Journal of Information Hiding and Multimedia Signal Processing ©2025 ISSN 2073-4212
Ubiquitous International Volume **16**, Number **2**, June **2025**

A Novel lightweight secure routing mechanism using Deep Reinforcement Learning for Body Area Sensor Networks

D. J. Jagannath[1], K. Martin Sagayam[1], D. Raveena Judie Dolly[1]

[1]Department of Electronics and Communication Engineering,
Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu, India
{jagan, martinsagayam, dollydinesh}@karunya.edu

J. Dinesh Peter[2]

[2]Department of Artificial Intelligence and Machine Learning,
Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu, India
dineshpeter@karunya.edu

Linh Dinh[3]

[3]Department of Information Systems,
Suffolk University, Boston, MA, USA
linhgdinh2412@gmail.com

Hien Dang[4,5,*]

[4]Faculty of Computer Science and Engineering, Thuyloi University, Hanoi, Vietnam
[5]Department of Mathematics and Computer Science, Molloy University, Rockville Centre, NY, USA
hiendt@tlu.edu.vn

[*]Corresponding author: Hien Dang

Abstract. *Smart healthcare technologies in the cutting-edge fields of the Internet of Things (IoT), Internet of Medical Things (IoMT), and engineering technology provide smart healthcare solutions using Body Area Sensor Networks (BASN). However, the efficient routing of information and information security are critical aspects of medical sensor networks (MSN) because of the sensitive nature of health data and the potential risks associated with unauthorized access, manipulation, or disclosure. Ensuring efficient routing with the security of medical sensor networks involves various measures and considerations including confidentiality, integrity, and availability. This study deals with an Artificial Intelligence (AI)-based lightweight secured routing mechanism using Deep Reinforcement Learning methodology. The proposed methodology is a reliable reinforcement-learning-based routing mechanism (RRR); hence, it was named R3. The prominence of this work includes a novel reliable reinforced routing mechanism, optimized decision-making by the proposed AI methodology, optimized routing of data, utilization of the least possible energy for data transmission, high reliability by integrating trust in the algorithm, and a faster shortest path algorithm with fewer than four hops to yield the shortest path for various traffic rates. The performance of the methodology was intensively evaluated against several potential attacks. The proposed R3-MedNet methodology was compared with four traditional techniques: (distributed energy-efficient clustering algorithm (DEEC), Stable Election Protocol (SEP), LEACH (low-energy adaptive clustering hierarchy (LEACH), and HEED (Hybrid Energy-Efficient Distributed clustering). Significant variations were observed in the energy levels of the nodes during various epoch stages. Simulation results prove that the proposed R3-MedNet methodology is resilient to various attacks.*

Keywords: IoMT-Internet of Medical Things, BASN-Body Area Sensor Networks, Deep Reinforcement Learning, MSN - medical sensor networks, Reliable Reinforced Routing, security measures.

---

1. **Introduction.** "Smart" is the key word for Technocrats and Engineers in the modern world of technology. The number of smart healthcare systems is increasing rapidly every day. Smart healthcare systems leverage advanced technologies to enhance efficiency, accessibility, and quality of services. These systems integrate various components, including sensors, devices, data analytics, and communication platforms to create a connected and intelligent healthcare environment.

The key aspects and components of smart healthcare systems include IoT-enabled medical and wearable devices. These include smartwatches, fitness trackers, and other wearable sensors that monitor vital signs, physical activity, and sleep patterns. The data from the BASN of all patients can be collected, deposited, examined, and appropriate treatments can be presented, over the Internet, anytime, anywhere, as shown in fig.1. Medical Sensors are implanted or external sensors that collect real-time data such as glucose levels, heart rate, and blood pressure. Health Information Exchange (HIE) through interoperability enables the seamless sharing of patient information among healthcare providers and systems, ensuring a comprehensive and up-to-date patient record. Electronic Health Records (EHR) are digitized and centralized patient health records that can be easily accessed by authorized healthcare professionals.

Telehealth and Telemedicine for Remote Consultations allow patients to consult healthcare providers remotely using video calls, thereby reducing the need for physical visits. Remote Monitoring of patients' health through connected devices and sensors enables proactive healthcare management. Big Data Analytics analyzes large volumes of healthcare data to derive meaningful insights for personalized treatment plans, population health management, and predictive analytics for disease prevention. Artificial Intelligence (AI)- and machine Learning (ML)-based diagnostic support. AI algorithms can assist in the analysis of medical images, pathology slides, and other diagnostic data to improve the

accuracy and efficiency. Predictive Analytics uses machine-learning models to predict disease outcomes, identify high-risk patients, and optimize treatment plans.

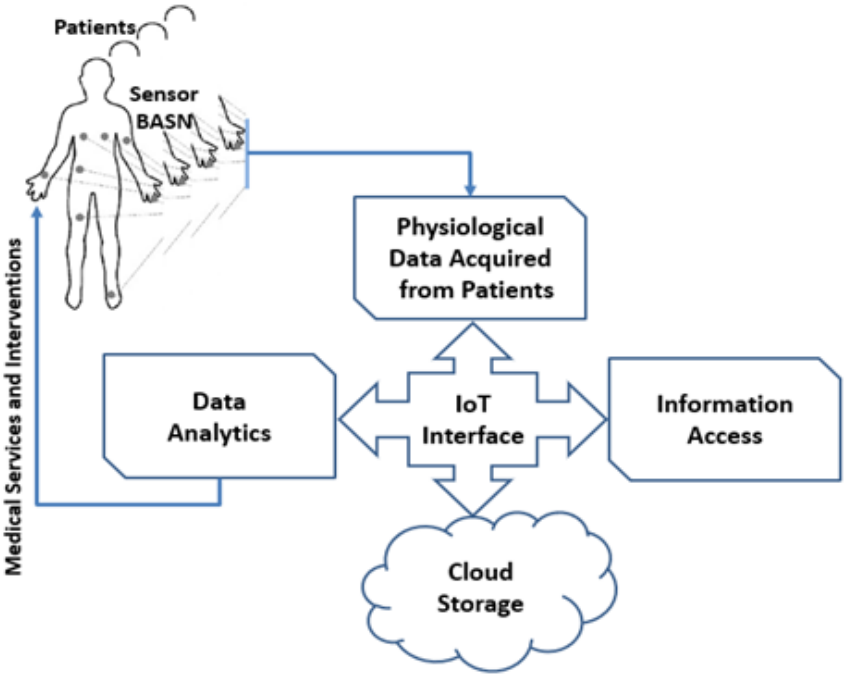

FIGURE 1. Smart Healthcare based on Body Area Sensor Network

The blockchain Technology ensures the security and integrity of healthcare data by employing a blockchain for secure and transparent data storage and sharing. Moreover, it facilitates patient consent management and data-access control. Mobile Health (mHealth) applications enable patients to manage their health through applications that provide medication reminders, health monitoring, and access to health-related information. Support preventive care and wellness programs through personalized recommendations. IoT-based Infrastructure for hospital asset tracking, environmental monitoring, and energy management. Automation and robotics for tasks such as surgery, medication dispensing, and logistics within healthcare facilities. Smart healthcare systems have the potential to improve patient outcomes, reduce healthcare costs, and enhance overall healthcare experience.

