International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

A Learning based Secure Routing Approach using Deep Reinforcement Learning in IoT Integrated Wireless Sensor Network

Battina Srinuvasu Kumar
Research scholar, Department of Computer Science & Engineering, Annamalai University 608002, India
and Associate professor, Department of Information Technology, SR Gudlavalleru Engineering College , Gudlavalleru-521356,India.
bskdata@gmail.com

Dr. S.G.Santhi
Associate Professor, Department of Computer Science & Engineering, Annamalai University-608002, India.
sgsau2009@gmail.com

Dr. S. Narayana
Professor, Department of Computer Science & Engineering, SR Gudlavalleru Engineering College, Gudlavalleru-521356, India.
satyala1976@gmail.com

Abstract—

The usage of Wireless Sensor Network (WSN) is ubiquitous in nature. With the emergence of Internet of Things (IoT) technology and its unprecedented use cases, the role of sensor networks as part of IoT application became crucial. WSN became backbone of IoT to realize integration between physical and digital worlds and connectivity to Internet. However, IoT devices are resource constrained with limited computational capabilities. The entire network is distributed in nature and has increased complexity. Routing in such WSN integrated IoT network plays an important role in achieving meaningful communication among objects. In this context, it is indispensable to have more energy efficient routing method. Since the IoT integrated sensor network is highly complicated, it is very dynamic in nature. Thus routing decisions are also dynamic leading to much importance to routing in such use cases. With the emergence of Artificial Intelligence (AI), it became possible to solve complex real world problems through learning based approach which acquires desired intelligence prior to making decisions. In this paper we proposed a deep reinforcement learning based routing mechanism for energy efficient routing in WSN-IoT integrated application. We proposed novel algorithms for network setup, formation of clusters and routing. Our method adapts to network changes due to energy levels, mobility and makes learning based routing decisions. We enhanced the method further with security to ensure its Qualityof Service (QoS) in presence of attacks. Our simulation study using MATLAB has revealed that the proposed secure routing approach outperforms existing protocols.

Keywords- Wireless Sensor Network, Internet of Things, Energy Efficient Routing, Deep Reinforcement Learning, Secure Routing.

1. INTRODUCTION

Sensor networks became ubiquitous in the real world applications. In every conceivable field, the usage of sensors is evident. For instance, in healthcare, in transportation, in water conservation and so on, sensors play vital role. Since sensors can work in even areas where human manual observation is not possible. In other words, sensors are also deployed in hostile environments. With the emergence of IoT technology, sensor networks became indispensable for IoT applications. However, the problem sensors remain that they are resource constrained devices. Therefore, energy efficiency is desired to be improved in sensor network. It is also important in IoT integrated sensor networks. As IoT networks use sensors in large scale, it is important to propose novel approaches exploiting AI. Towards this end, this paper focuses on learning based approach for routing.

IoT use cases became popular in the real world as IoT has revolutionized with unprecedented kind of applications such as smart home, smart city and smart healthcare to mention few. Many researchers attempted to improve routing in IoT integrated use cases as explored in [2], [9], [13], [14], [16] and [19]. Sarwesh et al. [2] proposed an architecture suitable for energy efficiency in IoT use cases. Krishnaraj and Smys [9] used IoT use cases for routing optimization towards energy efficiency. Shen et al. [13] proposed a centroid based routing mechanism for efficiency and energy conservation. Ikram et al. [14] focused on packet update caching mechanism in IoT use case for energy efficiency. Khan and Ali [16] implemented trust management in IoT network besides making it energy efficient. From the literature, it is observed there are plethora of routing approaches for energy efficiency in WSN and IoT networks. There are few researches that tried to use machine learning. There is need for learning based secure routing protocol by exploiting deep reinforcement learning. Our contributions in this paper are as follows.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

1. In this paper we proposed a deep reinforcement learning based routing mechanism for energy efficient routing in WSN-IoT integrated application.

2. We proposed novel algorithms for network setup, formation of clusters and routing. Our method adapts to network changes due to energy levels and mobility and makes learning based routing decisions.

3. We enhanced the method further with security to ensure its Quality of Service (QoS) in presence of attacks

4. Our simulation study using MATLAB has revealed that the proposed routing approach outperforms existing protocols.

The remainder of the paper is structured as follows. Section 2 threw light on existing routing methods. Section 3 presents the proposed learning based routing protocol. Section 4 presents results of experiments. Section 5 concludes our work.

2. RELATED WORK

This section reviews existing routing protocols to find useful insights. Yarinezhad and Azizi [1] proposed a routing protocol for sensor networks liked to IoT considering link quality and geographical location. Sarwesh et al. [2] proposed an architecture suitable for energy efficiency in IoT use cases. Soundari et al. [3] defined an algorithm for energy efficient data collection in sensor networks using ML approach. Bhardwaj and Kuma [4] proposed a bio-inspired algorithm for multiple-objective routing. Kaur et al. [5] has reviewed many QoS-aware routing protocols that exploit computational intelligence. Khalid et al. [6] proposed a cooperative routing protocol that is not only secure but also energy efficient. Merlin et al. [7] focused on a routing protocol for MANET with trust based approach leading to energy efficient routing. Mathebula et al. [8] proposed SDN based approach to control operations for energy efficiency in WSN. Krishnaraj and Smys [9] used IoT use cases for routing optimization towards energy efficiency. Saima et al. [10] proposed an algorithm for efficiency scheduling and joint routing in IoT network.

