SN Computer Science (2023) 4:808
https://doi.org/10.1007/s42979-023-02257-3

ORIGINAL RESEARCH

An Efficient ACO-based Routing and Data Fusion Approach for IoT Networks

Ishita Chakraborty[1] · Prodipto Das[1]

Received: 4 February 2023 / Accepted: 16 August 2023

Abstract
The ant colony optimization (ACO) is an evolutionary algorithm that tries to imitate the usual biological behavior of ants. Since Internet of Things (IoT) works by integrating and connecting devices of heterogeneous architecture, the size of the network increases rapidly. Therefore, in such situations ACO can be used to attain ideal solutions for large-scale optimization problems. As wireless sensors network (WSN) can integrate itself with IoT, the routing challenges faced by both of WSN and IoT are similar. To cope with the dynamics of the environment many intelligent routing algorithms have been designed. In this paper, an ACO-based routing algorithm for IoT networks has been proposed to analyze and enhance the scalability of the network, by minimizing the delay of the time critical applications. This would help in finding the optimal path for data transmission, and improve the efficiency of IoT communications. The proposed algorithm is simulated using network simulators (NS-2) that showed improvement in conserving energy when compared to the traditional ACO-based routing. Our proposed scheme prolonged the network lifetime and was found to have a 20% more packet delivery ratio, 19% reduced end-to-end delay and almost consumed 78% less energy.

Keywords IoT · ACO · Data fusion · WSN · Sensor selection · Optimization

Introduction

WSN has vast application in the field of medicine, military, home applications, environment, astronomy and many more. Every sensor node contains both processing as well as communication fundamentals that is intended to monitor the environment for different events that is specified by the designer of the network. Information regarding the environment is collected by the sensor node that is delivered to a central source node called the base station (BS) from where the user can extract or collect the desired data or information. Therefore, WSN must be easy to deploy, should have the possibility of multiple hop connections, be self-organizing in nature and should be able to support multimedia services. Since WSN contains large number of nodes that collect information from different sources, there exists a lot of redundancy [29]. Due to its low battery life, self-organizing nature and restricted transmission range, energy efficient routing remains a challenging research problem in WSN [20]. Node deployment is a process of setting up the sensor nodes in a particular region by finding out their relationship with each other. This process relies upon its intermediate node to find its next hop for delivering the packets from a source to a particular destination. Node deployment can either be self-organizing or deterministic [2]. Self-organizing systems has sensor nodes that are scattered randomly creating an adhoc like infrastructure, whereas in deterministic situation the sensors are placed physically and data are routed through predestined paths. As sensor nodes are deployed over larger physical area, it generates data that are highly redundant. If each and every sensor node begin to send their own data packet to the sink individually, then the time taken to traverse the data from source to destination would be very high which in turn will lead to high energy consumption and inconsistent result. Since sensor node operate on battery, there are chances of sensors to die midway [5]. Therefore, to increase the quality of service (QOS) of the network, data

This article is part of the topical collection "SWOT to AI-embraced Communication Systems (SWOT-AI)" guest edited by Somnath Mukhopadhyay, Debashis De, Sunita Sarkar and Celia Shahnaz.

Prodipto Das
prodiptodas@gmail.com

1 Department of Computer Science, Assam University, Silchar, India

Published online: 21 October 2023

aggregation is required. Data fusion is the combination of data to remove redundancy using some basic mathematical function such as summation, min–max, average, etc. [4]. The LEACH-MA protocol is introduced by the authors in [18]. They have incorporated a modified ACO foundation that employs the residual energy parameter in conjunction with the low-energy adaptive cluster hierarchy (LEACH) protocol to select the current cluster head. By combining energy and distance considerations, the algorithm minimizes energy consumption during cluster head selection. However, no specific details are furnished regarding the threshold value for cluster head designation. In [21], an ANT-BFS approach, a hybrid tree-based search method, is proposed to optimize information communication routes. It combines breadth-first search with ACO to curtail energy consumption. Potential challenges concerning memory and computational demands may arise when the cluster head and sink are distant, making it less suitable for large-scale applications. A content-based routing (CCR) protocol is introduced to ensure reliable information transmission in IoT applications [10]. The proposed scheme employs data aggregation and load balancing technique to transmit data based on message content while integrating link quality information. Although it achieves traffic reduction, details concerning reliability estimation and associated parameters remain unspecified. An energy-conscious ant-based routing algorithm is presented in [3] that employs reward and punishment mechanisms, along with pheromone update rules, to enhance network lifetime and energy efficiency. Despite its merits, dealing of various energy utilization factors is necessary. The delay parameter of a node is employed, but the factors effecting delay is not mentioned. An improved energy-saving ant colony-based routing protocol for sensor networks is presented by the authors in [25]. The paper discusses three phases: neighbor discovery via link information, packet transmission using the exponentially weighted moving average (EWMA) technique and efficient and reliable end-to-end delivery. Simulation results highlight its efficiency, especially in terms of energy efficiency and throughput, compared to standard and novel routing algorithms. A multi-constrained technique ensuring Quality of Service (QoS), known as IAMQER, is presented in [28], focusing on ant colony optimization. Simulation outcomes demonstrate reduced average energy consumption and increased packet delivery ratio. Moreover, a route evaluation function is introduced. However, the IAMQER strategy's reliance on traditional ACO methods may result in extended processing delays. The authors in [11] have addressed the energy consumption issues of IoT core network using ACO. They have used MATLAB to simulate their proposed algorithm to evaluate the scalability of the network. A cluster based energy efficient routing approach for IoT network is presented in [22]. Where the authors have used different network parameters like energy consumption, PDR, package loss rates and the number of alive nodes to evaluate proposed methodology. Different concerns and ongoing challenges within WSNs, along with their heuristic and meta-heuristic solutions is discussed by the authors in [19, 27, 31]. Addressing a subset of communication parameters alone is insufficient to achieve the desired communication quality in energy-constrained WSN applications. To effectively address routing challenges and enable efficient communication within WSNs, a comprehensive approach is necessary. Optimal data transmission routes with minimal energy consumption require consideration of various factors influencing energy parameters. Leveraging the benefits of the ant system, our proposed approach incorporates a fitness function which acts as route evaluation index for selecting the optimal paths by considering the existing energy levels and route distances. The network parameters like throughput, delay, energy consumption and packet delivery ratio are used to evaluate the performance of the proposed work (by incorporating the fitness function) while comparing it with the traditional ACO algorithm (without using the fitness function).

System Model

This section describes the process of data fusion and its simulation using a network simulator. The result of the proposed mechanism is compared with the work performed by the same authors in [5].

