SIBER-DELTAKE: AN IMPROVED ACO-KM-ECC BASED TRUST AWARE ROUTING TECHNIQUE IN WSN WITH OPTIMAL DATA RESOLUTION

V. NEELIMA, A. R. NASEER, G. NARSIMHA

1Assoc. Professor & Head, Department of CSE, JITS, Affiliated to JNTUH, Karimnagar, Telangana, India
2Professor & Head, Department of CSE, SOCIE, INHA University, Tashkent
3Professor & Head, Department of CSE, JNTUHCES, Sultanpur, Telangana, India

E-mail: neelima.jits@gmail.com, dr_arnaseer@hotmail.com, narsimha06@gmail.com

ABSTRACT

In this paper, our proposed Hybrid model SIBER-DELTAKE, an improved ACO-KM-ECC trust aware routing protocol is presented to provide secure routing in Wireless Sensor Networks by tackling both the identity and behavior related attacks launched by malicious nodes. The proposed SIBER-DELTAKE protocol is built on Ant Colony Optimization (ACO) which employs K-Medoids(KM) algorithm for formation of clusters and setting up of cluster heads to achieve optimal data resolution. It uses Elliptical Curve Cryptography (ECC) mechanism with key generation and management to secure data integrity as well as data privacy thereby establishing secure routing. It further takes into account Distance, Energy, Link Quality, Trust-Awareness metrics in the routing decision to provide trust enabled routing to prevent non-forwarding attacks by insider misbehaving nodes. The performance evaluation of our proposed hybrid model SIBER-DELTAKE was carried-out using NS-2 Simulator for varying network sizes. Our simulation results indicate that SIBER-DELTAKE performs extremely well in terms of malicious node mitigation, delivering of packets, delay in transferring packets, energy consumption and throughput thereby demonstrating the model effectiveness.

Keywords: Wireless Sensor Networks, Ant Colony Optimization, Secure Routing, Trust Aware Routing, K-Medoids Clustering Algorithm, Elliptic Curve Cryptography, Energy Balancing, Network life Maximization, Swarm Intelligence

1. INTRODUCTION

Due to the tiny inexpensive sensors deployed in Wireless Sensor Networks, to provide cost-effective solutions to a wide range of real world challenges, WSNs have gained immense popularity in industry, military, society and academia [1]. WSNs have been deployed in various scenarios to perform wide variety of functions including climate auditing, military surveillance, forest wildlife monitoring, earthquake monitoring, target tracking, infrastructure evaluation, health inspection, precision agriculture, and also Internet of Thing (IoT) [2-5]. In addition to having robust key management schemes to secure the network from external attacks [6], WSN requires strategies to mitigate the effect of launching non-forwarding attacks by insider nodes, that is, misbehaving nodes refusing to participate in the packet-delivery process. This calls for an efficient mechanism for monitoring the behaviour of the neighbouring forwarding nodes by assigning dynamic trust rating to nodes in the network based on the reputation they built over a period by being trustworthy in participating in the packet delivery and detecting and eliminating the misbehaving (untrustworthy) nodes in the packet forwarding path subsequently.

The past decade has witnessed intensive research to address the capacity of collaboration among sensors in data collection and processing, coordination and supervision of the sensing activity and data flow to the destination (sink) termed as base station. Wireless Sensor Networks are utilised for distributed and collusive sensing of physical occurrence, events of interests, compute local surrounding circumstances or some other metrics and further send the collected
information to a destination sink for suitable functioning. To perform this mission effectively, a proficient routing protocol is essential to set up communication paths connecting the sensor source nodes, and the destination sink. The path choice must be done in such a way that the network lifetime is increased. Due to the environmental characteristics in which the sensor nodes function, along with resource constraints such as on-board energy, processing capability, transmission power, storage restrictions, alert resource supervision and novel routing procedures are required that would compensate the network challenges and the differences. A critical situation to sustain the extensive operational life span of the individual sensor nodes and of the entire network is the efficient utilization of sensor node battery power. It is evident that the active lifetime of a sensor node has a strong dependence on its battery lifetime. The necessary life span of a specific sensor network range from few hours to a number of years based on the type of applications for which they are employed and has a very strong effect on the standard of energy efficiency and potentness of the sensor nodes. Hence WSN routing protocols should be designed with the main key factors of achieving Energy Efficiency and minimizing Network Lifetime (Life Span) by minimizing the entire number of message transmissions involved in the process of path discovery and data delivery, and to organise the forwarding of the packets throughout the multiple paths, so that all nodes can utilize their battery power at a comparable value with more focus on energy balancing among the paths to achieve overall gain in the lifetime of the network [7-10].

In mission critical applications such as battle field surveillance, forest wild life monitoring scenarios, it is required to deploy large number of sensor nodes ranging from thousands to millions in order to cover the large spatial regions at a high-resolution. The routing framework needs to be designed to exhibit scalable performance so as to handle such large and dense networks, and to cope effectively with the related challenges evolved from radio interference and from the requirement to find out, sustain, and use capable long multi-hop paths and also efficient handling of sensor nodes which generate a large amount of data [11]. To maintain the communication operating cost at tolerable levels during the distribution of a huge amount of knowledge, it becomes highly necessary to discover the finest routes in a large-scale network. Further, routing protocols must be designed to sustain the setup and employ data paths effectively for in-network data aggregation. For data intensive applications with the objective of maximizing lifespan of the network, application of hierarchical architecture for direction-finding such as clustering coupled with data aggregation has emerged as a significant and promising solution which makes the trade-off between energy efficacy and data deliverance. To reduce energy dissipation along the forwarding paths, many approaches have been proposed in the literature, but it is observed that nodes at the data aggregation points may get exhausted rapidly. By introducing a dynamical clustering approach, cluster heads are rotated dynamically so that the battery energy among the cluster nodes are balanced and some nodes in the clusters do not get drained rapidly. A variety of clustering approaches in different framework have been projected [12-16] and the majority of these clustering approaches intend at producing the cluster count and distance for transmission to be minimum and employ different techniques to elect cluster heads (CHs).

