beacon nodes in the following iterations and facilitate the localisation process further. The newly localised nodes have some more localisation error, which is a result of the error in their own localisation, and also the error carried over from the previous localisation. The localisation error, in this way, accumulates and increases through iterations, as shown in Figure 5. We use a combination of angle and distance to predict the location of any unknown node using only one beacon node, so the localisation error is small for initial iterations (less than 0.25m). The requirement of only one beacon node for localisation helps hasten the localisation process and converge the network faster. The result of this is that most localisations happen during the initial iterations, and the error is significantly minimised. A large number of localisation iterations would be counterproductive as the localisation errors accumulate with progressive iterations.

3.4 Network Localisability

The ground truth of the WSN network is depicted using a distance graph $G = (V, E)$, where V depicts the WSN nodes, and E represents an unweighted edge $(i, j) \in E$ between nodes i and

ALGORITHM 2: Successive localisability Algorithm**Preconditions:**

```

1:  $G = (V, E)$ 
2: function LOCALISABILITY( $G$ )
3:   Initially mark all  $v \in V$  in  $G$  as unlocalised
4:   Find a subgraph  $G_0 \in G$ ; mark all initially localised vertices (representing beacon nodes) as part of  $G_0$ 
5:   while  $v = (\emptyset G_0 \&& E_v \neq \emptyset)$  do
6:     if there are any unmarked vertices  $v$  that exist in  $G$  with at least one edge connected to one or more
      vertices in  $G_0$  then
7:       Add  $v$  into Graph  $G_0$  and mark it as localised
8:     else
9:       Move to another  $v$  in  $G$ 
10:    end if
11:   end while
12: end function
13: Count the total number of nodes localised

```

j. A graph/network is localisable if there is a unique location for every node in the graph and the distance between nodes i and j , $d(i, j): E \rightarrow \mathbb{R}$ conforms to certain conditions (in this case: the distance being less than the communication range (CR) of the WSN nodes). Network localisability can be widely used in network deployment, routing, energy management, and mobility control [28].

A successive localisable algorithm can be used to verify the localisability of a graph that harnesses iterative localisation. Existing solutions use the trilateration method to determine node localisability in a network, hence they use four connected components (K_4) to verify the localisability. The proposed algorithm, on the other hand, uses only one beacon node and thus there are only two connected components (K_2) involved in determining the location of an unknown node. Algorithm 2 is the successive localisability algorithm for our proposed method.

The necessary and sufficient condition for node localisability is that the node, represented by vertex v , must be one hop away from a beacon node.

4 Experimental Evaluation

This section is dedicated to examining the effectiveness of the proposed localisation approach through experiments. The evaluation process comprises comprehensive validation of the approach in a controlled environment using simulations. This is followed by an assessment of the approach in a real-world setting through a prototypical implementation. Section 4.1 contains information on the simulation environment and the parameters employed in the experiment. The results of the experiments conducted over the simulated environment and comparison with other state-of-the-art methods are discussed in Section 4.2. Finally, Section 4.3 demonstrates the efficacy of the approach in the real-world through a prototypical implementation.

4.1 Simulation Environment and Parameters

The proposed protocol is implemented using Python [42] and evaluated using a Kaggle Kernel on a system that comprises an Intel (R) Xeon (R) CPU from the Haswell CPU family, 16 GB RAM, and four CPUs with a frequency of 2.30 GHz each. A GPU is used to handle computationally intensive tasks, such as conversion of RSSI values to distance and Angle of Arrival calculations. The calculations involve large matrix values, and these can be efficiently computed in parallel using a GPU, especially since the calculations are independent of each other.

Table 1. Simulation Parameters

Parameters	Values for small regions (R_1)	Values for extensive regions (R_2)
Size of region	1000×1000 , 2000×2000 to 4000×4000	5000×5000 , 6000×6000 , and 7000×7000
IUN [*]	6000, 7000, and 8000	10000, 11000, 12000 to 15000
OBN [†]	1, 1.5, 2.5, and 5 %	1, and 1.5 %
CR [‡]	50, 75, and 100 m	75, and 100 m
EN [§]	E_0 , E_1 , E_2 , and E_3	E_0 , E_1 , E_2 , and E_3

Note: Size of the region is measured in square metres, IUN^{*} denotes the number of unknown nodes initially deployed, OBN[†] denotes the number of original beacon nodes (in percent of the number of IUN), and CR[‡] denotes the communication range of a node (in metres), and EN[§] denotes the effect of noise occurred due to various environmental factors like free space (E_0), sandy terrain (E_1), long grassy (E_2), and sparse tree (E_3).

Table 1 contains information on the simulation environment and the parameters used in the simulations to comprehensively assess the efficacy of the proposed localisation approach. The parameters used in the simulation are described in detail as follows:

- Region of Interest (R_x): Different sizes of square regions of interest are experimented with. The regions of interest are categorised into two types, R_1 and R_2 . The smaller regions of interest fall within R_1 , and the larger ones within R_2 . The different sizes of R_1 and R_2 used are shown in Table 1. Running the experiments on different sizes of the region of interest provides insight into the impact of the latter on the efficacy of our algorithm.
- Initial Unknown Nodes (IUN): **Initial Unknown Nodes (IUNs)** are nodes whose coordinates are not known initially and which are localised using the proposed approach. The number of IUNs is widely varied, as shown in the table for a thorough assessment of the approach. This is a good yardstick by which to assess the proposed approach.
- Original Beacon Nodes (OBN): There is a very small number of **Original Beacon Nodes (OBN)**, which are nodes whose locations are assumed to be known in advance. The number of OBNs is expressed as a percentage of unknown nodes, IUN, and varies between 1% and 5%. Experimenting with a small number of OBNs gives an idea of the utility of the approach as it appropriately demonstrates how such a small number of nodes forms the genesis of localising a large area. The distribution of the OBNs is also varied along the following lines: **Equidistant Deployment on the Boundary (EDB)**, where the OBNs are deployed on the boundary of the region of interest and are equidistant from each other; **Random Deployment at Boundary (RDB)**, where the OBNs are deployed randomly (not necessarily in an equidistant manner) on the boundary of the region of interest; and **Random Deployment (RD)**, where the OBNs are randomly deployed in the region of interest and not necessarily only at the boundary.
- Communication range (CR): The node's communication range implies the distance that a radio signal sent from a node travels. The magnitude and direction of these radio signals form the basis for assessing the distance and angle, respectively, of an unknown node from the known beacon node. Hence the CR plays a vital part in ensuring the efficacy of the approach. Varying the CR and assessing the effectiveness of the approach gives an idea of its utility in a heterogeneous deployment of nodes. For a better understanding of the effect of CR on localisation accuracy, we experiment with different values of communication range (as shown in Table 1) for both small and extensive regions.
- Environment Noise (E_x): Environment noise is a factor that impedes the smooth passage of radio signals between nodes. This is an important factor and needs to be analysed as the passage of radio signals plays an important role in assessing the relative positions of unknown

nodes and, subsequently, their localisation. A value of 0 for x implies no disturbance in the medium, and progressively larger values of x indicate noise and disturbance. The noise can be due to factors like irregular surfaces like sand, tall grass, trees and so on. A few values of x are experimented with in our simulations, and these correspond to surfaces with sand Equation (3), grass Equation (4), and sparse trees Equation (5). The effect of these on the localisation efficacy of the proposed approach is assessed.

4.2 Simulation Results

All possible combinations of parameters in Table 1 are experimented with, and the following factors are used to express and comprehend the outcome and efficacy of the proposed localisation method:

- Number of iterations (*NIT*): This factor specifies the number of times that the multi-iteration algorithm runs to localise the network to the maximum possible extent. Further iterations beyond *NIT* do not result in the localisation of any further node. The network localisation process depends critically on this value; the fewer the number of iterations, the quicker the network localisation process.
- Unknown nodes Remaining (*UNR*): *UNR* is the count of nodes that remain unlocalised at the end of the localisation process due mainly to not being within the communication range of any beacon node; the smaller this number, the better the algorithm.
- Localisation Error (*LE*): The Euclidean distance between the node's actual position and the position determined by the proposed approach expresses the localisation error.

$$\text{Localisation Error (LE)} = \sqrt{(X_{\text{actual}} - X_{\text{predicted}})^2 + (Y_{\text{actual}} - Y_{\text{predicted}})^2}, \quad (10)$$

- Average localisation error (*ALE*): The average localisation error is calculated as

$$\text{Total Localisation Error (TLE)} = \sum_{i=1}^n LE_i, \quad (11)$$

$$ALE = \frac{TLE}{(IUN - UNR)}. \quad (12)$$

- Localisation time (*LT*): The localisation time expresses the time (in seconds) expended to localise the network to the maximum possible extent. This factor provides an important insight, and values of *LT* beyond a certain limit may be unacceptable.

4.2.1 Experimental Environment. The experiments over the simulation environments were conducted with the intent of gaining insight into the efficacy of the proposed localisation approach under varying circumstances. The first factor that varied was the size of the region of interest (*ROI*). The sizes of the *ROI* that we experimented with vary between 1000×1000 square meters and 7000×7000 square meters. These are reasonably large outdoor locations, and any localisation approach effective in such an area can safely be assumed to be effective in outdoor locations of any size. The reason we restricted ourselves to simulation regions up to this size and not beyond is that the computation time was becoming exorbitantly long, and it was becoming difficult to conduct these with the computing resources available to us. The localisation error (*LE*), number of iterations (*NIT*) required to complete the localisation process, the number of nodes remaining unlocalised (*UNR*) at the end of the localisation exercise, and the localisation time (*LT*) were observed for different sizes of the region of interest. In addition to this, through the simulations, we tried to gain an understanding of the efficacy of our approach in various kinds of terrains. A factor called Environmental noise (E_x) discussed earlier, is introduced to represent different kinds of

terrains. E_0 represents a perfectly smooth terrain that is not realistic but useful as a benchmark; E_1 indicates a sandy terrain; E_2 is a grassy terrain; and finally, E_3 is a terrain with sparse trees.

Each of these terrains is progressively noisier than the earlier one and impedes the smooth passage of signals between the WSN nodes. It is important that the effect of the signal passage in various kinds of terrains be understood, as the latter forms the basis of our localisation exercise.

The deployment of the WSN nodes in the simulation environment is done in a manner that simulates a real-world deployment. The initially unknown nodes (*IUN*), the nodes whose locations are to be determined, are randomly deployed over the entire region of interest. The original beacon nodes (*OBN*), the nodes whose locations are known in advance, are very small in number and are deployed in three ways: (1) Randomly (*RD*) across the region of interest. This deployment scheme for beacon nodes conforms to parts of the region of interest that are accessible with GPS connectivity and where the node's original locations can be determined. These could correspond to zones in a forest area, for example, that are in clearings and have GPS connectivity; (2) Equidistant along the boundary (*EDB*), this is a realistic deployment for beacon nodes as the peripheries of the region of interest are usually accessible and the location of nodes can be determined in advance through GPS. In this scheme, the beacon nodes are equidistant from each other; and (3) Randomly on the boundary (*RDB*), the beacon nodes are again on the periphery and are randomly deployed at the periphery and are not necessarily equidistant. The three schemes of deployment give an idea of the nature and effectiveness of the proposed localisation approach for varied kinds of regions of interest.