Network Model

Two network scenarios called the mesh and the tree topology is considered for implementation purpose of the proposed algorithm. In the case of mesh topology, the nodes form a 20×20 Km^2 square region, centered around the sink node, with 37 additional sensor nodes uniformly distributed. For tree topology, the sink is placed at the upper center of a 25×25 Km^2 sensor field, organized hierarchically alongside 20 other nodes. The network configuration is presented in Table 1. The network also assumes the following:

i. Initially all the sensor nodes contain same energy.
ii. All nodes have similar characteristics and

Table 1 Network configuration

Network organization	Number of nodes	Range (Km)	Simulation area (Km^2)
Mesh	40	4.5	25×25
Tree	24	4	50×50

iii. Every node receive data from its neighboring sensor nodes

Energy Model

The radio model adopted for communication is proposed in Heinzelman, Chandrakasan and Balakrishnan, 2002. If the distance $d_{i,j}$ between the sensor nodes i and j are smaller than the threshold d_0 then the free-space model is used; else, the multi-path fading model is employed. The amount of energy consumed by the sensor node to transmit 1-bit data over a distance of $d_{i,j}$ is calculated using the following equation:

$$E^T_{i,j} = \begin{cases} E_{\text{elec}} + \xi_{fs} d^2_{i,j}, if d_{i,j} < d_0 \\ E_{\text{elec}} + \xi_{mp} d^4_{i,j}, if d_{i,j} \geq d_0 \end{cases}, \tag{1}$$

where E_{elec} is the power consumed by the transmitter, ξ_{fs} is the energy absorbed by the amplifier in free-space and ξ_{mp} is the energy absorbed in the multi-path model by the amplifier. The energy E_{RX} is assumed to be constant. d_0 is calculated using the following equation:

$$d_0 = \sqrt{\xi_{fs}/\xi_{mp}}. \tag{2}$$

The energy consumed, E_{con_i} of node i to broadcast a data packet of size D_{size_i} at time t is found using the following equation:

$$E_{\text{con}_i} = E^T_{i,j} * D_{\text{size}_i} * t. \tag{3}$$

E_{initial_i} is considered to be the initial energy of the sensor. The sensor residual energy of the sensor node SN_i is computed using the following equation:

$$E_{\text{res}_i} = E_{\text{initial}_i} - E_{\text{con}_i}. \tag{4}$$

Proposed ACO

There are many possible paths to route a packet from CHs to the BS; apart from the shortest, most efficient and predictable path, it is critical to convey aggregated data from CHs to the BS. This section presents an ant colony optimization-based probabilistic routing technique. This technique is encouraged by the effectiveness of ants to determine the shortened route from their colony to a food source [24]. It is used to find solutions to numerous optimization problems, such as obtaining the optimal route in a graph [12, 26]. Ant proceeds along a randomly selected path in the nearby communities to look for food sources. They communicate with one another with the release of a chemical signal known as pheromones. The quantity of pheromone released by the ants while returning to its colony is proportionate to the amount and quality of the food it has found, thereby increasing the probability of other ants following the pheromone-rich path [30]. The proposed ACO algorithm is depicted in Fig. 1.

While searching for an optimized path with reduced energy consumption from CHs to the BS, the qth ant at time t uses the probabilistic formula given in the following equation:

$$p^q_{i,j}(t) = \frac{(\tau_{i,j}(t))^\alpha (\eta_{i,j}(t))^\beta}{\sum_s (\tau_{i,s}(t))^\alpha (\eta_{i,s}(t))^\beta}, \tag{5}$$

where the heuristic constants $\alpha(\geq 0)$ and $\beta(\geq 1)$ control the influence of pheromone strength and heuristic information, respectively. On the edge i, j, $\tau_{i,j}(t)$ represents the strength of the pheromone and $\eta_{i,j}(t)$ represents the desirableness.$\eta_{i,j}(t)$ is calculated using the following equation:

$$\eta_{i,j}(t) = \frac{1}{d_{i,j}}, \tag{6}$$

$d_{i,j}$ is the length between i and j. The quantity of pheromone is calculated using the following equations:

$$\tau_{i,j}(t+1) = (1-\rho)\tau_{i,j}(t) + \rho\Delta\tau_{i,j}(t), \tag{7}$$

$$\Delta\tau_{i,j}(t) = \sum_{q=1}^{Q} \Delta\tau^q_{i,j}, \tag{8}$$

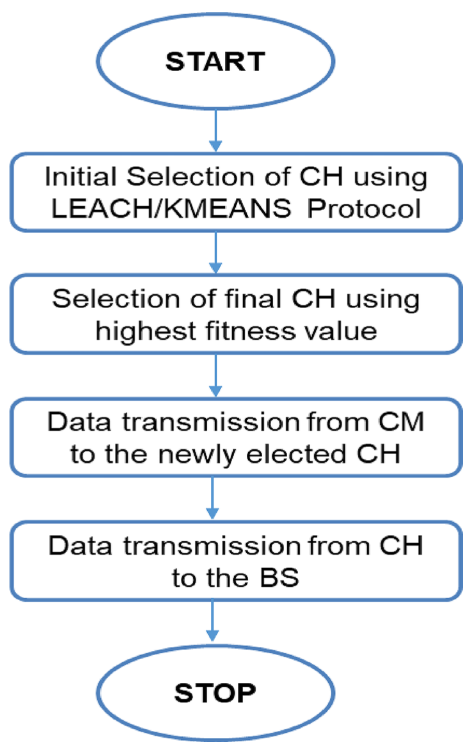


Fig. 1 Working of the proposed ACO algorithm

where ρ is the pheromone evaporation coefficient and $\Delta\tau_{i,j}(t)$ is the quantity of pheromone raised or deposited at edge i, j, which is given by the following equation:

$$\Delta\tau_{i,j}(t) = \begin{cases} M/L_q & \textit{if ant q uses an edge}(i,j) \\ 0 & \textit{otherwise} \end{cases}, \tag{9}$$

M denotes pheromone intensity and L_q denotes the length of the path taken by the qth ant to complete the tour. Additionally, the pheromone trail is globally updated when all ants have finished their route, i.e., a complete iteration. The pheromone is updated using the following equation. The process of modified ACO algorithm is explained with the help of Algorithm 1.

$$\tau_{i,j}(t+1) = \left(1 - \rho_g\right)\tau_{i,j}(t) + \rho_g \Delta\tau_{i,j}(t). \tag{10}$$

Algorithm 1 Modified ACO Routing Algorithm

Begin

Input: Cluster Heads (CH)

Output: Construct optimum path from cluster heads to base station

1. Initialise N_{ch}
2. Initialise the pheromone matrix as zero
3. Calculate $\eta_{i,j}(t)$ using equation (6)
4. **for** $CH_i \leftarrow$ 1 to N_{ch} **do**
    5. **while** $iter \leq M_{iter}$ **do**
        6. Set the ants at CH_i
        7. **for** $ant_q \leftarrow$ 1 to N_{ants} **do**
            8. Use equation (5) calculate to$p^{q}_{i,j}(t)$
            9. ant_q uses the roulette method to choose the next CH for CH_i
            10. The selected CH node is added to the taboo table of ant_q
            11. Use equation (7) to update $\tau_{ij}{}^{q}(t+1)$
        12. **end for**
        13. The fitness value of the route may be calculated by using equation (11) or (12)
        14. Update the best solution by comparing it with the previous best solution
        15. **if** the present solution is best, **then**
            16. Use equation (10) to update $\tau_{ij}(t+1)$
        17. **end if**
        18. Increment *iter* by 1
    19. **end while**
    20. Clear taboo table
21. **end for**
22. END

Time and Space Complexity

In the algorithm, there are three loops located at steps 4, 5 and 7 accordingly. According to the laws of algorithms, the time complexity will be regarded as $O(n)$ if there is a loop that runs from an initial value to a specific range. Time complexity increases with the increase of number of loops in nested loops. The time complexity, for instance, will be $O(n^2)$ if there are two loops in the nested loop. Time complexity for nested loops with three loops will be $O(n^3)$. Since there are three loops in nested structure in the proposed ACO method, the worst case time complexity is $O(n^3)$. The while loop has a condition iter $\leq M_{\text{iter}}$, which might not always be true depending on certain circumstances. This scenario may be considered as best case for the algorithm and the time complexity will be $O(n^2)$. But this situation is rare, so the average case time complexity is $O(n^3)$.

