

# Energy-Centric Route Planning using Machine Learning Algorithm for Data Intensive Secure Multi-Sink Sensor Networks

M.D.Vimalapriya, Vignesh Baalaji S, S.Sandhya

**Abstract:** *Wireless sensor network (WSN) is energy operated self-disciplined ad-hoc technology capable of sensing and actuating environmental phenomenon. The sensed information is shared between communication sinks for access and processing. Energy conservation and route planning are prominent in determining the efficiency of the sensor network. In this paper, energy-centric route planning (ECRP) technique is introduced to address the inequality in sensor node lifetime and routing along with security requirements. ECRP depends on the individual and co-operative energy expenses of the nodes to retain a balanced communication link. The expenses are monitored and a profitable route plan is designed using a machine learning algorithm that assists in identifying a non-deficient neighbor for routing. Both energy optimization and route refurbishing are controlled by the analysis and decisions of the learning algorithm. The Learning process is instigated with the node energy and current route neighbor information for discovering efficient communication paths to the sink. Security is administered using trust process in this path planning and routing for selecting reliable neighbors. This helps to retain security throughout the routing process. The impact of the proposed ECRP over sensor network is verified using the metrics: throughput, active nodes, transmitting energy, routing complexity and delay.*

**Keywords—**Energy Efficiency, Machine Learning, Node Selection, Route Discovery, WSN

## I. INTRODUCTION

Wireless sensor networks (WSNs) are deployed for detecting physical and environmental changes in real-world. It encompasses sensing devices known as nodes in different design and functions for sensing changes. The sophisticated design of sensor nodes enables multiple designs of sensing units besides processing, radio, and power units. The incorporation of recent wireless and communication technologies has projected WSN as an active research area due to their wide-spread applications. The size of a sensor network is variable and hence it is capable of monitoring and detecting events in a distributed large scale region. The detected events are then relayed to a processing center such as base station (BS) or hub-like collectors such as sink. The sink nodes are either fixed infrastructure units or gateway kind of nodes in the network that enables external user and network access. A fixed infrastructure unit possesses exceptional processing, storage, and power features and hence it is not resource constrained [1].

**Revised Manuscript Received on November 05, 2019.**

**M.D.Vimalapriya-** Department of Computer Science and Engineering, Women's Christian College, Chennai, India

**Vignesh Baalaji S-** Department of Computer Science and Engineering, R.M.K. Engineering College, Kavarapettai, Thiruvallur, India

**S.Sandhya-** Department of Computer Science and Engineering, R.M.K. Engineering College, Kavarapettai, Thiruvallur, India

Sensor nodes are physically alienated by distance factor and therefore they communicate through wireless links for information and packet data exchange. Routing and medium access control (MAC) protocols designed for sensor network facilitate route discovery and channel access for communication. Routing protocols designed for sensor network are focused to improve the outcome of the network by suppressing the communication issues originated due to resource limitations [2].

A routing protocol is a fundamental requirement of a sensor network for discovering transmission routes to the sink node. Due to the location and mobility of the nodes, the sink node is present in either single-hop or multi-hop distance from the transmitting source. Routing protocols are designed to alleviate the delays by identifying precise routes for data forwarding and more specifically achieving energy efficiency in WSN. The concept of energy potent routing inherits the advantages of both route discovery and energy conservation methods for generating reliable routes. It is adaptive to the varying network density and traffic patterns and hence it is suitable for data-intensive sensor applications [3]. Power conservation, routing decision and neighbor adaption are some of the functions of routing protocols to accustom scalability and flexibility in the network. Other than energy efficiency, specific routing metrics concentrate in delay, hop-count and network lifetime maximization. Considering these metrics helps to improve the design features of routing protocol and network outcome [4].

Machine learning has become a prominent solution approach to solving a complex problem and decision making. The robustness of machine learning algorithms is incorporated in WSN for analyzing and mitigating complex decision making in routing and energy management. Learning based intelligent solutions observe the environmental changes for analyzing constraints and classify the inputs accordingly to generate optimal results [5]. Machine learning process resolves complexity in routing by exploiting the intrinsic node attributes of environmental and event monitoring features. The decisions of the sensor nodes are dependent on the state of observation and decision making resulting in intelligent actions. The process of learning and decision making is extended in an iterative manner until a final optimal solution is achieved [6, 7]. Security is a promising concern in autonomous networks such as WSN. The autonomous nature of the nodes and wireless communication channel result in injection of multiple threats in the network. Injection of malicious nodes in the network degrades its performance and response to the end-users. Attackers/ adversaries cause resource depletion, data falsification, path failures, data loss, etc [8].

# Energy-Centric Route Planning using Machine Learning Algorithm for Data Intensive Secure Multi-Sink Sensor Networks

Therefore, it is necessary to mitigate them from the communication network in the absence of central administrator. Trust based security frameworks rely on the behavior of the nodes to ensure legitimate transmission. The participation of the nodes in a communication process is decided by the trust analysis of the node depending on the packet handling rate, delay, packet loss, etc [9, 10].

## II. RELATED WORKS

Huang et al. [11] proposed energy-aware dual-path geographic routing (EDGR) for addressing routing hole issues in WSNs. EDGR exploits remaining battery power, energy expenses and geographic position information of the nodes to perform routing decisions. It is dependent on the disjoint paths surpassing route holes for re-routing and hence is capable of achieving better packet delivery, network lifespan, and less delay. Routing decisions are instantaneous and lack continuous observations due to which independent overloading of nodes is high.

To improve transmission efficiency of the sensor network, Lai et al. [12] introduced a link balanced energy routing. This routing is designed based on predicted routing deliveries (PRD) metric for reducing the impact of resource constraints over node transmission. A leveraged network outcome is achieved by selecting weighted links as estimated using remaining battery power, delay, distance and link feature metrics. PRD based routing is proficient in improving the network's packet delivery and reducing energy expenses and delay.

Traffic shaper network coding-aware routing (TSCAR) is designed by Shao et al. [13] as a view of improving energy efficiency and routing in WSNs. TSCAR resolves the issues in decoding and traffic scheduling by observing the traffic flows in the network. It initiates network coding solutions for route discovery and neighbor adaption on the basis of transmission flows in the network. The solutions of the network coding are evaluated using the load and energy defined conditions for improving the chances of route availability in WSNs. Though, TSCAR achieves better throughput, delay, and network lifetime, the complexity in sensor node selection and re-routing increases routing complexity.

Qiao and Zhang [14] presented random projection-polar coordinate-chain routing (RPC) for improving the packet transmitting efficiency of the sensor network by alleviating data traffic. RPC is focused to locate nodes establish chained routing paths and random project-based data gathering for reliable transmission in WSN. Transmitting and data gathering nodes are located with the reference to sink node orientation. Chain routing is preceded by a greedy algorithm for reducing energy utilization and overhead in communication.