However, they also encounter challenges related to data privacy, interoperability, and ethical considerations, which must be carefully addressed for successful implementation. Securing a routing mechanism in medical sensor networks is crucial for ensuring the confidentiality, integrity, and availability of sensitive health data. Some of the strategies and mechanisms that can enhance the security of routing in medical sensor networks include Secure Routing Protocols designed with security. Protocols such as the Secure Efficient Ad hoc Distance Vector (SEAD) or Ariadne are specifically designed for secure ad hoc networks and can be adapted for medical sensor networks [1]. Implementation of protocols that support secure key management and distribution to protect communication channels [2]. Cryptographic techniques such as encryption can protect data confidentiality during transmission [3]. This helps to prevent eavesdropping attacks. Digital signatures can be employed for data integrity verification, ensuring that the data received by medical devices are not tampered with during transit [4]. The implementation of strong authentication mechanisms ensures that only authorized devices can participate in a network [5]. Access

control mechanisms are employed to restrict access to sensitive information based on the roles and privileges of the devices in the network [6]. Intrusion detection and prevention systems have been deployed to monitor networks for suspicious activities or anomalies [7]. These systems can help identify and mitigate potential security threats in real-time. Establishing trust models within the network to assess node reliability. Trust-based routing can be used to prefer routes through trustworthy nodes, thereby reducing the risk of malicious activity [8]. Implementing secure neighbor discovery mechanisms to verify the authenticity of neighboring nodes. This prevents attackers from impersonating legitimate nodes and disrupts the routing process [9]. This ensures that the nodes in a medical sensor network are synchronized in time [10]. Time synchronization helps prevent attacks such as replay attacks and contributes to accurate and reliable data correlation. If a medical sensor network relies on geographic routing, consider employing secure geographic routing protocols [11] that protect location information and prevent attackers from maliciously altering their routing decisions. Designing security mechanisms is always a challenging task considering the limited energy resources of medical sensors. Energy-efficient cryptographic algorithms and protocols [12] can help to minimize the impact on battery devices. Regular updates of the firmware and software of devices in the network are necessary to patch vulnerabilities and enhance the overall security. This includes maintaining the routing protocols and security mechanisms. Implementing physical security measures, such as securing access points and sensor nodes from physical tampering, is another way to protect the infrastructure of medical sensor networks.

Although, this research addresses challenges commonly found in Wireless Sensor Networks (WSNs), it is particularly relevant to Body Area Sensor Networks (BASNs) or Medical Sensor Networks (MSNs) which are specialized subcategories of WSNs. BASNs inherit the fundamental characteristics of WSNs, such as wireless communication, distributed sensing, and energy constraints, but are further challenged by stricter requirements on power consumption, latency, and reliability due to their vital use in continuous health monitoring. The proposed methodology in this article are evaluated within a general WSN context, are directly applicable to BASNs, especially in wearable and implantable sensor networks.

A combination of these security measures, tailored to the specific requirements and constraints of medical sensor networks, can help establish a robust and secure routing mechanism for the reliable and safe transmission of healthcare data. Regular security assessments and updates should be part of the ongoing maintenance plans for these networks. This study deals with an AI-based lightweight secured routing mechanism using Deep Reinforcement Learning methodology. The proposed R3-MedNet methodology is a Reliable Reinforcement learning-based routing mechanism. The performance of the methodology was intensively evaluated against several potential attacks. Simulation results prove that the proposed R3-MedNet methodology offers efficient routing and is resilient to various attacks.

1.1. **Related Works.** Medical sensor networks, like other networked systems, are vulnerable to various security attacks that can compromise the confidentiality, integrity, and availability of sensitive health data. There are numerous active, passive, and physical security attacks on medical sensor networks (MSNs). To address these security challenges, it is crucial to implement robust security measures, such as encryption, authentication, access control, and regular security audits, to protect the confidentiality and integrity of medical data in sensor networks. Additionally, ongoing research and development is

essential to avoid emerging threats and vulnerabilities in this rapidly evolving field. Artificial Intelligence (AI) can enhance the security of medical sensor networks by providing advanced threat detection, anomaly detection, and automated response capabilities.

Padmalaya et al. [13] reported the use of machine-learning techniques for routing in wireless sensor networks. All existing methodologies and their advantages and disadvantages have been reported. Qamar et al. [14] reported secure routing and monitoring schemes for sensor networks. To improve the secure routing process in wireless sensor networks, a deep routing protocol for low-power and lossy software-defined networks (DRPL_SDN) was proposed. The proposed DRPL_SDN exhibited a packet loss of 236 and energy consumption of 6%. Al-Jerew et al. [15] proposed a data-gathering reinforcement-learning algorithm. The proposed algorithm uses a reward function to select a set of Cluster Heads to balance the energy-saving and data-gathering latencies of a mobile Base Station. This work was reported to be the best in terms of the mean length of a mobile BS tour and the network's lifetime.

Chandrasekar et al. [16] reported a hybrid deep learning approach using a deep neural network and reinforcement learning algorithm to preserve network connectivity and improve wireless sensor network coverage. This technique provides a better balance between the network coverage and lifetime. Chen et al. [17] studied the optimization problem of computational offloading and resource allocation in edge computing for health care services. This study demonstrated low energy consumption and computation latency. Moreover, this technique demonstrates an increase in the utility of wireless body area networks. Prabhu et al. proposed a multiple agent-based reinforcement learning technique for energy-efficient routing [18]. The technique showed improvements in the packet delivery rate, average latency, energy consumption, and network lifetime. He et al. [19] proposed an effective system to unravel the routing optimization problem by adding a graph neural network structure to Deep Reinforcement Learning (DRL), called message-passing deep reinforcement learning. The aim was to achieve a load balance for network traffic and improve network performance.