In space complexity, if in the algorithm a matrix of size $m \times n$ exist space complexity is considered as $O(m \times n)$. If a k-dimensional array is used, where each dimension is considered as n, then the algorithm has a space complexity of $O(n^k)$. In the proposed algorithm, there are three main arrays namely CH for cluster heads, $\eta_{i,j}(t)$ and $p^q_{i,j}(t)$. All are linear arrays so the space complexity will be $O(n)$. Moreover, there are two matrices pheromone matrix and taboo table, each of them need a space in the order $m \times n$. Space complexity in this case will be $O(m \times n)$. Other than these space requirements a few variables are present in the proposed algorithm namely N, t, iter, ants, etc. Each of them will be considered as $O(1)$. Therefore, the final space complexity of the proposed algorithm is $O(m \times n)$ and the average case space complexity is $O(n)$.

Fitness Function

The fitness function evaluates the route by incorporating the residual energy of the node, length of the routing path and the angle formed by between the cluster Member (CM), CH and BS. Neglecting the evaluation of remaining energy, as seen in the conventional ACO algorithm, can impact the overall lifespan of the network. Each specific ant becomes associated with a routing path once all ants reach the destination node. The following are the two fitness functions that have been used to find the cost of an optimal path.

For the first fitness function, the length of the routing path and the angles formed between the CM, CH and BS are considered, and the fitness value is determined using the following equation:

$$f_{\text{fitness}^k_q} = \frac{1}{\theta_i * L^k_q}. \tag{11}$$

The second fitness function considers the residual energy of the node along with the parameters used in the first fitness function. The fitness value is calculated using the following equation:

$$f_{\text{fitness}^k_q} = \frac{E_{\text{res}_i}}{\theta_i * L^k_q}, \tag{12}$$

$$\theta_i = \cos^{-1}\left(\vec{v_1}.\vec{v_2}/\left|\vec{v_1}\right|.\left|\vec{v_2}\right|\right) * (180/\pi), \tag{13}$$

where the residual energy of the sensor node SN_i is given by E_{res_i}. θ_i is the angle formed between the CH, CM and BS calculated using Eq. (13). The length of the route for qth ant and kth iteration is indicated by the L^k_q. After all the ants have arrived at their destination, each ant compares their route to a prior one. The path with the highest level of fitness is considered.

Data Fusion

A simple topology with four nodes is considered to understand the basic workings of the algorithm. Network parameters like throughput, energy, delay, packet delivery ratio, etc. have shown better results on performing data fusion. The output of the same network without performing data fusion is also observed. We also attempt to reduce the number of sensors used for the proposed algorithm. A simple topology is depicted in Fig. 2 to illustrate the fusion process. From the figure, we notice that only three out of four nodes are required to generate the optimized path to reach the destination. The selected three nodes, called as the active nodes will participate during the process of fusion, by minimizing the overall energy consumed by the network. The simulation

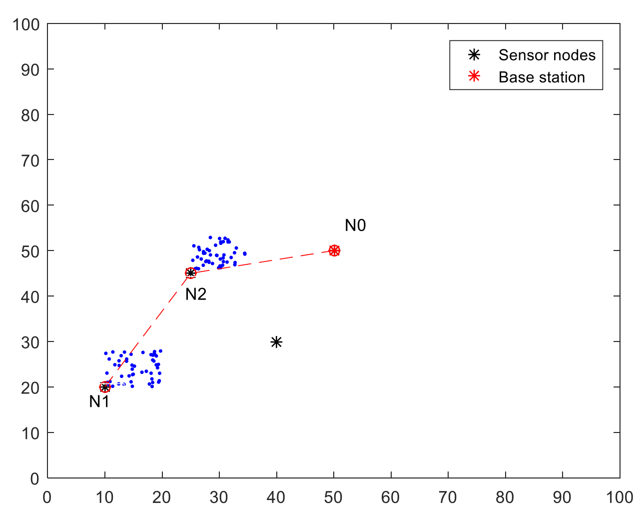


Fig. 2 Data fusion process

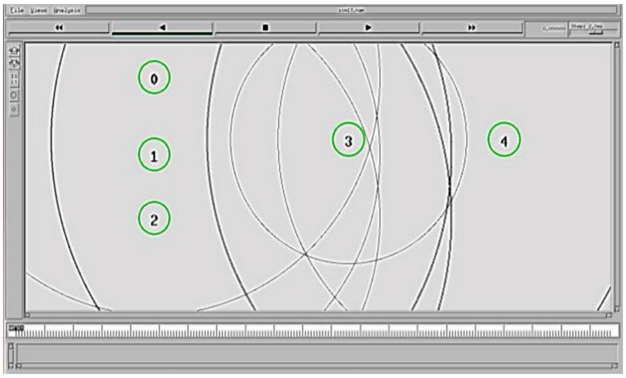


Fig. 3 Arithmetic data fusion

settings of Table 1 are kept intact, as shown in the previous work of the same author [5]. For this scenario, the length of the simulation is considered to be 6.1 s and the dimension of the network is modified to 100 × 100.

One intermediate node N2 is used to transfer data from source node N1 to the sink node N0. We assume that node N1 sends 512 bytes of data packets at the rate of 600 Kbps for three seconds. N2 begins its own transmission for another three seconds after N1 has finished transmitting its data.

To demonstrate the concept of arithmetic data fusion, a topology involving five nodes is examined as shown in Fig. 3. Among these, one node functions as the destination, three nodes serve as transmitters and an intermediate node, termed the fusion node, collects data from the transmitters. This intermediate node performs arithmetic data fusion before forwarding the data to the sink node. The arithmetic fusion entails transmitters operating for 5-s intervals, each transmitting data packets of 512 bytes at a rate of 600 Kbps. The assumption is that transmitting nodes dispatch fixed-size data packets. If we consider the node labeled '0' transmitting data packets for five seconds, given the specified packet attributes, this transmitter could send around 1172 packets per second, summing up to approximately 5860 packets over the 5-s duration, received by the intermediate node.

Equation (14) can be used to find the number of data packets conveyed by N1 in 3 s.

$$\begin{aligned}&\text{Max. Received Packets}\\&= S\text{imulation Time } * \text{ No. of Bytes sent per second.}\end{aligned} \tag{14}$$

Therefore, the maximum packet received by the destination node (N0) is calculated using the following equation:

$$\text{Max.Packet received by the fusion node} = \sum_{i=0}^{n-1} K_i, \tag{15}$$

where Ki is the data packets received from the predecessor nodes.