Distributed self-healing approach (DSHA) is projected by Elsayed et al. [15] for detecting and restoring communication faults in WSN. DSHA aids both cluster and node based fault detection and restoration. Sensor equipment defects, route errors, transmission, and congestion estimation faults are identified and are granted with reliable solutions in WSN. This augments to the self-disciplined characteristics of the nodes and its performance by achieving better fault tolerance and recovery.

Wang et al. [16] proposed an improvement of low energy adaptive clustering hierarchy (LEACH) protocol by

addressing clustering and single path restrictions. Conventional LEACH protocol is modified with the opportunistic routing model in the account of supporting even clustering and multipath routing. Improved LEACH supports WSN by selecting route nodes based on the hop, resource utilization and parallel transmission ability for improving network outcome. Henceforth, the cluster head and clusters are determined using energy and multipath ability of the nodes for achieving better energy efficiency and data transmission rate than conventional LEACH.

Forwarding area division and adaptive forwarding area selection (FADS) is introduced by Hong et al. [17] for leveraging the communication outcome of wireless sensor networks. FADS is designed to reduce inter and intra slot collisions and aids dynamic load handling of the sensor nodes. It segregates the communication area based on allocated time slots and hence the allocation of traffic loads to each sub-area. The quality of FADS is assessed using: delay, energy usage, and delivery ratio.

Huang et al. [18] proposed the annulus sector grid aided routing protocol (ASGRP) for improving the active time period and energy potency of WSNs. In an annulus grid clustering model, the network is classified into different clusters using arithmetic progression. From the base station perspective, the nodes in the clusters are partitioned to adopt a varying level of segments that are equidistant from each other. Cluster heads employ level based multi-hop routing with communication management feature for improving network life span and energy potency.

Darabkh et al. [19] projected energy-aware clustering and routing protocol (EA-CRP) for controlling transmit energy outlay of the nodes and to optimize data transmission in WSN. EA-CRP differentiates the base station, clusters and sensor nodes using a multi-layer architecture for reducing the communication distance. This helps to choose precise neighbors suitable for forming clusters with less routing overhead. More precisely this clustering approach balances the tradeoff between energy and distance to achieve better scalability and energy efficiency.

Attiah et al. [20] introduced an evolutionary stable strategy (ESS) for providing energy efficient routing in WSNs. ESS is modeled using replicator dynamics to gain knowledge of the energy requirement of the nodes from the previous state interactions. Route node adaption and transmission is decided on the basis of the dynamic strategy plans aided by evaluating the fairness index of the nodes. This game-theoretic approach achieves better fitness in selecting neighbors.

He et al. [21] exploited simultaneous wireless information and power transfer (SWIPT) in multi-hop energy-constrained wireless network (MECWN). SWIPT in MECWN employs an iterative energy allocation (IEA) algorithm for forwarder node-link selection. A specific energy cost model estimates the energy requirement of the links in both information transmission and SWIPT modes for achieving. Defined by the energy cost metric, energy-aware SWIPT routing (ESWIPT) is designed for leveraging the energy effectiveness and transmission capacity of the network.

Sun and D. Li [22] introduced trust aware routing protocol with multi-attributes for improving the security in WSN. This trust method achieves better energy efficiency along with

communication and data efficiency by mitigating the impact of malicious nodes. Specifically, routing attacks launched across the source and sink nodes is evaded using this trust model.

Leach Trust (LEACH-T) is designed by Yang et al. [23] for detecting and mitigating attacks in WSN through continuous monitoring. This clustering method is designed as a game theoretic approach for reducing the tradeoff between energy efficiency and security. Trust assessment of the nodes is carried out through continuous monitoring and cluster based routing.

### Energy-Centric Route Planning (ECRP)

ECRP is designed to leverage the conduct of WSN equating the resource constraint behavior of the sensor nodes. The fundamental operations such as routing and transmission are optimized by balancing energy expenses and link endurance in the network. The network representations and its energy requirements are described in the following sub-sections

### Network Representation

Considering a set of  $N$  sensor autonomous node that are placed in a distributed network. The nodes are separated by a physical distance ' $d$ ' through wireless links  $E$ . The nodes are capable of transmitting and receiving packet data and are hence regarded to from a graph  $G = (N, E)$ . The transmission range  $R$  of a node is fixed. The network consists of multiple sink nodes that serve as ad-hoc destination. Sensor nodes possess random mobility and it is assumed to have less impact over network outcome. A schematic WSN representation is presented in Figure 1.

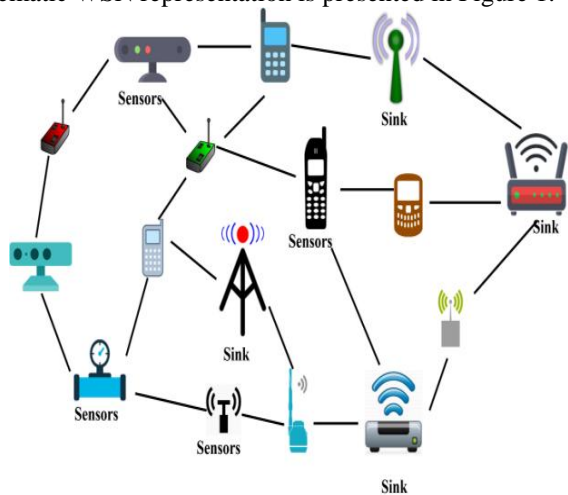


Figure 1 Schematic WSN Representation

The autonomous nodes operations are powered by self-equipped battery sources. The initial backup energy ( $E_b$ ) is spent by the node for transmitting and receiving packets. The node segregates energy for transmitting ( $E_t$ ) and ( $E_r$ ) the energy usage ( $E_u$ ) of a node is

$$E_u = E_t + E_r \tag{1}$$

Where

$$\left. \begin{aligned} E_t &= d_t \cdot p_t \cdot t_t \\ E_r &= d_r \cdot p_r \cdot t_r \end{aligned} \right\} \tag{2}$$

Here,  $d_t$  and  $d_r$  are the packet data transmitted and received at time  $t_t$  and  $t_r$  respectively. The radio unit of the sensor

node utilizes  $p_t$  and  $p_r$  power for transmitting and receiving  $d_t$  and  $d_r$ .

Sensor node's energy level decreases as it performs network operations and hence its lasting energy ( $E_l$ ) is estimated as

$$\left. \begin{aligned} E_l &= E_b - E_u \\ \text{and} \\ N_{ls} &= \frac{E_l}{E_b} \end{aligned} \right\} \tag{3}$$

Where,  $N_{ls}$  is the node's life span.