Numerous research works have explored BSN-specific solutions, particularly in health monitoring, focusing on energy efficiency, data reliability, and secure communication. This research work aligns with these goals by addressing reliability, energy efficiency and security which are crucial not only for general WSNs but also for BSNs where node placement on the human body and power limitations demand high performance under strict constraints. By positioning the proposed methodology in the broader WSN landscape, we ensure scalability and transferability to body-centric use cases.

Bangotra et al. [20] proposed a routing protocol mechanism using a machine-learning technique to select a relay node from a list of potential forwarder nodes to achieve energy efficiency and reliability in the network. The methodology showed that it saves energy; hence, remote patients could connect with healthcare services for a longer duration with the integration of IoT services. Li et al. [21] proposed a machine-learning-based approach for collecting data from multiple sensor devices in IoT. A genetic algorithm and deep reinforcement learning were used in this investigation. The aim was to improve the coverage ratio of large data collections and reduce the collection costs in smart IoT-enabled systems. Frikha et al. [22] conducted a comprehensive survey of the existing applications of reinforcement learning and deep reinforcement learning techniques in IoT communication technologies and networking. The study also analyzed other techniques that apply these methodologies to wireless IoT to resolve issues related to routing, scheduling, resource allocation, dynamic spectrum access, energy, mobility, and caching. Haiyun Ma et.al [23] presented a data aggregation scheme tailored for wireless sensor networks (WSNs).

It influences digital signatures and homomorphic encryption to ensure data privacy and integrity during aggregation processes, addressing key security concerns in WSNs.

2. **System Modelling.** Modelling a medical Wireless Body Area Sensor Network (WBASN) involves creating a representation of a system that captures its key components, interactions, and behaviors. This can be helpful in designing and optimizing a network, assessing its performance, and identifying potential challenges. This study involved randomly organized heterogeneous sensors within a geographic area (a medical centre), as depicted in figure 2. The sensor network is a multihop star topology network composed of three node variants: reduced function nodes, full function nodes, and cluster sink nodes. In figure 2, the orange nodes are the reduced function nodes, green nodes are the full function nodes, and blue nodes represent the cluster sink nodes. The reduced function nodes (orange) are the nodes that sense several biosignals and transmit these data to the green nodes. The full-function nodes (green) are information consolidators and transmitters to the cluster sink nodes. The cluster sink nodes (blue) are the network organizers and information transmitters at the base station. These heterogeneous sensors communicated with the central base station at various initial energies.

The sensor nodes transmit information to the cluster sink nodes, which, in turn, transmit information to the base station directly or indirectly through a mobile device. All nodes provide information to the base station about their location-global Positioning System (GPS), heterogeneous energy type, and energy level. The geographic area in our experiments was limited to a 100m x 60 m indoor environment, which consisted of a medical wireless sensor network. In compliance with the IEEE 802.15.6 standard the maximum number of sensor nodes that can be accommodated within this environment was 128. The range for sensor node communication was 5m using relay frames. Hence, a more sophisticated routing protocol is imperative to transmit frames from sensor nodes through the base station to the final server for Physiological Data Acquisition.

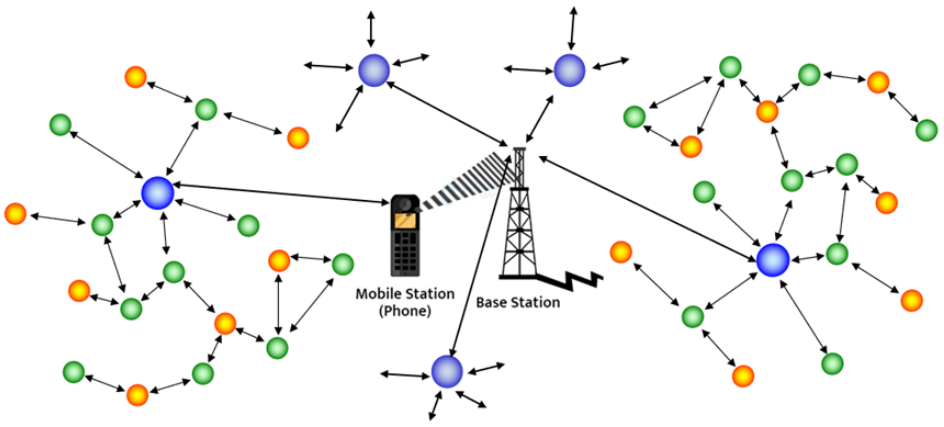

FIGURE 2. Medical Wireless Body Area Sensor Network

The energy consumption in a wireless body area sensor network (WBASN) is often modelled based on the communication and sensing activities of individual nodes. One common model is the energy consumption for transmitting and receiving data as well as the energy consumed during sensing. The energy consumed by a node in a WBASN can be expressed as:

$$E^{total} = E^{tx} + E^{rx} + E^{S} + E^{I} \tag{1}$$

where E^{total} is the total energy consumed by node. E^{tx} is the energy consumed during data transmission. E^{rx} is the energy consumed during data reception. E^S is the energy consumed during sensing or data acquisition. E^I represents the energy consumed when a node is idle. The specific mathematical expressions for each component may depend on the characteristics of the wireless communication hardware and sensing equipment used in the WBASN. Hence, in our work;

$$E^{tx} = E_{amp} \cdot d^{\alpha} \cdot L \tag{2}$$

Where: E_{amp} is the energy required to transmit a bit per unit distance. d denotes the transmission distance. α is the path loss exponent. L is the size of the transmitted packet.

$$E^{rx} = E_{amp} \cdot L \tag{3}$$

$$E^S = E_S \cdot N \tag{4}$$

Where E_S is the energy required for sensing per sample. N is the number of samples obtained during the sensing process.

$$E^I = P_I \cdot T_I \tag{5}$$

Where P_I is power consumption in the idle state. T_I is the time spent in the idle state. The goal is to estimate the energy consumption based on the specific characteristics and requirements of the WBASN in question. Find the shortest path in a network for any network tree T with n nodes. The amount of energy dissipated by a node for a single data collection, Ed_{Tn} can be estimated using equation (6).

$$Ed_{Tn} = L \cdot E^{rx} \cdot I_g \cdot (CT_n - 1) + L \cdot E^{tx} \cdot I_g \cdot CT_n \tag{6}$$

Where CT_n is the number of nodes that represent a subtree along with a root node n. The information generation rate is denoted as I_g.