Path Optimization

Path optimization techniques are required for obtaining the desired goals like reduced energy consumption, increased lifetime of network, security, etc. In this paper, ant colony optimization (ACO) algorithm, developed by Marco Dorigo in 1997, is implemented to route the packets in the WSN meant for various applications [16]. This ant colony system is also used to resolve the travelling salesman problem and quadratic assignment problems [8, 9, 13, 23]. ACO is based on the point that ants are capable of find the shortest route between their nest and the food source [17]. They do this without any visual evidences, but by only exploiting pheromone information [14]. Pheromone is a chemical released by ants which is deposited in the way they move or travel. While moving to food source, pheromones are deposited on the ground by the ants, which in turn help the other ants to construct the path. Pheromone vaporizes with time, so the path having less pheromone is discarded by the follower ants and the path that has more concentration of pheromone is chosen. Finally an optimized route is obtained by the food source. ACO algorithm has the positive feedback system, uses greedy searching technique and is distributed and robust in nature. This makes ACO popular to be widely used in various optimization problems [15]. ACO is used only where predefined source and destination exits. Pheromone value is calculated based on the number of hops travelled from source to destination node by synthetic ant. Pheromone is inversely proportionate to the distance, that is, there are more chances for the artificial ants to move through the edge having less distance [1]. IoT consist of large collection of sensor nodes. Hence, in this paper, we have used the ACO algorithm to selectively choose few sensor nodes to take part in data transmission, to reduce the energy consumption. The nodes will be active only when it has data to transfer or else it will be in sleep mode for improved performance. In our proposed algorithm, we assume that the sensor nodes will operate with different modulations and code rates as required. We also assume that each sensor node must contain burst profiles with different modulation and code frequencies, as shown in Table 2.

The IEEE 802.16 standard offers the flexibility of employing diverse modulations and coding schemes across separate transmissions to achieve varying bitrates. As signal

Table 2 Bit rates and transmission ranges

Modulation	Coding rate	Transmission range (km)	Bit rate (Mbps)
QPSK	1/2	5	2
16-QAM	1/2	3.5	5.5
64-QAM	3/4	2	11

attenuation increases, the utilization of a robust modulation scheme with a lower data rate becomes crucial. Establishing a connection between a subscriber station and a base station can potentially yield enhanced throughput through the use of multiple short hops instead of a single lengthy hop. For instance, if the distance between a source and a destination is 5 km, a QPSK 1/2 (Quadrature Phase Shift Keying) modulation scheme operating at a data rate of approximately 2 Mbps could be employed. Conversely, if another subscriber station is positioned in between, say at a distance of 2.5 km, then employing a 16-QAM 3/4 (Quadrature Amplitude Modulation) scheme could yield around 11 Mbps on each link. Consequently, the overall end-to-end data rate would be approximately 5.5 Mbps. However, the introduction of a larger number of hops toward the base station can lead to increased control and scheduling overhead. The burst profiles of the nodes describe a range of modulation techniques along with their code rates over the link distances, to achieve different bit rates in the wireless medium. This offers higher data rates and a high end-to-end throughput for nodes with smaller distances. We try to optimize the network by evaluating the network parameters like energy (E), throughput (S), packet delivery ratio (PDR), and delay (D). The performance of the network is analyzed by performing data fusion upon them and also without it. The working of traditional ACO algorithm is shown in Fig. 4. In this paper, we are comparing the path generated by MA/MD and MD/MA in [6, 7] with that of the path generated by the ACO. We also compare the network parameters obtained for each of the above algorithms.

Two topologies are taken into account in determining the efficiency of the said algorithms. The initial topology consists of 40 nodes with a base station located at the

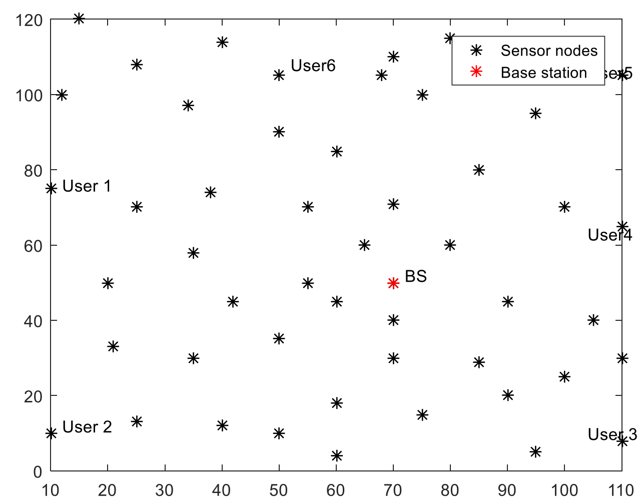
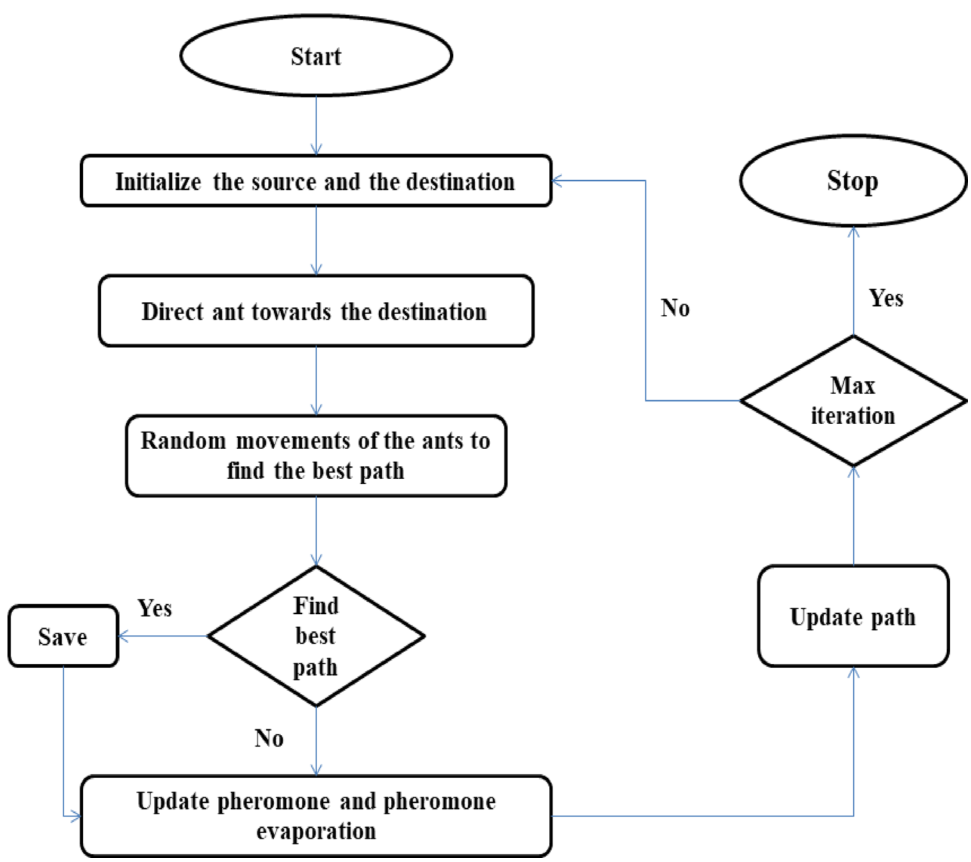