### III. METHODOLOGY

In this route planning, multi-path routing is employed for discovering more than one route to the sink. The routes are correlated with the bi-directed graph  $G$  such that, the routes are the connections between the root and leaf nodes. The path nodes represent the levels  $G$ , and the level is confined to  $R$ . Data packet transmission instigates with the knowledge of  $E_l$  and  $N_{ls}$  of the nodes in a level. If two or more sensor nodes share a common level, transmission is executed by considering  $d$ . It is obvious that a node with shortest distance to the neighbor is opted for transmission. Let the data transmission be segregated for  $r$  rounds. The energy deficient nodes at each level is identified by evaluating equation (3) and are rotated to refurbish the performed, from the intermediate level nodes other than initiating from the root. From the available set of routes passing through each level a balanced energy concentrated path is selected. The adjustment of routing and transmission relies on the node lifetime and path refurbishing time. There are three stages of routing and route selection in ECRP namely:

- (i) Route Construction
- (ii) Pruning and
- (iii) Refurbishing Transmission Routes

The above stages are classified to co-operatively retain balanced and energy efficient transmission. The stages are explained as follows

#### Route Construction

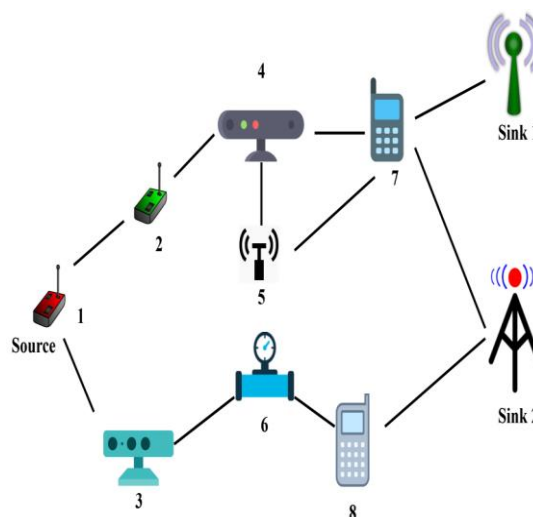
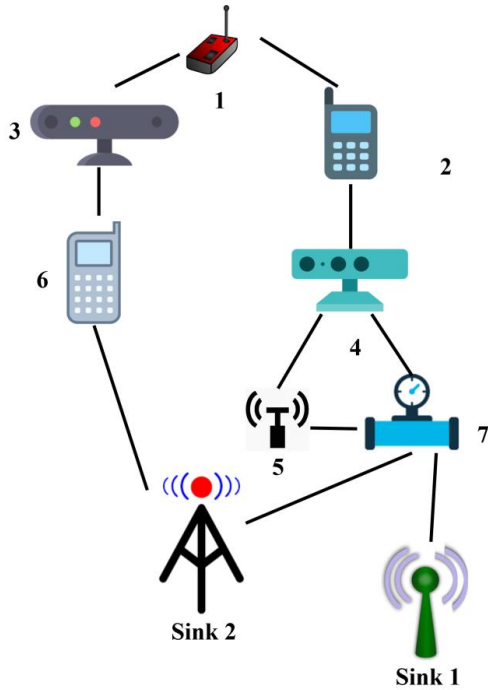


Figure 2(a) WSN with Multiple-Sinks

# Energy-Centric Route Planning using Machine Learning Algorithm for Data Intensive Secure Multi-Sink Sensor Networks

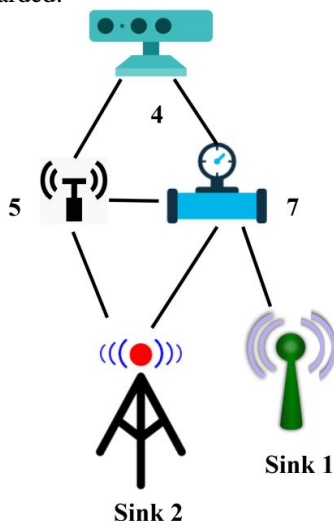


**Figure 2(b) Available Routes in Multi-Sink WSN**

Post the conventional route discovery processes, the routes from the root to the sink are established. Consider a source communicating with more than one sink in a single round (figure 2(a)), the level of the routing path is same as the number of hops between source and sink. The routes constructed consist of both linear (1-3-6-8-sink 2) Figure 2 (a) and disjoint nodes (1-2-4-7-sink1 and 1-2-4-5-7- sink 1) figure 2(b) in the route. Disjoint node causes replication and excess energy expenses at the time of transmission and level possessing a same successor (in the next level) is refined to form a single route. A tree based structure illustrated in Figure 2b shows the presence of two nodes (5 and 7) in the same level. There are two cases of forming a route in this condition:

### a) Either of the nodes alone connects the sink

In this case, the connecting node to the sink forms the route from the root node. From figure 2(b), as node 7 connects atleast two sink nodes, it is opted for packet transmission. The node 5 is considered as an intermediate leaf whose path is initially discarded.



**Figure 2(c) Disjoint Routes Illustration**

### b) Nodes present at the same level connecting the sink.

In this scenario the communication time ( $c_t$ ) between level nodes is estimated. In figure 2(b) say of node 5 and 7 connects to the sink as in Figure 2(c)

Now, the communication between 5 to sink 2 and 7 to sink 2 is computed as

$$c_t = \left[ \frac{2^b - 1}{R} \times d \right] \quad (4)$$

Here,  $b$  is the buffer capacity of the level nodes that performs transmission. The node that grants  $\max\{c_t\}$  is then opted for concluding the routing path. The communication time relies on equation (3) metrics.

The above-discussed cases are valid for the first round or until the node encounters deficiency in energy.

### Pruning

In the pruning stage, feasible transmission is achieved by discarding deficient level nodes and to retain the existing communicating links. Communication links between the nodes are balanced in accordance with energy expenses and transmission rate. Discarding deficient nodes from the transmission routes causes a change in packet delay and also increases routing complexity. Therefore to analyze the energy efficiency and consented transmission, regression machine learning is introduced. The path nodes that are classified under condition (a) and (b) as discussed in the route construction process is considered for regressive learning. This learning process assists route pruning and refurbishing to retain and to enhance the operations of the sensor network. The learning process is modeled for analyzing the transmitting condition (i.e.) {yes or no}. The transmission state of a node depends on  $N_{Is}$  is modeled as energy outcome and value of  $b$  relies on  $(d_t/t_t)$  of the ancestor node.