Tables 2, 3, 4, 5, 6, and 7 show the results of the average values for various parameters described above to assess the efficacy of the proposed approach in different sizes of the region of interest with different terrains and with different deployment schemes for the beacon nodes. The comprehensive results are very large in number and cannot be shown here but are available for the interested reader at the following [url](#).

4.2.2 Visualisation of the Localisation Process. In this section, we visualise the localisation process in a simulation set-up comprising a large number of initially unlocalised nodes (*IUN*) and a very small number of original beacon nodes (*OBN*). There are three visualisations (for *ROI*, *IUN*, and *OBN* is $3000 \times 3000\text{ m}^2$, 4000, and 1%, respectively), one for each of the three deployment schemes for the *OBNs*. The visualisations depict the process of localising unknown nodes through localisation epochs, ultimately converging to the localisation of all nodes or a majority of the nodes. Figure 6 depicts the localisation of nodes when the *OBNs* are randomly deployed over the entire region of interest. The *OBNs* are coloured blue and are very small in number in comparison to the unlocalised nodes, *IUNs*, shown in red. As the *IUNs* are localised, their colour changes from red to green. Figure 7 depicts the localisation process when the beacon nodes are deployed at the boundary of the region of interest and equidistant from each other. Finally, Figure 8 depicts the localisation process for *OBNs* deployed randomly at the boundary of the region of interest. The intent of the visualisations is to help the reader appreciate the process and better comprehend its progress.

The number of iterations required to localise the *UNs* varies owing to differences in deployment type of the original beacon nodes (*OBNs*). Random deployment (*RD*) of *OBNs* across the entire region mostly leads to fewer iterations when compared to deployment of *OBNs* on the boundaries (equidistant (*EDB*) or random (*RDB*)). Random deployment (*RD*) of beacon nodes over the entire region makes them proximate to a much larger number of unknown nodes per iteration, thus localising them immediately and leads to fewer iterations. Similarly, *OBNs* deployed randomly on the boundary (*RDB*) require fewer iterations than when *OBNs* are deployed in a strictly equidistant manner on the boundary (*EDB*). In Figures 6, 7, and 8, the deployment of the initial unknown nodes (*IUN*) and the *OBNs* is random, and visualisation is shown for an area of size 3000×3000

Table 2. Beacon Nodes Deployed Randomly in Network (RD)

EN*	Size[†]	A-LF[‡]	A-NIT[§]	A-UNR[¶]	A-LT
E_0	$1000 \times 1000 \text{ m}^2$	0.0495	2.89	0.00	38
	$2000 \times 2000 \text{ m}^2$	0.0498	5.64	0.08	67
	$3000 \times 3000 \text{ m}^2$	0.0526	11.56	40.44	136
	$4000 \times 4000 \text{ m}^2$	0.0579	23.00	1333.08	634
	$1000 \times 1000 \text{ m}^2$	0.0643	3.76	0.00	41.88
	$2000 \times 2000 \text{ m}^2$	0.0648	7.33	0.09	73
E_1	$3000 \times 3000 \text{ m}^2$	0.0684	15.02	42.47	150
	$4000 \times 4000 \text{ m}^2$	0.0752	29.90	1399.74	697
	$1000 \times 1000 \text{ m}^2$	0.1712	5.63	1.52	51
	$2000 \times 2000 \text{ m}^2$	0.1811	11.00	1.75	103
	$3000 \times 3000 \text{ m}^2$	0.1856	22.53	34.02	200
	$4000 \times 4000 \text{ m}^2$	0.1911	65.88	1413.64	942
E_2	$1000 \times 1000 \text{ m}^2$	0.3413	7.04	2.90	63
	$2000 \times 2000 \text{ m}^2$	0.3719	15.28	3.75	144
	$3000 \times 3000 \text{ m}^2$	0.4123	31.77	71.78	266
	$4000 \times 4000 \text{ m}^2$	0.4454	87.82	1540.87	1271
	$1000 \times 1000 \text{ m}^2$	0.4955	10.00	0.00	213
	$2000 \times 2000 \text{ m}^2$	0.5797	16.97	0.00	286
E_3	$3000 \times 3000 \text{ m}^2$	0.576	30.81	30.11	507
	$4000 \times 4000 \text{ m}^2$	0.591	35.97	1747.83	1108
	$1000 \times 1000 \text{ m}^2$	0.851	9.81	0.00	235
	$2000 \times 2000 \text{ m}^2$	0.753	20.88	0.00	309
	$3000 \times 3000 \text{ m}^2$	0.748	37.89	31.62	542
	$4000 \times 4000 \text{ m}^2$	0.769	44.25	1835.23	1178
E_4	$1000 \times 1000 \text{ m}^2$	0.1874	14.71	1.97	290
	$2000 \times 2000 \text{ m}^2$	0.1924	31.31	2.18	432
	$3000 \times 3000 \text{ m}^2$	0.1966	56.84	44.92	721
	$4000 \times 4000 \text{ m}^2$	0.1989	65.88	1908.02	1591
	$1000 \times 1000 \text{ m}^2$	0.3718	18.39	3.75	356
	$2000 \times 2000 \text{ m}^2$	0.4019	43.53	4.66	605
E_5	$3000 \times 3000 \text{ m}^2$	0.4383	80.14	94.78	958
	$4000 \times 4000 \text{ m}^2$	0.4513	87.82	2079.74	2148
	$1000 \times 1000 \text{ m}^2$	0.601	8.00	0.00	209
	$2000 \times 2000 \text{ m}^2$	0.567	16.92	0.06	300
	$3000 \times 3000 \text{ m}^2$	0.565	31.17	38.11	529
	$4000 \times 4000 \text{ m}^2$	0.593	33.97	1892.64	956

Table 3. Beacon Nodes Deployed Equidistance on Boundary (EDB)

Table 4. Beacon Nodes Deployed Randomly on Boundary (RDB)

	A-LF[‡]	A-NIT[§]	A-UNR[¶]	A-LT
	0.0601	8.00	0.00	209
	0.0655	7.97	0.00	213
	0.0579	16.97	0.00	286
	0.0576	30.81	30.11	507
	0.0591	35.97	1747.83	1108
	0.0851	9.81	0.00	235
	0.0753	20.88	0.00	309
	0.0737	21.99	0.06	330
	0.0735	40.52	40.02	582
	0.0770	44.16	1987.27	1052
	0.1874	15.60	2.10	283
	0.1914	32.99	2.78	462
	0.1966	60.78	48.91	774
	0.2111	65.88	1951.56	1420
	0.3824	19.50	3.99	348
	0.4183	45.85	5.95	646
	0.4423	85.69	103.20	1030
	0.4609	87.82	2127.20	1917

Note: EN^{*} denotes the noise included by external environmental factors while calculating the RSSI value such as for free space (E_0), sandy terrain (E_1), long grassy terrain (E_2), and sparse tree terrain (E_3). Size[†] denotes the area of the network in the square of metres, A-LF[‡] denotes Average error per node for localisation (in metres), A-NIT[§] denotes the average number of iterations, A-UNR[¶] denotes the average number of nodes remains unknown, A-LT^{||} denotes Average execution time (in seconds).

Table 5. Beacon Nodes Deployed Randomly in Network (RD)

EN*	Size[†]	A-LE[‡]	A-NIT[§]	A-UNR[¶]	A-LT
E_0	$5000 \times 5000 \text{ m}^2$	0.0740	12.75	8.42	650
	$6000 \times 6000 \text{ m}^2$	0.0740	18.63	150.96	1043
	$7000 \times 7000 \text{ m}^2$	0.0740	27.63	1188.29	2130
E_1	$5000 \times 5000 \text{ m}^2$	0.0962	16.58	8.84	715
	$6000 \times 6000 \text{ m}^2$	0.0962	24.21	158.51	1147
	$7000 \times 7000 \text{ m}^2$	0.0962	35.91	1247.71	2343
E_2	$5000 \times 5000 \text{ m}^2$	0.2089	24.86	7.74	943
	$6000 \times 6000 \text{ m}^2$	0.2109	32.44	91.36	1560
	$7000 \times 7000 \text{ m}^2$	0.2292	48.12	1517.19	3234
E_3	$5000 \times 5000 \text{ m}^2$	0.5012	32.82	9.52	1245
	$6000 \times 6000 \text{ m}^2$	0.5901	44.12	103.23	2122
	$7000 \times 7000 \text{ m}^2$	0.6442	68.33	1618.84	4463

Table 6. Beacon Nodes Deployed Equidistance on Boundary (EDB)

EN*	Size[†]	A-LE[‡]	A-NIT[§]	A-UNR[¶]	A-LT
E_0	$5000 \times 5000 \text{ m}^2$	0.0756	38.08	5.33	2946
	$6000 \times 6000 \text{ m}^2$	0.0745	52.42	110.79	4170
	$7000 \times 7000 \text{ m}^2$	0.0740	71.00	1883.75	5912
E_1	$5000 \times 5000 \text{ m}^2$	0.0983	46.84	5.60	3103
	$6000 \times 6000 \text{ m}^2$	0.0969	64.47	116.33	4396
	$7000 \times 7000 \text{ m}^2$	0.0962	87.33	1977.94	6227
E_2	$5000 \times 5000 \text{ m}^2$	0.2112	70.26	9.19	4096
	$6000 \times 6000 \text{ m}^2$	0.2289	86.39	129.02	5978
	$7000 \times 7000 \text{ m}^2$	0.2268	117.02	2011.94	8593
E_3	$5000 \times 5000 \text{ m}^2$	0.5231	92.75	11.31	5407
	$6000 \times 6000 \text{ m}^2$	0.6501	117.49	145.79	8131
	$7000 \times 7000 \text{ m}^2$	0.6834	166.17	2146.74	11859

Table 7. Beacon Nodes Deployed Randomly on Boundary (RDB)

EN*	Size[†]	A-LE[‡]	A-NIT[§]	A-UNR[¶]	A-LT
E_0	$5000 \times 5000 \text{ m}^2$	0.0758	37.96	6.46	2939
	$6000 \times 6000 \text{ m}^2$	0.0749	54.46	136.63	4318
	$7000 \times 7000 \text{ m}^2$	0.0748	73.58	1985.17	6701
E_1	$5000 \times 5000 \text{ m}^2$	0.0985	49.35	6.78	3233
	$6000 \times 6000 \text{ m}^2$	0.0974	70.80	143.46	4750
	$7000 \times 7000 \text{ m}^2$	0.0973	95.66	2084.43	7372
E_2	$5000 \times 5000 \text{ m}^2$	0.2232	74.02	11.56	4267
	$6000 \times 6000 \text{ m}^2$	0.2289	94.87	135.02	6460
	$7000 \times 7000 \text{ m}^2$	0.2312	128.18	2034.16	10173
E_3	$5000 \times 5000 \text{ m}^2$	0.5313	97.70	14.22	5633
	$6000 \times 6000 \text{ m}^2$	0.6622	129.02	152.57	8785
	$7000 \times 7000 \text{ m}^2$	0.6914	182.02	2170.45	14039

Note: EN^{*} denotes the noise included by external environmental factors while calculating the RSSI value such as for free space (E_0), sandy terrain (E_1), long grassy terrain (E_2), and sparse tree terrain (E_3). Size[†] denotes the area of the network in the square of metres, A-LE[‡] denotes Average error per node for localisation (in metres), A-NIT[§] denotes the average number of iterations, A-UNR[¶] denotes the average number of nodes remains unknown, A-LT^{||} denotes Average execution time (in seconds).

