RI-RPL: a new high-quality RPL-based routing protocol using Q-learning algorithm

Niloofar Zahedy[1] · Behrang Barekatain[1,2] · Alfonso Ariza Quintana[1,3]

Abstract

The lack of a central controller, severe resource constraints, and multi-path data routing have turned data exchanges into one of the fundamental challenges of the Internet of Things. Despite numerous research efforts on various aspects of routing and data exchanges, some fundamental challenges such as the instant negative impacts of selecting the best possible path and the absence of measures to observe the dynamic conditions of nodes still exist. This study introduces a method called RI-RPL, based on the development of the RPL routing protocol, along with the use of reinforcement learning to address these challenges effectively. To achieve this, RI-RPL is designed in three general stages. In the first stage, routers are aligned with optimizing the RPL protocol with a focus on the Q-learning algorithm. In the second stage, based on learning and convergence, changes in the parents' learning in different network conditions are supported. In the third stage, control and management changes are coordinated. The reason for choosing this algorithm is its ability to address the desired challenges effectively without wasting network resources for calculations. Simulation results using the Cooja software show that the proposed RI-RPL method, compared to similar recent methods such as ELBRP, RLQRPL, and RPL, has improved successful delivery rates by 4.03%, 13.26%, and 28.87%, respectively, for end-to-end delay by 3.04%, 9.82%, and 13.12%, respectively, for energy consumption optimization by 10.43%, 28.91%, and 36.35%, respectively, for throughput by 10.23%, 28.45%, and 46.88%, respectively, and for network data loss rate by 15.06%, 34.95%, and 49.66%, respectively.

✉ Behrang Barekatain
  Behrang_Barekatain@iaun.ac.ir

[1] Faculty of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran

[2] Big Data Research Center, Najafabad Branch, Islamic Azad University, Najafabad, Iran

[3] E.T.S.I. Telecomunicación, Universidad de Málaga, Málaga, Spain

1 Introduction

Internet of Things (IoT) networks are the most recent communication technology in the current era and have gained extensive attention from researchers and wide-ranging research in today's world. This technology consists of a large number of nodes, known as wireless sensors, and one or more sink nodes, all equipped with radio communications capabilities, enabling them to communicate with each other through electromagnetic waves [1]. Among the prominent features of IoT, we can mention aspects such as multi-path routing, dynamic topology, wireless network communication, the absence of central infrastructure and controller, complete distribution, and limitations, especially in energy and node communication resources [2, 3]. These characteristics provide unique advantages for IoT and have led to its use in various applications. The most important of these applications include military, industrial, research, medical applications, smart cities, maritime, transportation system automation, and crisis management [4–7]. Despite the advantages and diverse applications of IoT, it faces unique challenges and complexities, such as security, routing, energy consumption, the development of suitable equipment, and more. In IoT, all network members have individual electronic identifiers that allow them to be identified and distinguished [8]. The sensor nodes, due to their extensive distance from the sink node and the radio range limitation, are unable to communicate directly with the sink node. In these circumstances, multi-path communications with the participation of other nodes are used to achieve communication between network members and the sink node [9]. These multi-path communications are performed using routing protocols. The presence of such conditions has made routing and service discussions (including security, quality, stability, reliability, etc.) in IoT particularly relevant [10, 11]. The distributed nature of network operations, particularly in terms of communication and data exchanges, necessitates the cooperation of all nodes. This issue, along with other challenges related to IoT, such as network topology dynamics, wireless communication interference, and severe resource constraints, has turned the design of routing algorithms for these networks into a serious challenge [12, 13]. It can be stated that routing in IoT, due to its extensive challenges and unique characteristics, is a highly sensitive and complex issue. Furthermore, the distributed nature of IoT networks has made routing protocols the foundation of their operation. The presence of these two factors clearly emphasizes the need for the design of efficient routing protocols to ensure continuous and stable network performance.

Various methods and algorithms have been introduced for use in this field so far, but among them, only the RPL[1] protocol has been widely accepted as the routing standard for IoT [14]. This protocol was introduced by the ROLL[2] working group for low-power and lossy networks (especially IoT) and was adopted by the IETF[3] as the routing standard for IoT [15]. Currently, RPL is the most important and widely used

---

[1] Routing Protocol for Low-Power and Lossy Networks.

[2] Routing Over Low Power and Lossy.

