Adequate Sparse Secure and Minkowski Distance Based Location Privacy Approach in Wireless Sensor Network

Uma Meena¹*  Anand Sharma²

¹ College of Engineering and Technology, Mody University, Rajasthan, India
* Corresponding author’s Email: umameenaphd@gmail.com

Abstract: Wireless sensor networks (WSNs) are susceptible to various issues such as privacy and security. To overcome these issues, our work first validates the location privacy issues and security protection. So that introduce a novel method known as Adequate Sparse Secure and Minkowski distance based Location Privacy (ASSMLP), in this first phase concentrates on location privacy which is improved by presenting a technique known as Fake Source and Fake Sink (FSFS). In this the location of fake node is separated by Modified Pillar k means based Minkowski distance. Second phase deliberate on security and here confidentiality of message can be ensured through sparse matrix. Thus the result shows throughput, energy consumption, energy efficiency, delay and packet delivery ratio and it achieves 2%, 7%, 36%, 8% and 8% better performance than existing work. Clustering accuracy is showed by means of minimum silhouette coefficient. Since real node cannot identify with an intruder.

Keywords: Minkowski distance, Extended Euclidean algorithm, Fake source and fake sink, Sparse matrix, Adequate secure and location privacy preserving.

1. Introduction

Wireless sensor systems (WSNs) have got such an extraordinary measure of thought among a group of their different applications, for instance, managing imperatives plants, logistics and stock, combat areas, and remedial watching [1]. A sensor is a simple, battery fuelled apparatus with required computational power and memory that is outfitted with distinguishing radio transmission units. WSNs are generally connected with the outside world through a computationally executable centre called the sink that is in like manner responsible for data gathering and data join [2, 3]. Remote sensors need to talk with each other through remote transmission. Remote communications are definitely not hard to be taken after or listened silently by intruder [4].

To suppress this kind of issue, area security is thus vital, especially in undermining circumstances. Failure to secure such information can absolutely destabilize the normal explanations behind sensor framework applications. Area security measures should be created to keep the intruder from deciding physical areas of source sensors and sinks [5]. Initially, an intruder can intrude broadcast medium for coordinating data. Second, sensors as a general rule have required taking care of pace and imperativeness supplies. It is extremely costly to apply standard unknown transmission frameworks for disguising the transmission between sensor center points and sinks. We need to find range security those records to the advantage hindrances of sensor center nodes [6].

However, it should be seen that, the thoughts of security and credibility are orthogonal and large in adjusted transmission (with symmetric-key encryption for assurance protection and message approval codes for authenticity [7, 8]). Regardless, these transitional points should not take any information about the last aggregate or individual data in a faultless arrangement. It gives these two goals are in strife. Appropriately, consider study on the security definitions and exhaustive examinations [9, 10].

In the meantime, the security must be considered in the data aggregation operation, since remote
sensor frameworks are considered on unverifiable destabilization circumstances [11 - 14]. There is a plan of security approaches that could improve the robustness and dedication of WSNs and achieve secure data collection. For instance, cryptography-based advances, receipt instruments, key exchange and affirmation traditions, trust-based frameworks and assurance defending innovations [15, 16].

In case of alternative, cryptographic framework gives some customer affirmation measures for security enhancement. Various experts convey how to arrange the cryptographic approval for WSNs. As demonstrated by the unmistakable cryptographic primitives, we can use the most mastermind arrangements into individuals when all is said in key-based arrangements, and the symmetric key-based arrangements [17]. An extensive bit of the proposed plots subsequently contain security defect which affects reality. In addition, security issues are inefficient with respect to the interest of cryptographic counts, secure storage space, and correspondence data [18, 19]. In this way illustrating of secure and assurance protecting WSN with territory acceptance technique is a basic undertaking in framework arrangement.