In Wireless Sensor Networks, routing in a secure fashion is an extra challenging feature as a result of the type of the direction-finding process in infrastructure-less environment. This calls for a cooperative behaviour among neighbouring nodes to route the packets. In such an environment, the selection of a forwarding node should facilitate the more secure path to transmit the data. A secure routing approach must deal with security inconvenience that may arise out of undesired behaviour related to forwarding node’s identity or behaviour [6]. If the forwarding node is recognised as an unwanted node based on its behaviour, then it is treated as a misbehaving node which is responsible for severe attacks. Those serious attacks can be eliminated by offering security forces like efficient cryptography approaches that validate the forwarding node’s identity. On the other hand, a forwarding node may misbehave in the network by refusing to participate in forwarding packets by launching non-forwarding attacks. Such misbehaving nodes can be avoided by models that validate and estimate the node’s behaviour based on trust metric in the network. These two issues emphasize two main responsibilities namely securing content of the packet and securing delivery of the packet need to be taken care while designing a safe and secure path-
routing resolution. Securing Content of the packet deals with the problems in security that relate to their identity. The aim of this mission guarantees that the unauthorized nodes have no access to the packet as it travels from the source node to the destination node. This duty can be performed by providing Data Confidentiality and Data Integrity services. In Data Confidentiality service, the content of the packet initiated from the source node should be accessed only by the destination node and any other intermediate forwarding node must not have any authority to access the packet. Thus, if an intermediate node handles the packet, then it implies that the sink node individuality has been compromised. In Data Integrity service, the sink (destination) node should identify any alterations that could happen in the data package when it receives a data package from the source node. Securing Content of the packet is acquired normally on the property of individuality verification where a node undergoes verification for its identity following certain criteria and then a routing decision is taken. This process is attained earlier using crypto-based frameworks. Securing Delivery of the packet mainly deals with behaviour associated security issues. The main goal is to ensure that the target sink node will be ultimately receiving the transmitted packet. Thus, a misbehaving or compromising forwarding node ought not to misroute a data packet, drop a data packet or refuse the routing to all the nodes in the network by Service Denial attacks. This particular type of a security service can be interpreted as Data Availability which implies that if any node has authorization to retrieve knowledge from any other nodes, next that node must attain this knowledge at whatever time and with no excessive interruption. A resolution for securing delivery of the packet task along with energy efficacy relies on the conception of behaviour trust where trusted nodes can only forward the packet. This technique is so called trust enabled routing.

To acknowledge, trust enabled routing significance, one needs to consider scenarios which justify the immediate need for emphasizing the implementation of Trust enabled routing. Misbehaving nodes in Wireless Sensor Network (WSN) can indulge in misrouting packets to awry targets pointing to misreport or can deny totally data forwarding to their destination node which leads to loss of much needed valid information. Mission critical applications such as military, health or commercial applications can be very sensitive to these attacks where WSN nodes have the utmost responsibility to collect and distribute very biting and extremely sensitive data. Hence, it becomes highly essential to design a Trust aware routing protocol along with data security features to preserve data confidentiality, and data immunity during information distribution, to provide proper data delivery without any corruption, to defend the value of communicated knowledge and above all, to tackle the behaviour related attacks resulting in non-forwarding of data packets. The trust aware or trust enabled routing is mainly concerned with the opinion of neighbouring nodes regarding the behaviour of a next forwarding node in order to make a proper routing decision to forward a packet. This opinion is termed as Trust metric as it determines the trustworthiness of a node to participate in data forwarding task to transfer data when it gets from its preceding node in the direction from the source to the sink. The challenging task of trust computation is carried out by a distributed reputation system [17] associated with each node which performs several operational tasks such as observation of forwarding node’s behaviour by neighbouring nodes of interest (which are awake), exchanging node’s opinion and experience in addition to modelling the collected observations and exchanged information to dynamically evaluate nodes confidence levels and assign ultimately suitable trust rating for nodes. In the perspective of wireless sensor networks, distributed reputation system uses a cooperative filtering algorithm where in every node gives a rating to other nodes based on the collection of knowledge that the neighbouring nodes carry about every other nodes based on their observations on each node’s behaviour during an interval of time to calculate ratings for a group of nodes in the neighbourhood. Reputation systems have been employed very extensively in a variety of domains such as distributed artificial intelligence, evolutionary biology, economics, e-commerce applications and online auctioning, ad hoc and wireless sensor networking [18-20]. In reputation systems, most of the concepts rely on social networks analogy. In general, Reputation System consists of three components – Monitoring, Rating and Response. Monitoring component is in charge for watching the performance of the neighbouring nodes. Rating component will permit the nodes to rate their neighbours depending on the self node observation, observations of other nodes that
exchange among themselves, past history of the watching node and positive threshold rate. Response component has the responsibility of deciding with reference to different capable reactions it could carry out, like eliminating misbehaving nodes or yet killing them by excluding them totally from the network based on the knowledge built by nodes on others’ reputations.

Swarm Intelligence (SI) is a relatively novel framework which refers more commonly to the survey of the collective behaviour of multi-component schemes that organise using decentralized controls and self-organization. Most of the research work in different domains has been inspired by collective behaviours observed in natural systems such as ant colonies (ACO), flocks of birds (PSO), and schools of fishes [21]. Ant Colony systems have successfully tackled the challenges posed by the nature using their inherent appealing characteristics such as adapting to varying surrounding conditions, strong and flexible to the failures produced by inner or outer factors, achieving multifaceted behaviours and collaborative action on the basis of a restricted set of rules and effective supervision of controlled resources coupled with global intelligence which is greater than individual capabilities [22]. Similarities could be drawn with ant colony systems when one considers many of the major challenges to be handled in practical recognition of wireless sensor networking solutions such as resource constraints, absence of centralized authority and infrastructure, complexity and dynamicity of large scale networks, need for survivability and self-configurability, and lastly unattended resolution of potential failures.