Fig. 5 Mesh configuration

Fig. 4 Working of ACO algorithm



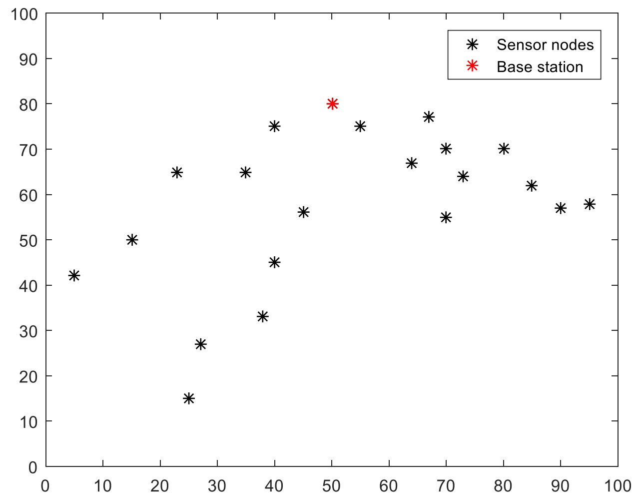


Fig. 6 Tree configuration

middle called 'Mesh Topology' and the second topology consists of 24 nodes with the base station placed at the topmost level, creating a hierarchical architecture called 'Tree Topology.' Both the mesh and the tree topologies are shown in Figs. 5 and 6. The network is configured in its entirety by optimizing each source–destination pair. Routes produced by ants taking into account the different number of ants are shown in Figs. 7, 8, 9, 10, 11 and 12.

The simulation of data fusion on the active nodes for both the mesh and tree topology is shown in Fig. 13a and b, respectively.

After the optimized path is generated by ACO, data fusion is performed on them. The mesh topology of 40 nodes and the tree topology with 24 nodes is optimized to 26 nodes and 15 nodes, respectively. Each of the user transmits their own data by combining the data taken from the earlier node.

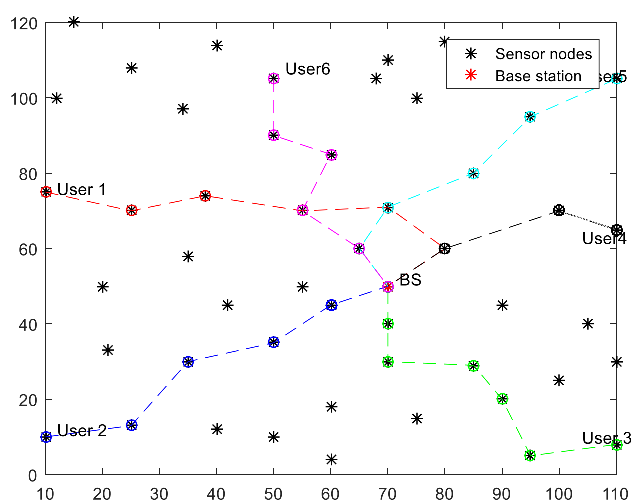


Fig. 7 Path construction using 50 ants

Performance Evaluation

Performance Metrics

The 26 active nodes of mesh and 15 nodes of tree topology are chosen for data fusion and the remaining nodes will be in inactive mode. The sensors which are active is selected to increase the overall routine of the network with regards to energy, throughput and delay. The nodes in our topology are varied only to perform data fusion, while the rest of the network features, such as antenna gain, transmit power,

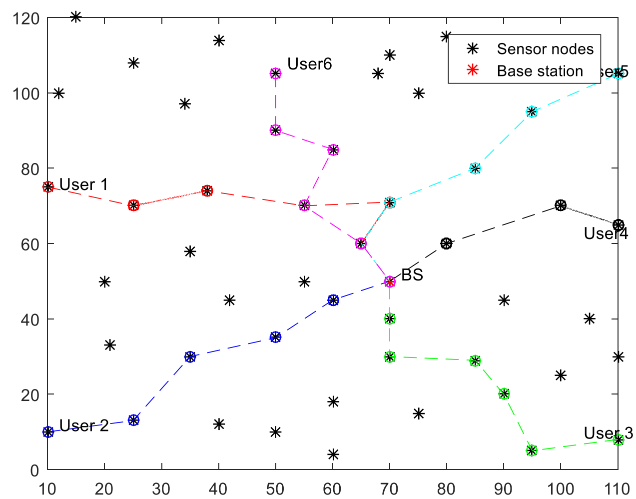


Fig. 8 Path construction using 100 ants

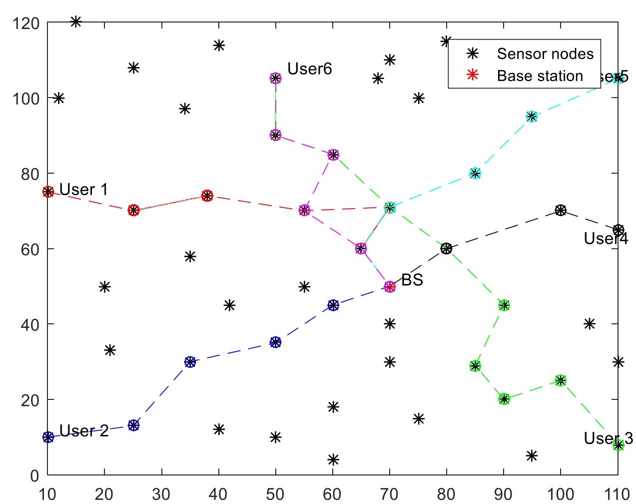


Fig. 9 Path construction using 150 ants

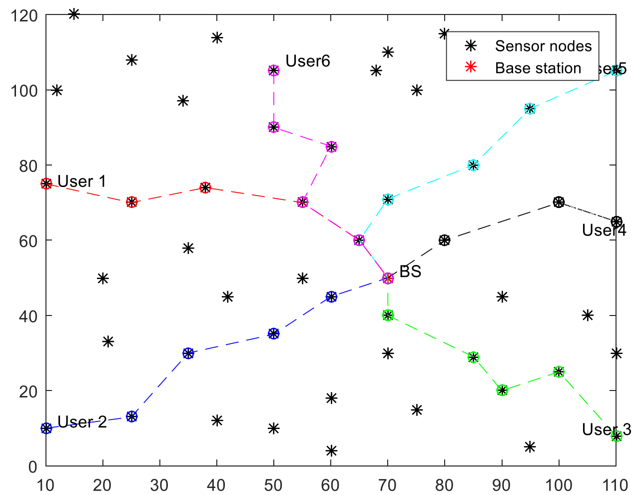


Fig. 10 Path construction using 200 ants

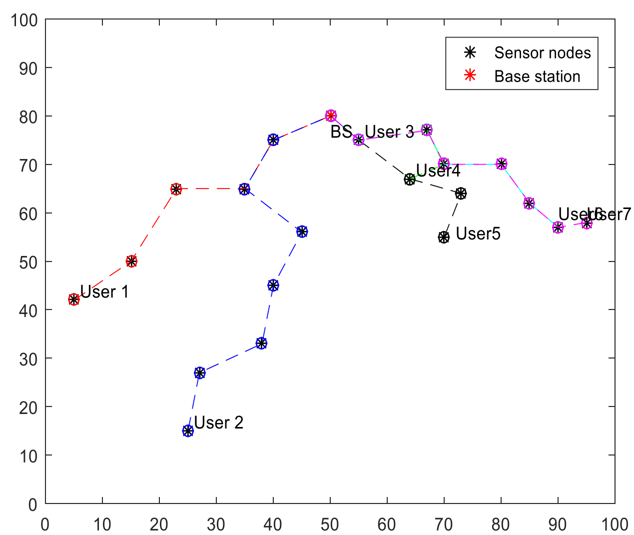


Fig. 11 Tree topology with 50 ants

wireless channel, routing protocol, etc., remain unchanged. The details regarding the number of active nodes is shown in Table 3.