In order to overcome the complexities of past paper, we presenting a technique for secure source and sink area security assurance strategy. So we present one technique to be specific Adequate Sparse Secure and Minkowski distance based Location Privacy (ASSMLP) approach. The significant contribution of the paper comprised of two stages. In this first stage, source and sink location privacy is improved by using Fake source and Fake sink (FSFS). In this process, Fake source and sink is localized based on Modified Pillar k means clustering algorithm through this process cluster head is assigned as the real sensor node and other nodes assigned as the fake nodes. So the packet transmission is takes place from cluster head to cluster head. Since the message cannot be retrieved by more than one intruder at a time. The significance second stage is security enhancement by sparse matrix encryption and decryption algorithm and in which one time key is produced by Extended Euclidean algorithm (EEA). The private key could be a one-time key since it is utilized to encode plain content. Conjointly this can be partner key that is extreme for intruder to look out the key. This stage, encoded record is send to the approved client or keeps it locally then the approved client recovers the document at the area of real sink.

The remaining section of the manuscript as follows: Section 2 displays the works identified with the remote sensor system's security and location privacy. The Section 3 reveals the issues depiction in the wireless sensor system. Section 4 clarifies the proposed approach. Area 5 gives test results and discussions. Finally conclusion of this work is explained in the section 6.

2. Related works

Some of the recent research work related to the security and location privacy preserving in wireless sensor network is listed below:

Zhou, J et al. [20] introduces the secure location privacy preserving methodology in the cloud-based IoT. In this technique location privacy of particular data can be achieved without using homomorphic encryption. Finally secure data forwarding and efficient location privacy can be achieved with this technique. Drawback is cipher text access control is not present. But in our work cipher text control is possible with one time key.

Luo, E et al. [21] projected the technique for privacy preserving in PMSNs. In which matching outcomes are more relevant to the mutual authentication. The next think is instead of focusing single hops multiple hops were considered for further processing. Own matching was preferred to customize their protocol. Computation time, was not efficient. Since, delay is high compared to our work.

Li, F et al. [22] presented a technique known as EECDRA (energy-efficient cluster-based dynamic routes adjustment approach). This technique mainly concentrated on a cost minimization on a route construction. Thus data delivery within the routes was enhanced. Compared to our work energy consumption is high.

Chaudhry, S.A et al. [23] framed the technique known as SIP protocol and which was based on Elliptic Curve Cryptography (ECC). This technique evades all the attacks and it gives the privacy authentication during packet transmission. Thus SIP provides authentication related with HTTP digest. It is not adequate to deal with all types of attacks. Since, it is not secure.

Kwon et al. [24] proposed a novel D2D confirmation convention with a safe introductory key foundation utilizing cipher text-policy attribute-based encryption (CP-ABE). By utilizing CP-ABE, the anticipated plan permits the imparting gatherings to commonly validate and infer the connection key in an expressive and secure way. We demonstrate that the arranged plan is secure against MITM and replay assault in D2D versatile multi-bounce systems. Computation cost is relatively high and compared to our work it is not efficient.
3. Problem formulations

Let \( N \) be arbitrarily organized sensor nodes in a checked range. In this section, deliberate the location privacy and security issues. The parametric performance such as packet delivery ratio, energy consumption, energy efficiency, throughput and average end to end delay are evaluated to enhance the efficiency of the proposed method. These measurements are utilized to assess the vitality utilization and security upgrade plans. The fundamental target of this paper is to accomplish secure and location privacy and less vitality expended transmission over the whole system.

In periodic assortment all the detectors send the packets to the sink node level once only sensor has object in its vary. During this methodology the overhead is extremely high. The fundamental target of this paper is to accomplish secure and location privacy and less vitality expended transmission over the whole system. For that, introduces the network with the network nodes and the cluster heads. Consider a WSN with \( N \) number of nodes in which sensor nodes are consistently and arbitrarily distributed. WSN is indicated by an attached graph \( G(V, E) \) where \( V \) is the total number of sensor nodes and \( E \) is the total number of edges that are communicating the vertex \( V \). Sensor nodes are compactly placed. Sensor nodes are restricted in memory, computational volume and less energy and power consumption. Because of a large number of sensors and large amount of overhead, these types of nodes have global identification (ID).

4. Proposed methodology

Adequate Sparse Secure and Minkowski distance based Location Privacy (ASSMLP) method comprise of two phase: (i) Location privacy improvement phase, (ii) Security improvement phase. Figure 1 shows the process flow for the suggested system shows fake node localization is carried out by Modified pillar k means based Minkowski distance. Then the onetime key is produced by Extended Euclidean algorithm (EEA). Then the encryption and decryption is done with the help of sparse matrix cryptography.