In this paper, we present our proposed Hybrid model SIBER-DELTAKE, an improved ACO-KM-ECC trust aware routing protocol based on Ant colony optimization technique that uses K-Medoids (KM) algorithm for formation of clusters and setting up of cluster heads, employs Elliptical Curve Cryptography (ECC) mechanism for secure routing with key generation and management, further taking into account Distance, Energy, Link Quality, Trust Awareness in the routing decision. This is an extension to our model SIBER-DELTA [23] which deals with trust aware routing. Generally, the routing techniques proposed in the literature either consider the problem of secure routing which tackles only identity attacks, i.e., attacks against data security and data confidentiality or consider the problem of trust enabled routing which tackles only behaviour attacks, i.e., non-forwarding attacks due to misbehaviour of nodes. Because of this, the results reported in these models do not actually represent the impact of both the attacks on the network performance. In this proposed hybrid model, we consider both the identity and behaviour related attacks. Our model is designed specifically to suit the harsh and hostile environment such as battlefield, disaster prone, forest area and unattended regions where environmental circumstances keep changing drastically and exposure to various types of threats and attacks keeps increasing. Depending on the environment where they are deployed and the prevailing surrounding environmental and networking conditions, it is noticed that the link quality and other related parameters may vary which are not taken into account when selecting the next forwarder by various ant colony based routing algorithms for WSN reported in the literature. Taking these into account, our approach suggests an improved Forwarder Selection Function to select the best next neighbour to forward the packet to the sink node. It is also observed that the Pheromone Update Model varies from one algorithm to another as the parameters used in the computation of the amount of pheromone concentration to be placed on the path traversed by the backward ant differ. Further, it is found that the amount of pheromone computed to be placed on the path during return journey is not proper to reflect that path as the optimal during the simulation period. Keeping these in mind, a new Pheromone Update Model has been designed by taking into account the parameters like trust rating of the path, nodes having minimum energy along the path, available average Energy, number of hops (i.e., distance indicating shortest path), and link quality of the path to strengthen a path with adequate pheromone to select that path as best path to arrive at the destination node from the source node and at the same time maintain data security and data confidentiality.

The rest of the paper is organized as follows. In Section 2, we present some of the previous work related to Secure routing approaches for Wireless Sensor Networks. In section 3, we briefly describe our existing Trust Aware routing model SIBER-DELTA. Section 4 provides detailed discussion on our proposed hybrid model.
SIBER-DELTAKE (an extension to SIBER-DELTA model), provides data security and incorporates Distance, Energy, Link Quality, and Trust-awareness metrics in the routing decision. The simulation setup, performance evaluation metrics used, results and discussion are presented in section 5, followed by concluding remarks.

2. RELATED WORK

In this section, we present some of the related work carried out in the area of secure routing for wireless sensor networks.

In the area of Trust Aware routing, several approaches for tackling the behavioural attacks were proposed with different objectives. The RFSN approach proposed by S. Ganeriwal and M. Srivastava in [24] employs Bayesian framework which uses beta distribution model [25] for representing reputation, in which every node maintains reputation of other nodes and uses it to estimate their trustworthiness. The Distributed Reputation-based Beacon Trust System (DRBTS) proposed work in [26] computes trust of a beacon node using quorum voting technique where a beacon sensor must get votes for its trustworthiness from minimum half of their common neighbouring nodes so that faulty beacon nodes providing false location knowledge are removed. The TIBFIT (Trust Index Based Fault Tolerance for Arbitrary Data Faults in Sensor Networks) protocol proposed work in [27] detects the events and finds the location where the misbehaving or malicious sensor nodes exist, diagnoses and finally isolates the malicious nodes. Parameterized and Localized trUst management Scheme (PLUS) for WSN proposed work in [28] adopts a localized distributed framework where trust is designed based on either direct or indirect observations. In Locally Aware Reputation System (LARS) proposed in [29], every node maintains the reputation data values of all its single-hop neighbour-nodes using direct observation and the malicious node is never removed from the wireless sensor network, but it is given a chance to build its reputation by good cooperation over a time-out period to be accepted for routing. Trust-Aware Routing Framework (TARF) proposed work in [30] uses an approach that identifies the vulnerable nodes which abuse “stolen” identities to misroute packages by their least trust ratings thereby serving the nodes to evade adversary nodes which misroute packets. A resilient trust model, SensorTrust with a focus on data integrity for hierarchical WSN proposed work in [31] uses the aggregator to maintain trust estimations for children nodes by integrating their long-term reputation and short-term risk and taking into consideration both communication robustness and data integrity using a Gaussian model.

A Reputation system based framework for Energy Efficient, Trust-enabled Secure Routing for wireless Sensor Network is proposed in [17, 29-33]. A customized reputation system (SNARE) proposed work in [32,33] projects the study of algorithms and protocols that communicates with the network layer directly by adopting the geographical routing principle to cope with large network dimensions and relies on a distributed trust management system for the detection of malicious nodes. Three main components of the system are – monitoring, rating and response. The monitoring element, proposed in [34], notifies the activities occurring in forwarding a packet where a watching node will be periodically and probabilistically watching neighbourhood to detect misbehaving events and report these to rating element. Next, the rating element, proposed in [35], estimates the threat a watching node would supply, used for routing decision. Threat is calculated based on their previous misbehaving activities of that malicious node which is further based on the direct observation and the indirect operation. In this proposed work- Geographic, Energy, Trust Aware Routing protocol (GETAR)[6, 36], system follows the defensive response framework wherein based on the trust relations, a node avoids faulty or malicious nodes based on the routing procedure for increasing network lifetime and balancing the load. A simple but robust and independent system to evaluate reputation systems called REputation Systems-Independent Scale for Trust On Routing (RESISTOR) is also proposed in [35].