For performance evaluation, we have considered 2 network scenarios which have been optimized by ACO algorithm. The following parameters are considered to evaluate the performance of the network.

1. **Throughput (S)**—It is calculated as the amount of data packet that has been successfully transferred from a single source to a particular destination at a given period of time. It is evaluated using the following equation:

$$\text{Throughput (S)} = (T_{\text{PR}}/1000 * 512)/1024, \quad (16)$$

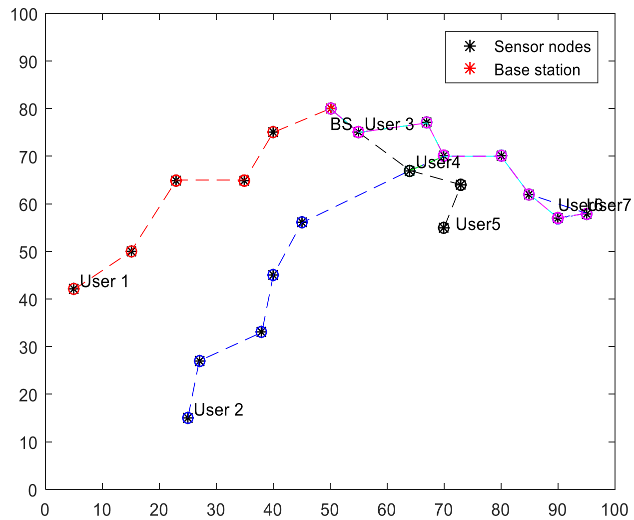


Fig. 12 Tree topology with 20 ants

where T_{PR} is the number of total packets received.

2. **Energy (E_N) –** This metric is used to measure the average energy and the total energy consumed by the network. It is calculated using the following equation:

$$\text{Energy } (E_N) = (\text{N } * 10000) - T_{EC}, \quad (17)$$

where, N is the total number of nodes and T_{EC} is the overall energy consumption by the network.

3. **Delay (D_N) –** This parameter describes the time taken a data packet takes to travel through the network. It is calculated using the following equation:

$$\text{Delay}(D_N) = P_R - P_S, \quad (18)$$

where P_R is the time at which receiver receives the packet and P_S is the time taken to send each packet.

4. **Packet delivery ratio (P_{DR})**—It is the ratio between the number of successful packets received by the receiver to that sent by the sender. It is evaluated using the following equation:

$$\text{Packet Delivery Ratio}(P_{\text{DR}}) = (P_{\text{TR}}/P_{\text{TS}}) * 100, \quad (19)$$

where P_{TR} is the total number of packets received and P_{TS} is the total number of packets sent.

Other than these four parameters, there are two parameters for testing the performance of the proposed mechanism using data analysis approaches. The network scalability and the database scalability of the network are important and are to be discussed for testing its performance. The network has two topology scenarios namely mesh and tree topology. In the experiment, number of nodes in mesh topology and

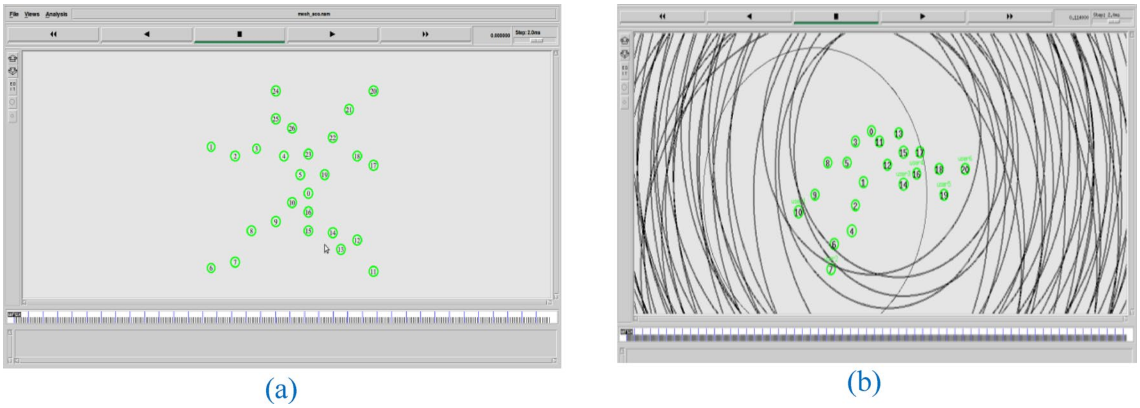


Fig. 13 Arithmetic data fusion on the optimized path generated by ACO for mesh and tree topology

tree topology 40 and 24, respectively. The expansion of the network in terms of nodes is a common requirement for any network. The network scalability Eq. (20) is to be measured to ensure its performance in the expanded form of network.

$$\text{Network scalability} = \left(\frac{N\prime}{N}\right) * \text{performance parameter}, \tag{20}$$

where N is the number of nodes in the existing network and N' is the number of nodes in the expanded network. Performance parameters can be any of parameters namely throughput, energy, delay and PDR of the existing network. Naturally the network scalability factor will be increased according to the increase of the number of nodes in the expanded network.

The database scalability is an important testing parameter for increasing number of nodes. Since each node of the network work according to the route information available in the database. The cluster information, cluster head and other route information are stored in the database. The overall performance of the network also depends on the database scalability factor. When number of nodes increases, the route information increases exponentially.

$$\text{Database scalability} = (V)^d/100, \tag{21}$$

where V is the volume of data in the database of the existing network and d is the scale factor $\frac{N'}{N}$.

Table 3 Active nodes during data fusion

Topology	No. of nodes	No. of active nodes
Mesh topology	40	26
Tree topology	24	15

Results and Discussion

The path achieved using the traditional ACO algorithm is compared with the path generated by the proposed ACO algorithm. The comparison between the traditional ACO, Minimum Distance/Minimum Angle (MD/MA) and Minimum Angle/Minimum Distance (MA/MD) proposed in [5] is also made in this paper to check the efficiency of the route which is obtained by the minimum angle routing. The authors in [5] have reported minimum angle routing to be beneficial when the number of nodes in the network is larger.They have also reported that using minimum angle routing reduces the delay of the network as the subscriber node does not deviate much from the reference line which is being drawn to form the angle between the subscriber node, base station and the intermediate node. In "Time and Space Complexity", the proposed ACO algorithm requires the length of the routing path, the residual energy of the node and the angles formed between the CM, CH and BS to evaluate both the fitness functions. The angle considered between the CM, CH and BS is calculated as proposed in [5]. By simulating both the topologies, we find that in mesh topology the traditional ACO produces better throughput when compared to both the algorithms MD/MA and MA/MD given in [5]. Delay and the energy consumption are also found to be higher in the traditional ACO algorithm than MD/MA and MA/MD. However, PDR is found to be higher in MA/MD than in traditional ACO algorithm. The throughput, energy consumption and packet delivery ratio of the mesh topology on performing data fusion is advanced in the traditional ACO algorithm when compared with MA/MD. After performing data fusion in tree topology, traditional ACO algorithm yields better result when compared to both MA/MD & MD/MA.