**Step 1: Fake Source and fake sink technique (FSFS):**

Fake Source and Fake Sink (FSFS) techniques mostly used for refining location privacy by introducing fake source and sink in addition with real source and sink.

\[
f(x) = d^{MKD}(i, j) = \sqrt[\lambda]{\sum_{k=0}^{n-1} |y^i_k - y^j_k|^\lambda}
\]

In Eq. (1), \( d^{MKD} \) is the Minkowski distance between the real object \( i \) and fake object \( j \), \( n \) is the total number of nodes in the network and \( \lambda \) is the order of the Minkowski metric. Below part of this manuscript represents the two phases for improving overall performance. The first phase is the location privacy improvement phase and the next phase is the security improvement phase.

4.1 Location privacy improvement phase:

Location privacy is enhanced by using Fake Source and Fake Sink technique. In this key is created with Extended Euclidean algorithm (EEA). Then the encryption and decryption is done with the help of sparse matrix cryptography.
If assured sink is not operative the sensor network will failure. So it is crucial to use these techniques along. The schematic representation of the location privacy improvement stage is given in fig. 2 in which real and fake object are separated by the Modified pillar k means based Minkowski distance. Once actual object is about to the sink it transmits this to the sink node by mistreatment packet. Attacker has its peculiar sensor network to examine the sensor network worked by commander and will see the alterations in communication pattern and will establish the location of object. But we have a tendency to area mistreatment of fake objects here; once sensor has information connecting to the vital object, fake object together sends the packet to the destination. This confuses the challenger relating to the presence of real object. In periodic assortment all the detectors send the packets to the sink node level once only sensor has object in its vary. During this methodology the overhead is extremely high. First stage of this method is sensor nodes initialization. Then the real and fake objects and sinks are prepared by Modified pillar k means algorithm in which the distance between the real and fake nodes are divided by Minkowski distance.

Modified pillar k means algorithm:

Initial starting points are generated randomly using K-means algorithm it is difficult to reach global optimum which will lead to incorrect clustering results [25]. These obstacles in K-means have been addressed by specifying a procedure to initialize the cluster centers before proceeding with the standard k-means optimization iterations. Modified pillar k Means algorithm is applied to calculate smallest distance between centroid and object for forming cluster. During this cluster, cluster head is selected as the real object and adjacent nodes are selected as the fake nodes. So the standardisation of the node parameter can be achieved by Eq. (2).

\[ X_{new} = \frac{x - \mu}{\sigma} \]  

(2)

Where \( \mu \) is mean value, \( \sigma \) is standard deviation. Mean and standard deviation calculated using following equation:

\[ \mu = \frac{1}{n} \sum_{i=1}^{n} x_i \]  

(3)

\[ \sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}} \]  

(4)

Where \( x_i \) represents a variable at the index of a node, and \( N \) is total number of nodes in the network. Following formula for using normalization by decimal scaling and this is for normalising nodal points.

\[ X_{new} = \left( \frac{x}{10^7} \right) \]  

(5)

Where, \( j \) is the smallest integer then the initial centroid is calculated by the following formula,

\[ v_i = \left( \frac{1}{c_i} \right) \sum_{j=1}^{c_i} x_i \]  

(6)

Where \( x_i \) represents a variable at the index of a node, \( c_i \) is number of attributes to ensure fair distribution of cluster. Then the pillars are formed based on Eq. (7).

\[ p = \Sigma_{i=1}^{k} (\|c_i - r_i\|) \]  

(7)

In Eq. (8), \( r_i \) represents the ith real centroid of the cluster. Following Minkowski distance formula to calculate With-in cluster sum of distance between centroid and object for forming clusters.