Let  $\rho_h$  and  $\rho_l$  represent the chances of a node having high and low  $N_{Is}$ . Similarly, the probability of accepting and discarding a packet be  $\rho_{b-a}$  and  $\rho_{b-d}$  respectively. The maximum lifespan and minimum lifespan of a node is estimated by verifying its  $E_u$  in the successive rounds. For assessing the chances of occurrence of either of the above for energy and buffer is computed using an index  $\Delta i_e$  and  $\Delta i_b$ .

$$\left. \begin{aligned} \Delta i_e &= 1 - \left( \frac{\rho_h}{\rho_h + \rho_l} \right)^2 - \left( \frac{\rho_l}{\rho_h + \rho_l} \right)^2 \\ \Delta i_b &= 1 - \left( \frac{\rho_{b-a}}{\rho_{b-a} + \rho_{b-d}} \right)^2 - \left( \frac{\rho_{b-d}}{\rho_{b-a} + \rho_{b-d}} \right)^2 \end{aligned} \right\} \quad (5)$$

Post the estimation of this index, the weight of a level node  $\omega_n, n \in N$  is computed using equation (6)

$$\omega_n = \frac{\rho_h + \rho_{b-a}}{\rho_h + \rho_l + \rho_{b-a} + \rho_{b-d}} (\Delta i_e \oplus \Delta i_b) \quad (6)$$

The weight estimation as in equation (6) is used for making a decision based on energy and transmission. Unlike conventional weight methods, preference weights are not added in equation (6). This means that the routes neighbors are equally given priority on lifetime and transmission basis. Figure 3 illustrates the learning decision process

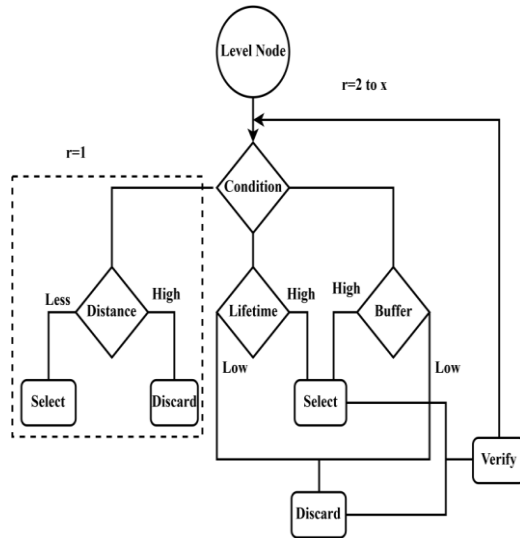


Figure 3 Learning Decision Making

In this decision making process, the conditions based on distance, lifetime and buffer are estimated. For a reliable solution, distance is less, lifetime and buffer space is high. This decision consists of two outputs differentiating distance entity and lifetime, buffer entity. If the distance condition is not satisfied by the node, then it is discarded. If distance condition is satisfied, then both the lifetime and buffer conditions are verified. Now, the pruning process is instigated. In the first transmission, the distance factor is considered for selecting a neighbor. This is the same as in the route construction process. The pruning process is instigated for disjoint and same sink connecting routes. It continues for all the transmission rounds  $r$ . In Figure 3, the regression process occurs from  $r = 2$  to  $r = x$ . In the case of  $r = 1$ , the nodes possess high energy and hence distance factor is deliberated for selecting a transmission node. Level nodes failing the condition in equation (7) are pruned.

$$\left. \begin{aligned} \frac{\rho_h}{(\rho_h + \rho_l + \rho_{b-a} + \rho_{b-d})} + \Delta i_e > \omega_n + \frac{\rho_l}{(\rho_h + \rho_l + \rho_{b-a} + \rho_{b-d})} \\ \frac{\rho_{b-a}}{(\rho_h + \rho_l + \rho_{b-a} + \rho_{b-d})} + \Delta i_b > \omega_n + \frac{\rho_{b-d}}{(\rho_h + \rho_l + \rho_{b-a} + \rho_{b-d})} \end{aligned} \right\} \quad (7)$$

The level nodes that do not satisfy equation (7) are discarded from the routes considering them as intermediate leaf node. Therefore, the depth of the sub-tree (at a specific level) is decreased from  $\Delta_n * \sum_{i=1}^{n \in level} d_i + 2\lambda$  to  $\sum_{i=1}^{n-m} d_i + 2\lambda$ . Here,  $m$  is the number of nodes satisfying equation (7) and  $\lambda$  is the maximum communication pairs observed in the level.  $\Delta_n$  is the probability of disjoint  $\Delta_n = 1$  if disjoint  $\Delta_n = 0$ , if not disjoint. Therefore, the routes are discarded if they intercept  $\lambda$  in any level due to mobility or routing protocol. Here, route plane is retained to satisfy both energy and transmission constraints.

### Refurbishing Transmission Routes

Route failures are common in WSN as the nodes rely on the power source and  $E_b$  for communication purposes, Replacement or recharging drained power sources is less feasible due to deployment in a hostile and non-hostile environment. Therefore an error in routing path occurs due to early energy exhaustion and hence link failures. This impacts the performance of the communicating nodes resulting in network degradation. To address this issue, the routing protocols re-discover routes to the destination by opting new neighbor. Neighbor replacement and route

formulation is modeled using energy and distance metrics. In this proposal, refurbishing transmission routes is aided by machine learning decisions. The learning process is operated in a step-by step manner to ensure path availability. For this purpose, the least preferable cases of the learning process is considered. In the learning process as in Figure 3, The “select” action is performed by considering a node satisfying “high” in terms of life time and buffer. The other possible combinations of the above conditions are tabulated in Table 1.

Table 1 Possible Combination of Lifetime and Buffer

	Lifetime	Buffer	Solution
Case 1	High	High	Select
Case 2	High	Low	Discard
Case 3	Low	High	Select
Case 4	Low	Low	Discard

In the above table, the least feasible case of lifetime and buffer are not considered for selecting a level node. Rather the other two sub-optimal cases cannot be discarded completely. Instead, the more near-to- reliable condition satisfying level node is opted for pursuing transmission and refurbishing transmission path.