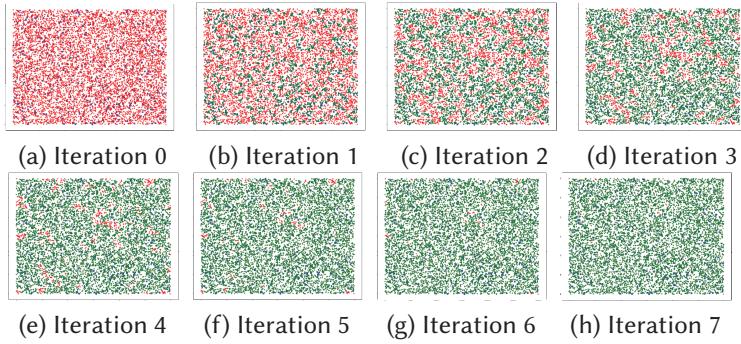


Fig. 6. Node localisation for OBNs deployment. Here, blue nodes represent OBNs, while red and green nodes represent unlocalised and localised nodes in that iteration, respectively.

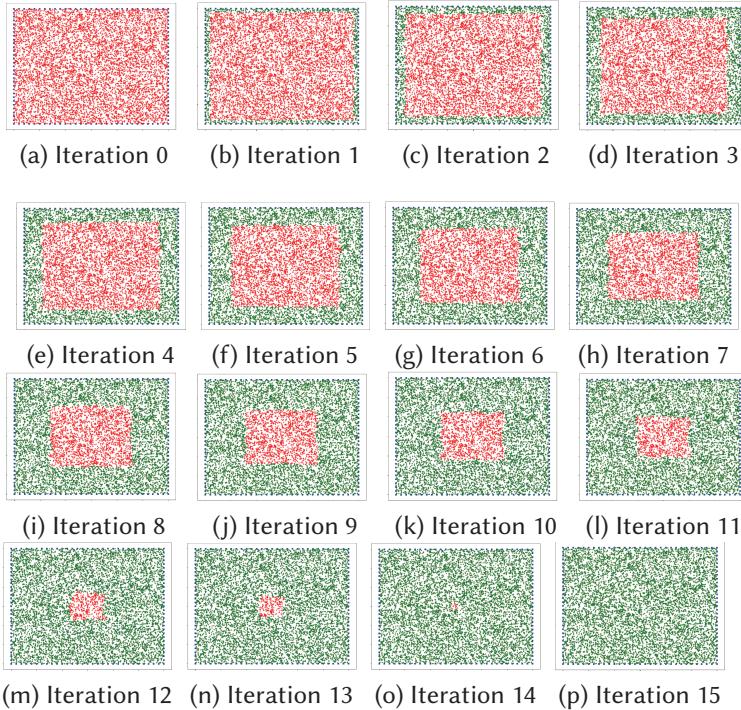


Fig. 7. Node localisation for OBNs deployment at boundary. Here, blue nodes represent OBNs, while red and green nodes represent unlocalised and localised nodes in that iteration, respectively.

m^2 . The EDB deployment follows a particular pattern for localising unknown nodes (as seen in Figure 7), resulting in fewer UNRs (seen as red nodes) than the RDB. In the case of RDB, however, the localisation follows irregular patterns and may end up with more UNRs (seen as red nodes), which adversely varies the number of iterations required to localise the entire region.

4.2.3 Impact of Varying the Area of the Region of Interest. In this section and a few subsequent ones, we study the impact of varying the characteristics of the region of interest on the efficacy of the proposed method. This is important to confidently establish the proposed method as a universal

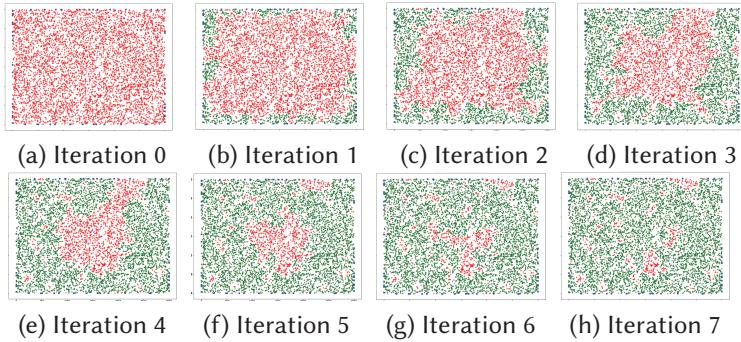


Fig. 8. Node localisation for OBNs deployed randomly at the boundary. Here, blue nodes represent initial beacon nodes, while red and green nodes represent unlocalised and localised nodes in that iteration, respectively.

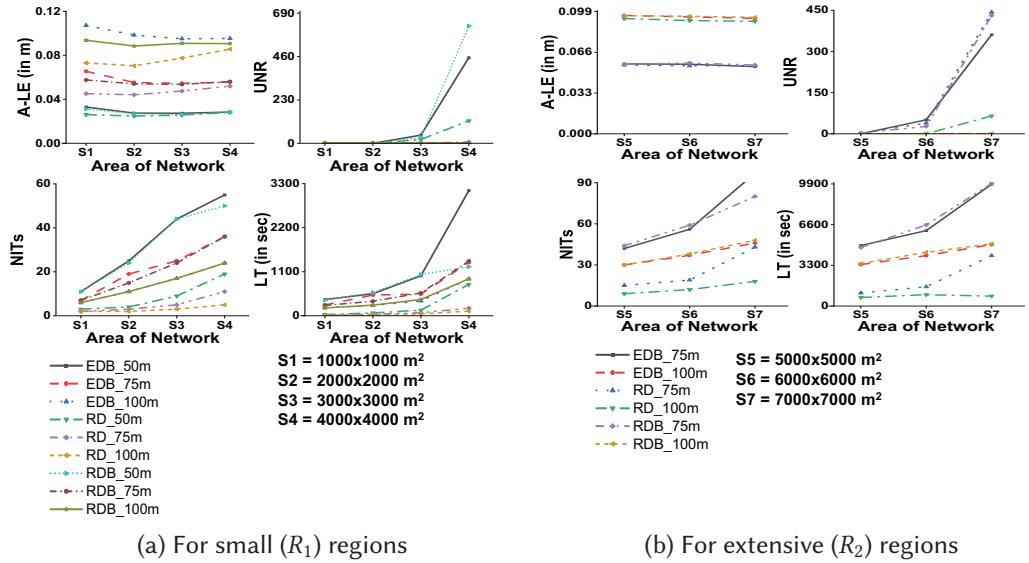


Fig. 9. Impact of varying area. The plot labels represents the communication range and deployment type (like 50m_EDB represents the EDB type of deployment with 50m communication range of a node).

solution for localisation irrespective of the nature of the region. The first factor that we study is the size of the region of interest (ROI). Figure 9 shows the effect of variation of the area of the region of interest on: the average localisation error ($A - LE$), the number of unknown nodes remaining (UNR), the number of iterations ($NITs$) required to localise most of the nodes in the ROI , and the localisation time (LT) taken to localise the entire ROI .

To provide a more precise visualisation of the effects, the following configurations were used for plotting the figures. Two ROI sizes are experimented with: small regions (R_1) with size ranging from 1000×1000 to 4000×4000 square meters; and extensive regions with size ranging from 5000×5000 to 7000×7000 square meters. The number of unknown nodes, IUN , to start with is 15,000. The number of beacon nodes, $OBNs$ is 1% of the known beacon nodes which works out to 150. The $OBNs$ are deployed along all three schemes, randomly throughout the region (RD), randomly at the boundaries (RDB), and at the boundaries equidistant from each other (EDB). The

communication range of WSN nodes also varies in real deployments and the same is simulated by varying the communication range of radio signals between nodes. The communication ranges experimented with are 50 m, 75 m, and 100 m, for small regions (R_1); and 75 m and 100 m for extensive regions (R_2). We do not experiment with 50 m in extensive regions as the nodes in these regions are sparse and 50 m becomes too small a range for effective localisation. The effect of change in area of the ROI on each factor is discussed as follows:

- Average Localisation Error ($A - LE$): The average localisation error varies very slightly as the size of the ROI increases. This validates the claim that the proposed model can be used for localisation in outdoor locations of any size with a satisfactory degree of accuracy. In fact, there is a very slight reduction in the localisation error with an increase in the ROI size which bodes well for the performance of the system in even larger $ROIs$.
- Unknown Nodes Remaining (UNR): The number of initial unknown nodes, $IUNs$, that remain unlocalised at the end of the localisation process is near zero for smaller $ROIs$ but for large $ROIs$ this number increases suddenly. This is because a larger ROI leads to several nodes falling beyond the communication range, CR of the nearest known node. The Unknown Nodes Remaining $UNRs$ are mostly near zero when the communication range being considered is large (75 m and 100 m). This is because a large communication range leads to almost all unknown nodes falling within the range of at least one beacon node and thus being localised. This is especially so when the distribution of $IUNs$ is random. The number of $UNRs$ is large when the communication range is small (50 m). For R_1 regions, the number of $UNRs$ is small for S_1 , S_2 , and S_3 , and has an average value of 0.02, 0.28, and 36.91 nodes. This value is higher for the S_4 region with the average being 317.90. There is a large difference in $UNRs$ in areas (S_1 , S_2 , and S_3) and S_4 because the same number of IUN is distributed in different area sizes. The deployment thus becomes dense for S_1 , S_2 , and S_3 as these are smaller in size compared to S_4 . Similarly, for R_2 regions, when the communication range is small (75 m), the number of UNR for area sizes S_5 , S_6 , and S_7 has an average value of 8.42, 30.96, and 318.73. Just like the case of R_1 region, the $A - UNR$ is low for areas S_5 and S_6 compared to that for S_7 because the deployment of IUN is dense in the former. These results are for free space environment (E_0) when the beacon nodes are placed randomly throughout the regions. In a similar manner, the results are computed for three other environments with two other deployment types and can be accessed at the following [URL](#).
- Number of Iterations ($NITs$): The $NITs$ needed to localise the entire ROI in general increases with the region size. Yet, in some cases (as seen in Figure 9 (for $NITs$)), the $NITs$ remain constant beyond a certain size of the ROI . This is because the UNR , as explained in the last point, increases for larger $ROIs$, and the number of iterations required for localisation is unchanged.
- Localisation Time (LT): The time required to localise the entire ROI increases with an increase in size of the region. This is along expected lines as a larger ROI leads to a larger number of iterations ($NITs$) and hence time. There is a slight decrease in the LT , however, when the assumed communication range, CR , of the nodes is large. This is because a larger CR results in further nodes, and in effect, a larger number of nodes being localised in each iteration and hence a smaller LT . Another phenomenon, not as common, is the decrease in LT with an increase in the size of the ROI . The reason for this could be due to a larger number of nodes remaining unlocalised (UNR), leading to a smaller number of nodes localised, and effectively smaller LT .