Minkowski distance:

Fake source and sink localization is depends on the Minkowski distance. In case of Manhattan distance \( \lambda=1 \) and hence it can be only used for \( L_1 \)-norm. In case of Euclidean algorithm \( \lambda=2 \) and it is appropriate for \( L_2 \)-norm. Nevertheless these two techniques have specific limitation in distance matrix. In order to exhaust these tasks tend to introduce one technique mentioned to as Minkowski distance that is generalised metric that has different cases of the universal type. The Minkowski distance is usually used once variables are measured on magnitude relation scales with associate zero value. From the objective function, Eq. (8) is given as

\[ d^{MKD}(i,j) = \lambda \sqrt{\sum_{k=0}^{n-1} |y_i - y_j|^2} \]  

(8)

In Eq. (8), \( d^{MKD} \) is the Minkowski distance between the real object and fake object (centroid), \( n \) is the total number of nodes in the network and \( \lambda \) is the order of the Minkowski metric. Same method is used for calculating the distance between real and fake sink. Which means selected distance measure between a data point and the cluster center is intended which is an indicator of the distance of the
data points from their separate cluster centres. Even if it is defined for any \( \lambda > 0 \), it is rarely used for values further than 1, 2, and \( \infty \) [26]. Minkowski metric is transformed for \( \lambda = \infty \) and it becomes:

\[
d_{M,K,D}(i,j) = \lim_{\lambda \to \infty} \left( \sum_{k=0}^{\infty} |y_i - y_j|^\lambda \right)^{1/\lambda} = \max |y_i - y_j|
\]

(9)

Minkowski metric of the order \( \lambda \) returns the space alongside that axis on which the two objects show the greatest entire difference. Sensors having real object in its range will send the real packet to the real as well as fake sink node. Sensors taking fake objects in its range will send fake packet to the real as well as fake sink node. Next to the location privacy enhancement the security of the information is carried out sparse matrix encryption.

4.1.2 Security improvement phase:

In this segment, the security of our ASSMLP plan for WSNs is dissected. It demonstrates the proposed ASSMLP plan is substantial and useful. Point by point examination likewise demonstrates the proposed plan could withstand and fulfil security prerequisites in WSNs.

Encryption/decryption using sparse matrix:

The plain text is encrypted by utilizing Sparse Matrices algorithm. This algorithm is a multi-organized encryption and decryption. By conveying encryption algorithm at the sender side message is encrypted before sending message at the beneficiary side and it is held at real source. At the collector side decryption operation is performed and this procedure is held at real sink as opposed to fake one. We consider a succession of positive numbers. We consider the succession of message then the key is produced by Euclidean algorithm.

Extended Euclidean algorithm:

The extended Euclidean algorithm is fundamentally used for finding GCD. The extended Euclidean algorithm is an allowance to the Euclidean algorithm, which computes, besides the greatest common divisor of two integers. Key era by the source sensor is as follows. Similar to the common algorithm, one starts with isolating a by b, trailed by successive divisions with remainders. The algorithm stops when a division has zero leftover portions for the first run through. The greatest common divisor \( gcd(a, b) \) is the last nonzero leftover portion.

Extended Euclidean algorithm (EEA) is used to calculate GCD of \( g \) and \( h \) such that the value of \( gcd(a,b)=1 \) such that the key is generated by Eq. (10)

\[
ag + bh = gcd(a,b)
\]

(10)

If we considering \( a=e \) and \( b=\Phi(n) \) then the key value of the sparse matrix is given by the following Eq. (11). Then Eq. (5) by description (they ought to be co-prime in place of the converse to occur). Then the termination condition is given in Eq. (12).

\[
gcd(e,\phi(n)) = 1
\]

(11)

\[
gcd(e,0) = e = T_0
\]

(12)

From Eq. (12), secure key \( T_0 \) is used for the sparse matrix encryption and decryption process. Thus the secure key is generated with the Extended Euclidean algorithm (EED).

The encryption algorithm is as follows:

The encryption algorithmic program is predicated on prime numbers that are arbitrarily selected. The private key \( T_0 \) could be a one-time key because it is employed to encrypt plain text. Conjointly this can be an associate key that is tough for associate entrant to search out the key.

1) Select arbitrary integers \( \{v_1,v_2,v_m\} \) of a few prime numbers that remain not indistinguishable with the prime factors of all.

2) Select the public key of our algorithm from the matrix \( T^m \) and we express the matrix \( W^m \), wherever the diagonal matrix \( v^m \) has highlights \( \{v_1^T,v_2^T,..v_m^T\} \) and \( v_m \) is the private key of the projected system.