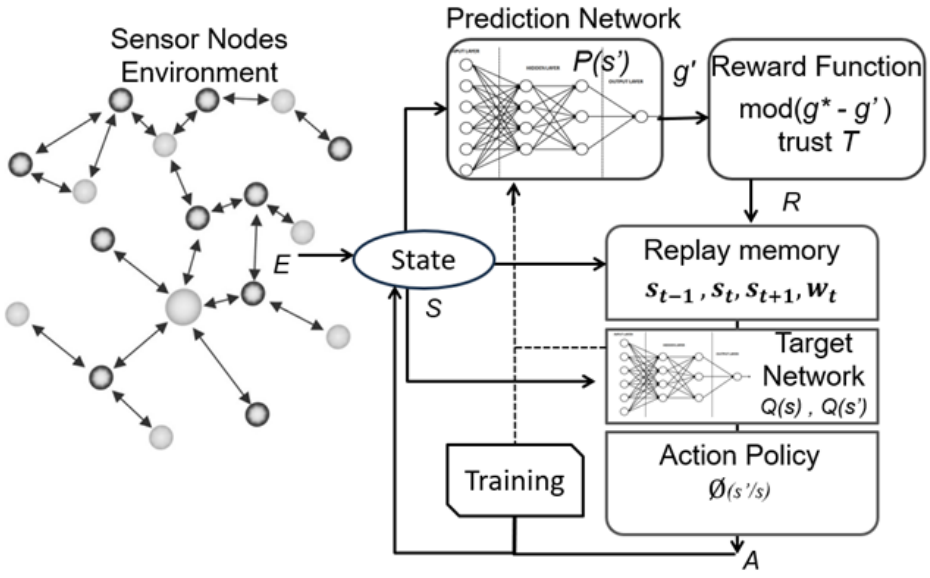

FIGURE 3. Proposed novel Reliable Reinforced Routing- R3-MedNet

3. **The Proposed R3-MedNet Methodology.** The proposed Reliable Reinforcement Routing- R3-MedNet methodology is illustrated in Figure 3. The methodology was implemented by applying Deep Q-Networks (DQN) to Wireless Body Area Sensor Networks (WBASN) using reinforcement learning techniques to optimize decision-making processes

within the context of WSNs. Wireless Sensor Networks (WSN) typically consist of numerous small energy-constrained sensor nodes, and their goal is often to optimize energy efficiency, data transmission, or sensing tasks. The ultimate goal of this study was to optimize the routing of data by utilizing the least possible energy for data transmission in a WBASN. That is, finding possible paths that consume less energy by connecting each node to the cluster sink nodes in the sensing environment. Energy efficiency was achieved by considering the available and utilized energy levels of the nodes.

The WBASN represents the sensor node environment E, which consists of reduced-function nodes, full-function nodes, and cluster sink nodes. The state space is denoted by S, which represents all nodes of the network, their energy levels, and the number of hops to the cluster sink node. The reward function R evaluates the difference between the estimated optimum path and the target path. Action space A estimates next-hop routing. The replay memory stores all necessary past and present data for training. The Q-function approximation technique has been used in studies using the DQN model, as reported by Mnih et al. [24]. A high-dimensional environment (E) offers an initial state (g) that is observable to some extent over the state observation (S). Deep Reinforcement Learning using the DQN model estimates $Q(s, s')$ given the input state observation (s) and the possible optimum path state (s'). The optimum path $Q^{\lambda}(s, s')$ is responsible for the rewards based on the action policy $\emptyset(s'/s)$. The estimate $Q^{\lambda}(s, s')$ is the updated Q value of node i for state S^i and action A^i as future rewards. The criterion for an optimum path state (s') given an observation state (s) depends on the action policy $\emptyset$. The replay memory holds the dataset for training the DQN that records the entire data history. The prediction network provides an estimate of the optimum path $g' = R(s)$ from which the rewards and difference between the target and estimated routing path can be computed as the reward function. Reliability is ensured by integrating trust in the reward function, as shown in Equation (7), which makes the training/learning agent choose the optimum reliable routing path.

$$R_{t+1} = \begin{cases} -(1-T) \text{ if } s_o \neq null \\ -(1-T) \text{ if } s_o = null \text{ for all } s_o \\ 0 \text{ otherwise} \end{cases} \tag{7}$$

where, R_{t+1} is the reward function for a node 'a' connecting to a node 'b' as the next hop node after a time 'T.' The 'Trust value for the node 'a' over the node 'b is denoted as 'T.' The state observations of node 'a' for node 'b is denoted by 's_o' at time 't.' We considered that all the sensor nodes possessed equal initial energy levels. The optimization of the energy and energy efficiency can be obtained using Equation (8).

$$E_{opt} = \frac{E^r}{E^i} + \frac{E^r}{E^a}(J) + (E^{tx} + E^{rx})(1-J) \tag{8}$$

where Er denotes the energy remaining in a node and is denoted by E^r. The initial energy of the node is denoted by E^i. The average energy of the optimum path connecting all nodes to the cluster sink node is denoted by E^a. The Reliable Reinforced Routing- R3 algorithm is summarized in Table 1. Energy optimization is accomplished based on the estimated optimum routing path such that the overall energy consumption is the lowest when using a deep reinforcement learning algorithm. Hence, the Reliable Reinforcement Routing- R3 algorithm evaluates the best possible route with minimum energy without disturbing the network.