4.2.4 Effect of Change in Number of Initial Unknown Nodes (IUN). In this section, we understand the impact of varying the number of Initial Unknown Nodes ($IUNs$) on: the Average Localisation

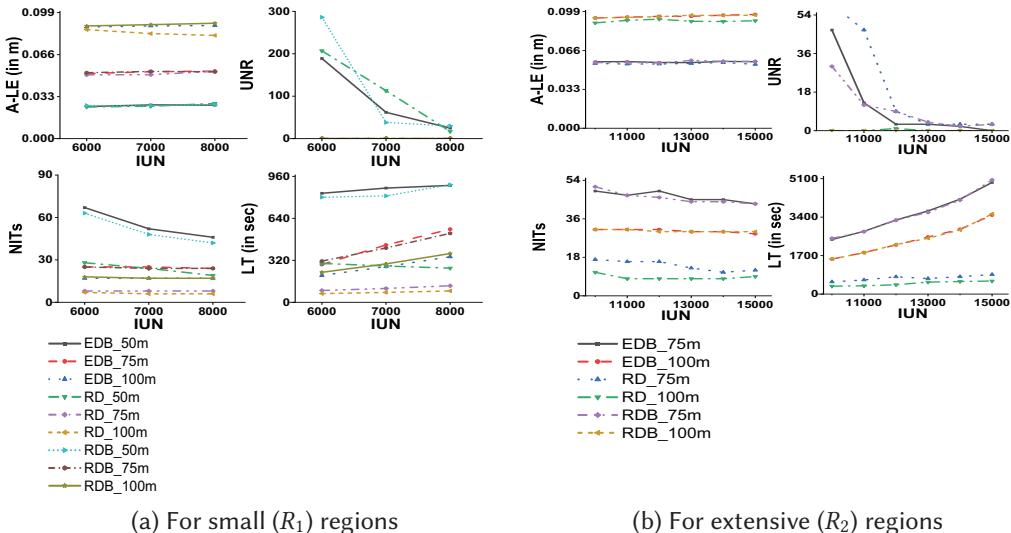


Fig. 10. Impact of varying the initial unknown nodes with respect to other parameters. The plot labels represents the communication range and deployment type (like 50m_EDB represents the EDB type of deployment with 50m communication range of a node).

Error ($A - LE$); the number of unlocalised nodes remaining (UNR); the total number of iterations ($NITs$) required to localise the entire ROI ; and the localisation time (LT) to localise the ROI (LT). The graphs in Figure 10 demonstrate the effect of varying the number of $IUNs$ on these factors. Here, we also look at two different ROI sizes. The small region (R_1) measures 3000×3000 square meters, and the number of original beacon nodes ($OBNs$), nodes whose locations are known in advance, is 1.5% of the $IUNs$ (ie. 90 when the number of $IUNs$ is 6000). The number of $IUNs$ that we experiment with in small regions is 6000, 7000, and 8000. This gives us a good idea of the behaviour of the model as the number of $IUNs$ changes. The communication range of the nodes in small regions that we experiment with are 50 m, 75 m, and 100 m.

Similarly, we conduct the experiments in extensive regions (R_2) as well that measure 5000×5000 square meters. The number of $OBNs$ here is also 1.5% of the number of $IUNs$. The number of $IUNs$ in these extensive regions varies between 10,000 and 15,000. The communication range of the nodes in extensive regions is 75 m, and 100 m. As mentioned earlier, a communication range of 50 m proves to be too small for large regions with nodes far apart.

The specific observations for each of the factors as the number of $IUNs$ varies is as follows:

- Average Localisation Error ($A - LE$): The $A - LE$ does not vary sharply but does increase gradually as the number of $IUNs$ increase, both in small and extensive regions. This is along expected lines; a more significant number of $IUNs$ would compromise the localisation accuracy a little, given that the known $OBNs$ do not increase much.
- Unknown Nodes Remaining (UNR): The number of nodes that remain unknown at the end of the localisation process (UNR) falls sharply with an increase in the number of $IUNs$. This is interesting and is attributed to the fact that a larger number of IUN within a fixed size of ROI leads to a higher density of $IUNs$. In other words, the number of $IUNs$ within a unit area of the ROI becomes more extensive and thus, the nodes are closely packed together. This leads to a sharp reduction in the number of nodes that are not within the communication range of known (already localised) nodes, and thus, very few are left out at the end as $UBNs$.

- Number of Iterations (*NITs*): The number of iterations remains relatively stable for large communication ranges like 75 m and 100 m. However, when the communication range is 50 m, there is a sharp fall in the *NITs* required to localise all nodes. This is because larger communication ranges already facilitate communication between most neighbouring nodes and do not benefit much from nodes being closely packed to each other (which is a consequence of an increased number of *IUNs* in the same sized *ROI*). Systems with small communication ranges (like 50 m), however, significantly benefit from an increased number of *IUNs* as most nodes are more closely packed and most nodes fall within the communication range of each other.
- Localisation Time (*LT*): The time required to localise the entire network, or the expected localisation time, increases with a larger number of *IUNs* and hence node density. This rise is attributed to an increased number of computations required to localise a larger number of nodes.

4.2.5 Effect of Varying Deployment Type. In this section, we explore the effect of the deployment strategy on various parameters characterising our localisation approach. Three localisation strategies are studied: (1) Equidistant Deployment at the Boundary (*EDB*), where the original beacon nodes (*OBNs*), nodes whose locations are known at the beginning, are deployed along the peripheries of the region of interest and are at equal distance from each other; (2) Random Deployment (*RD*), where the (*OBNs*) are deployed randomly across the *ROI*; and (3) Randomly at the Boundary (*RDB*), where the *OBNs* are deployed at the periphery and randomly and not necessarily equidistant from each other. *EDB* and *RDB* deployments are common as the *ROIs* are usually inhospitable terrains, and establishing the location of a node in advance (using GPS devices or other common means) is only possible at the periphery of the *ROI*. *RD* across the *ROI* is less common, but in terrains like forests, there are random points that can be reached and which have feeble GPS connectivity, and thus deployment of *OBNs* is possible. These deployment strategies were chosen for a comprehensive study of all possibilities. Section 4.2.2 has nice visualisations of the localisation process for these three deployments.

The experiments to study the effect of the deployment strategy were conducted over small (R_1) regions and extensive (R_2) regions. In R_1 , there were 8000 *IUNs* and the *OBNs* numbered 1% (or 80) of the *IUNs*. For the larger R_2 regions, the number of *IUNs* was 15000, and *OBNs* were again 1% (or 150 in number) of the *IUNs*. The communication range of the nodes in both regions was fixed at 150 m. Figure 11 shows the changes in the characteristics of the localisation approach with a change in the area of the network, the *ROI*.

- Average Localisation Error (*A-LE*): The *A-LE* remains largely constant and has a small value for the three deployment types as the size of the *ROI* increases. This is true for both small and large deployment regions. This result reiterates our claim that the proposed approach is an effective localisation tool irrespective of the type and size of the *ROI*.
- Unknown Nodes Remaining (*UNR*): The *UNR*, the number of nodes that remain unlocalised at the end of the localisation process, is near zero for relatively small *ROIs*, which is desirable. As the *ROI* size increases for the same number of *IUNs* (Initially Unlocalised Nodes), the deployment density of the *IUNs* drops and certain *IUNs* end up quite far away from the nearest known node. This results in such *IUNs* getting missed and adds to the number of *UNRs*. This leads to a sudden and rapid spike in the graph.
- Number of Iterations (*NIT*): The number of iterations (*NITs*) required to localise the entire network uniformly increases as the size of the *ROI* increases. This is along expected lines as the localisation process gradually moves from the peripheries to the centre (in the case of *EDB* and *RDB*). In the case of *RD*, the *NITs* required are much smaller because the existence

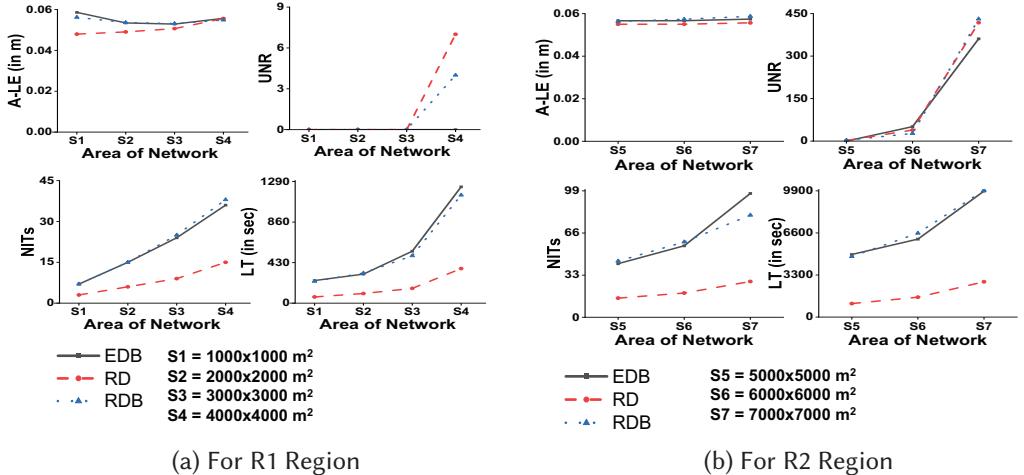


Fig. 11. Impact of different deployment type while varying the Area.