In the area of ACO based routing algorithms for WSN, several approaches with different constraints were proposed using variants of the Basic Ant Colony Optimization (ACO) based routing approach proposed in [22]. The Energy Efficient Ant Based Routing (EEABR) Protocol proposed in [37] uses pheromone distribution in such a way that nodes nearer to the destination have high pheromone when compared to the other nodes. But it does not take into account link quality resulting in excessive delay in packet
An improved protocol IEEABR [38], a variant of EEABR allows non-optimal paths to be selected for packet transmission, increasing network lifetime and preserving network connectivity. This algorithm shows better performance when the network is dynamic and for higher network density. Three ant colony based routing algorithms for WSN – SC, FF, FP were proposed in [39]. Sensor Driven and Cost-Aware Ant Routing (SC), as the name implies, assumes that ants use GPS to determine the sink location at the initial routing phase so that ants can select initially the best direction towards sink to travel and each node maintains cost to the sink from each of its neighbours. The SC algorithm is energy efficient but suffers from a low success rate. Flooded Forward Ant Routing, FF Protocol is a multipath routing protocol which uses broadcast method to route packets to the sink by flooding forward ants to the sink. The FF algorithm has shorter time delays; however, the algorithm creates a significant amount of traffic. Flooded Piggybacked Ant Routing FP Protocol uses constrained flooding of both forward and data ants to route the data and to discover optimal paths which minimize energy usage in the network along with the data ants holding the forward list. It outperforms SC & FF with high success rate, but incurs high energy consumption, hence it is not energy efficient. It has been seen from the detailed analysis of various-reported ant colony based routing algorithms for WSN in the literature [40, 41], most of the ant colony based routing techniques do not consider all the parameters to select the best quality path in terms of energy, distance, link quality and other metrics thereby leading to the selection of sub-optimal paths. SIBER-XLP which takes into account link quality of the path along with energy, distance to select the best quality path from source to sink for packet forwarding is presented in [42]. This proposed work selects the best next neighbour to forward the packet to the sink node with the sole objective of improving the Network Lifetime by balancing the energy among the nodes in the network to ensure that some nodes along the path do not get depleted fast (resulting in Network disconnections or partitioning) and at the same time selecting good quality links along the path to guarantee that node energy is not wasted due to too frequent retransmissions.

Authentication and key management schemes are the most important security services to provide data security and data confidentiality. Due to the extreme resource constraints in wireless sensor networks, there was extensive research in the last two decades aimed at developing techniques such as random key pre-distribution for pair wise key establishment[43, 44] and broadcast authentication [45,46] to provide security without the expensive Public key cryptography operations deployed in traditional networks. However, random key pre-distribution techniques cannot ensure key establishment among any two nodes and endure arbitrary node compromises at the same time. Moreover, it has become highly challenging task to achieve loose time synchronization required by all broadcast authentication schemes based on TESLA [47] in wireless sensor networks. It has been observed that Public Key Cryptography(PKC) techniques such as Diffie-Hellman (DH) key exchange protocol [48] can be used to achieve Pairwise key establishment without suffering from the node compromise problem and ECDSA digital signature scheme [49] can be used to provide broadcast authentication without requiring time synchronization. In recent years, there has been greater efforts to study the application of Public key cryptography on resource-restricted sensor networks and in this direction, Elliptic Curve Cryptography (ECC) has emerged as highest preference among several PKC options as a result of its fast computation, small key size, and compact signatures[50–53]. For example, an ECC technique requires 160 bits on different parameters, such as 160-bit limited field operations and 160-bit key length [54] to provide security equivalent to 1024-bit RSA. There has been growing research in this direction to develop optimization techniques to speed up the ECC operations to make easier the adoption of (Elliptic Curve Cryptography) ECC-based PKC in Wireless Sensor Networks by designing and developing techniques such as random key pre-distribution and broadcast authentication [45,46] to provide security without the expensive Public key cryptography operations deployed in traditional networks. However, due to the decryption at the aggregators both these approaches suffer leakage of data privacy. Several public-key-based privacy homomorphic
encryption (PH) techniques were proposed in [66], but these techniques fail to resist node compromise attacks. Secure-Enhanced Data Aggregation based on Elliptic Curve Cryptography (SEDA-ECC) proposed work in [67] is built on the basic truth of confidentiality homomorphic encryption (PH) and the approach of divide & conquer. A cryptographic policy based on the ECC map and an Chaotic map discussed in [68] employs Elliptical curve points for the adoption of the communicative nodes and a Chaotic map parameter for the production of pseudo-random bit sequences used in XOR transitions and encryption of the data set. EECBKHM proposed in [69] is a Group based key management framework for authentication in WSN (wireless sensor networks). In this approach, groupings are organized into the grid and heads of groups (CHs) are carefully selected based on the cost of energy, coverage and computing capability. The Base station entrusts a set of EBS keys that contain pairwise keys for the intra connection of the cluster and the inter-cluster of each cluster head (CH). When a malfunction node is detected in its cluster, CH notifies this to the BS which in turn performs re-keying operation to recover the node in a secured manner. This approach increases packet delivery ratio with low power consumption and reduces node capture attacks. A New cryptographic coding system based on the Oval Curve and similar encoding to secure the packet transfer in WSN was proposed in [70]. This method is based on the GASONeC algorithm [71] which constructs the optimal network architecture in a cluster form. This approach employs Elliptic Curve Cryptography to commute public and private keys using a 176-bit encryption key consisting of combining the node ID, Elliptic curve encryption key, and the distance to its cluster head. Homomorphic encryption is used to allow cluster head to aggregate the encrypted data without having to decrypt them thereby reducing the energy consumption of cluster heads. This proposed technique greatly improves the network lifetime, memory requirements, communication overhead, and energy consumption.

The next section briefly describes our SIBER-DELTA model [23].

3. SIBER-DELTA APPROACH

SIBER-DELTA model is an extension to our model SIBER-VLP [42] which does not take into account Trust Awareness. The current SIBER-DELTA model is shown in Figure 1.