3) Sparse product of the matrix \( W^m \) can be processed by Eq. (14).

\[
X_n = \prod_{k=1}^{m} (1 + \chi^k v_m^T T_0^m)
\]

(13)

At this, we arranged a cipher which is focused on sparse product of matrix \( T^m \) and based on this matrix, sparse product will be calculate. The technique comprises encryption of a typical content into cipher content \( X_n \) over a onetime key called \( T_0 \). Here the packet decryption is carried at the real sink node.

The decryption process is as follows:

The decryption procedure is built by picking arbitrarily chosen sparse matrix and this dense matrix enhances the security. The arrangement for the decrypted text is gotten from the figure text of
proposed technique and it can be find out by the condition (14) and (15).
1) Compute the sequence
\[ z_1 = X_n (X_n = \text{max}) \]  (14)
\[ z_{n+1} = X_{n+1} - z_1 \]  (15)

Where \( z_1 \) and \( z_{n+1} \) are the sequences which is the maximum value of cipher text. These sequences mentioned in Eqs. (14) and (15) are calculated to find out the plain text.

2) Analyse the prime factorization of the each \( z_i \) where \( \{i=1, 2...m\} \)

3) Estimate \( v_i \) and \( T_0 \) from Eq. (17) then this equation proves that \( v_i \) is not identical to prime factors. From Eq. (16) inverse of sparse matrix is related with the sequence of message and the prime factors.

\[ X(U^m)^{-1}_m = 1 + \chi_1 v_1 (n = 1) \]  (16)

4) Calculate prime numbers \( \{v_2, v_m\} \) from Eq. (17) and from the definition of \( T_0 \) we can get the plain text.

\[ \frac{t(U^m)^{-1}}{x_n} = X_s v_n^{T_0} \text{ For } n > 1 \]  (17)
\[ X_s = \frac{z_n}{v_n^{T_0}} (if \ n = 1) \]  (18)
\[ X_s = \frac{z_n}{v_n^{T_0}} (if \ n > 1) \]  (19)

We get the plain text \( \chi(i=1,2,...m) \). Consequently the first plain text is recovered at the real sink node.

Since the Sparse matrix cryptography is uneven and therefore the cryptography methodology is focused on prime numbers that territory unit indiscriminately chose. The private key could be a one-time key since it is utilized to cipher plain text. Furthermore this is frequently machine key that is intense for an interloper to look out the key. In decipherment the cipher is decrypted thrice that is difficult to unscramble for an individual.

5. Experimental results and discussions

From these processes the proposed approach will perform very well to maintain better security and source and sink location privacy in WSN. The proposed Adequate Sparse Secure and Minkowski distance based Location Privacy (ASSMLP) approach simulated using MATLAB simulator and the enactment are compared with the prevailing Biology inspired Self-organized Secure Autonomous Routing Protocol (BIOSARP) technique [27].

5.1 Simulation environment

We discovered the simulation exploration operating MATLAB with a hundred nodes systematically distributed. At first the nodes area unit placed at random within the explicit space. Additional simulation factors are registered in Table 2.

5.2 Performance evaluation

Packet Delivery Ratio analysis (PDR):

It is the proportion of packets received with success to the entire packets transferred with n number of nodes.

\[ \text{Packet delivery ratio} = \frac{D_p}{G_p} \]  (20)

\( D_p \)-Number of Packet received by the receiver
\( G_p \)-Number of Packet generated by the source
Figure 4 represented the performance comparison of PDR in which extent data gives the guideline about how unequivocally the packets in the convention got to the less than receiving end. The most astonishing estimation of this proportion speaks to the better standard of the proposed calculation's presentation furthermore dedicated that more measure of bundles is conveyed to the higher layers.

Energy consumption analysis:

It gives the relation of the energy consumed for total packets received and the amount of energy consumed by the nodes to transfer the packets.

\[ \text{Energy Consumption} = (E_R \times N) + E_T \]  \hspace{1cm} (21)

- \(E_R\)-Received Energy
- \(N\)-Number of Nodes
- \(E_T\)-Transmitter Energy

Figure 5 explains the comparison analysis of energy consumption analysis in which it compares the proposed technique with the existing BIOSARP technique. From this it achieves less energy consumption than the prevailing technique.