Farman et al. [11] focused in improving WBAN stability period in routing with energy usage optimization. Rui et al. [12] investigated on energy efficiency in WSN using SDN approach and also trust management. Shen et al. [13] proposed a centroid based routing mechanism for efficiency and energy conservation. Ikram et al. [14] focused on packet update caching mechanism in IoT use case for energy efficiency. Nivedhitha et al. [15] proposed a protocol for dynamic multi-hop routing with energy efficient mechanism. Khan and Ali [16] implemented trust management in IoT network besides making it energy efficient. Other important contributions include network coding based routing [17], lion whale optimization [18], hybrid multipath routing [19] and game theory based routing [20]. From the literature, it is observed there are plethora of routing approaches for energy efficiency in WSN and IoT networks. There are few researches that tried to use machine learning. There is need for learning based routing protocol by exploiting deep reinforcement learning.

3. PROPOSED ROUTING PROTOCOL

We proposed a novel routing protocol for an IoT use case. Instead of traditional approaches, we exploited learning based approach which is based on runtime situations.

3.1 Reinforcement Learning

Our protocol is based on Reinforcement Learning (RL)which is nothing but a Markov Decision Process which involves different states (S), actions (A), transition probability (P) and reward (R). Figure 1 illustrates typical agent based RL approach which has an iterative process to make optimal decisions based on the interactions between agent and environment. Agents response to the runtime environmental state is known as action. State information can have details that help in decision making. State information includes signal strength, hop count, residual energy etc. that are used for decision making. Action refers to the routing decision that improves energy efficiency in the IoT network. Reward is nothing but cost of action.

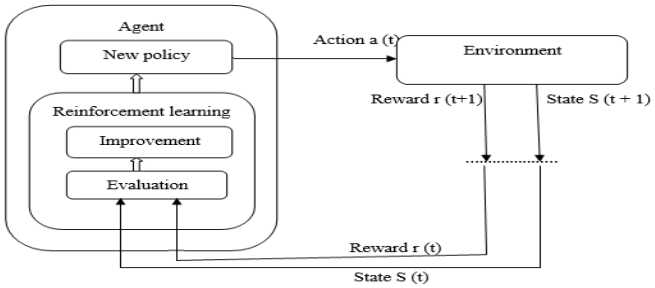

Figure 1: Reinforcement learning used in the proposed routing protocol

IoT network consists of number of devices. Each device acts as an agent. The state space includes different possible routes to the base station. Action space reflects set of possible routing decisions. Policy is nothing but specific behaviour of agent. Policy thus contains a pair reflecting action and state. This pair is made either through deterministic approach or stochastic approach which gets optimized over time. The aim of the RL used in the proposed routing protocol is to achieve optimal policy leading to energy efficiency.

3.2 Our Protocol

We proposed routing protocol based on RL discussed in Section 3.1. It exploits cluster formation and designed for energy efficiency. It enables devices in the IoT network to learn through RL prior to making best routing decisions. The sender device fills packet header with local information, information about other devices in the neighbourhood besides updating routing table. The local information contains hop count, position coordinates, residual energy and device id. The protocol has three important phases known as setup, formation of clusters and transmission of data.

3.2.1 Network Setup

It is the process of setting up IoT based network and also elect initial cluster head. Each node in the network computes Q-value, as in Eq. 1, based on its local information. A heartbeat message is broadcasted by base station which consists of its position details. On receiving such details, every device

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

computes Q-value and hop count using Eq. 1 and Eq. 2 respectively. It is assumed that devices do have different levels of energy. We consider a distance threshold for distance between CH and base station to get rid of network overhead.

$$Q = \begin{cases} \frac{1}{N_h}, & if\ E_{min} = E_{max}, \\ p \times \left(\frac{E_r - E_{min}}{E_{max} - E_{min}}\right) + (1-p) \times \frac{1}{N_h}, & if\ E_{min} \neq E_{max}, \end{cases} \quad (1)$$

$$N_h \cong \frac{D_{link}}{TX_{range}}. \quad (2)$$

Once this process is done, CH selection process follows. This process is carried out along with network setup as defined in the Algorithm 1.

Algorithm 1: Setup and CH Election

Input: Nodes of IoT network N, sink node s, min distance threshold th1, max distance threshold th2

```
Begin
For each node n in N
Compute distance between n and s
  Compute hop count  //Eq. 2
  Compute Q value  //Eq. 1
End For
While(chCount<=chTotal)
 Qmax←Qmax(N)
 For each node n in N
   IF th1<=distance of n<th2 Then
   IF chCount=0 Then
     Add n to chCount
     Remove n from N
   Else
     For h==1 to chCount
       distance←findDistance(n, CH(h))
       IF distance >= th1 Then
         C←true
       Else
         C←false
         Break
       End If
     End For
   IF C==true Then
     Add n to chCount
     Remove n from N
   End If
 End If
End For
End While
```

Algorithm 1: Setup and CH Election

As presented in Algorithm 1, it takes nodes of IoT network N, sink node s, min distance threshold th1, max distance threshold th2 and results in network setup and initial CH election.

3.2.2 Cluster Formation

Once CH election process is computed there is need for cluster formation. In the cluster formation phase, each CH notifies its neighbours that it is the elected cluster head. The notification value contains its location details, id and Q-value. On receiving the notification, each non-CH node makes decision to join a cluster based on the distance and conveys the same message to CH. The message also contains its local information. Once this process is done by all non CH nodes corresponding CH node accepts the request for including as member of the cluster.