The nodes are assessed from the pre-pruning process to verify the availability of transmission paths. Preference is given to the sink (nodes in the same level). In Figure 2(c), the link 4-5-sink 2 is considered in the refurbishing process other than 4-7-sink 2. Here, the level node is rotated to ensure un-interrupted transmission using the new neighbor. The level node rotation occurs in any of the  $r$  rounds from 1 to  $x$ . Therefore the rotation in  $(x - r)^{th}$  or  $(x - r + 1)^{th}$  round with lag. The rotation  $\gamma$  of the node  $n \in N$  in  $(x - r)$  is expected as

$$\gamma_{n(x-r)} = (x - r)|n|, n \in level \text{ and } \Delta_n = 1 \quad (8)$$

This is feasible only if two or more nodes in the transmitting path are disjoint at the same point/at any level. The ancestor node now skips the current neighbor entry from its routing table to opt a new neighbor such that

$$n_{new} = \begin{cases} \text{leaf } n, \text{ if } \gamma_{n(x-r)} + \Delta h_{(x-r)} \Delta t_{(x-r)} \leq \gamma_{n(x-r)} < \Delta h_{(x-r)} + \Delta t_{(x-r)} \text{ and} \\ \gamma_{n(x-r)} + \Delta t_{(x-r)} < \Delta h_{(x-r-1)} \\ n, \text{ otherwise} \end{cases} \quad (9)$$

Replacing a level node is not accomplished until another decision is imposed on case 2 and case 3. A new decision is imposed to ensure maximum likely factors favoring transmission. In this case, the latency between the level nodes is assessed. Unlike the other re-route discovering procedures, in ECRP, the latency between predecessor and a level node (different levels) is alone verified as it requires more routing messages to verify delay until the destination sink. Time latency experienced by the next level node  $t_{lat}$  is computed using equation (10)

$$t_{lat} = \left( \frac{2^a - 1}{R} * d^* \right) \quad (10)$$

Where  $d^*$  is the distance between  $n_{new}$  and next the level node.

# Energy-Centric Route Planning using Machine Learning Algorithm for Data Intensive Secure Multi-Sink Sensor Networks

Henceforth, the decision is forwarded with two stages of estimation and analysis (Figure 4)

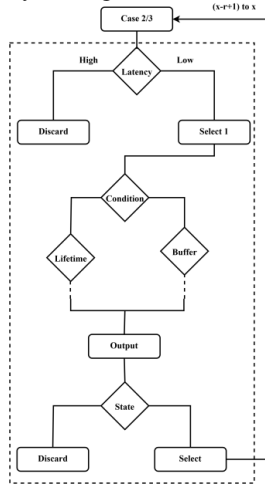


Figure 4 Two-Stage Learning Decision

The analysis is recurrent for  $r = 1$  to  $x$ . The decision splitting process/ stages in learning is instigated if a level node fails to satisfy equation (7) for the metrics estimated in equation (3) and (10).

## A. Contrary Analysis

If two nodes are said to be present at the same level the distance from either of the next level node is the same. As the nodes possess mobility and the locations of nodes are not the same, there is an inequality in delay estimation. Therefore, delay estimated in equation (10) satisfies an inequality

$$\sum_{j=1}^h \phi_n^j > N \quad (11)$$

Where  $h$  the hop is count and  $\phi_n^j$  is the degree of the nodes in the level. Therefore the distance  $d$  and  $d^*$  varies with a small error.

## B. Security Analysis

In the security analysis subsection, selection of legitimate nodes in the energy-concentrated routing process is described. This process is achieved after the pruning process as described in equation (7). Refurbishing routes is aided by trust dependent neighbor selection other than path availability. The trust ( $d_t$ ) of a node is compute on the basis of packet it handles in  $(x-r)$  to  $r$  rounds. Therefore, the trust of a node in the previous and current state is estimated as

$$d_t = \frac{\sum_{i=1}^{x-r} \frac{p_{s_i}}{p_{r_i}}}{(x-r)} + \frac{\sum_{i=x-r}^x \frac{p_{s_i}}{p_{r_i}}}{x} \quad (12)$$

Where,  $p_{s_i}$  and  $p_{r_i}$  is the packet sent and received at a round  $i$  provided  $i \in (x-r)$  or  $i \in x$  but  $(x-r) \neq x$ . The above  $d_t$ , is estimated for all  $\phi_n^j$  to identify the neighbor with maximum  $d_t$ . The refurbishing process instigated if the network throughput decreases and routing complexity increases. Therefore, the  $n_{new}$  selected in each level must satisfy  $\max\{d_t\}$  to ensure optimal transmission. If an eaves dropper node/adversary communicates in the wireless channel, the rate of data transfer decreases. In this process, security is partially concentrated and is integrated with the route discovery and neighbor selection process. The

advantages of this security in ECRP are its less computation and swift detection. This is because,  $\max\{d_t\}$  is assessed for  $n_n$  in each level at the time of refurbishing routes. Besides, the leaf nodes are not induced in the trust verification process.

## C. Performance Analysis

ECRP is evaluated using network simulator experiments for verifying its consistency in WSN. A network consisting of 100 sensor nodes with a transmission radius of 100m is deployed in a network with 2500m<sup>2</sup> dimension. Table 2 presents the detailed simulation configuration and its values.

Table 2 Simulation Configuration and Values

Simulation Configuration	Value
Network Dimension	500m x 500m
Sensor Nodes	100
Transmission Range	100m
Data Packet Size	124bytes
Buffer Capacity	80packets
Backup Energy	5J
$p_t$	50nJ
Transmission Rounds	300-2000

To assess the performance of ECRP, a comparative analysis of the results are made with the existing EA-CRP [19], TSCAR [13] and ESWIPTR [21] methods. For comparison, the metrics accounted are throughput, end-to-end latency, average energy utilization, number of alive nodes and routing complexity.

## D. Throughput Analysis

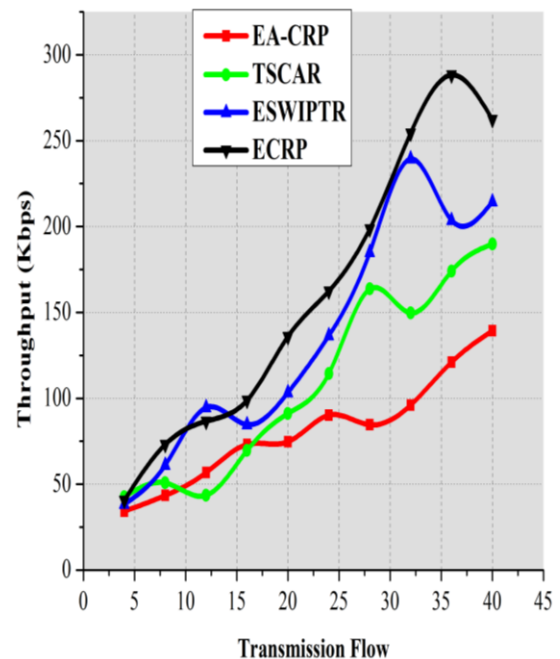


Figure 5 Throughput Comparisons

A comparative analysis of throughput between the existing methods and the discussed ECRP is presented in Figure 5. Rate of transmission flow increases the packet data transmission and hence the congestion factor. Route discovery in ECRP is aided by the multi-path routing protocol and then is refined on the basis of node lifetime and transmission acceptance ability. In  $(x - r)$  transmission round, the level node is rotated if it satisfies two optimistic metrics among delay, lifetime and transmission acceptance capacity. Hence, the availability of path post energy failure is ensured by opting nodes that satisfies maximum conditions and does not converge to equation (7). Therefore, the transmission is aided by case 1 i.e. both node lifetime and packet acceptance ability is "high" on the basis of two conditions. In the successive transmission rounds, the path is further refined by learning from previous communications and augmenting delay condition. The inequality in distance after selecting a new level node in  $(x - r)$  also eases precise packet data transmission.