TABLE 1. The R3 Algorithm

1	**Input:**
2	Reward Function $mod(g^* - g')$ trust T
3	T - The trust table
4	Q table – the Q values
5	**Output:** best possible network route with minimum energy
6	Initialization of Prediction network and Target network
7	initial state (g)
8	state observation (S)
9	estimates $Q(s, s')$ optimum path $Q^{\lambda}(s, s')$
10	rewards based on the action policy $\emptyset(s'/s)$
11	updated Q value of node i, for the state S_t^i and the action A_t^i
12	R_{t+1} reward function
13	E_{opt} energy optimization
14	While True do
15	Loop and wait 't'
16	Q_{max} choose and transmit
17	Proceed to next iteration for Q update
18	If $Q(s, s')$ is true
19	optimum path $Q^{\lambda}(s, s')$ estimate trust T $R_{t+1} = \begin{cases} -(1-T) \text{ if } s_o \neq null \\ -(1-T) \text{ if } s_o = null \text{ for all } s_o \\ 0 \text{ otherwise} \end{cases}$
20	else
21	updated state S_t^i and the action A_t^i
22	end
23	end

4. **Results and Discussions.** The proposed reliable reinforced routing- (R3-MedNet) methodology was tested, with the implications mentioned in Sections 3 and 4. The investigation was carried out for five different heterogeneous network sizes of 48, 68, 88, 108, and 128 in compliance with the IEEE 802.15.6 protocol standard. The range for sensor node communication was 5m using relay frames.

The simulation environment of a 48-node heterogeneous R3-MedNet network with three cluster sink nodes is illustrated in Fig. 4. All the parameters of the implications mentioned in Sections 4 and 5 and the simulation values are tabulated in Table 2. The investigation was also extended using four other commonly used algorithms by the majority of the latest research: distributed energy-efficient clustering algorithm (DEEC), Stable Election Protocol (SEP), low-energy adaptive clustering hierarchy (LEACH), and Hybrid Energy-Efficient Distributed clustering (HEED) [25]. All the algorithms were implemented and tested in the same environment, standards, and test parameters.

The proposed R3-MedNet methodology was compared with all the techniques mentioned in the previous paragraph. Significant variations were observed in the energy levels of the nodes during various epoch stages. The maximum number of rounds was limited to 1600 epochs during the investigation phase. The initial energy of all the nodes

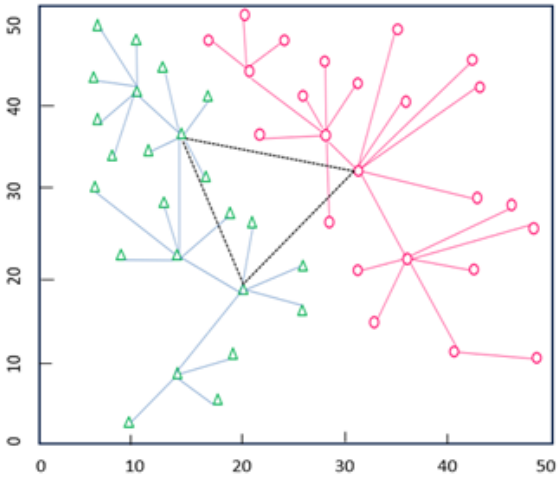

FIGURE 4. Illustration of a 48-node R3-MedNet heterogenous sensor network with three cluster sink nodes

TABLE 2. The R3-MedNet Simulation Parameters

Parameters	**Values**
Medical Sensor Network Area Size (m^2)	100 X 60
Base Station location	50, 50
Number of sensor nodes	48, 68, 88, 108, 128
Heterogeneous nodes variants	heterogeneous sensors composed of three variants of nodes; reduced function nodes, full function nodes and cluster sink nodes.
Ratio of heterogeneous nodes variants	01:07:13
Ideal - reduced function nodes initial energy Ei (j)	200
Ideal - full function nodes initial energy Ei (j)	400
Ideal - cluster sink nodes initial energy Ei (j)	800
Energy to transmit a bit per unit distance E_{amp} (j)	0.05
Energy required for sensing per sample E_S (j)	0.03
Power consumption in the idle state P_I (j)	0.002
Radio Range (m)	5
Traffic rate - exponential (packets/sec)	2,4,6,8,10 and 12
Weighting factor	0.5
Epochs	1600
Learning rate α	0.01

was 100%, which started to depreciate as the number of epochs increased. This phenomenon is graphically plotted in figure 5, with the average percentage of energy remaining in the nodes after every 100 epochs for all the five algorithms investigated.

The proposed R3-MedNet methodology outperformed other traditional algorithms in terms of energy dissipation. The proposed R3-MedNet methodology resulted in 37.84% of the energy remaining after 1600 epochs. The HEED, LEACH, SEP, and DEEC algorithms resulted in 31.89, 30.214, 19.81, and 14.983% of the energy remaining after 1600 epochs, respectively. This provides substantial proof of the significance of the proposed R3-MedNet methodology.

Conversely, the energy utilization of the nodes was calculated with respect to the number of epochs. This phenomenon is graphically plotted in figure 6, with the average percentage of energy utilized by the nodes after every 100 epochs for all the five algorithms investigated. The proposed R3-MedNet methodology outperformed other traditional algorithms in terms of energy utilization. The average energy utilized by the proposed R3-MedNet methodology was 62.16% after 1600 epochs. The HEED, LEACH, SEP, and DEEC algorithms resulted in 68.11%, 69.786%, 80.19%, and 85.017% energy utilization, respectively, after 1600 epochs. The energy efficiency study provided substantial proof of the significance of the proposed R3-MedNet methodology.