It is to be observed that some devices that have base station in the range of transmission do not involve in cluster formation. However, they are capable of directly communicating with base station and that saves energy.

3.2.3 Data Transmission

In this phase, RL plays crucial role while making data transmission decisions. In this phase each node in the network acts as an agent involved in RL. The node follows learning based approach in making optimal routing decisions that lead to energy conservation. Learning involves in computing Q-value or updating it based on the intermediate action and reward dynamics besides observing at finding best policy that reflects highest reward. The best decisions are made based on Q-value update function, reward function and energy consumption dynamics. Energy consumption is computed as in Eq. 3.

$$\begin{cases} E_{TX}(k,d) = E_{elec} \times k + E_{amp} \times k \times d^m, \\ E_{Rx}(k) = E_{elec} \times k, \end{cases} \quad (3)$$

Where the energy consumption made at transmitter is denoted as $E_{TX}(k,d)$ while the same at receiver is denoted as $E_{Rx}(k)$. It results in residual energy. Then Q-value gets updated by passing reward as an argument. Each node (learning agent) gets reward for each action. It reflects the cost of the action which helps in understanding whether the action is appropriate. Here an action is nothing but choosing a neighbour node towards next hop in order to route packet. Reward function is associated with hop count (denoted as Nh) and energy saving (denoted as Er). Dlink which refers to the distance of agent (node) and base station through other node is computed as expressed in Eq. 4 to Eq. 6.

$$D_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \quad (4)$$

$$D_{j,sink} = \sqrt{(x_j - x_{sink})^2 + (y_j - y_{sink})^2}, \quad (5)$$

$$D_{link} = D_{i,j} + D_{j,sink}, \quad (6)$$

Eq. 2 helps in computing hop count and the actual reward function is computed as expressed in Eq. 7.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

$$r_{t+1} = \begin{cases} \frac{1}{N_h}, & if\ E_{min} = E_{max}, \\ p \times \left(\frac{E_r - E_{min}}{E_{max} - E_{min}}\right) + (1-p) \times \frac{1}{N_h}, & if\ E_{min} \neq E_{max} \\ -100, & if\ E_r < 0, \end{cases} \quad (7)$$

Where p denotes probabilistic value which is used to reflect the influence of Er. Higher p value determines a neighbour consisting of high energy levels and suitable as next-hop selection by agent. On the other hand, q denotes closest neighbour to high probability node. There is need for trade-off computation between the two variables to optimize the rouging decisions. In the process of RL reward value is kept in packet header to get rid of overhead. The neighbouring nodes can overhear the reward value in order to update their routing table.

4. SECURITY ENHANCEMENT

We enhanced our protocol further to ensure secure routing mechanisms in the underlying sensor network associated with IoT. Our system model has number of nodes (N), number of edges (E) and an edge is denoted as $\{w_{ij}|(v_i, v_j)\}$. The whole network is denoted as G where source node is denoted as src while sink node is denoted as sink. Given a source and destination nodes (s and d), a path between the nodes is $l_s = \langle src, \dots, sink \rangle$. It is a multi-hop path in which data is transmitted. In this scenario, there might beattacker capturing certain nodes and such nodes are known as malicious nodes. There is need for devising a mechanism to handle such attacks. We considered an attack model associated with selective forwarding attack as discussed in [21]. It is an attack where malicious node intentionally chooses to drop certain packets. Thus the attack causes incomplete data transmission. This kind of attack can be detected by maintaining trust value for nodes and updating the values from time to time. Based on the nodes behaviours, the trust value gets updated. Based on this, a node can be identified to have either normal or abnormal behaviour. Figure 2 shows procedure for computation of attack probability.

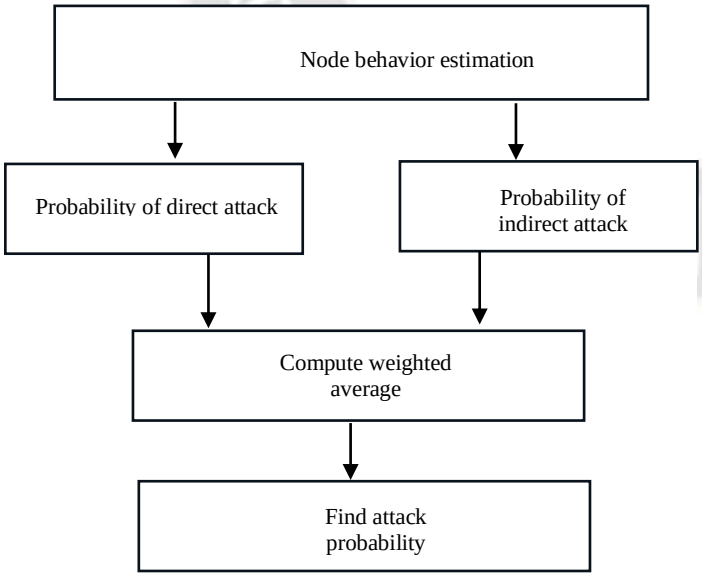

Figure 2: Shows procedure to compute attack probability

With respect to node forwarding attack, attack probability is computed as a weighted average of direct and indirect attack probabilities. Figure 3 shows nodes in network to understand direct and indirect attack probabilities.