**Trust Model**

In our Trust Model, nodes rate each other by using the information of their own direct and indirect interactions with their neighbours. The direct interaction of a node with its neighbouring node is termed as First-Hand Information (FHI). The nodes also collect their neighbours' interactions with that node being rated termed as Second-Hand Information (SHI) which is used to make the rating unbiased. To facilitate these node activities, the simulation period consists of ‘n’ slots where each slot is divided into two sub-periods termed as Forwarding and Monitoring Interval (TFMI) followed by Update Interval (TUPI) as shown in fig 2. During the Forwarding and Monitoring Interval (TFMI), nodes forward their packets, record the transmission and reception of packets to and from their neighbours. During the Update Interval (TUPI), each node computes the Forwarding Misbehaviour Index of their neighbours based on First Hand Information (FHI) and Second-Hand Information (SHI) about their neighbours monitored during the TFMI period. The Forwarding Misbehaviour Index of the neighbour nodes are used then to determine the Mistrust Index of their neighbours. Finally, Trust ratings of all neighbour nodes participating in packet forwarding are computed.

![Figure 1: SIBER-DELTA Framework](image-url)
arrangement of nodes that can monitor the same activities. It is enough to have only one monitoring node active at a time as the observed knowledge can be shared among each other. When an active monitor goes to sleep, another neighbouring node wakes up using a suitable MAC scheduling approach. It is assumed that the network has bidirectional links which is necessary to guarantee packets overhearing during node’s active period.

Forwarder Selection Function

Every node along the path from the source node to the destination (sink) node selects the next best neighbor to redirect the packet to the sink node using a Forwarder selection function. This probability function must always choose an optimal path from a source to the sink to forward packets with the sole objective to enhance the Network Lifespan by balancing the energy levels on all the nodes in the network to ensure that some nodes along the path do not get depleted fast and at the same time selecting good quality links along the path to guarantee that node energy is not wasted due to too frequent retransmissions. This forwarder selection function mainly depends on the Pheromone Trail (PT) and heuristic function indicating node Trust Rating(TR), node’s energy(EN) and node’s link probability(LP). The Pheromone Trail (PT) represents the concentration of pheromone depositions on the path to the nodes considering Energy, distance and link quality along the path to the destination. In other words, higher PT represents the better good quality path in terms of energy, distance and link quality from a source node to the destination.

The Forwarder Selection Function (FSF) \((n_i, n_j)\) used by the current node \(n_i\) to choose the most excellent forward node denoted as \(n_j\) in the neighbourhood is defined by as

\[
\text{FSF} (n_i, n_j) = \frac{\left(\sum_{h=1}^{n_{\text{NS}}} \frac{\Delta P T}{E R(n_j)} \right)_{n_j \in NBS(n_i)}}{\sum_{n_{\text{NS}} \in NBS(n_i)} \frac{\Delta P T}{E R(n_j)}},
\]

where \(NBS(n_i)\) represents the set of neighboring nodes of \(n_i\), the concentration of pheromone deposited onthe path between the nodes \(n_i\) and \(n_j\) represented as \(P T(n_i, n_j)\), the energy level of the neighbouring node \(n_i\) is represented as \(E N(n_i)\), TR\((n_i, n_j)\) represents the Trust rating of the neighbour node \(n_j\) as given by node \(n_i\), LP\((n_i, n_j)\) is the link probability between nodes \(n_i\) and \(n_j\), and specified by the expression:

\[
LP(n_i, n_j) = \frac{1}{ETX(n_i, n_j)}
\]

where the ETX is an Expected Transmission Count which measures the transmission link by considering the past events happened on that link. The influence or impact of a pheromone depositions on the path, energy level of the node, quality of the link between nodes and node trust rating are controlled by the parameters \(\alpha, \beta, \gamma, \delta\).

The Node Energy level, \(EN(n_j)\) is defined as

\[
EN(n_j) = \frac{ER(n_j)}{ER(n_j)} \text{Where} ER(n_j) > E_{th}
\]

where the initial node energy of the node \(n_j\) is denoted as \(EI(n_j)\) and the remaining node energy level of the node \(n_j\) is denoted as \(ER(n_j)\) and threshold node energy is denoted as \(E_{th}\).

Pheromone Model

The Pheromone update model aims to select the optimal path during simulation considering the following requirements. The pheromone concentration should be such that the strongest path should have the largest amount of pheromone whereas the weakest path should have the least amount of pheromone or almost zero. Further, among the competing stronger paths for selection, variations in pheromone concentration should be such that always the strongest path is selected.

The concentration of additional pheromone to be deposited or the Pheromone update by the backward ant during its return journey is specified by the below expression:

\[
\Delta P T = (\text{Path Energy Quality}) \ast (\text{Path Link Quality}) \ast (\text{Path Trust Rating})
\]

\[
= \left(\frac{E_{\text{avg}}}{E_{\text{min}}} \right) \cdot \left(\frac{\text{LP}(Ptk)}{\text{NS}(Ptk)} \right) \cdot \left(\frac{\text{TR}(Ptk)}{\text{PT}(Ptk)} \right)
\]

where \(E_{\text{avg}}, E_{\text{min}}, \text{NS}, \text{LP}, \text{TR}\) are the parameters collected by the forward ant during its journey from the source node to the destination.
node which are compiled and evaluated, after the forward ant reaches the target. \( E_{avg} \) is the average energy level of the nodes in the path Ptk, \( E_{min} \) minimum energy level of the nodes in the path Ptk, Nhsd, number of hops along the path Ptk from the source node to the destination node, LP, average link probability of the path Ptk and TR, trust rating of the nodes along the path Ptk. As soon as the forward ant arrive at the destination node, \( \Delta PT \) is calculated.

4. HYBRID MODEL - SIBER-DELTAKE

In this part, our proposed hybrid model SIBER-DELTAKE Routing protocol for WSN which is an extension to our trust aware routing model SIBER-DELTA [23] is presented.