The performance of the traditional ACO algorithm is also compared with the proposed ACO algorithm by

Table 4 Network performance

Topology	S (Kbps)		D (ms)		E (J)		PDR (%)	
	Proposed ACO	Traditional ACO	Proposed ACO	Traditional ACO	Proposed ACO	Traditional ACO	Proposed ACO	Traditional ACO
Mesh topology with fusion	2.1253	1.329	0.0169	0.0208	98.305	126	98.05%	81.67
Tree topology with fusion	1.6568	0.9982	0.0236	0.0292	96.3548	123	92.4142%	72.74

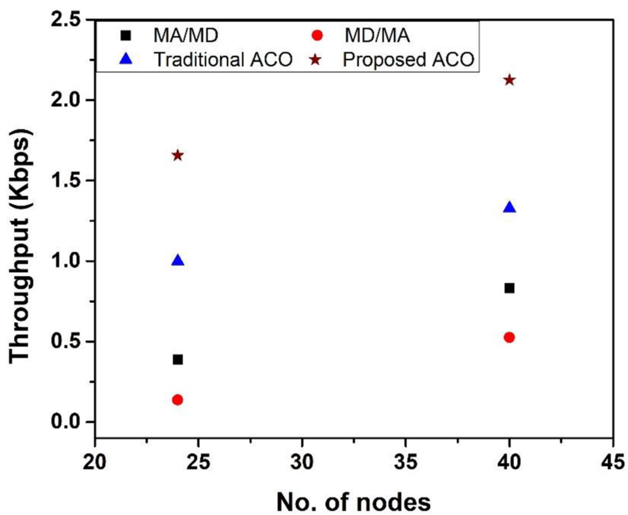


Fig. 14 Representation of throughput (Kbps)

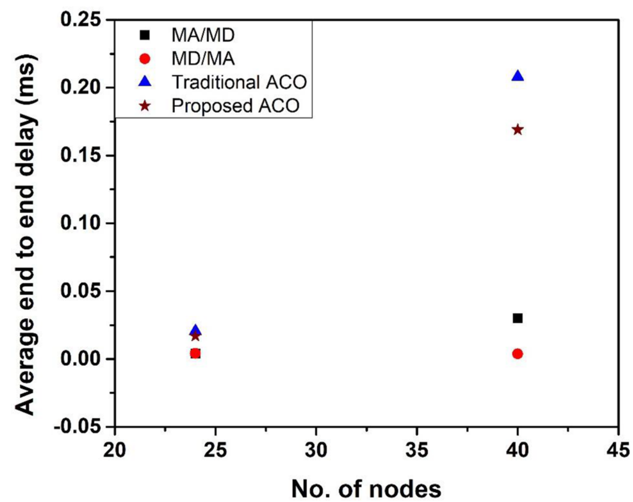


Fig. 15 Representation of delay (ms)

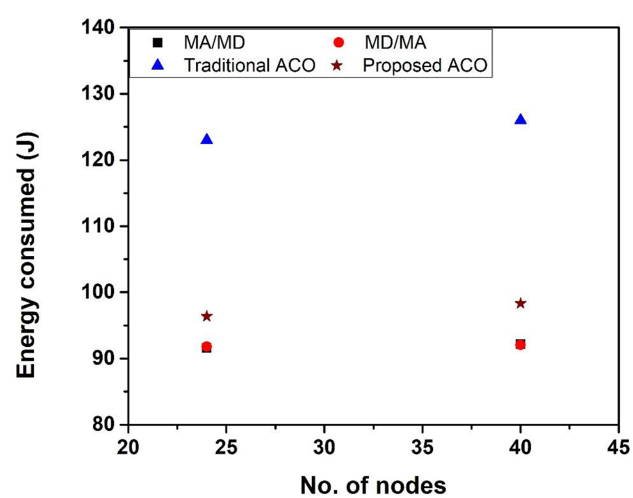


Fig. 16 Representation of energy (J)

incorporating the fitness function. Table 4 shows the performance of the network in terms of throughput, energy, delay and PDR when the fitness function is used. The proposed ACO algorithm shows 20% more packet delivery ratio than the existing ACO algorithm. The energy consumption in case of the proposed ACO has been found to 78% less than the traditional approach. In case of end-to-end delay, the traditional ACO algorithm experiences 19% more delay then the proposed one. The performance of the network in terms of throughput in the proposed scenario is also better than the conventional approach. In the proposed method since the best path is found by taking into account the length of the routing path, the residual energy of the node and the angles formed between the CM, CH and BS it improves the overall performance of the network. The following graphs (Figs. 14, 15, 16 and 17) show the performance of the network for both the traditional and proposed ACO algorithm.

The network scalability of the Eq. (20) for throughput value (2.1253) for the proposed algorithm ranges from 2.66 to 122.88. And the database scalability is calculated

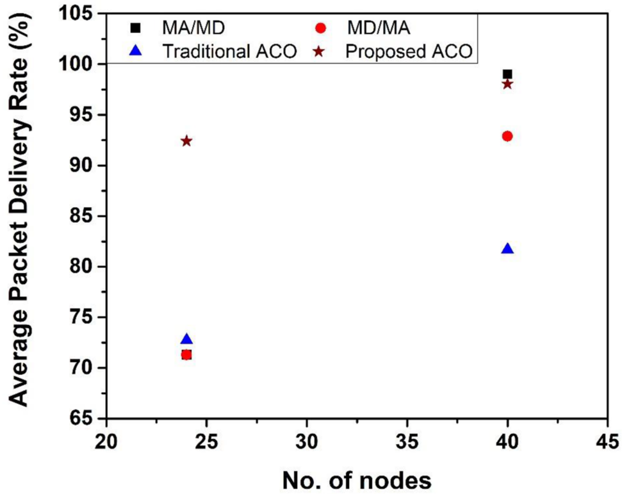


Fig. 17 Representation of PDR (%)

as approximately 3.16 when the volume of the database in the existing network is 100 and number of nodes increased from 40 to 50 nodes.

Conclusion and Future Work

This paper elaborates the working of the ant colony optimization method to perform data fusion and compares the results with the work done by the authors previously. It has been observed that the throughput of the network is higher when packets are transferred from a particular source to certain destination by applying data fusion on them. Data fusion increases the overall throughput of the network. When ACO is applied on the mesh topology the results of parameter is better on performing data fusion. However, we see that energy consumption of ACO with fusion is slightly higher than that of without fusion. This higher energy consumption can be ignored as PDR is higher where data fusion is applied. When the data is transferred from source to destination without fusion the delay is found to be more in ACO. The throughput, delay of the tree topology is found to be better when data fusion is applied. We also notice that data fusion produces slightly higher energy consumption.