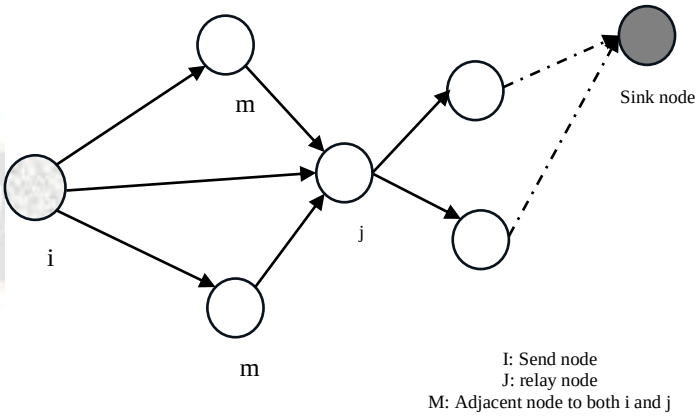

Figure 3: Nodes in network to understand direct and indirect attack probabilities

Attack behaviour of a node is evident, when it is not forwarding a packet but dropping it selectively. Instead of using traditional promiscuous monitoring, we proposed a two-hop ack method. ACK packet plays crucial role in attack detection. Here is the formal procedure of our method. Consider nodes denoted as *i, j* and *k*. When a packet is sent by *i* to *j* and then *j* forwards it to *k*. In this case node *i* must receive ACK from *j* and *k*. It indicates that *j* has normal behaviour. If not, it establishes the fact that *j* is malicious. The attack probability of node j to node *i* ($Attack_{\langle i,j \rangle}$) has direct and indirect probabilities denoted as $Attack_{d\langle i,j \rangle}$ and $Attack_{id\langle i,j \rangle}$ respectively. The probability of direct attack is computed as in Eq. 8.

$$Attack_{d\langle i,j \rangle} = \frac{|P_{ij} - P_j|}{P_{ij}} \quad (8)$$

where

P_{ij} denotes total packets sent to j from i

P_j denotes total packets forwarded successfully by j

It reflects the packet loss ratio pertaining to node *j*. The probability of indirect attack is expressed as in Eq. 9.

$$Attack_{id\langle i,j \rangle} = \frac{\sum_{m=1}^{r} \alpha_m Attack_{d\langle m,j \rangle}}{\sum_{m=1}^{r} \alpha_m}, \left(\alpha_m = 1 - Attack_{d\langle i,m \rangle,}\ m \neq i\right) \quad (9)$$

where

m denotes an adjacent node to nodes i and j

$\alpha_m$denotes trust of node i on m

r denotes count of adjacent nodes to i and j

Nodes j's indirect attack probability is computed by the finding attack probability of node j to node m.

Thus indirect attack probability is computed as weighted average of j's attack probabilities to adjacent nodes of i and j.

It is possible to combine direct and indirect attack probabilities associated with *j* as in Eq. 10.

$$Attack_{\langle i,j \rangle} = \frac{\sum_{m=1}^{r} \alpha_m Attack_{d\langle m,j \rangle}}{\sum_{m=1}^{r} \alpha_m}, \left(\alpha_m = \begin{cases} 1, & m = i, \\ 1 - Attack_{d\langle i,m \rangle,} & m \neq i\ . \end{cases}\right) \quad (10)$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

If all the nodes in a path from source to sink behave normally, there is successful communication in sensor network. As the nodes in the path are in a serial system, attack probability can be computed as in Eq. 11.

$$A_L = 1 - \prod_{k=1}^{K}(1 - \text{Attack}_{\langle m_k, m_{k+1}\rangle}), \qquad (11)$$

Here the number of hops in the path is denoted as K.
The node *j* in the above scenario has two associated values such as distance to sink and residual energy. The node which has less distance to sink node and has higher residual energy is best used as next-hop. Status of each node is thus computed as in Eq. 12.

$$E_j = \frac{\eta e_{\langle j,sink\rangle}}{\eta e_{\langle j,sink\rangle} + e_{rj}}, \qquad (12)$$

Where η is a balancing factor, e_{rj} denotes node j's residual energy and $e_{\langle j,sink\rangle}$ indicates the energy needed to send packet to sink. The status value for entire path is computed as in Eq. 13.

$$E_L = \sum_{k=1}^{K}\left(\frac{\eta e_{\langle m_{k+1},sink\rangle}}{\eta e_{\langle m_{k+1},sink\rangle} + e_{rm_{k+1}}} \times \frac{1}{K}\right). \qquad (13)$$

It adds status value of each node in the path in order to obtain the value for entire path. The resultant value is normalized to have a value less than 1. A trust value which encapsulates status and attack probability is considered to make secure routing decisions. This phenomenon is influenced by the work in [21]. Given a node n wants to send data to sink, it needs to choose a relay node denoted as S_n. A cost function is computed based on status values and attack probabilities for entire path.

$$C_{n(1)} = A_L$$

$$= 1 - \prod_{k=1}^{K}(1 - \text{Attack}_{\langle m_k, m_{k+1}\rangle})$$

$$= \text{Attack}_{\langle n,s_n\rangle} - (\text{Attack}_{\langle n,s_n\rangle} - 1)[1 - \prod_{k=1}^{K}(1 - \text{Attack}_{\langle m_k, m_{k+1}\rangle})] \qquad (14)$$

$$= \text{Attack}_{\langle n,s_n\rangle} - (\text{Attack}_{\langle n,s_n\rangle} - 1)C_{S_n(1)}$$

$$C_{n(2)} = E_L$$

$$= \sum_{k=1}^{K}\left(\frac{\eta e_{\langle m_{k+1},sink\rangle}}{\eta e_{\langle m_{k+1},sink\rangle} + e_{rm_{k+1}}} \times \frac{1}{K}\right)$$

$$= \frac{\eta e_{\langle S_n,sink\rangle}}{\eta e_{\langle S_n,sink\rangle} + e_{rs_n}} \times \frac{1}{K} + \sum_{k=2}^{K}\left(\frac{\eta e_{\langle m_{k+1},sink\rangle}}{\eta e_{\langle m_{k+1},sink\rangle} + e_{rm_{k+1}}} \times \frac{1}{K}\right) \qquad (15)$$