### E. End-to-End Latency Analysis

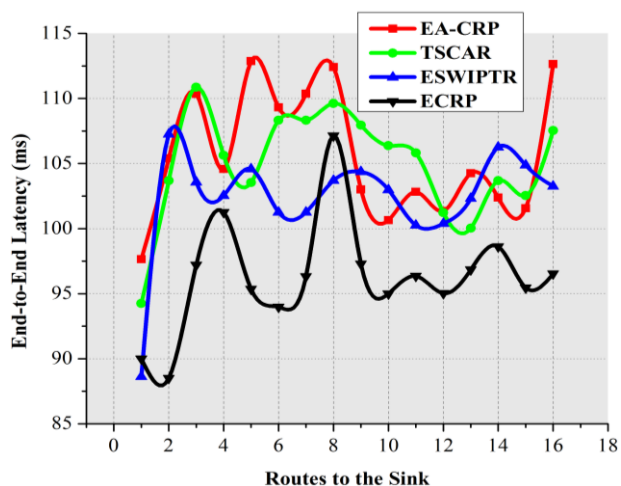


Figure 6 End-to-End Latency Comparisons

Figure 6 portrays the latency observed in ECRP compared with the existing methods. As the number of routes to the sink node increases, latency varies due to varying hop-count and distance between the nodes. More specifically, if a node rotation is pursued in a level due to energy deficiency, latency is observed. In ECRP, the routes are refurbished either in  $(x - r)$  or the successive transmission round,  $(x - r + 1)$  wherein the delay is high (Routes 4, 8 and 14, in Figure 6). Contrarily, the overall delay in reaching the destination in the account of  $\theta_n^l$  (as in equation (11)) is less in ECRP as the refined nodes (based on case 1) and exempted intermediate leaf nodes (in route the construction process) are accounted for refurbishing new routes to the sink node. Level nodes with maximum  $\omega_n$  and that passes equation (7) are directly assessed for latency (equation (10)) to form the refurbished route to the sink. Therefore, the conventional process of re-broadcast and control message exchange time are confined in ECRP. This helps to lessen end-to-end latency.

### F. Average Energy Utilization Analysis

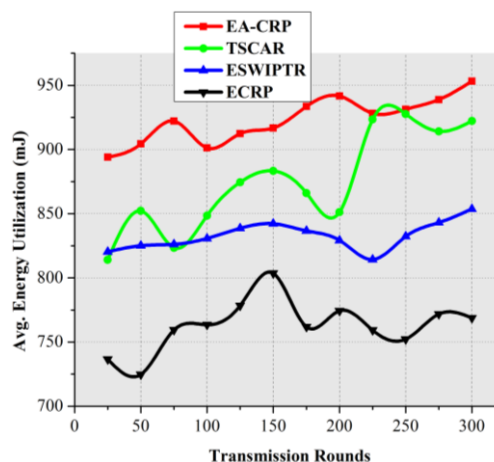


Figure 7 Average Energy Utilization Comparisons

In energy efficient routing process, transmissions of the energy detained node are processed using a second shortest distance and high remaining energy neighbor. Packet data acceptance capacity of a node relies on its available energy and node its lifetime. ECRP exploits this feature in selecting a level node. Similarly, the route node that is deficient of energy is rotated and hence the path remains the same. This reduces the energy expenses in generating additional control messages. Similarly, the node with high packet handling capacity is selected for receiving and forwarding packets. Therefore, redundant node overloading is prevented that reduces the energy expenses of the inter-level neighboring node. These additive features help to reduce the energy expenses of the nodes and hence retains the utilization level of the nodes. Learning assisted evaluation improves the chances of selecting a proficient level node satisfying both energy and delay constraints. This process is iterative for all the  $r$  rounds of transmission. Besides, the discarding process avoids intermediate leaf nodes and energy deficient nodes to participate in the transmission process for which the energy utilized is less (Figure 7).

### G. Alive Nodes Analysis

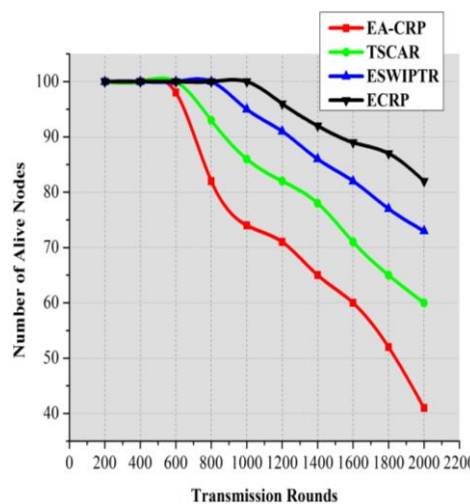


Figure 8 Live Nodes Comparisons

# Energy-Centric Route Planning using Machine Learning Algorithm for Data Intensive Secure Multi-Sink Sensor Networks

As discussed before, energy expenses in ECRP are controlled that helps to retain the residual energy of the nodes at a high level. If the remaining node energy is high, the lifetime of the node is retained at an appreciable level. The active node is capable of surviving better rounds of transmission before and after  $(x - r)$ . In both the cases, the number of active nodes is retained by pruning intermediate leaf nodes and rotating weight-deficient nodes. The retained nodes are suspended from communication and hence the transmission imposed state of the nodes is not met. Transmissions are imposed on energy deficient nodes if a majority of nodes are converged to equation (7) and equation (11) does not hold for the nodes in the same level. ECRP prevents this condition by simplifying the learning based evaluation in two stages. Both the learning stages aim at selecting and discarding optimal and deficient nodes correspondingly. This helps to retain sufficient nodes in the network with at least less remaining energy keeping them alive (Figure 8).