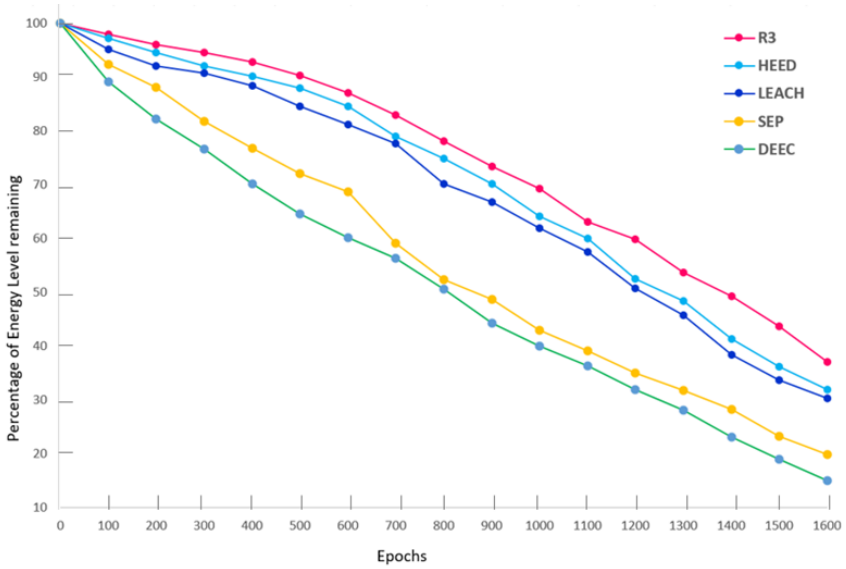

FIGURE 5. Percentage of energy level remaining

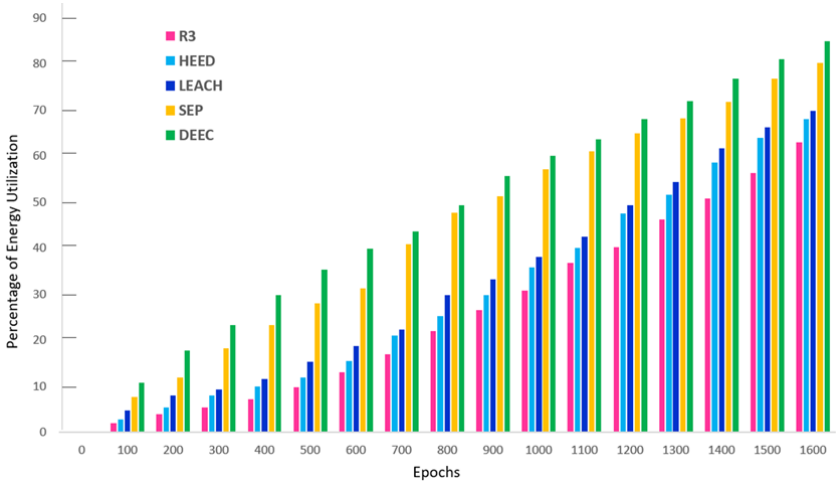

FIGURE 6. Energy Utilization

The experiments were further evaluated using the delivery ratio parameters for all algorithms that were investigated. The algorithms were tested at traffic rates of 2, 4, 6, 8, 10, and 12 packets per second (packets/s). The delivery ratio in each case was evaluated for all the test traffic rates, as depicted in figure 7. The proposed R3-MedNet methodology revealed an average delivery ratio of 95.15% for a traffic rate of 2 packets/s, and 98.89%

for a traffic rate of 12 packets/s. This technique showed a very good convergence for all traffic rates.

None of the other algorithms exhibited convergence comparable to that of the proposed R3-MedNet methodology. The HEED algorithm revealed 63.32% to 82.895% for 2 and 12 packets/s, respectively. The convergence was good after 4 packets/s. The LEACH algorithm showed 41.72%–69.36333% accuracy for 2 and 12 packets/s, respectively. However, the convergence was found to be unsatisfactory. The SEP algorithm showed 18–47.955% accuracy for 2 and 12 packets/s, respectively. However, the convergence was found to be unsatisfactory. DEEC ranged from 11% to 35.78833% for 2 and 12 packets/s, respectively. The convergence was good for traffic rates of more than 6 packets/s. The average delivery ratio for each of the algorithms tested; R3-MedNet, HEED, LEACH, SEP and DEEC were, 97.562%, 82.895%, 69.364%, 47.955% and 35.789% respectively.

Moreover, the number of hops required to yield the shortest path was evaluated for all the algorithms tested. The results of this evaluation are shown in Figure 8. The proposed R3-MedNet methodology was again found to be significant compared with the other algorithms. The shortest path estimation was achievable with fewer than four hops for various traffic rates.

Wireless Body Area Sensor Networks (WBASNs) face several internal threats that can compromise their operation, security, and reliability. Internal threats typically arise within a network, and can be intentional or unintentional. Numerous internal attacks can occur in a WBASN. In our investigations, we intentionally created a few malicious nodes, and experiments were conducted. The number of malicious nodes increases in the following order: 8, 16, 32, 40, and 48. The outcome of this investigation was the reliability factor plotted in figure 9. The proposed R3-MedNet methodology is significantly more reliable even with more malicious nodes. The reliability factor is approximately 95% when there are no malicious nodes. When approximately 48 malicious nodes are present, the reliability factor is 82%. The reliability factors of the other algorithms decreased significantly as the number of malicious nodes increased, as shown in figure 9. Hence, the proposed R3-MedNet methodology proved to be more reliable than the other traditional algorithms.

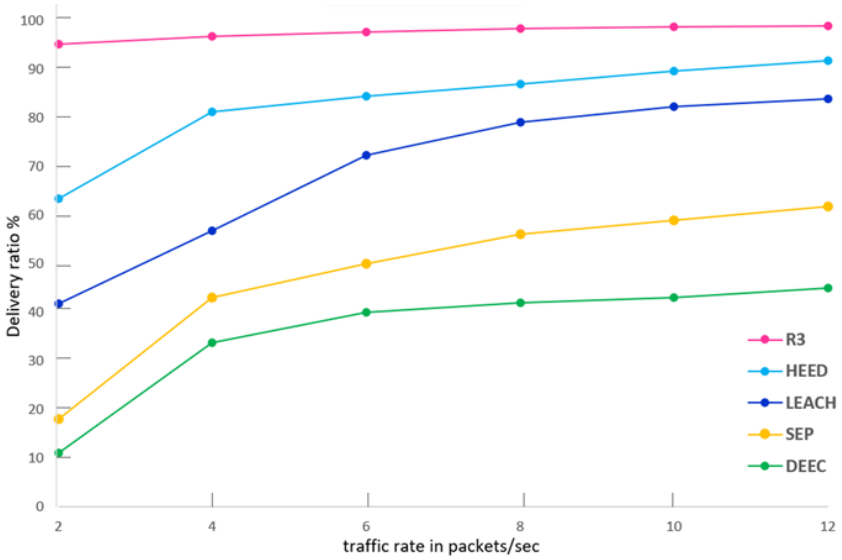

FIGURE 7. Delivery Ratio

Moreover, Deep Reinforcement Learning has emerged as a promising technique for secure routing in BASNs, owing to its ability to adaptively select optimal routes while