The cost function of n is

$$c_n = c_{n(1)} + c_{n(2)}. \qquad (16)$$

As expressed from Eq. 14 to Eq. 16, cost of n is computed by considering trust of all nodes in the path in order to arrive at global optimization. Based on this a routing decision is made to avoid selective forwarding attack.

5. RESULTS AND DISCUSSION

We have made a simulation study using MATLAB where 100 x 100 m area is used to deploy 100 nodes involved in IoT network. In the middle of the sensing area base station is positioned. The nodes are heterogeneous in nature in terms of energy resources. Transmission range is set to 20m. Initial energy of every node is set to either 1 or 2 joules to ensure heterogeneity. Data size considered is 4000 bits. Value of α variable is set to 1 while γvariable is set to 0.95. The value of E_{amp} is set to $100\text{x}10^{-12}$joules/bit/m^2. And E_{elec} is set to $50\text{x}10^{-9}$ joules/bit.

5.1 Results

Performance of the proposed protocol is provided in this section in terms of cluster formation, energy consumption dynamics, dead nodes and on.

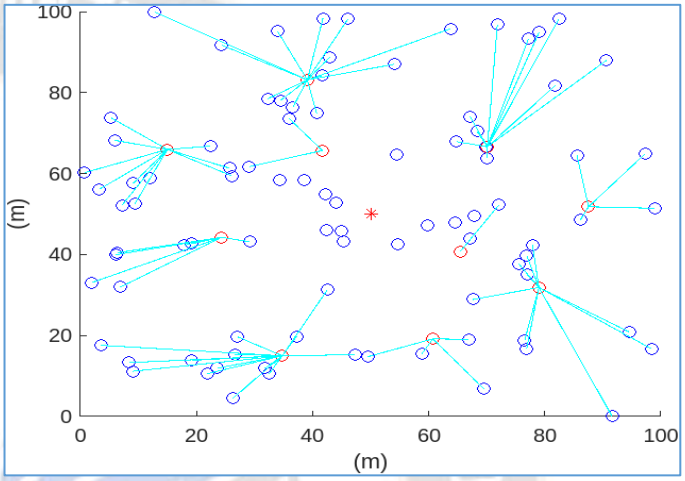

Figure 4: Result of cluster formation with 100 nodes in sensing area covering 100 x 100 m

As presented in Figure 4, 100x100 m simulation area is used to simulate an IoT network with sensor nodes and base station. The base station is positioned at (50,50) coordinates at the centre of the simulation area. Many clusters are formed as per the methodology described in Section 3. There are some nodes nearby the base station that are not part of any cluster. Every cluster has an elected cluster head. In the simulation experiment, many observations are made in terms of operating nodes, energy consumption and dead nodes. Besides the proposed protocol is compared against state of the art.

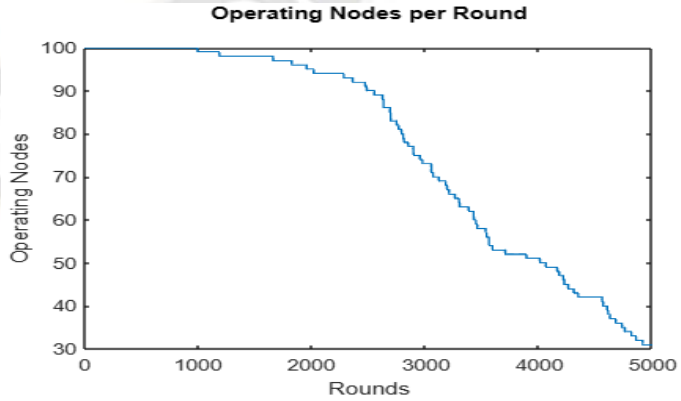

Figure 5: Shows operating nodes against number of rounds

As presented in Figure 5, the number of operating nodes are taken in vertical axis from 30 to 100 while the horizontal axis shows up to 5000 rounds incremented by 1000 rounds. An important observation is that as the number of rounds in

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

increased, gradually the number of operating nodes is decreased.

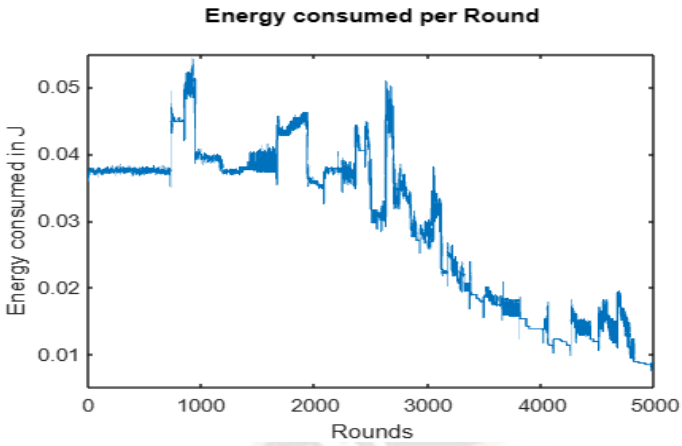

Figure 6: Energy consumption against number of rounds

As presented in Figure 6, energy consumption dynamics is visualized against number of rounds of simulation. As the number of rounds is increased the energy consumed in the network is gradually decreased.