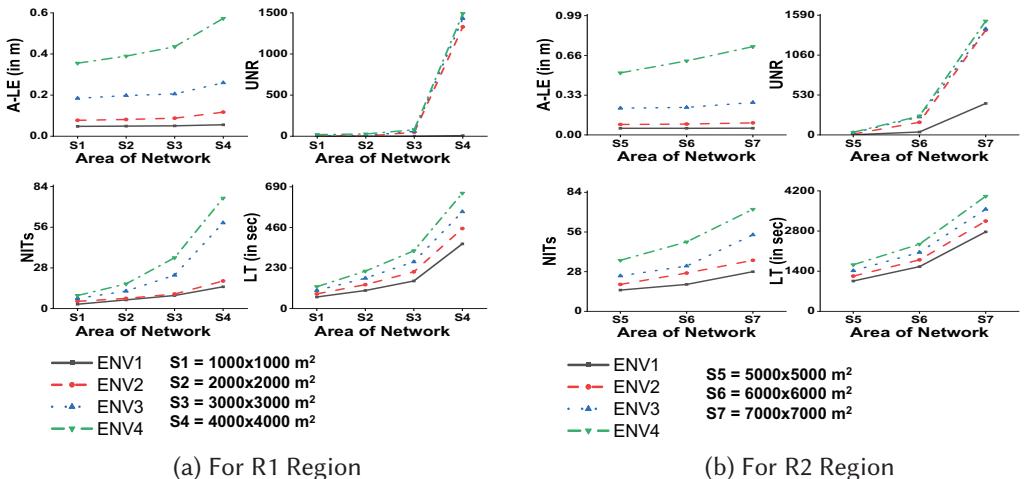


Fig. 12. Impact of varying the environments.

of known beacon nodes (*OBNs*) across the *ROI* enables the convergence of the localisation process from multiple points and thus requires much fewer iterations.

- Localisation Time (*LT*): The localisation time (*LT*) for similar reasons as *NITs* uniformly increases with increase in *ROI*. For random deployment, the localisation time is much smaller as the localisation progresses from multiple points.

4.2.6 Effect of Varying Terrain. In this section, we endeavour to get insight into the efficacy of the proposed localisation approach in various kinds of terrain. We investigate four distinct terrains: free space (*ENV1*), sandy terrain (*ENV2*), grassy terrain (*ENV3*), and terrain with sparse trees (*ENV4*). The results (Figure 12) were obtained through simulations with deployments over both small regions (*R1*) and extensive regions (*R2*). In the *R1* region, 8000 initial unknown nodes (*IUN*) were deployed, with the number of *OBNs* being 1% of *IUNs* i.e., 80 with a communication range (*CR*) of 75 m. For the extensive *R2* regions, the number of *IUNs* was increased to 15, 000, and

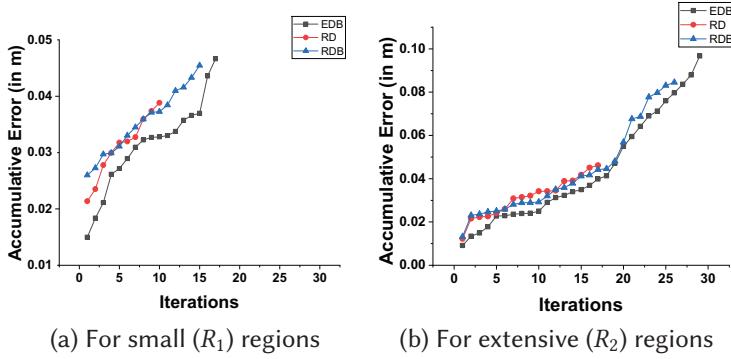


Fig. 13. Effect of Accumulative Error with respect to the number of iterations. For R_1 and R_2 regions, area sizes $4000 \times 4000\text{ m}^2$ and $7000 \times 7000\text{ m}^2$ are considered, respectively, with CR of 100 m. Here EBD, RDB means the beacon nodes are deployed equidistance and randomly on the boundary of the region. Where RD represents the beacon nodes are deployed randomly throughout the region.

the OBNs were again 1% or 150 in number, and the CR was 75 m. The observations were recorded for all the terrain kinds described above:

- Average Localisation Error (A-LE): The localisation error rises as the amount of obstacles (environmental noise) in the environment increases. The presence of environmental noise makes it challenging for the RSSI method to calculate the distance accurately.
- Number of Iterations (NITs): As environmental noise increases, the signal strength decreases (for instance, a signal that can communicate up to 75 m in free space may only reach 25 to 30 m in sparse trees). Consequently, the decrease in CR leads to an automatic increase in the NITs required.
- Localisation Time (LT): The time taken to localise the network increases with the level of environmental noise. When the NITs increase, the LT also increases.
- Unknown Nodes Remaining (UNR): The presence of obstacles significantly increases the number of nodes that remain unknown (UNR) as environmental noise intensifies.

4.2.7 Effect of Accumulative Error. In this section, we explore the variation of accumulative error in localisation as the process progresses through iterations. We conduct simulations for both small (R_1) and extensive (R_2) regions. These comprise regions of size varying between $4000 \times 4000\text{ m}^2$ and $7000 \times 7000\text{ m}^2$; the number of OBNs is taken as 1% of the IUNs with a communication range (CR) of 100 m. The largest values for region size and CR are considered to best understand the impact of accumulative error. The results (Figure 13) show that the accumulative error increases with the progress of iteration but the rise is not very marked.

4.2.8 Node Localisability. In a WSN network, if the number of nodes for a fixed area is large, i.e., the network is dense the network is easily localisable. However, a dense deployment is difficult in larger areas such as those ranging in size between $1000 \times 1000\text{ m}^2$ and $7000 \times 7000\text{ m}^2$. For such large areas, it is necessary to check the node localisability before moving forward with the actual deployment. Doing this makes it possible to estimate the minimum number of nodes required to cover the entire area, how many of the nodes can be localised, and which ones are such nodes (that can be localised). To establish the network's localisability, we find the minimum number of nodes required to localise 95% of unknown nodes with the communication range (CR) of the WSN nodes being 100 m. We run simulations with the area of deployment varying between $1000 \times 1000\text{ m}^2$ and $7000 \times 7000\text{ m}^2$ for the three types of OBN deployment methods (Random (RD), Equidistant on the

Table 8. Network Localisability for Varies Deployment and Area Sizes

Size [†]	EDB			RD			RDB		
	MNR [‡]	OBN [§]	PL [¶]	MNR [‡]	OBN [§]	PL [¶]	MNR [‡]	OBN [§]	PL [¶]
1000 × 1000 m ²	2273	80	95.78%	1210	80	95.88%	1650	80	95.83%
2000 × 2000 m ²	5551	80	95.99%	2951	80	97.44%	4627	80	96.64%
3000 × 3000 m ²	6837	80	95.36%	4553	80	96.42%	5910	80	95.58%
4000 × 4000 m ²	7050	80	95.12%	5732	80	95.76%	6307	80	95.43%
5000 × 5000 m ²	8527	100	95.31%	6845	100	96.06%	7839	100	95.65%
6000 × 6000 m ²	9786	100	95.05%	7918	100	96.03%	8798	100	95.55%
7000 × 7000 m ²	11003	100	95.56%	9423	100	96.53%	10385	100	96.35%

Note: The above table shows the minimum number of nodes required to localise the network. Here, RD, EDB, and RDB represent the deployment of OBN randomly, equidistant on the boundary, and randomly on the boundary, respectively. The Size[†] and MNR[‡] imply the area of the network, and the minimum nodes (IUN) required to localise the network, respectively. OBN[§], PL[¶] stand for the original beacon nodes deployed and percentage of IUN localised, respectively.

Table 9. Simulation Parameters for Comparison

Parameters	Values for small regions (R_1)	Values for extensive regions (R_2)
Size of region	1000 × 1000, 2000 × 2000 to 4000 × 4000	5000 × 5000, 6000 × 6000, and 7000 × 7000
IUN*	8000	15000
OBN [†]	1, 1.5, 2.5, and 5 %	1, 1.5, 2.5, and 5 %
CR [‡]	50, 75, and 100 m	75, and 100 m
EN [§]	E_0, E_1, E_2 , and E_3	E_0, E_1, E_2 , and E_3

Note: The size of the region is measured in square metres, IUN* denotes the number of unknown nodes initially deployed, OBN[†] denotes the number of original beacon nodes (in terms of percent of the number of IUNs), CR[‡] denotes the communication range of a node (in metres), and EN[§] denotes the effect of noise due to environmental factors like free space (E_0), sandy terrain (E_1), long grassy terrain (E_2), and terrains with sparse trees (E_3).

Boundary (EDB), and Randomly on the Boundary (RDB)). Table 8 shows the minimum number of nodes required for localising the network.

The simulation results are computed for the environment E_0 with no interference from outside objects while keeping the OBN constant at 80 and 100 for R_1 and R_2 regions, respectively. The OBNs are deliberately kept constant to facilitate better comprehension of the simulation results. The results clearly show that the proposed method performs well for both regions with a high percentage of node localisability.

4.2.9 Comparison with SOTA. In this section, we compare simulation results based on the localisation accuracy of the proposed method with well-known methods like trilateration [47], Angle of Arrival [23], and state-of-the-art machine learning-based methods [17] like neural networks, random forest, and support vector regression. The received signal strength values are used in the methods based on **trilateration (TL)**, **random forest (RF)**, **neural networks (NN)**, and **support vector regression (SVR)**. For approaches utilising AoA, two angles from different beacon nodes are used to localise an unknown node. An iterative approach is used to compute the results. Table 9 shows the simulation parameters used for computing results for comparison with state-of-the-art methods.

Table 10. Average Localisation Error (A-LE) for R_1 Region when Beacon Nodes Randomly Deployed in Network

EN*	Size[†]	Proposed[‡]	RF[§]	NN[¶]	SVR	TL^{**}	AoA^{††}
E_0	$1000 \times 1000 m^2$	0.0332	0.0359	0.0376	0.0407	0.4524	0.4691
	$2000 \times 2000 m^2$	0.0345	0.0366	0.0389	0.0414	0.5198	0.6026
	$3000 \times 3000 m^2$	0.0383	0.0414	0.0427	0.0459	0.5130	0.5393
E_1	$4000 \times 4000 m^2$	0.0466	0.0498	0.0528	0.0569	0.5325	0.6256
	$1000 \times 1000 m^2$	0.0470	0.0502	0.0532	0.0575	0.5918	0.7409
	$2000 \times 2000 m^2$	0.0485	0.0526	0.0550	0.0590	0.6527	0.7645
E_2	$3000 \times 3000 m^2$	0.0571	0.0615	0.0636	0.0689	0.6975	0.8405
	$4000 \times 4000 m^2$	0.0609	0.0658	0.0689	0.0748	0.7026	0.7641
	$1000 \times 1000 m^2$	0.1599	0.1734	0.1820	0.1989	0.8564	0.8645
E_3	$2000 \times 2000 m^2$	0.1648	0.1789	0.1875	0.2060	0.9157	0.9972
	$3000 \times 3000 m^2$	0.1733	0.1887	0.1972	0.2174	0.9356	1.0382
	$4000 \times 4000 m^2$	0.1768	0.1922	0.2016	0.2213	0.9646	1.0567
E_3	$1000 \times 1000 m^2$	0.3240	0.3530	0.3700	0.4073	1.1580	1.2435
	$2000 \times 2000 m^2$	0.3566	0.3881	0.4075	0.4476	1.1709	1.3130
	$3000 \times 3000 m^2$	0.3920	0.4272	0.4476	0.4919	1.3652	1.4929
		0.4341	0.4724	0.4960	0.5443	1.5772	1.5939

Note: EN*, and Size[†] denotes the environmental noise and size of the network. Proposed[‡], RF[§], NN[¶], SVR^{||}, TL^{**}, AoA^{††} denotes proposed method, random forest, neural network, support vector regression, trilateration, and angle of arrival, respectively.