![Figure 3: Framework of Hybrid algorithm](image)

This proposed system is based on Swarm Ant Colony Optimization with K-medoids algorithm for cluster formation and cluster head selection. It employs secure mechanism using Elliptic Curve Cryptography (SACO-KM-ECC) to provide data security and data confidentiality. It selects the optimal route from source to destination by taking into account the metrics Distance, Energy, Link quality, Trust awareness in the routing decision.

**K-medoids Algorithm(KM)**

K-medoids algorithm is used in our model SIBER-DELTAKE for the formation of Clusters and selection of cluster heads. K-medoids clustering [72] is a variant to K-means approach which is more robust to noises and outlier. K-medoids approach uses an actual point in the cluster to represent the centre of a cluster instead of using the mean point. The object, medoids with the minimum sum of distances to other points is most centrally located.

The K-medoids algorithm is used in our model SIBER-DELTAKE for the formation of Clusters and selection of cluster heads. K-medoids clustering [72] is a variant to K-means approach which is more robust to noises and outlier. K-medoids approach uses an actual point in the cluster to represent the centre of a cluster instead of using the mean point. The object, medoids with the minimum sum of distances to other points is most centrally located.

The algorithm proceeds in two steps [73] termed as BUILD and SWAP steps. In the BUILD-step, \( k \) "centrally located" objects, to be used as initial medoids are sequentially selected. In the SWAP-step, swap is carried out when the objective function can be minimized by interchanging (swapping) a selected object with an unselected object. This process is continued until the objective function can no longer be reduced as illustrated in figure 4. The various steps used in the K-Medoids algorithm are described in fig.5.
Fig. 4 - K-Medoids main algorithm

The four situations encountered in this process are – (i) shift-out membership where an object $P_j$ is shifted into another cluster from currently considered cluster of $C_i$, (ii) update current medoid wherein the current medoid $C_i$ is replaced while finding a new medoid $C_m$, (iii) no change, i.e., objects in the current cluster result have the same or even smaller square error criterion (SEC) measure for all the possible redistributions considered; and lastly (iv) Shift-in membership - assign an outside object $P_j$ to the current cluster with the new medoid $C_m$.

Step1: Selection of Initial medoids

1.1 The distance between each pair of all objects computed using Euclidean distance as a dissimilar measure is as follows:

$$d_{ij} = \sqrt{\sum_{a=1}^{n}(X_{ja} - X_{ia})^2}, i, j = 1, ..., n$$

1.2 To build an initial guess at the centres of the clusters, $P_{ij}$ is calculated

$$P_{ij} = \frac{\sum_{i=1}^{n}P_{ij}}{\sum_{i=1}^{n}d_{ij}}, i = 1, ..., n; j = 1, ..., n$$

1.3 Calculate $\sum_{i=1}^{n}d_{ij}$ at every object and arrange them in ascending order. The k objects having the least value are selected as first group medoids.

1.4 Every object is assigned to the closest medoid.

1.5 The sum of distance from each and every object to their medoids and the current best possible value is to be calculated.

Step2: Determine new medoids

The sum of the distance to other objects in its cluster is minimized by replacing the current medoid in every cluster by the object.

Step3: Computation of new optimal value

3.1 Select every object to the closest new medoid.

3.2 The total distance from all objects to their new medoids and new best value is calculated. Stop the algorithm when the optimal value is one and the same to the preceding one. Otherwise, go back to the Step 2.

Figure 5: Steps of the K-Medoids algorithm

Secure ECC Mechanism

Elliptic curve cryptography (ECC) has become most desired approach to public key cryptography which is based on the algebraic structure of elliptic curves over limited fields [49-52]. An elliptic curve over prime field $\mathbb{F}_p$, where $p$ is a large prime number, is defined by a cubic equation of the form $y^2 = x^3 + ax + b$ where $a, b \in \mathbb{F}_p$ are integers that satisfy the equation $4a^3 + 27b^2 \neq 0$. To have ECC based secure communication, every sensor node in the network must know an elliptic curve in addition to base
point p which lies on the curve. It is assumed here that during the initial setup or the initialization phase, the elliptic curve parameters and also the base point p are loaded before only into the memory of every sensor node. Every node chooses a random prime integer as its private key and generates its public key by multiplying the private key by the base point p in order to have a secure communication between a pair of nodes. For example, to setup communication between two nodes A and B, a secret key generated among nodes A and B is shared – node A choosing an arbitrary prime numeral PRIK_A and node B choosing an arbitrary prime numeral PRIK_B as their private keys. Then generating their public keys using the following equations, PUBK_A = PRIK_A * p and PUBK_B = PRIK_B * p. Curve points represent the public key for node A as well as node B. PRIK_A and PRIK_B are multiplied by base point p to produce the public keys. Generated Public keys help in producing a shared secret key SSK= PRIVK_A * PUBK_B= PRIVK_B * PUBK_A = PRIVK_A * PRIVK_B * p. It becomes computationally difficult for a hacker to determine node A private key and node B private key, PRIK_A and PRIK_B by knowing the values of PUBK_B; PUBK_A, and p. The major acts of point - addition & doubling of the elliptic curve are used by the sensor node to produce its own public key. If node A wants to transmit a message m to node B which is encoded on the elliptic Curve E into point M. Node A uses Node B’s public key to encrypt the message. The cipher text obtained is given by C = {k*p, M + k*PUBKB}, where k is a random integer and this random number k ensures that cipher text generated will be different for the same message each time which will make it hard for the adversary to decrypt the message. Let C_1 and C_2 be the two cipher texts that are generated: C_1 = k*p, C_2 = M + k*PUBKB. Once the message is received, node B decrypts the encrypted message by the below decryption step:

\[ M_1 = C_2 - PRIK_B * C_1 = M_1 + k*PUBK_B - PRIK_B * k * p = M_1 + k*PUBK_B - k*PUBK_B = M_1 \] (original Message)

As cluster heads are involved in receiving the encrypted data from their members of cluster, then processing the data to perform data aggregation and finally forwarding the aggregated data to the base station, they consume more energy when compared to the member nodes. In order to reduce the energy consumption by cluster heads, cluster heads combine the encrypted message arriving from the members of the cluster and performs no decryption of the message at the CH. Homomorphic encryption allows cluster heads to perform aggregation of the encrypted data with no decryption thereby reducing the energy consumption of cluster heads. This results in saving of more energy and much stronger privacy of data as attackers will not be capable to hack data from intermediary nodes.