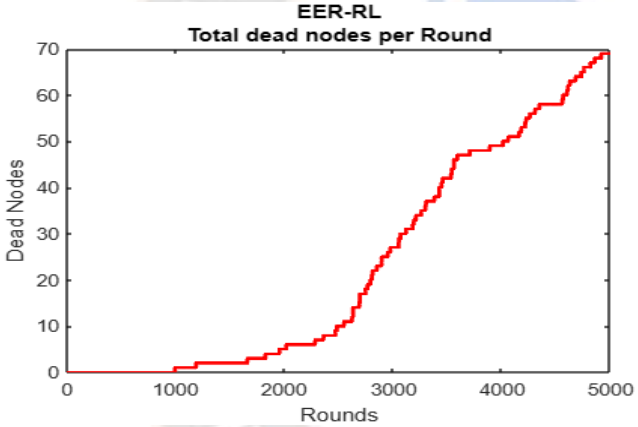

Figure 7: Dead nodes against number of rounds

As presented in Figure 7, number of dead nodes observation started from the round number 1000. From round 1000 onwards, there is gradual increase in total number of dead nodes that consume complete energy.

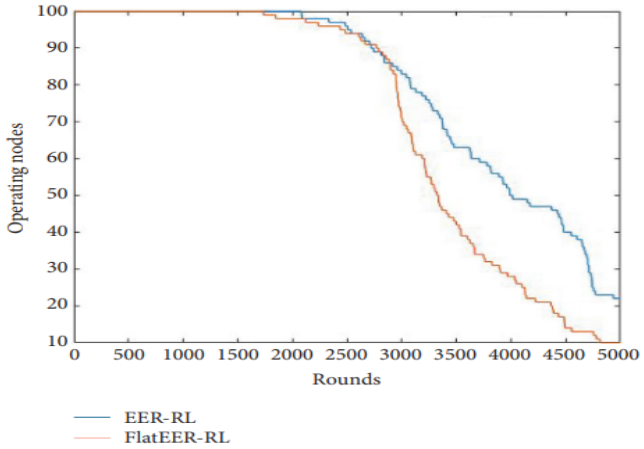

Figure 8: Shows operating nodes comparison between the existing and proposed routing protocols

As presented in Figure 8, performance of the proposed routing protocol shown in blue colour data series is compared against existing routing protocol shown in red colour data series. As the number of rounds is increased, the operating nodes is gradually decreased. Less in operating nodes in any given round indicates less performance. In the same fashion, higher in number of operating nodes in any given round indicates better performance. Based on this, it is observed that the proposed RL based protocol achieved superior performance over its predecessor.

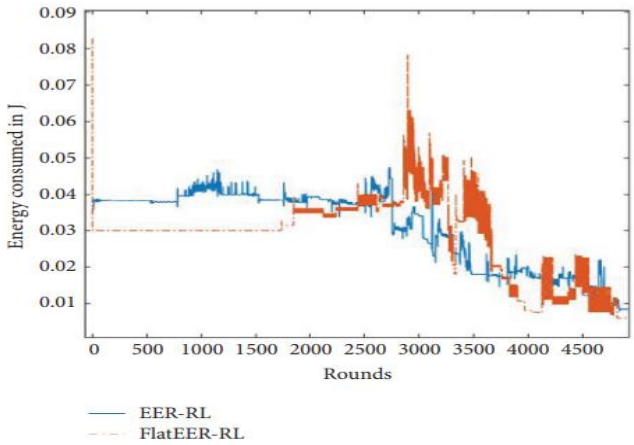

Figure 9: Energy consumption comparison between the existing and proposed routing protocols

As presented in Figure 9, performance of the proposed routing protocol shown in blue colour data series is compared against existing routing protocol shown in red colour data series. As the number of rounds is increased, the energy consumption is gradually decreased except at some rounds. Less in energy in any given round indicates better performance. In the same fashion, higher in energy consumption in any given round indicates less performance. Based on this, it is observed that the proposed RL based protocol achieved superior performance over its predecessor.

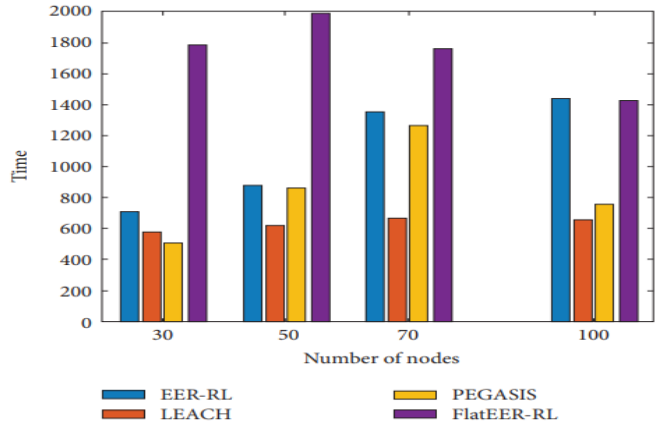

Figure 10: Performance comparison in terms of network lifetime

As presented in Figure 10, simulated experiments are made to ascertain network lifetime when the number of nodes is 30, 50, 70 and 100. It does mean that 4 experiments are made with different routing protocols. The proposed RL based protocol is shown in blue colour data series while other colours show different existing protocols including widely used LEACH protocol. It is observed that network lifetime increases when

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

there is increase in number of nodes in the IoT network. Another observation is that different protocols showed different network lifetime dynamics. This is due to variation in energy consumption while routing protocols in operation. The proposed routing protocol exhibits more energy efficiency leading to increased network lifetime. In fact, it outperforms existing protocols like PEGASIS and LEACH.