The experiments are performed by varying the parameters in Table 9. Unlike the proposed method, the other five methods are hardly able to localise initial unknown nodes (*IUN*) in a setting where the original beacon nodes (*OBN*) are only 1%, 1.5%, and 2.5% of the total number with a communication range (*CR*) of 50 and 75 m. In these cases, the number of unknown nodes remaining (*UNR*) is very large because these methods require at least two or more beacon nodes to localise an unknown node. The results are compared based on four parameters: average localisation error, average number of iterations, average unknown nodes remaining, and average localisation time for both R_1 and R_2 regions. The results are evaluated for four different environments: free space (E_0), sandy terrain (E_1), long grassy terrain (E_2), and sparse tree terrain (E_3).

4.2.9.1 Average Localisation Error. The average localisation error is an important parameter for comparing localisation algorithms. In 2-D space, the parameter is computed as the Euclidean distance between the actual coordinates and the predicted coordinate values of the nodes to be localised. Tables 10, and 11 show the average localisation error for R_1 , and R_2 regions, respectively. The proposed method outperforms the five state-of-the-art methods by significant margins. For the R_1 region, 8000 *IUNs* were deployed with 5% of these serving as *OBNs* with an *CR* of 100m. As stated earlier, we are only able to compare with such a liberal number of *OBNs* because the existing methods were unable to make a mark with smaller numbers. Similarly, for R_2 regions, 15000

Table 11. Average Localisation Error (A-LE) for R_2 Region when Beacon Nodes Randomly Deployed in Network

EN*	Size[†]	Proposed[‡]	RF[§]	NN[¶]	SVR	TL^{**}	AoA^{††}
E_0	$5000 \times 5000 m^2$	0.0357	0.0430	0.0787	0.0487	0.5705	0.6606
	$6000 \times 6000 m^2$	0.0396	0.0481	0.0877	0.0536	0.5632	0.5918
	$7000 \times 7000 m^2$	0.0482	0.0570	0.1052	0.0653	0.5846	0.6859
E_1	$5000 \times 5000 m^2$	0.0505	0.0599	0.1105	0.0672	0.7150	0.8366
	$6000 \times 6000 m^2$	0.0595	0.0694	0.1289	0.0778	0.7639	0.9193
	$7000 \times 7000 m^2$	0.0635	0.0740	0.1375	0.0842	0.7696	0.8365
E_2	$5000 \times 5000 m^2$	0.1739	0.1939	0.3678	0.2218	1.0011	1.0897
	$6000 \times 6000 m^2$	0.1828	0.2043	0.3872	0.2339	1.0228	1.1344
	$7000 \times 7000 m^2$	0.1865	0.2083	0.3948	0.2382	1.0546	1.1548
E_3	$5000 \times 5000 m^2$	0.3785	0.4171	0.7955	0.4759	1.2786	1.4331
	$6000 \times 6000 m^2$	0.4160	0.4586	0.8747	0.5226	1.4901	1.6289
	$7000 \times 7000 m^2$	0.4607	0.5068	0.9675	0.5780	1.7209	1.7390

Note: EN*, and Size[†] denotes the environmental noise and size of the network. Proposed[‡], RF[§], NN[¶], SVR^{||}, TL^{**}, AoA^{††} denotes proposed method, random forest, neural network, support vector regression, trilateration, and angle of arrival, respectively.

IUNs are deployed in different regions of various sizes, with the *OBN* constituting 5% of the total nodes and with a *CR* of 100m. The proposed method performs better than existing state-of-the-art methods for regions of all sizes and environments.

4.2.9.2 Average Number of Iterations. This parameter indicates the number of iterations required to complete a network's localisation process; the fewer the iterations, the better the algorithm. This is critical in a real-world implementation because topological information such as the number of known and unknown nodes change after every iteration, and maintaining this information is an additional overhead (in terms of energy) on the network. It also helps in analysing the number of nodes localised per iteration and is a measure of the localisability of the network.

Experiments were conducted with 8000 *IUNs* of which 5% were *OBN* with a *CR* of 100m for all four environments in regions falling in the category of R_1 in terms of size. Similarly, for R_2 regions, 15000 *IUNs* were deployed with 5% of these being *OBN* with a *CR* of 100m for all four environments. The experimental results clearly indicate that the proposed method outperforms existing approaches in the terms of number of iterations. We do not include all the results here but interested readers can access the same at the following [URL](#).

4.2.9.3 Average Unknown Nodes Remaining. The Average Unknown Nodes Remaining is a parameter that indicates the average number of initially unknown nodes (*IUN*) that remain unlocalised at the end of the localisation process. An *IUN* can remain unknown if the nodes are not within the communication range of beacon nodes or are outliers in the network. If the number of unknown nodes in a network is small, the algorithm works better. The network density is inversely proportional to the number of unknown nodes remaining; here, the density implies the number of *IUNs* deployed throughout the network. The quantitative results for the average number of unknown nodes remaining for R_1 regions are computed by deploying 8000 *IUNs* of which 5% are *OBNs*

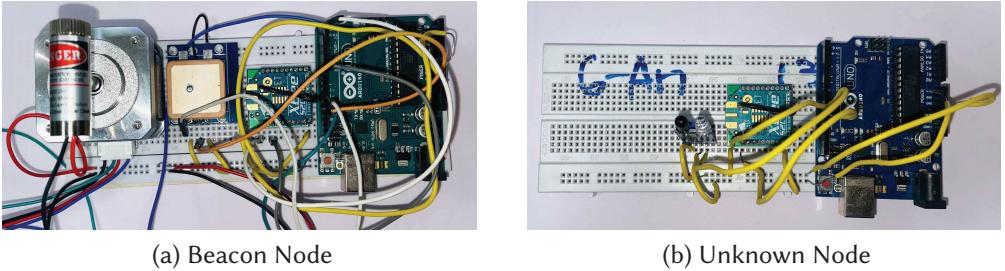


Fig. 14. The configuration of beacon node and unknown node.

with a *CR* of $100m$. Similarly, for R_2 regions the 15000 *IUNs* are deployed with 5% of *IUNs* being *OBNs* with a *CR* of $100m$. The results in both regions are computed on four different simulation environments, affirming that the proposed method works better than the existing state-of-the-art methods. Detailed quantitative results on this factor are available at the following [URL](#).

4.2.9.4 Average Localisation Time. The average localisation time represents the time required to localise all possible nodes in a region of a given size and a given number of nodes. Localisation time is an important parameter that helps to determine the time approximately required to execute the algorithm and localise nodes in challenging real-world applications like forest fire detection, battlefield surveillance, and so on. The average localisation time is calculated by deploying 8000 *IUNs* in regions of sizes varying between (1000×1000 and $4000 \times 4000 m^2$) with 5% of the *IUN* being *OBN* having a communication range of $100m$. The other region experimented with has an area between 5000×5000 and $7000 \times 7000 m^2$, with 15000 *IUNs* and 5% of the *IUN* being *OBN* with a *CR* of $100m$. The proposed work is clearly seen to outperform existing state-of-the-art methods in all settings. The detailed results are available at the following [URL](#).

4.3 Real-world Prototypical Implementation

To assess the feasibility of the proposed localisation approach in a real-world outdoor environment localisation, we made a simple prototypical implementation of a wireless sensor network in the sports ground of our Institute. The deployment comprised two types of nodes: the nodes with known locations to start with, called original beacon nodes (*OBN*), and the initially unknown nodes (*IUN*). The *OBN* was configured using an Uno microcontroller [6], an XBee S2C (for measuring the received signal strength) [34], a Neo 6M GPS module (for determining the location) [33], a unipolar stepper motor [10], a laser light (mounted on the stepper motor for measuring the angle of arrival of the signal) [13], and a 9-volt zinc-carbon battery. The *IUN* included an Uno microcontroller, an XBee S2C, and a 9-volt zinc-carbon battery. The Xbees were configured using the ZIGBEE TH Reg firmware and API 2 mode. The setup of the *OBN* and the *IUN* is shown in Figure 14.

Two factors are computed for establishing the location of unknown nodes as per the proposed approach: the distance between the unknown *IUN* and known *OBN* node and the angle of arrival of a signal from the *OBN* to the *IUN*. Computing the distance is relatively straightforward through direct mapping between the received signal strength and the distance. Computing the angle of arrival of the signal is a little more involved. The norm for computing the angle of arrival has been the use of antenna arrays. This approach, however, is limited to only a few unknown nodes, owing largely to the use of signals vulnerable to interference and noise. To overcome these issues, we have developed a simple yet novel technique for measuring the angle of arrival. This technique involves mounting a laser light on the unipolar stepper motor of the *OBN* nodes. The stepper

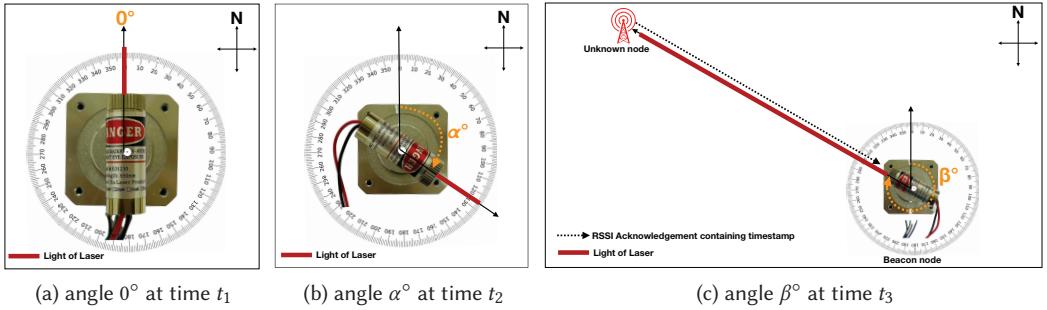


Fig. 15. The above figure facilitates the calculation of the AoA between the anchor and unknown node, which are synchronised using a global time clock. (a) At a specific time t_1 , the laser is directed towards the north, indicating a zero-degree angle. (b) The motor rotates by an angle α , within a time interval t_2 . (c) The light from the laser strikes the unknown node at time t_3 . The Unknown node then transmits this time information to the Beacon node using RSSI communication. This information is utilised to determine the angle between the nodes. This is the same setup we used for our AoA calculations.

motor by default rotates 360° in 200 steps, with each step, therefore, covering 1.8° ($360/200$). For our deployment, however, we adjusted the step size of the stepper motor to 360, so that each rotation step rotates the motor by 1° . The stepper motor is aligned such that zero degrees coincide with the North direction of the compass and rotates clockwise until the laser light shines on the unknown node. At this point, the stepper motor is stopped, and the angle is measured by counting the number of steps rotated. The technique is pictorially demonstrated in Figure 15, illustrating the computing of the angle of arrival. We are working towards developing an automated mechanism for stopping the stepper motor when the laser beam falls on the unknown node and measuring the angle.