[3] Internet Engineering Task Force.

routing protocol for the Internet of Things, with almost all research related to IoT routing based on this protocol's design and development. However, although RPL effectively covers the routing needs of the Internet of Things and addresses them, it still faces various challenges, including increased control overhead, lack of trust establishment capability, absence of measures to observe the dynamic conditions of nodes, negative effects of momentary parent selection, and low communication quality [16]. Among these challenges, the quality and quality of service support in RPL have been a major focus of recent researchers [16]. In general, the RPL standard does not provide any capability to meet the quality requirements of transmitted data, making it an inefficient protocol in this regard. The main goal of the proposed research is to address this issue as one of the most important issues related to the field of IoT routing. RPL, in addition to routing, allows users to optimize and redefine it according to their needs [17]. This capability is provided through the concept of the objective function in the RPL protocol. The current research mainly focuses on optimizing and improving this function. The objective function is responsible for selecting intermediate nodes to form the network communication graph and transmit data through them. In general, RPL has two objective functions, namely, OF0 [18] and ETX [19]. However, it is important to note that this single-criteria assessment of the objective function and the selection of parents based on it not only do not meet the quality needs of transmitted data but also in some cases lead to the selection of inefficient nodes and paths [20].

Numerous studies have been conducted to improve this issue, each in a unique way and with a focus on specific techniques aimed at enhancing this problem [21–40]. However, some critical challenges in this regard have not been adequately covered in previous studies or have not received attention. The existence of these challenges will lead to the performance degradation of previous methods and have negative effects on network performance. Among the most important of these challenges are: (1) momentary negative effects, (2) the absence of measures to observe the dynamic conditions of nodes, and (3) low communication quality.

In the upcoming research, the aim is to optimize and enhance the objective function of the RPL protocol based on the utilization of artificial intelligence techniques, specifically Q-learning. To improve Quality of Service (QoS), the Q-learning criterion, which determines decision-making based on penalties and rewards, is defined based on a multi-criteria evaluation to effectively ensure and support QoS. To address two other issues (the negative effects of momentary computations on objective function evaluation and considering the dynamic conditions of nodes), Q-learning will be employed in a training and convergence step. This will not only enhance data exchanges but also preserve network performance efficiency. Furthermore, it leads to improvements in successful reception parameters, end-to-end delay, energy consumption, data loss rate, and network throughput.

In summary, the contributions of the proposed method can be listed as follows:

- Optimization and improvement of the RPL objective function with the efficiency of reinforcement learning algorithms (specifically and for the first time, the Q-learning algorithm). The reason for choosing this algorithm is the ability to solve the two challenges without wasting network resources on calculations.

- Based on the efficiency of this technique, in addition to improving QoS, it has been tried to eliminate the negative effects of real-time calculations on the evaluation of the objective function and to include the dynamic conditions of the nodes in the calculations of the objective function. To improve QoS, the learning criterion is defined based on multi-criteria evaluation, so that QoS can be provided and supported effectively.
- To solve two important problems (the negative effects of instantaneous calculations on the evaluation of the objective function and considering the dynamic conditions of the nodes), the step of training and convergence of learning Q will be used. With the efficiency of the training step, the selection of parents is not only limited to the moment of sending DIOs but also after the selection of parents, there will be the ability to evaluate the objective function and replace the parents.
- The training step provides the ability to adapt and adapt to changing network conditions, and the convergence step will guarantee the optimal performance of the proposed method.
- In the proposed method, in addition to solving the past challenges related to the momentary negative effects, the lack of measures to consider the dynamic conditions of the nodes, and the low quality of communications, it will be designed and implemented in such a way that the optimal performance of the network will also be within the negative radius of its practical measures.

The structure of the article is as follows: In the second section, we will delve into the research background, followed by an overview of the Q-learning algorithm in the third section. Sections 4, 5 and 6, respectively, discuss the problem statement, details of the proposed method, and evaluation and analysis of results. Finally, the article's findings are summarized in the conclusion section.

2 Preliminary

Based on the previous sections, the RPL objective function is going to be enhanced using artificial intelligence techniques, more precisely, the Q-learning technique. Below, we provide detailed explanations of the key aspects related to this technique.