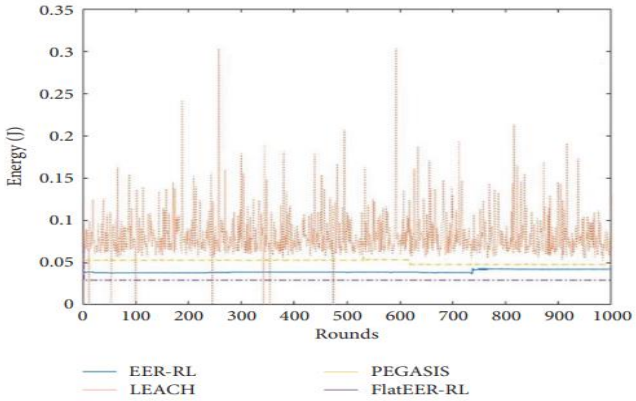

Figure 11: Performance comparison in terms of energy consumption

As presented in Figure 11, simulated experiments are made to ascertain energy consumption for 1000 rounds of simulation. The proposed RL based protocol is shown in blue colour data series while other colours show different existing protocols including widely used LEACH protocol. It is observed that each protocol has varied energy consumption. With respect to energy consumption, our protocol outperforms PEGASIS and LEACH.

5.2 Security Analysis

Our security enhancement made our protocol robust against attacks. In our empirical study, it is observed that QoS of the protocol is better when compared with the one which has no security enhancement.

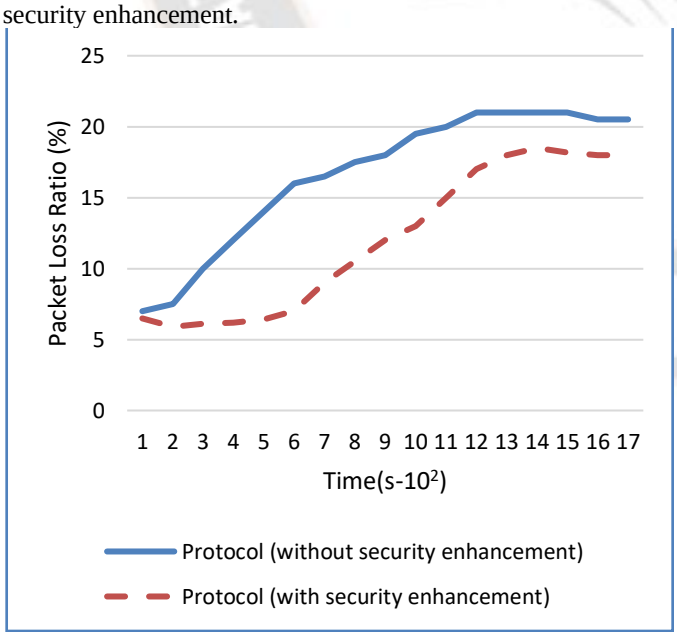

Figure 12: QoS comparison in terms of packet loss ratio when simulation is made with 20% malicious nodes

As presented in Figure 12, packet loss ratio is the metric considered to evaluate the protocol with and without security enhancement. Low in packet loss ratio indicates better performance. With security enhancement, in presence of attacks, the protocol is able to achieve least packet loss ratio reflecting increased QoS in routing. As the simulation time elapsed is increased, there is gradual increase in the packet loss ratio. However, it is evident that the protocol with security enhancement has least packet loss ratio showing the significance of security in routing.

6. CONCLUSION AND FUTURE WORK

In this paper we proposed a deep reinforcement learning based routing protocol for energy efficient routing in WSN-IoT integrated application. It exploits cluster formation and designed for energy efficiency. It enables devices in the IoT network to learn through RL prior to making best routing decisions. The sender device fills packet header with local information, information about other devices in the neighbourhood besides updating routing table. The local information contains hop count, position coordinates, residual energy and device id. The protocol has three important phases known as setup, formation of clusters and transmission of data. We proposed novel algorithms for network setup, formation of clusters and routing. Our method adapts to network changes due to energy levels and mobility and makes learning based routing decisions. Our simulation study using MATLAB has revealed that the proposed routing approach outperforms existing protocols. We enhanced the method further with security to ensure its Quality of Service (QoS) in presence of attacks. In future, we improve our routing protocol with further investigations on its efficiency and security.

References

[1] RaminYarinezhad and Sadoon Azizi; (2021). An energy-efficient routing protocol for the Internet of Things networks based on geographical location and link quality . Computer Networks. http://doi:10.1016/j.comnet.2021.108116.

[2] Sarwesh, P., Shet, N. S. V., &Chandrasekaran, K. (2017). Energy-Efficient Network Architecture for IoT Applications. Beyond the Internet of Things, 119–144. http://doi:10.1007/978-3-319-50758-3_5.

[3]Soundari, A. Gnana and Jyothi, V. L. (2019). Energy Efficient Machine Learning Technique for Smart Data Collection in Wireless Sensor Networks. Circuits, Systems, and Signal Processing. http://doi:10.1007/s00034-019-01181-3.