Our deployment comprised positioning the OBNs equidistant from each other at the periphery of a 60×38 square meter sports ground. The unknown IUN nodes were subsequently, randomly deployed across the ground. Figure 16 provides a “bird’s eye view” of the setup used on the ground using a drone camera.

Figure 17 shows the same set-up with the ROI divided into grids. The nodes circled in red represent the unknown IUNs, while the known OBNs are circled in yellow. The points marked with crosses (Pred points) indicate the locations computed using our proposed method.

The deployment comprised a total of 11 nodes, 7 IUNs, 3 OBNs, and 1 base station. The actual and computed positions of IUNs are shown in Table 12. The average localisation error over the prototypical deployment was 2.509 m , with the ROI being $60 \times 38\text{ m}^2$ (i.e., 2280 m^2).

In our simulation, the localisation error was around 1 m , whereas in real-world deployments the error went up to 2.5 m . The stark difference in these values between simulations and the real-world is due largely to the inaccuracies of low-cost range-based devices (whose precision is as low as 2 m) and the environmental conditions (minor grassy terrains). As the localisation process progresses through iterations, with the newly localised nodes serving as beacon nodes for subsequent iterations, the localisation error accumulates. It is difficult, in a real-world setting to scale our deployment to measure the localisation error after a large number of iterations, but based on our limited deployment we propose the expression in Equation (13) to calculate the accumulative error.

$$A_{Err} = D_{Err}^{Iter} + \prod_i^n L_{Err}(i). \quad (13)$$

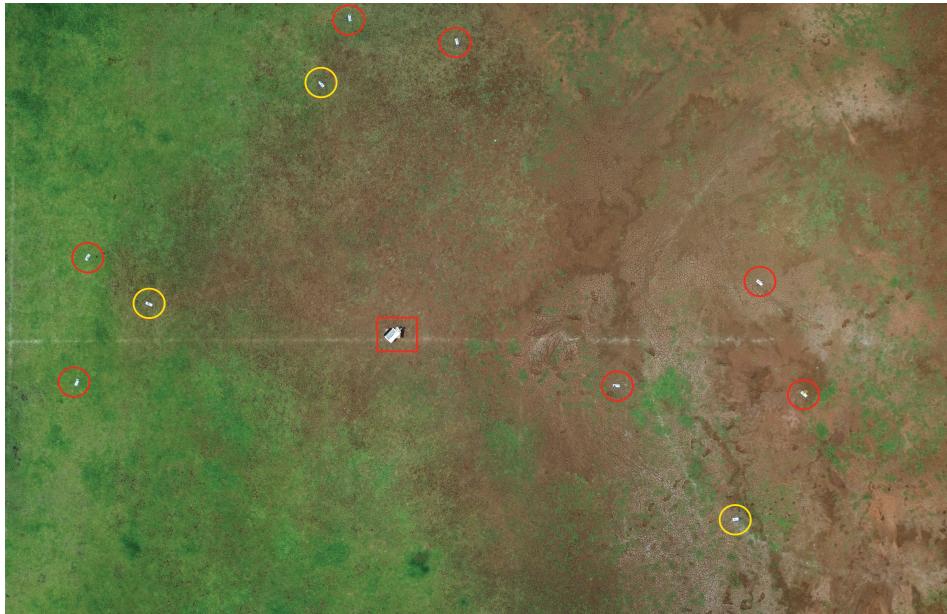


Fig. 16. Nodes are deployed in real-life scenarios, with the Unknown nodes represented by red circles and the Beacon nodes represented by yellow circles.

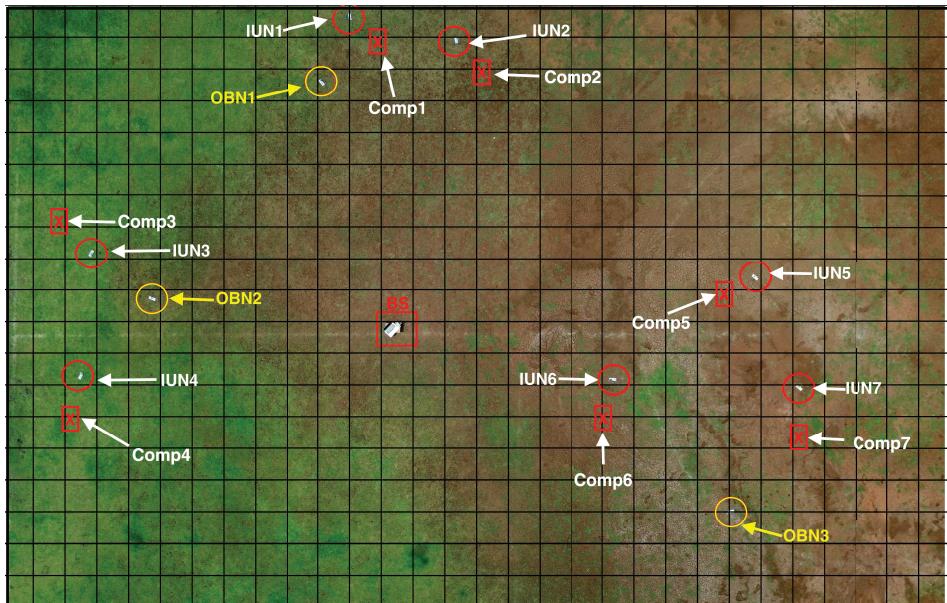


Fig. 17. This depiction illustrates the top view of the deployment represented in a grid format, with each grid cell measuring $2 \times 2 \text{ m}^2$. The grid's topmost left and bottom right corners correspond to the (0,0) and (60,38) coordinates, respectively. Grid lines have been added to the image to aid in understanding the measurements. In this representation, "IUN" denotes the initial unknown node, "OBN" stands for the original beacon node and "Comp" represents the computed location of IUN.

Table 12. Localisation Error Calculation in Above Scenario

	Actual		Computed		Error
	X	Y	X	Y	
Node 1	21.948	0.758	23.375	2.141	1.988
Node 2	28.550	2.230	29.933	4.238	2.438
Node 3	5.487	15.524	3.435	13.472	2.902
Node 4	4.684	23.375	4.283	25.918	2.574
Node 5	47.242	17.130	45.413	18.245	2.142
Node 6	38.587	23.509	37.784	26.007	2.624
Node 7	50.097	24.089	50.186	26.989	2.901
		Average Error		2.509m	

Note: The **Actual** indicates the real x and y relative coordinate values on a scale from (0,0) to (60,38). The **Computed** displays the calculated values of these points using the proposed algorithm. The **Error** represents the Euclidean distance between the actual and predicted values.

Where D_{Err} is the error due to the device; the exponent $Iter$ denotes the number of iterations or, in other words, the number of hops from the initial node to reach the unknown node; and $L_{Err}(i)$ is the localisation error for iteration i . As the experiments are conducted over relatively small areas with a much smaller number of beacons and unknown nodes as compared to the simulations, only a limited understanding of the propagation of accumulative error in real-world environments is possible. In addition to inaccuracies due to the devices, environmental factors also affect the accuracy of localisation in real-world deployments. Generalised results obtained using the Equation (13) could be beneficial in understanding the effect of accumulative error in harsh terrains like flood and fire detection before actual deployment, where the deployment cost of devices is high.

It is important to note that the intention of the real-world deployment is not to assess the effectiveness of the system in terms of localisation accuracy. The results are quite compromised owing to the use of inexpensive and basic equipment that was easily accessible to us. Rather the intent of the deployment is to demonstrate the feasibility of the system in a real-world environment.

5 Conclusion

This article proposed a novel technique for identifying the location of nodes in a WSN without the need for GPS and using the coordinates of only one beacon node whose location is known in advance. The technique proposed comprises the use of a hybrid combination of the AoA and RSSI to identify the location of unknown nodes. This hybrid localisation approach permeates through the entire region of interest, irrespective of its size, through an iterative approach. Here, the initial beacon nodes localise unknown nodes, which in turn serve as beacon nodes in the subsequent iteration. The uniqueness of our approach is that it is the first of its kind that only requires one beacon node to localise an unknown node. This is made possible through the use of a combination of RSSI and AoA. Furthermore, our approach overcomes the unrealistic assumption that all unknown nodes are within the communication range of all beacon nodes to start with. This is made possible through adopting an iterative approach to cover the entire region of interest. The effectiveness of the proposed localisation approach was thoroughly validated through extensive simulation studies. This was necessitated owing to the lack of access to a sufficiently large WSN deployment. Subsequently, the real-world efficacy of the approach was validated through a prototypical implementation.

Acknowledgments

We want to thank the Defries-Bajpai Foundation for funding the research. The authors would like to acknowledge the **Ubiquitous Computing (UbiComp)** Lab of IIT Indore for providing the research facilities to conduct the presented work. We thank Ms. Sneha Shukla and Ms. Drishti Sharma for providing valuable insights into some figures.