5. Experimental Results

In this section, we present the simulation setup, performance evaluation metrics used in our simulation and the simulation results showing the desired behavior of the hybrid model.

5.1 Simulation Setup

Our proposed hybrid model SIBER-DELTAKE was simulated using NS-2 simulator by considering static network scenarios with network sizes of 25, 50 and 100 nodes randomly distributed in the network area of 1000x500 m². Our proposed SACO-KM-ECC based SIBER-DELTAKE system is compared with SIBER-DELTA [23] with trust awareness and SIBER-VLP [42] without trust awareness for varying network sizes by introducing 10%, 20%, and 30% attackers in the network. The performance of the network is evaluated using the following metrics - Packet Delivery Ratio, End to End Delay, Dropping Ratio, Energy Consumption and Throughput. Table 1 shows the system parameters used in our simulations.

5.2 Performance Evaluation Metrics

In this section, we present the performance metrics used in the evaluation of our proposed hybrid approach: SIBER-DELTAKE.

(i) Packet Delivery Ratio is defined as

\[ PDR = \frac{\text{Delivered packets at the destination node}}{\text{Generated packets at the source node}} \] …..(7)

(ii) Packet drop ratio is defined as

\[ PLR = \frac{\text{Generated packets} - \text{Received packets}}{\text{Generated packets}} \] .... (8)

(iii) End2End Delay or Latency is defined as
Delay (i) = receiving time (i) – sending time (i)
Total delay = Total delay + Delay (i)

\[
\text{Average delay} = \frac{\text{Total Delay}}{\text{Count}} \quad \ldots (9)
\]

where \( \text{Count} = \text{Total packet count} \) and \( i = \text{Packet sequence number} \)

(iv) Throughput is defined as \( \text{Thr} \)

\[
\text{Throughput} = \frac{(\text{Received data} \times 8)}{\text{Data transmission period}} \quad \ldots (10)
\]

Table 1: Static Scenario - Simulation parameters

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>VALUE</th>
<th>PARAMETER</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>CBR</td>
<td>Propagation</td>
<td>Two-way ground</td>
</tr>
<tr>
<td>Traffic</td>
<td></td>
<td>model</td>
<td></td>
</tr>
<tr>
<td>Transmission</td>
<td>10 packets/sec</td>
<td>Packet size</td>
<td>512 bytes</td>
</tr>
<tr>
<td>rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radio range</td>
<td>250m</td>
<td>Routing</td>
<td>SIBER-DELTAKE,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>protocols</td>
<td>SIBER-DELTA,</td>
</tr>
<tr>
<td>Topology</td>
<td>Random</td>
<td></td>
<td>SIBER-VLP</td>
</tr>
<tr>
<td>Number of nodes</td>
<td>25, 50, 100</td>
<td>Simulation time</td>
<td>9000ms</td>
</tr>
<tr>
<td>Area</td>
<td>1000x500 m²</td>
<td>Bandwidth</td>
<td>10Mbps</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maximum packets</td>
<td>10000</td>
</tr>
<tr>
<td>Clusters</td>
<td>8</td>
<td>Malicious nodes</td>
<td>10%, 20%, 30%</td>
</tr>
<tr>
<td>Initial energy</td>
<td>100J</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.2 Results and Discussion

In this part, we exhibit the screenshots simulated, results and analysis of the simulation findings of hybrid model SIBER-DELTAKE when compared with SIBER-DELTA with trust awareness and SIBER-VLP without trust awareness.

Figures 6 to 29 show the screenshots showing some of the simulation scenarios. Figure 6 shows the random deployment of 25 sensor nodes in the network of area 1000x500m². The broadcasting process shown in figure 7 clearly demonstrates the start of broadcasting and how the nodes establish communication with remaining nodes which are within the communication range to identify neighboring nodes. The processing of packets in the network depends on the neighboring nodes identified in the broadcasting step. Screenshot in figure 8 depicts the start of communication between nodes with Constant Bit Rate (CBR) for traffic level purpose and User Datagram (UDP) protocol used for data transmission purpose.

Screenshots in Figures 9, 10, 11 and 12 show the attack level scenarios wherein dropping of packets due to the presence of malicious nodes are indicated. Screenshot in fig 13 illustrates broadcasting in our proposed model. All nodes are represented based on the topology values in this network. Screenshot in fig 14 shows the delay, bandwidth and link between communicating nodes. Figure 15 depicts the routing process in the network. As shown in the screenshot, depending upon the distance among the nodes, clustering process starts inside the network as well as all nodes are differentiated based on clusters. Screenshots in figures 16 and 17 show the introduction of malicious nodes into the proposed model.

In the screenshot in figure 18, average energy levels of different nodes are shown along with distance among the neighboring nodes. From the figure, it is easy to identify the nodes having more energy levels that can be used for communication. Previous hop node, next hop node, index node, source node, destination node and associated pheromone values based on communication in the network are shown in figure 19. The pheromone values indicate available optimal paths in the network.

Screenshot in figure 20 shows the deployment of nodes in our proposed hybrid model SIBER-DELTAKE. Here, all nodes are represented based on their topology values. In this network, 1000x500m² area is considered and all nodes are setup into different clusters. Finally, eight clusters are formed. In every cluster, one of the nodes is chosen as a cluster head and these cluster heads communicate with remaining cluster nodes. The start of key generation process is shown in figure 21. The ECC mechanism is used to generate private and public keys for all nodes to establish secure communication. Figure 22 shows the broadcasting in the network based on the ECC keys generated. SIBER-DELTAKE approach results in optimal paths and secure communication levels.