## H. Routing Complexity Analysis

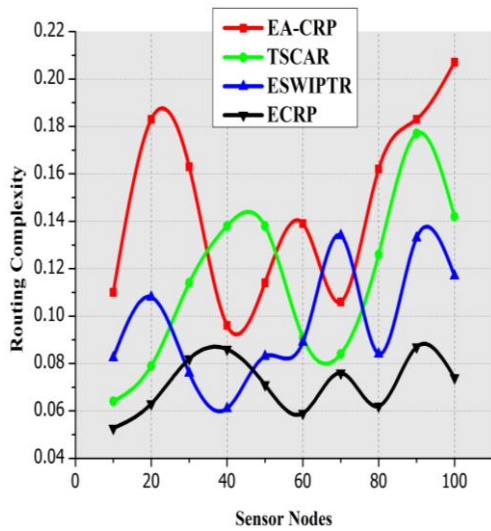


Figure 9 Routing Complexity Comparisons

Table 3 Comparative Analysis of Experimental Values of EA-CRP, TSCAR, ESWIPTR, and ECRP

Metric	EA-CRP	TSCAR	ESWIPTR	ECRP
Throughput (Kbps)	139.36	189.84	214.49	262.29
End-to-End Latency (ms)	112.64	107.54	103.26	96.54
Avg. Energy Utilization (mJ)	953.3	922.27	853.7	768.75
Alive Nodes Count	41	60	73	82
Routing Complexity	0.207	0.142	0.117	0.074

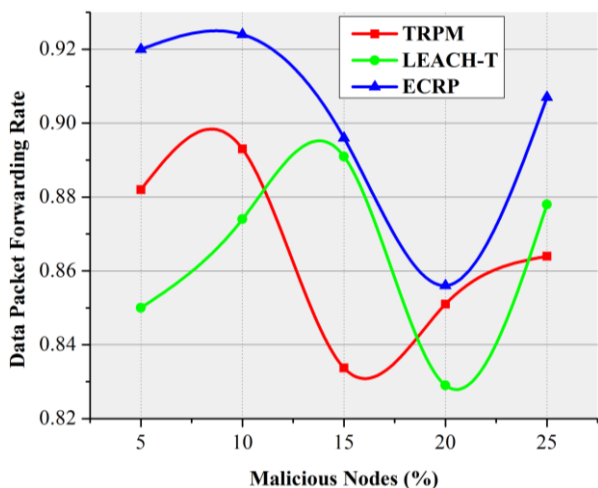
## I. Security Performance Analysis

The performance of the proposed ECRP is analyzed for data forwarding rate and detection rate as a comparative study with respect to security. In the comparative analysis, the existing TRPM [22] and LEACH-T [23] are considered along with ECRP.

Figure 9 depicts the comparative analysis of routing complexity metric between the existing and proposed methods. In ECRP, additional control messages are discarded by pruning the multi-path discovered and selecting best-fit nodes for transmission. In the refurbishing process, additional re-broadcast is not used instead, the pruned nodes are given preference for transmission. The pruned nodes have to satisfy the additional learning condition of delay to ensure reliable delivery. For this purpose  $\phi_n^j$  is computed in each level for assessing its weight and delay factor. The  $\omega_n$  node available in  $\phi_n^j$  is assessed for  $t_{lat}$  for improving the stability of route. A stabilized route sustains multiple transmission rounds and a rotation is required in  $(x - r)$ , at which the frequency of neighbor selection time is high. This frequency is high than the conventional re-broadcasting and re-routing time in the existing methods. Therefore, the number of control messages exchanged in this interval is less, reducing routing complexity. The comparative analysis values are presented in Table 3.



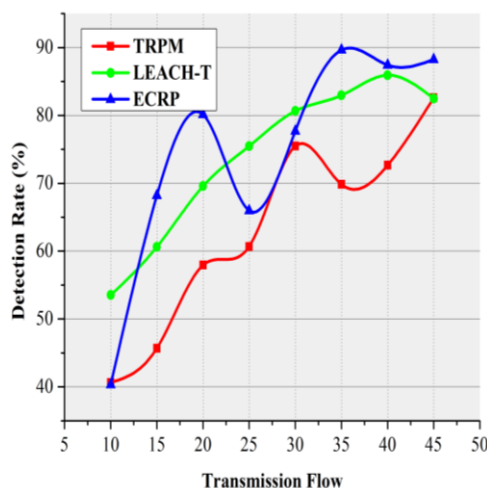
**J. Data Forwarding Rate Analysis**



**Figure 10 Data Forwarding Rate Analysis**

The rate of reliable neighbor selection is refined using leaf node elimination and trust based filtering. This helps to improve the rate of detection. High is the detection rate, high is the rate of legitimate nodes joining the routing and transmission process. The nodes participating in the transmission process on the basis of filtering (leaf node and trusted node) in the path refurbishing process assists maximum data forwarding in the proposed ECRP.

**K. Detection Rate Analysis**



**Figure 11 Detection Rate Analysis**

In Figure 11, the detection rate analysis is compared between the existing and proposed methods. The detection in the proposed method concerns with the newly selection nodes ( $n_{new}$ ) at the time of route refurbishing. The nodes participating in the successive communication are verified for its data handling rate. This data handling rate determines its trust and the node with high trust is capable of participating in the communication process. The analysis of the nodes for trust is fewer compared to the other methods, retaining higher detection rate.

**Table 4 Security Comparative Analysis of Experimental Values of TRPM, LEACH-T, and ECRP**

Metric	TRPM	LEACH-T	ECRP
Data Forwarding Rate	0.864	0.878	0.907
Detection Ratio (%)	82.61	82.47	88.25

**IV. CONCLUSION**

In this paper, energy-centric route planning (ECRP) technique for optimizing wireless sensor network performance and security is discussed. ECRP is designed on the basis of two-stage machine learning for verifying node adaptability based on energy and packet acceptance capacity and delay. The stages in learning refine level node selection for route pruning and refurbishing based on the sequential condition estimation. The communication rounds are segregated on the basis of identifying deficient nodes to swiftly refurbish the routes with less latency. The trust based security model introduced for ECRP helps to identify reliable neighbors for data transmission at the time of route refurbishing. This prevents the participation of illegitimate nodes in the routing and transmission process other than leaf nodes. Performance evaluation of ECRP is verified using a comparative analysis of throughput, energy utilization live nodes and routing complexity observed in WSN.

**REFERENCES**

1. K. Cengiz and T. Dag, "Energy Aware Multi-Hop Routing Protocol for WSNs," in *IEEE Access*, vol. 6, pp. 2622-2633, 2018.
2. M. Sefuba and T. Walingo, "Energy-efficient medium access control and routing protocol for multihop wireless sensor networks," in *IET Wireless Sensor Systems*, vol. 8, no. 3, pp. 99-108, 6 2018.
3. X. Ding, Y. Tian, and Y. Yu, "A Real-Time Big Data Gathering Algorithm Based on Indoor Wireless Sensor Networks for Risk Analysis of Industrial Operations," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 3, pp. 1232-1242, 2016.
4. J. Yan, M. Zhou, and Z. Ding, "Recent advances in energy-efficient routing protocols for wireless sensor networks: A review," *IEEE Access*, vol. 4, pp. 5673-5686, 2016.
5. S. Zidi, T. Moulahi and B. Alaya, "Fault Detection in Wireless Sensor Networks Through SVM Classifier," in *IEEE Sensors Journal*, vol. 18, no. 1, pp. 340-347, 1 Jan.1, 2018.
6. G. Künzel, G. P. Cainelli, I. Müller, and C. E. Pereira, "Weight Adjustments in a Routing Algorithm for Wireless Sensor and Actuator Networks Using Q-Learning," *IFAC-PapersOnLine*, vol. 51, no. 10, pp. 58-63, 2018.