[4] Reeta Bhardwaj and Dinesh Kuma. (2019). MOFPL: Multi-objective fractional particle lion algorithm for the energy aware routing in the WSN. Elsevier, pp.1-34. https://doi.org/10.1016/j.pmcj.2019.05.010.

[5] Kaur, Tarunpreet and Kumar, Dilip (2019). A survey on QoS mechanisms in WSN for computational intelligence based routing protocols. Wireless Networks. http://doi:10.1007/s11276-019-01978-9.

[6] Saeed, Khalid; Khalil, Wajeeha; Ahmed, Sheeraz; Ahmad, Iftikhar; Khattak, Muhammad Naeem Khan (2020). SEECR: Secure Energy Efficient and Cooperative Routing Protocol for Underwater Wireless Sensor Networks. IEEE Access, 1–1. http://doi:10.1109/ACCESS.2020.3000863.

[7] Merlin, R. Tino and Ravi, R. (2019). Novel Trust Based Energy Aware Routing Mechanism for Mitigation of Black Hole Attacks in MANET. Wireless Personal Communications. http://doi:10.1007/s11277-019-06120-8.

[8] Ishmael Mathebula;BasseyIsong;NaisonGasela and Adnan M. Abu-Mahfouz; (2020). Analysis of Energy-efficient Techniques for SDWSN Energy Usage Optimization .2nd International Multidisciplinary Information Technology and Engineering Conference (IMITEC). http://doi:10.1109/imitec50163.2020.9334084.

[9] Krishnaraj, N. and Smys, S. (2019). A Multihoming ACO-MDV Routing for Maximum Power Efficiency in an IoT Environment. Wireless Personal Communications. http://doi:10.1007/s11277-019-06562-0.

[10] Abdullah, Saima; Asghar, Mamoona N.; Ashraf, Mashavia and Abbas, Naila (2020). An Energy-Efficient Message Scheduling Algorithm with Joint Routing Mechanism at Network Layer in Internet of Things Environment. Wireless Personal Communications. http://doi:10.1007/s11277-019-06959-x.

[11] Ullah, Farman; Zahid Khan, M.; Faisal, Mohammad; Rehman, Haseeb Ur; Abbas, Sohail and Mubarek, Foad S. (2020). An Energy Efficient and Reliable Routing Scheme to enhance the stability period in Wireless Body Area Networks. Computer Communications,S0140366420319654. http://doi:10.1016/j.comcom.2020.10.017.

[12] Wang, Rui; Zhang, Zhiyong; Zhang, Zhiwei and Jia, Zhiping (2018). ETMRM: An Energy-efficient Trust Management and Routing Mechanism for SDWSNs. Computer Networks, S1389128618301725. http://doi:10.1016/j.comnet.2018.04.009.

[13] Jian Shen,Anxi Wang, Chen Wang, Patrick C. K. Hung and Chin-Feng Lai,. (2017). An Efficient Centroid-Based Routing Protocol for Energy Management in WSN-Assisted IoT. IEEE. 5, pp.1-10. http://doi:10.1109/ACCESS.2017.2749606.

[14] Din, IkramUd; Hassan, Suhaidi; Almogren, Ahmad; Ayub, Farrukh and Guizani, Mohsen (2019). PUC: Packet Update Caching for energy efficient IoT-based Information-Centric Networking. Future Generation Computer Systems, S0167739X19325087. http://doi:10.1016/j.future.2019.11.022.

[15] Nivedhitha, V.; Saminathan, A. Gopi and Thirumurugan, P. (2020). DMEERP: A dynamic multi-hop energy efficient routing protocol for WSN. Microprocessors and Microsystems, 79, 103291. http://doi:10.1016/j.micpro.2020.103291.

[16] Khan and Zeeshan Ali (2018). Using Energy-efficient Trust Management to Protect IoT Networks for Smart Cities. Sustainable Cities and Society, S2210670717313136. http://doi:10.1016/j.scs.2018.03.026.

[17] Singh, Amit and Nagaraju, A. (2020). Low latency and energy efficient routing-aware network coding-based data transmission in multi-hop and multi-sink WSN. Ad Hoc Networks, 107(), 102182. http://doi:10.1016/j.adhoc.2020.102182.

[18] SureshKumar, K., &Vimala, P. (2021). Energy efficient routing protocol using exponentially-ant lion whale optimization algorithm in wireless sensor networks. Computer Networks, 197, 108250. http://doi:10.1016/j.comnet.2021.108250.

[19] Jabbar, Waheb A.; Saad, WasanKadhim and Ismail, Mahamod (2018). MEQSA-OLSRv2: A Multicriteria-based Hybrid Multipath Protocol for Energy-Efficient and QoS-Aware Data Routing in MANET-WSN Convergence Scenarios of IoT. IEEE Access, 1–1. http://doi:10.1109/ACCESS.2018.2882853.

[20] Sarma, Hiren Kumar Deva; Borah, Samarjeet and Dutta, Nitul (2019). Advances in Communication, Cloud, and Big Data Volume 31 || Survey on Energy-Efficient Routing Protocols in Wireless Sensor Networks Using Game Theory. , 10.1007/978-981-10-8911-4(Chapter 1), 1–9. http://doi:10.1007/978-981-10-8911-4_1.

[21] Xiao, B.; Yu, B.; Gao, C. CHEMAS: Identify suspect nodes in selective forwarding attacks. J. Parallel Distrib. Comput. 2002, 67, 1218–1230