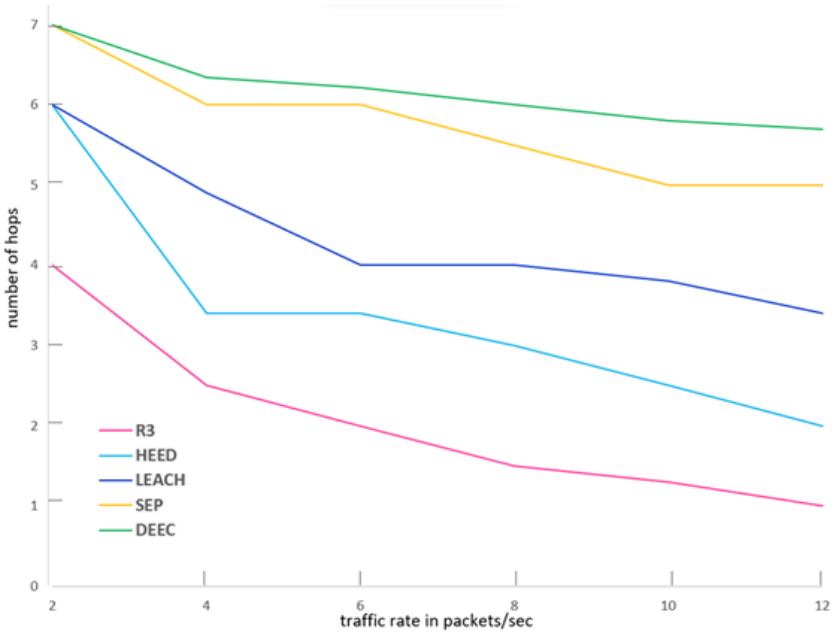

FIGURE 8. Hops to shortest path

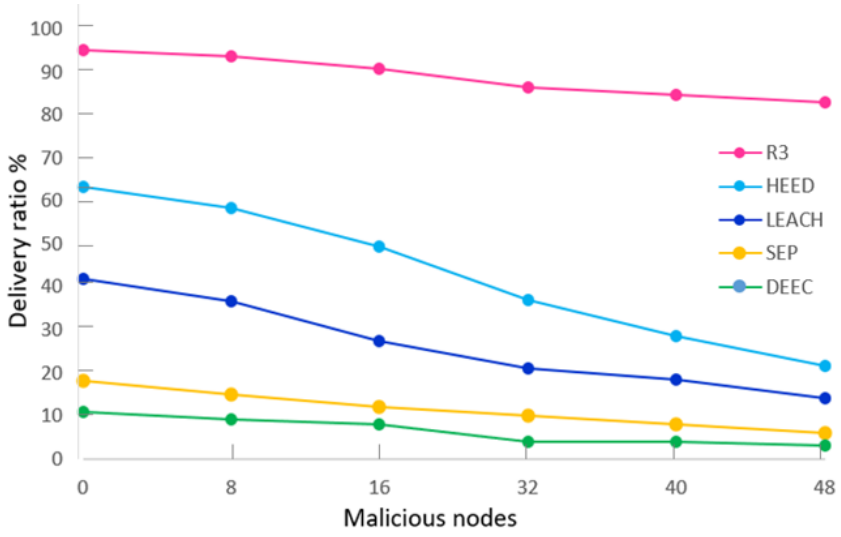

FIGURE 9. Reliability factor

considering factors such as security, energy efficiency, and network congestion. Several techniques that use DRL for secure routing have been proposed, each with various strategies and advantages. (1) Double Deep Q-Network (DDQN) for Secure Routing. The use of two Q-networks can increase computational load, particularly in resource-constrained BASNs. Training a DDQN is computationally intensive and requires careful tuning of the hyperparameters. (2) Proximal Policy Optimization (PPO) for Secure Routing. The performance of PPO is sensitive to the choice of hyperparameters, which may require extensive tuning. A PPO can struggle in very large networks with many nodes because it may require significant computational resources for training. (3) Deep Deterministic Policy Gradient (DDPG) for Secure Routing. DDPG may require a large number of samples to converge to an optimal solution, particularly in a dynamic network. Similar to other reinforcement learning models, DDPG can be prone to overfitting if not properly

regularized. (4) Multi-Agent Deep Reinforcement Learning (MADRL) for Secure Routing. The coordination between multiple agents can be complex, particularly when dealing with adversarial attacks or network failure. MADRL requires substantial communication between agents, which may add overhead in terms of the network traffic and latency. (5) Deep Convolutional Neural Networks (CNNs) for Secure Routing. CNNs require a large amount of labelled data for training, which may be difficult to collect in a dynamic BASNs. CNNs are better suited for classification tasks than continuous decision-making processes such as routing. The ultimate goal of the proposed method is to optimize the routing of data by utilizing the least possible energy for data transmission in a WBASN. The proposed novel R3-MedNet methodology proved to be significantly more reliable and optimized by the proposed AI methodology, with optimized routing of data utilizing the least possible energy for data transmission. Moreover, it offers high reliability by integrating trust into the algorithm. The proposed methodology yielded the shortest path for various traffic rates with fewer than four hops.

5. **Conclusion.** The proposed reliable reinforced routing- (R3-MedNet) methodology was investigated using randomly organized heterogeneous sensor networks within a geographic area of 100m x 60m size indoor environment. The sensor network uses a multi-hop star topology network with three node variants: reduced function nodes, full function nodes, and cluster sink nodes. In compliance with the IEEE 802.15.6 standard the maximum number of sensor nodes that can be accommodated within this environment was 128. The range for sensor node communication was 5m using relay frames.

The methodology was implemented by applying Deep Q-Networks (DQN) to Wireless Body Area Sensor Networks (WBASN) using reinforcement learning techniques to optimize decision-making processes within the context of WSNs. The ultimate goal of this study is to optimize the routing of data by utilizing the least possible energy for data transmission in a WBASN. That is, finding possible paths that consume less energy by connecting each node to the cluster sink nodes in the sensing environment. Energy efficiency was achieved by considering the available and utilized energy levels of the nodes. Energy optimization is accomplished based on the estimated optimum routing path such that the overall energy consumption is the lowest when using a deep reinforcement learning algorithm.

Hence the proposed novel Reliable Reinforcement Routing- R3 algorithm was able to estimate the best possible route with minimum energy without disturbing the network energy balance. Reliability was ensured by integrating trust in the reward function, which made the training/learning agent choose the optimal reliable routing path. The proposed R3-MedNet methodology outperformed other traditional algorithms and proved to be the best in terms of energy dissipation and utilization. The proposed R3-MedNet methodology is significantly more reliable even with a greater number of malicious nodes. The reliability factor is approximately 95% when there are no malicious nodes. When approximately 48 malicious nodes are present, the reliability factor is 82%. The reliability factors of other algorithms decreased significantly as the number of malicious nodes increased. Moreover, the number of hops required to yield the shortest path was achievable in less than four hops for various traffic rates, which was much faster with fewer hops compared with the other algorithms. The proposed R3-MedNet methodology was evaluated to be much more reliable; hence, the name R3-MedNet is a reliable reinforced routing mechanism. The future work of this research is to investigate a security methodology with possible external attacks on medical sensor networks.