Therefore, for optimal application of network resources, wireless sensor networks can employ data fusion. We can also perform data fusion on other evolutionary algorithm and compare their performances by taking into consideration various performance evaluation parameter.

The work may be extended to next generation of IoT (NGIoT) with additional features for faster routing and path optimization may be implemented in future.

Data availability Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Aggarwal R, Mittal A, Kaur R. Various optimization techniques used in Wireless Sensor Networks. Int Res J Eng Technol (IRJET). 2016;3(6):2085–90.
2. Akkaya K, Younis M. A survey on routing protocols for wireless sensor networks. Ad Hoc Netw. 2005;3(3):325–49. https://doi.org/10.1016/j.adhoc.2003.09.010.
3. Bao R, Pan H, Dong Q, Yu L, Shao L. Ant colony-based routing algorithm for wireless sensor networks. Chinese J Sens Actuat. 2011;24(11):1644–8.
4. Bijarbooneh FH, et al. Cloud-assisted data fusion and sensor selection for internet of things. IEEE Internet Things J. 2016;3(3):257–68. https://doi.org/10.1109/JIOT.2015.2502182.
5. Chakraborty I, Chakraborty A, Das P. Sensor selection and data fusion approach for IoT applications. Adv Intell Syst Comput. 2019. https://doi.org/10.1007/978-981-13-1280-9_2.
6. Chakraborty I and Hussain MA. A simple joint routing and scheduling algorithm for a multi-hop wireless network. In: 2012 International Conference on Computer Systems and Industrial Informatics, ICCSII 2012. 2012. https://doi.org/10.1109/ICCSII.2012.6454620.
7. Chakraborty I and Sarmah U. A simple routing algorithm for Multi-hop Wireless Network. In: Proceedings of 2015 IEEE 9th International Conference on Intelligent Systems and Control, ISCO 2015. 2015. https://doi.org/10.1109/ISCO.2015.7282314.
8. Colorni A, et al. Distributed optimization by ant colonies. In: Proc. on European conference on artificial life. Elsevier Publishing; 1991. p. 134–42.
9. Colorni A, Dorigo M, Maniezzo V. An investigation of some properties of an "Ant algorithm." PPSN'92. Belgium: Elsevier Publishing; 1992. p. 509–20.
10. Devi MD, Geetha K, Saranyadevi K. Content based routing using information centric network for IoT. Procedia Comput Sci. 2017;115:707–14. https://doi.org/10.1016/j.procs.2017.09.145.
11. Djukanovic G, Popovic G, Kanellopoulos D. Scaling complexity comparison of an ACO-based routing algorithm used as an IoT network core. J Inf Technol Appl (Banja Luka) - APEIRON. 2020;20(2):73–80. https://doi.org/10.7251/jit2002073dj.
12. Dorigo M, Birattari M, Stutzle T. Ant colony optimization. IEEE Comput Intell Mag. 2006;1(4):28–39.
13. Dorigo M, Maniezzo V, Colorni A. Ant system: optimization by a colony of cooperating agents. IEEE Trans Syst Man Cybern Part B (Cybernetics). 1996;26(1):29–41. https://doi.org/10.1109/3477.484436.
14. Dorigo M, Maria L. Ant colonies for the travelling salesman problem. Biosystems. 1997;43:73–81.
15. Duan P, Al Y. Research on an improved ant colony optimization algorithm and its application. Int J Hybrid Inf Technol. 2016;9(4):223–34. https://doi.org/10.14257/ijhit.2016.9.4.20.
16. Elbeltagi E, Hegazy T, Grierson D. Comparison among five evolutionary-based optimization algorithms. Adv Eng Inform. 2005;19(1):43–53. https://doi.org/10.1016/j.aei.2005.01.004.

17. Gambardella LM, Taillard D, Dorigo M. Ant colonies for the quadratic assignment problem. J Oper Res Soc. 1999;50(2):167–76. https://doi.org/10.1057/palgrave.jors.2600676.
18. Gupta V, Sharma SK, et al. Cluster head selection using modified ACO BT. In: Das KN, et al., editors. Proceedings of Fourth International Conference on soft computing for problem solving. New Delhi: Springer India; 2015. p. 11–20.
19. Heinzelman WB, Chandrakasan AP, Balakrishnan H. An application-specific protocol architecture for wireless microsensor networks. IEEE Trans Wirel Commun. 2002;1(4):660–70. https://doi.org/10.1109/TWC.2002.804190.
20. Hu H. Trust based secure and energy-efficient protocol for wireless sensor networks. In: IEEE access, vol. 10. 2022. p. 10585–96.
21. Khoshkangini R and Zaboli S. Efficient routing protocol via Ant Colony Optimization (ACO) and Breadth First Search (BFS). 2015, p. 374–81. https://doi.org/10.1109/iThings.2014.69.
22. Kooshari A, et al. An optimization method in wireless sensor network routing and IoT with water strider algorithm and ant colony optimization algorithm. Evolut Intell. 2023. https://doi.org/10.1007/s12065-023-00847-x. (**0123456789**).
23. Maniezzo V, Colorni A. The ant system applied to the quadratic assignment problem. IEEE Trans Knowl Data Eng. 1999;11(5):769–78. https://doi.org/10.1109/69.806935.
24. Nayyar A, Singh R. Ant colony optimization (ACO) based routing protocols for wireless sensor networks (WSN): a survey. Int J Adv Comput Sci Appl. 2017;8(2):148–55.
25. Nayyar A, Singh R. IEEMARP—a novel energy efficient multipath routing protocol based on ant Colony optimization (ACO) for dynamic sensor networks. Multimed Tools Appl. 2020. https://doi.org/10.1007/s11042-019-7627-z.
26. Sharmin A, Anwar F and Motakabber SMA. A noble approach of ACO Algorithm for WSN. In: Proceedings of the 2018 7th International Conference on Computer and Communication Engineering, ICCCE 2018, 2018, p. 152–6. https://doi.org/10.1109/ICCCE.2018.8539295.
27. Srivastava A, Mishra PK. A survey on WSN Issues with its heuristics and meta-heuristics solutions, wireless personal communications. Springer; 2021. https://doi.org/10.1007/s11277-021-08659-x.
28. Wang Y, et al. Improved ant colony-based multi-constrained QoS energy-saving routing and throughput optimization in wireless Ad-hoc networks. J China Univ Posts Telecommun. 2014;21(1):43–59. https://doi.org/10.1016/S1005-8885(14)60267-3.
29. Liu X, Li S, Wang M. An ant colony based routing algorithm for wireless sensor network. Int J Future Gener Commun Netw. 2016;9(6):75–86. https://doi.org/10.14257/ijfgcn.2016.9.6.0.
30. Xue X et al. The basic principle and application of ant colony optimization algorithm. In: 2010 International Conference on Artificial Intelligence and Education (ICAIE), 2010, p. 358–60.
31. Yadav RK, Mahapatra RP. Energy aware optimized clustering for hierarchical routing in wireless sensor network. Comput Sci Rev. 2021;41: 100417. https://doi.org/10.1016/j.cosrev.2021.100417.