# Energy-Centric Route Planning using Machine Learning Algorithm for Data Intensive Secure Multi-Sink Sensor Networks

7. Shakeel PM. Neural Networks Based Prediction Of Wind Energy Using Pitch Angle Control. International Journal of Innovations in Scientific and Engineering Research (IJISER). 2014;1(1):33-7.
8. F. Ishmanov and Y. B. Zikria, "Trust Mechanisms to Secure Routing in Wireless Sensor Networks: Current State of the Research and Open Research Issues," *Journal of Sensors*, vol. 2017, pp. 1–16, 2017.
9. Y. Liu, M. Dong, K. Ota, and A. Liu, "ActiveTrust: Secure and Trustable Routing in Wireless Sensor Networks," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 9, pp. 2013–2027, 2016.
10. A. Ahmed, K. A. Bakar, M. I. Channa, A. W. Khan, and K. Haseeb, "Energy-aware and secure routing with trust for disaster response wireless sensor network," *Peer-to-Peer Networking and Applications*, vol. 10, no. 1, pp. 216–237, 2015.
11. H. Huang, H. Yin, G. Min, J. Zhang, Y. Wu and X. Zhang, "Energy-Aware Dual-Path Geographic Routing to Bypass Routing Holes in Wireless Sensor Networks," in *IEEE Transactions on Mobile Computing*, vol. 17, no. 6, pp. 1339-1352, 1 June 2018.
12. X. Lai, X. Ji, X. Zhou and L. Chen, "Energy Efficient Link-Delay Aware Routing in Wireless Sensor Networks," in *IEEE Sensors Journal*, vol. 18, no. 2, pp. 837-848, 15 Jan.15, 2018.
13. X. Shao, C. Wang, C. Zhao and J. Gao, "Traffic Shaped Network Coding Aware Routing for Wireless Sensor Networks," in *IEEE Access*, vol. 6, pp. 71767-71782, 2018.
14. J. Qiao and X. Zhang, "Polar Coordinate-Based Energy-Efficient-Chain Routing in Wireless Sensor Networks Using Random Projection," in *IEEE Access*, vol. 6, pp. 21275-21286, 2018.
15. W. M. Elsayed, S. F. Sabbeh, and A. M. Riad, "A Distributed Fault Tolerance Mechanism for Self-Maintenance of Clusters in Wireless Sensor Networks," *Arabian Journal for Science and Engineering*, vol. 43, no. 12, pp. 6891–6907, 2017.
16. Z.-X. Wang, M. Zhang, X. Gao, W. Wang, and X. Li, "A clustering WSN routing protocol based on node energy and multipath," *Cluster Computing*, 2017.
17. C. Hong, Y. Zhang, Z. Xiong, A. Xu, H. Chen, and W. Ding, "FADS : Circular/Spherical Sector based Forwarding Area Division and Adaptive Forwarding Area Selection routing protocol in WSNs," *Ad Hoc Networks*, vol. 70, pp. 121–134, 2018.
18. J. Huang, D. Ruan, and W. Meng, "An annulus sector grid aided energy-efficient multi-hop routing protocol for wireless sensor networks," *Computer Networks*, vol. 147, pp. 38–48, 2018.
19. K. A. Darabkh, N. J. Al-Maaitah, I. F. Jafar, and A. F. Khalifeh, "EA-CRP: A Novel Energy-aware Clustering and Routing Protocol in Wireless Sensor Networks," *Computers & Electrical Engineering*, vol. 72, pp. 702–718, 2018.
20. A. Attiah, M. F. Amjad, M. Chatterjee, and C. Zou, "An evolutionary routing game for energy balance in Wireless Sensor Networks," *Computer Networks*, vol. 138, pp. 31–43, 2018.
21. S. He, K. Xie, W. Chen, D. Zhang and J. Wen, "Energy-Aware Routing for SWIPT in Multi-Hop Energy-Constrained Wireless Network," in *IEEE Access*, vol. 6, pp. 17996-18008, 2018.
22. B. Sun and D. Li, "A Comprehensive Trust-Aware Routing Protocol With Multi-Attributes for WSNs," *IEEE Access*, vol. 6, pp. 4725–4741, 2018.
23. L. Yang, Y. Lu, S. Liu, T. Guo, and Z. Liang, "A Dynamic Behavior Monitoring Game-Based Trust Evaluation Scheme for Clustering in Wireless Sensor Networks," *IEEE Access*, vol. 6, pp. 71404–71412, 2018.

Services, He has published a paper based on Machine Learning in the Conference. He has got Certifications from Udemy, Oracle Academy, CourseEra, NPTEL, SOLO Learn. He has undergone Training and has attended many workshops. His research areas of Interest includes Computer Networks, Machine Learning, Deep Learning.

**S. Sandhya** has completed her Bachelors in Computer Science and Engineering (B.E-CSE) at R.M.K Engineering College, Kavarapettai. She has done projects, An Online Quiz and Campus Interaction Application in Android for Virtusa Consulting Services and Bus Tracker Android Mobile Application. She has published a paper based on Machine Learning in the Conference. She has got Certifications from Oracle Academy, NPTEL, SOLO Learn. She has undergone Training and Internships, has attended many workshops. Her research areas of Interest includes Computer Networks, Machine Learning, Database Management System.

## AUTHORS PROFILE

**Dr.M.D.Vimalapriya** is an Assistant Professor in the Department of Computer Science and Technology at Women's Christian College, Chennai-6, having completed MCA., MPhil., Ph.D. She has around 18 years of Experience and has been awarded with cash and gold medals for her achievements. Adding to her accolades, She has published three books, has published research papers in Reputed Journals and presented papers in International and National Conferences. She has attended many workshops, has chaired Technical session in Conference and has been resource person in Faculty Development Programme. Her research areas of Interest but not limited to includes Computer Networks, Machine Learning and Data Mining.

**Vignesh Baalaji.S** has completed his Bachelors in Computer Science and Engineering (B.E-CSE) at R.M.K Engineering College, Kavarapettai. He has been an Application Developer - Intern for 6 (Six) months, CZ Smart Mobility, Application Developer - Intern for 6 (Six) months at Virtusa Consulting Services, He has done RideShare project Carpooling Mobile Application for CZ Smart Mobility and An Online Quiz and Campus Interaction Application in Android for Virtusa Consulting