References

- [1] Abdulaziz Alsayyari and Abdallah Aldosary. 2018. Path loss results for wireless sensor network deployment in a long grass environment. In *Proceedings of the 2018 IEEE Conference on Wireless Sensors*. IEEE, Langkawi, Kedah, Malaysia, 50–55.
- [2] Abdulaziz Alsayyari, Ivica Kostanic, Carlos Otero, Mohammed Almeer, and Kusay Rukieh. 2014. An empirical path loss model for wireless sensor network deployment in a sand terrain environment. In *Proceedings of the 2014 IEEE World Forum on Internet of Things*. IEEE, Seoul, Korea (South), 218–223.
- [3] Th Arampatzis, John Lygeros, and Stamatis Manesis. 2005. A survey of applications of wireless sensors and wireless sensor networks. In *Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation Intelligent Control, 2005*. IEEE, Limassol, Cyprus, 719–724.
- [4] Abdalkarim Awad, Thorsten Frunzke, and Falko Dressler. 2007. Adaptive distance estimation and localization in WSN using RSSI measures. In *Proceedings of the 10th Euromicro Conference on Digital System Design Architectures, Methods and Tools*. IEEE, Lübeck, Germany, 471–478.
- [5] Salah Azzouzi, Markus Cremer, Uwe Dettmar, Thomas Knie, and Rainer Kronberger. 2011. Improved AoA based localization of UHF RFID tags using spatial diversity. In *Proceedings of the 2011 IEEE International Conference on RFID-Technologies and Applications*. IEEE, Sitges, Spain, 174–180.
- [6] Yusuf Abdullahi Badamasi. 2014. The working principle of an arduino. In *Proceedings of the 2014 11th International Conference on Electronics, Computer and Computation*. IEEE, Abuja, Nigeria, 1–4.
- [7] Constantine A Balanis. 2016. *Antenna Theory: Analysis and Design*. John wiley & sons.
- [8] Suvankar Barai, Debajyoti Biswas, and Buddhadeb Sau. 2020. Sensors positioning for reliable RSSI-based outdoor localization using CFT. In *Proceedings of the 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security*. IEEE, Gunupur Odisha, India, 1–5.
- [9] Paolo Barsocchi, Stefano Lenzi, Stefano Chessa, and Gaetano Giunta. 2009. A novel approach to indoor RSSI localization by automatic calibration of the wireless propagation model. In *Proceedings of the VTC Spring 2009-IEEE 69th Vehicular Technology Conference*. IEEE, 1–5.
- [10] Zhang Benhua, Li Chenghua, Sun Shiming, and Gan Lu. 2011. Design on a unipolar and unidirectional stepper motor circuit. In *Proceedings of the 2011 International Conference on Electronic and Mechanical Engineering and Information Technology*. IEEE, Harbin, China, 1795–1797.
- [11] Soumya J. Bhat and Santhosh K. Venkata. 2020. An optimization based localization with area minimization for heterogeneous wireless sensor networks in anisotropic fields. *Computer Networks* 179, 1 (2020), 107371.
- [12] Ghulam Bhatti, Yasir Javed, Munir Naveed, and Sayed Asif. 2020. Out-door localization in large-scale wireless sensor networks by using virtual nodes. In *Proceedings of the 2020 Advances in Science and Engineering Technology International Conferences*. IEEE, Dubai, United Arab Emirates, 1–7.
- [13] Benjamin Chu. 2007. *Laser Light Scattering: Basic Principles and Practice*. Courier Corporation.
- [14] Bram Dil, Stefan Dulman, and Paul Havinga. 2006. Range-based localization in mobile sensor networks. In *Proceedings of the European Workshop on Wireless Sensor Networks*. Springer, Zurich, Switzerland, 164–179.
- [15] Bugra Gedik and Ling Liu. 2007. Protecting location privacy with personalized k-anonymity: Architecture and algorithms. *IEEE Transactions on Mobile Computing* 7, 1 (2007), 1–18.
- [16] David Kiyoshi Goldenberg, Arvind Krishnamurthy, Wesley C. Maness, Yang Richard Yang, Anthony Young, A Stephen Morse, and Andreas Savvides. 2005. Network localization in partially localizable networks. In *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies*. IEEE, Miami, FL, USA, 313–326.
- [17] Rupendra Pratap Singh Hada, Uttkarsh Aggarwal, and Abhishek Srivastava. 2023. A study and analysis of a new hybrid approach for localization in wireless sensor networks. *Journal of Web Engineering* 22, 2 (2023), 279–302.
- [18] Rupendra Pratap Singh Hada and Abhishek Srivastava. 2024. Dynamic cluster head selection in WSN. *ACM Transactions on Embedded Computing Systems* 23, 4 (2024), 1–27.
- [19] Shinsuke Hara and Daisuke Anzai. 2008. Experimental performance comparison of RSSI-and TDOA-based location estimation methods. In *Proceedings of the VTC Spring 2008-IEEE Vehicular Technology Conference*. IEEE, 2651–2655.

- [20] Tian He, Chengdu Huang, Brian M Blum, John A Stankovic, and Tarek Abdelzaher. 2003. Range-free localization schemes for large scale sensor networks. In *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking*. ACM, San Diego CA USA, 81–95.
- [21] Tian He, Chengdu Huang, Brian M Blum, John A Stankovic, and Tarek F Abdelzaher. 2005. Range-free localization and its impact on large scale sensor networks. *ACM Transactions on Embedded Computing Systems* 4, 4 (2005), 877–906.
- [22] Lingxuan Hu and David Evans. 2004. Localization for mobile sensor networks. In *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking*. ACM New York, NY, USA, Philadelphia, Pennsylvania, USA, 45–57.
- [23] Jehn-Ruey Jiang, Chih-Ming Lin, Feng-Yi Lin, and Shing-Tsaan Huang. 2013. ALRD: AoA localization with RSSI differences of directional antennas for wireless sensor networks. *International Journal of Distributed Sensor Networks* 9, 3 (2013), 529489.
- [24] Paweł Kulakowski, Javier Vales-Alonso, Esteban Egea-López, Wiesław Ludwin, and Joan García-Haro. 2010. Angle-of-arrival localization based on antenna arrays for wireless sensor networks. *Computers and Electrical Engineering* 36, 6 (2010), 1181–1186.
- [25] Anh Tuyen Le, Le Chung Tran, Xiaojing Huang, Christian Ritz, Eryk Dutkiewicz, Son Lam Phung, Abdesselam Bouzerdoum, and Daniel Franklin. 2020. Unbalanced hybrid AOA/RSSI localization for simplified wireless sensor networks. *Sensors* 20, 14 (2020), 3838.
- [26] Deyu Lin, Quan Wang, Weidong Min, Jianfeng Xu, and Zhiqiang Zhang. 2020. A survey on energy-efficient strategies in static wireless sensor networks. *ACM Transactions on Sensor Networks* 17, 1 (2020), 1–48.
- [27] Xuan Liu, Jiangjin Yin, Shigeng Zhang, Bo Ding, Song Guo, and Kun Wang. 2018. Range-based localization for sparse 3-D sensor networks. *IEEE Internet of Things Journal* 6, 1 (2018), 753–764.
- [28] Yunhao Liu and Zheng Yang. 2024. Localizability. In *Proceedings of the Location, Localization, and Localizability: Location-awareness Technology for Wireless Networks*. Springer, 111–130.
- [29] Marcin Luckner, Maciej Grzenda, Robert Kunicki, and Jarosław Legierski. 2020. IoT architecture for urban data-centric services and applications. *ACM Transactions on Internet Technology* 20, 3 (2020), 1–30.
- [30] Jari Luomala and Ismo Hakala. 2022. Adaptive range-based localization algorithm based on trilateration and reference node selection for outdoor wireless sensor networks. *Computer Networks* 210 (2022), 108865.
- [31] Shaghayegh Monfared, Trung-Hien Nguyen, Luca Petrillo, Philippe De Doncker, and François Horlin. 2018. Experimental demonstration of BLE transmitter positioning based on AOA estimation. In *Proceedings of the 2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications*. IEEE, Bologna, Italy, 856–859.
- [32] M Shahkhir Mozamir, Rohani Binti Abu Bakar, Wan Isn'i Sofiah Wan Din, and Zalili Musa. 2020. GbLN-PSO algorithm for indoor localization in wireless sensor network. In *Proceedings of the IOP Conference Series: Materials Science and Engineering*. IOP Publishing, Pahang, Malaysia, 012033.
- [33] Maria Paola Panetta, M. Abbrescia, C. Avanzini, R. Baldini Ferroli, L. Baldini, G. Batignani, M. Battaglieri, S. Boi, E. Bossini, F. Carnesecchi, et al. 2019. The new trigger/GPS module for the EEE project. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 936, 1 (2019), 376–377.
- [34] Rajeev Piyare and Seong-ro Lee. 2013. Performance analysis of XBee ZB module based wireless sensor networks. *International Journal of Scientific and Engineering Research* 4, 4 (2013), 1615–1621.
- [35] Masoomeh Rudafshani and Suprakash Datta. 2007. Localization in wireless sensor networks. In *Proceedings of the 6th International Conference on Information Processing in Sensor Networks*. ACM, Cambridge Massachusetts, USA, 51–60.
- [36] Nasir Saeed, Haewoon Nam, Tareq Y. Al-Naffouri, and Mohamed-Slim Alouini. 2019. A state-of-the-art survey on multidimensional scaling-based localization techniques. *IEEE Communications Surveys and Tutorials* 21, 4 (2019), 3565–3583.
- [37] Andreas Savvides, Chih-Chieh Han, and Mani B Srivastava. 2001. Dynamic fine-grained localization in ad-hoc networks of sensors. In *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*. ACM New York, NY, USA, Rome, Italy, 166–179.
- [38] Sukhwinder Sharma, Rakesh Kumar Bansal, and Savina Bansal. 2013. Issues and challenges in wireless sensor networks. In *Proceedings of the 2013 International Conference on Machine Intelligence and Research Advancement*. IEEE, KATRA, JAMMU, INDIA, 58–62.
- [39] Shilpi and Arvind Kumar. 2023. A localization algorithm using reliable anchor pair selection and Jaya algorithm for wireless sensor networks. *Telecommunication Systems* 82, 2 (2023), 277–289.
- [40] Munesh Singh, Sourav Kumar Bhoi, and Sanjaya Kumar Panda. 2021. Geometric least square curve fitting method for localization of wireless sensor network. *Ad Hoc Networks* 116, 1 (2021), 102456.
- [41] Slavisa Tomic, Marko Beko, and Milan Tuba. 2019. A linear estimator for network localization using integrated RSS and AOA measurements. *IEEE Signal Processing Letters* 26, 3 (2019), 405–409.

- [42] Guido Van Rossum, et al. 2007. Python programming language.. In *USENIX Annual Technical Conference*, Vol. 41. USENIX Association, Santa Clara, CA, 1–36.
- [43] Shengchu Wang, Xianbo Jiang, and Henk Wyneersch. 2022. Cooperative localization in wireless sensor networks with AOA measurements. *IEEE Transactions on Wireless Communications* 21, 8 (2022), 6760–6773.
- [44] Enyang Xu, Zhi Ding, and Soura Dasgupta. 2011. Source localization in wireless sensor networks from signal time-of-arrival measurements. *IEEE Transactions on Signal Processing* 59, 6 (2011), 2887–2897.
- [45] Zheng Yang, Lirong Jian, Chenshu Wu, and Yunhao Liu. 2013. Beyond triangle inequality: Sifting noisy and outlier distance measurements for localization. *ACM Transactions on Sensor Networks* 9, 2 (2013), 1–20.
- [46] Zheng Yang and Yunhao Liu. 2011. Understanding node localizability of wireless ad hoc and sensor networks. *IEEE Transactions on Mobile Computing* 11, 8 (2011), 1249–1260.
- [47] Zheng Yang, Yunhao Liu, and X.-Y. Li. 2009. Beyond trilateration: On the localizability of wireless ad-hoc networks. In *Proceedings of the IEEE INFOCOM 2009*. IEEE, Rio de Janeiro, Brazil, 2392–2400.
- [48] Slim Zaidi, Ahmad El Assaf, Sofiene Affes, and Nahi Kandil. 2016. Accurate range-free localization in multi-hop wireless sensor networks. *IEEE Transactions on Communications* 64, 9 (2016), 3886–3900.
- [49] Deyue Zou, Shuyi Chen, Shuai Han, Weixiao Meng, Di An, Jinqiang Li, and Wanlong Zhao. 2020. Design of a practical WSN based fingerprint localization system. *Mobile Networks and Applications* 25, 1 (2020), 806–818.

Received 19 March 2024; revised 7 September 2024; accepted 21 January 2025