The screenshot in Fig. 23 shows the routing process between clusters in the network. Every cluster head communicates with the nodes within the clusters for data aggregation. All nodes properly participate in data forwarding thereby improving the network performance.
previous hop node, next hop node, index node, source node, destination node and pheromone values associated with the nodes are shown in figure 24. The pheromone values indicate available optimal paths in the network. In this simulation process, pheromone values shown represent higher level of communication compared to previous approaches.

The screenshot in figure 25 shows the final trust rating values of each node in the network. The final trust values are computed using the SIBER-DELTA approach based on direct and indirect trust values associated with the nodes. In the SIBER-DELTA approach, direct trust values are computed using node’s direct observation known as firsthand information and indirect trust values are computed using the secondhand information provided by the neighboring nodes based on their observation. The final trust ratings show the trustworthiness of the nodes in forwarding the packets and nodes are selected based on their high trust rating for packet forwarding.

The Public keys and Private keys generated based on ECC and associated with the nodes to provide data security and data confidentiality and to establish secure communication are shown in figure 26 and 27 respectively. The screenshot in figure 28 shows the distance of nodes to their neighbour nodes and midpoint position of nodes to the base station to form clusters and cluster head selection in the network. Here midpoints of nodes are calculated using the K-Medoids algorithm to setup efficient clusters to achieve better performance in the network. The screenshot in figure 29 shows the trace file of the network. The trace file represents all attributes, variables, values and data levels. Here every parameter is considered and evaluated to know the performance of network. In this screenshot, simulation time, node values, node positions, energy levels of nodes, packet size, process of request and receiving of data, delay time and trust values are displayed.

Next, we present the performance of our proposed hybrid model SIBER-DELTAKE, and the existing models SIBER-DELTA and SIBER-VLP in the presence of 10%, 20% and 30% malicious nodes in the network of size=25, 50 and 100 nodes.
Figure 10: Packet drops at malicious node

Figure 11: Packet drops at Attack Level

Figure 12: Data transmission at attack level

Figure 13: Broadcasting in our model

Figure 14: Link, Bandwidth and Delay between nodes

Figure 15: Routing Process in the network

Figure 16: Broadcasting of data during malicious nodes entry

Figure 17: Malicious nodes introduced into our model
Figure 18: Average Energy levels of individual nodes

Figure 19: Pheromone values with Hops Information

Figure 20: Nodes deployment in SIBER-DELTAKE model

Figure 21: Key generation for all the nodes

Figure 22: Broadcasting in network based on keys

Figure 23: Routing Process between Clusters in the Network
As it is seen from figs. 30a), 31a) and 32a), SIBER-VLP model exhibits performance degradation as malicious nodes are introduced in the network. As the number of malicious nodes increase, an increase in packet drops is observed due to the presence of more malicious nodes in the paths selected by the ants. It is evident from the plots that SIBER-DELTAKE and SIBER-DELTA models exhibit high packet delivery ratio and SIBER-DELTAKE performing better than SIBER-DELTA as more trusted and secure optimal paths are selected to forward the packets resulting in higher performance. There may be slight decrease in the packet delivery ratio as the percentage of malicious nodes introduced into the
network increases. This situation would happen when more malicious nodes appear along the selected paths and packets get lost due to selection of paths with higher hop count.

It can be seen from figs. 30b), 31b), 32b) that SIBER-DELTAKE consume little more energy when compared to SIBER-VLP and SIBER-DELTA but it is reasonable considering the fact that hybrid model needs to perform ECC computation to provide data confidentiality and data Integrity in the presence of trust awareness. Though packet delivery ratio is less in SIBER-VLP, but the comparable energy consumption in this case may be due to both packet routing and packet retransmissions.
Fig 31a) Effect of Malicious nodes on Packet Delivery Ratio – 50 nodes

Fig 31b) Effect of Malicious nodes on Energy Consumption – 50 nodes

Fig 31c) Effect of Malicious nodes on Latency – 50 nodes

Fig 31d) Effect of Malicious nodes on Throughput – 50 nodes

Fig 32a) Effect of Malicious nodes on Packet Delivery Ratio – 100 nodes

Fig 32b) Effect of Malicious nodes on Energy Consumption – 100 nodes

Fig 32c) Effect of Malicious nodes on Latency – 100 nodes

Fig 32d) Effect of Malicious nodes on Throughput – 100 nodes
As far as the end to end delay is considered, it can be seen from figures 30c), 31c), 32c) that hybrid model has low delay when compared to other models as it selects always the most trusted and secure optimal paths. Moreover, as the number of nodes increases, there will be more number of alternate paths available to route the packets so that the malicious nodes along the selected paths can be avoided. It is clear from the figures 30d), 31d), 32d) that SIBER-DELTAKE has higher throughput when compared to SIBER-DELTA and SIBER-VLP. SIBER-VLP performs very poorly in the existence of larger malicious or faulty nodes in the network.

6. CONCLUSION

Secure Routing in Wireless Sensor Networks mainly deals with the security problems that may arise due to both the identity and behavioral related attacks by the malicious nodes. Considering these critical issues, we have developed a Hybrid model SIBER-DELTAKE, an improved ACO-KM-ECC trust aware routing protocol based on Ant colony optimization technique to provide secure routing in WSN. Identity related attacks are refrained by employing Elliptical Curve Cryptography (ECC) based mechanism with key generation and management to secure data integrity as well as data privacy. Behavioral level attacks are mitigated by taking into account distance, energy, link quality, trust-awareness metrics in the routing decision thereby providing trust enabled routing to prevent non-forwarding attacks by
insider misbehaving nodes. SIBER-DELTAKE model uses K-Medoids(KM) algorithm for formation of clusters and setting up of cluster heads to achieve optimal data resolution. The proposed hybrid model was simulated using NS-2 simulator for its performance evaluation and compared with the SIBER-DELT and SIBER-VLP protocols with varying network sizes. Our simulation results indicate that SIBER-DELTAKE performs extremely well in detecting and avoiding malicious nodes thereby achieving high Packet Delivery Ratio, low Latency and greater Energy Efficiency.

REFERENCES