Average end to end delay analysis:

End to end delay is denoted as the period reserved for a packet to be transferred between networks from sender to receiver. The end to end packet delay is measured because the part of total end-to-end delays within the entire communication once related to the quantity of packets well offer to the receiver end nodes throughout the whole recursive run.

Figure 6 means the end to end delay comparison in which delay is measured because the part of total end-to-end delays within the entire communication. Once related to the quantity of packets well offer to the receiver end nodes throughout the whole recursive run.

Energy efficiency analysis:
It is the ratio of useful energy conveyed by the network to the total energy provided to the network.

\[
\text{Energy Efficiency} = \frac{E_{\text{transferred}}}{E_{\text{supplied}}} \tag{22}
\]

\(E_{\text{transferred}}\) Useful Energy Transferred by the Network

\(E_{\text{supplied}}\) Total Energy Supplied to the device

Figure 7 deliberated the energy efficient analysis in which efficiency of the proposed technology achieves better performance while comparing to the existing methodology.

**Throughput analysis:**

Throughput is characterized as the rate at information is completely transmitted for each packet sent. It is the entire scope of packets conveyed by the receiver.

\[
\text{Throughput} = \frac{\text{Packet Received}}{\text{Delay}} \tag{23}
\]

Figure 8 represented the throughput analysis and in this performance of ASSMLP achieves maximum throughput while comparing to the BIOSARP.

**Mean Silhouette Coefficient Analysis:**

Silhouette coefficient measures the clustering accuracy and it does not depend on number of clusters

\[
\text{MSC} = \frac{1}{m} \left[ \frac{a(i) - d(i)}{\max[a(i), d(i)]} \right] \tag{24}
\]

\[a(i) = \min[A(i, k)] \tag{25}\]

\(d(i)\) is the average distance from one point \((i)\) to another point with in a single cluster. \(A(i, k)\) is the distance from one point to another point in another cluster.

The performance of Mean silhouette distance is high when it reaches the value nearer to one.

Table 3 represented the performance comparison for the existing and the proposed methodology. Since the overall performance of the proposed Adequate Sparse Secure and Minkowski distance based Location Privacy (ASSMLP) approach is compared with the Biology inspired Self-organized Secure Autonomous Routing Protocol (BIOSARP) technique. Parametric comparison of existing system has less efficient in compared with the proposed technique. Energy efficiency of an existing system has 1.08 percentages. Average Delay has 8.2 ns. Percentage representation of Packet delivery ratio and throughput of prevailed technique have 8.4 and 8.2 respectively. In case of proposed technique Energy efficiency has 5.2 percentages. Average Delay has 7.6 ns. Packet delivery ratio and throughput of proposed technique have 7.7 percentages. Hence the proposed ASLP approach has efficient performance when compared to the prevailing BIOSARP approach.
6. Conclusion

Adequate Sparse Secure and Minkowski distance based Location Privacy approach concentrated on location protection is enhanced by Fake Source and Fake Sink system. Moreover with real source, fake object is localized by Modified pillar k means based Minkowski distance in which the distances. So the data is send from the real source can't be recovered by the aggressor. Security improvement is performed by utilizing sparse matrix encryption method and here onetime key is produced by Extended Euclidean Algorithm is utilized to encrypt plain content. Thus decrypting procedure is just held at the area of original sink. The proposed arrangement can be exhibited to satisfy the indispensable essentials through security investigation taking into record profitable and convincing formal technique and it can be executed in a MATLAB tool. Moreover, we made an execution investigation using the improvement in perspective of packet delivery ratio, energy consumption, energy efficiency, throughput, average end to end delay and mean silhouette coefficient. Thus the efficient clustering is achieved with this modified pillar-K means clustering and the interruption of adversaries is cannot be possible in this network with the novel sparse matrix encryption.

Future work related to the WMN should discover secure policy to transmit the packet with more throughput as possible. Future work will focus on the every system should provide perfect link condition and sets the nodes with different module as possible. So such optimization technologies like Ant lion optimization, dragon fly, lion optimization and other routing algorithms may use to optimize the network complexity in terms of energy and throughput.

References