Acknowledgements. This study was funded in part by a Molloy University research grant. We would like to thank all of our universities, institutes, and organizations for their time and support in this study.

REFERENCES

[1] Y.-C. Hu, D. B. Johnson, and A. Perrig, "Secure Efficient Ad hoc Distance Vector Routing," in *Proc. Fourth IEEE Workshop on Mobile Computing Systems and Applications (WMCSA'02)*, 2002.
[2] Q. Lin *et al.*, "Secure Internet of Medical Things (IoMT) Based on ECMQV-MAC Authentication Protocol and EKMC-SCP Blockchain Networking," *Information Sciences*, vol. 654, Art. 119783, 2024.
[3] W. Zhang *et al.*, "A Lightweight Security Model for Ensuring Patient Privacy and Confidentiality in Telehealth Applications," *Computers in Human Behavior*, Art. 108134, 2024.
[4] A. S. Nadhan and I. J. Jacob, "Enhancing Healthcare Security in the Digital Era: Safeguarding Medical Images with Lightweight Cryptographic Techniques in IoT Healthcare Applications," *Biomedical Signal Processing and Control*, vol. 88, Art. 105511, 2024.
[5] T. Arpitha, D. Chouhan, and J. Shreyas, "Anonymous and Robust Biometric Authentication Scheme for Secure Social IoT Healthcare Applications," *Journal of Engineering and Applied Science*, vol. 71, no. 1, pp. 1–23, 2024.
[6] Q. Lin *et al.*, "Secure Internet of Medical Things (IoMT) Based on ECMQV-MAC Authentication Protocol and EKMC-SCP Blockchain Networking," *Information Sciences*, vol. 654, Art. 119783, 2024.
[7] A. Paya *et al.*, "Apollon: A Robust Defense System against Adversarial Machine Learning Attacks in Intrusion Detection Systems," *Computers & Security*, vol. 136, Art. 103546, 2024.
[8] A. Habbal, M. K. Ali, and M. A. Abuzaraida, "Artificial Intelligence Trust, Risk and Security Management (AI TRiSM): Frameworks, Applications, Challenges and Future Research Directions," *Expert Systems with Applications*, vol. 240, Art. 122442, 2024.
[9] V.-V. Vo *et al.*, "Active Neighbor Exploitation for Fast Data Aggregation in IoT Sensor Networks," *IEEE Internet of Things Journal*, 2024.
[10] H. Wang *et al.*, "Clock Synchronization with Partial Timestamp Information for Wireless Sensor Networks," *Signal Processing*, vol. 209, Art. 109036, 2023.
[11] D. W. Wajgi and J. V. Tembhurne, "Localization in Wireless Sensor Networks and Wireless Multimedia Sensor Networks Using Clustering Techniques," *Multimedia Tools and Applications*, pp. 1–51, 2023.
[12] V. Kalaivani, "Enhanced BB84 Quantum Cryptography Protocol for Secure Communication in Wireless Body Sensor Networks for Medical Applications," *Personal and Ubiquitous Computing*, vol. 27, no. 3, p. 875, 2023.
[13] G. K. Swetha, G. Padmalaya Nayak, S. Gupta, and K. Madhavi, "Routing in Wireless Sensor Networks Using Machine Learning Techniques: Challenges and Opportunities," *Measurement*, vol. 178, 2021.
[14] S. Qamar, "Optimal Sensor Network Routing with Secure Network Monitoring Using Deep Learning Architectures," *Neural Computing & Applications*, vol. 35, pp. 19039–19050, 2023.
[15] O. Al-Jerew, N. Al Bassam, and A. Alsadoon, "Reinforcement Learning for Delay Tolerance and Energy Saving in Mobile Wireless Sensor Networks," *IEEE Access*, vol. 11, pp. 19819–19835, 2023.
[16] V. Chandrasekar *et al.*, "Hybrid Deep Learning Approach for Improved Network Connectivity in Wireless Sensor Networks," *Wireless Personal Communications*, vol. 128, no. 4, pp. 2473–2488, 2023.
[17] Y. Chen, S. Han, G. Chen *et al.*, "A Deep Reinforcement Learning-Based Wireless Body Area Network Offloading Optimization Strategy for Healthcare Services," *Health Information Science and Systems*, vol. 11, Art. 8, 2023.
[18] D. Prabhu, R. Alageswaran, and S. M. J. Amali, "Multiple Agent-Based Reinforcement Learning for Energy Efficient Routing in WSN," *Wireless Networks*, Art. 111, 2023.
[19] Q. He *et al.*, "Routing Optimization with Deep Reinforcement Learning in Knowledge Defined Networking," *IEEE Transactions on Mobile Computing*, vol. 23, no. 2, 2023.
[20] D. K. Bangotra *et al.*, "An Intelligent Opportunistic Routing Algorithm for Wireless Sensor Networks and Its Application towards E-Healthcare," *Sensors*, vol. 20, no. 14, p. 3887, 2020.

[21] T. Li *et al.*, "DRLR: A Deep-Reinforcement-Learning-Based Recruitment Scheme for Massive Data Collections in 6G-Based IoT Networks," *IEEE Internet of Things Journal*, vol. 9, no. 16, pp. 14595–14609, 2021.

[22] M. S. Frikha *et al.*, "Reinforcement and Deep Reinforcement Learning for Wireless Internet of Things: A Survey," *Computer Communications*, vol. 178, pp. 98–113, 2021.

[23] H. Ma *et al.*, "A Provable Private Data Aggregation Scheme Based on Digital Signatures and Homomorphic Encryption for Wireless Sensor Networks," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 8, no. 3, pp. 536–543, 2017.

[24] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. A. Riedmiller, "Playing Atari with Deep Reinforcement Learning," *CoRR*, abs/1312.5602, 2013.

[25] S. Hassan and M. Ahmad, "Energy Heterogeneity Analysis of Heterogeneous Clustering Protocols," *International Arab Journal of Information Technology*, vol. 19, no. 1, pp. 45–54, 2022.