- Reinforcement learning

Various learning techniques have been proposed for leveraging the concept of learning. One of the most effective and efficient techniques among them is reinforcement learning. This technique will be employed in the development of the proposed protocol. In reinforcement learning, learning is based on interacting with the environment. In this context, the learning process of agents (IoT nodes) will continuously improve through actions taken in each state. Each chosen action by the agent in a given state results in either a reward or a penalty. The policy involved in this process aims to achieve the highest level of rewards [41]. In this technique, some terms are used as follows:

- Agents: These refer to the members of the IoT network that are expected to perform suitable routing and data communication actions.
- States: These denote specific conditions that network members can monitor.
- Actions: These refer to activities that network members decide to take (such as selecting preferred parents and transmitting data). The actions that sensors can choose are limited to a specific number of activities.
- Rewards: After each activity performed by nodes, the respective node (IoT member) receives a reward or punishment based on its behavior.
- Q-learning

The Q-learning mechanism is one of the most common reinforcement learning techniques. This mechanism will be utilized for the development of quality routing in the proposed protocol. This mechanism is environment-independent and operates based on designed rewards and penalties. In this mechanism, a Q value is assigned to each action performed by IoT members, representing the positive impact of the desired activity. In essence, based on each node's activity, a reward is considered, and the Q value is updated accordingly. Over time and through Q value updates, learning improves, enabling nodes to perform optimal and appropriate behaviors [41].

3 Related work

Given the high importance of routing in IoT, numerous methods and protocols have been introduced for use in this field, especially in recent years, reflecting the significance and value of this research topic among other research areas. In this section, some of the most important and recent research has been reviewed and evaluated from various perspectives, particularly their capabilities and limitations.

In the proposed research, a protocol called EGDAS-RPL[4] has been introduced by Sankar et al. to enhance routing and data exchange management for IoT [21]. This protocol, based on RPL and utilizing a technique called virtual partitioning, aims to optimize energy consumption while managing data exchanges. This protocol is generally divided into three main steps. In the first step, the network is segmented, in the second step, a leader is selected for each segment, and finally, parents are chosen for the leaders, and data are communicated by segment leaders and parents. The result of this design and optimization is improved data exchanges and energy consumption efficiency. However, on the flip side, it lacks measures to support quality of service, does not account for the dynamic conditions of nodes, and is susceptible to the momentary negative effects on routing decisions.

In studies conducted by Touzene et al., a protocol named EARPL[5] has been introduced to improve IoT routing [22]. This protocol is focused on optimizing and enhancing the RPL objective function and aims to develop energy-efficient routing.

---

[4] Energy-Aware Grid-based Data Aggregation Scheme in Routing.

[5] Energy-Aware RPL.

Based on this performance, network exchanges are improved, and energy consumption is optimized. To achieve this, EARPL introduces a new objective function that, in addition to link quality (ETX index), also considers the energy status of nodes in the objective function calculation, which guides routing decisions. The most significant capability of this method can be seen in energy consumption optimization. However, its limitations include low quality of exchanges, increased congestion in some network nodes, and susceptibility to momentary negative effects on routing decisions.

Rana et al. introduced a protocol called EBOF[6] in 2020 to improve data exchange quality while maintaining load balance and traffic equilibrium [23]. In this protocol, routing processes are based on the evaluation of factors derived from the aggregate ETX index and an assessment of node congestion and intermediate path statuses. The result of this design is an enhancement in data transactions and congestion control. However, this method also lacks effective measures to support unforeseen quality of service in data exchanges. Additionally, it does not account for the dynamic conditions of nodes in routing decisions.

In reference [24], Mirshra et al. introduced an enhanced protocol called EHA-RPL[7] based on the extension of RPL to improve its performance. EHA-RPL introduces a combined metric for routing improvement, wherein nodes are evaluated and intermediate paths are selected based on the outcome of three parameters: remaining energy, expected transmission count (ETX index), and hop count to the sink node. Simulation results indicate improvements in successful receptions at sink nodes and reduced transaction delays. However, this method imposes high overhead on the network and is also affected by issues related to the lack of consideration for the dynamic conditions of nodes and momentary negative impacts on routing decisions.

In research conducted by Seyfollahi et al., a protocol named L2RMR[8] was proposed [25]. The goal of this protocol is to enhance service quality while maintaining graph balance and network load equilibrium. L2RMR focuses on the RPL objective function and aims to improve routing accordingly. The proposed L2RMR objective function is composed of an evaluation of parent conditions, the number of children, and rank. Based on this function, preferred parents are selected. Its key capability lies in congestion control and support for load balancing. However, L2RMR lacks measures to cover other routing aspects, particularly quality of service support, making it inefficient in this regard. It also strongly suffers from momentary negative effects on routing decisions. In research conducted by Hassani et al. in 2020 [26] and 2021 [27], protocols named FTC-OF[9] and IRH-OF[10] were introduced, respectively, to improve IoT service quality. In both of these studies, they addressed the issues arising from the single-metric nature of the RPL protocol's objective function

---

[6] A New Load Balancing Objective Function for Low-Power and Lossy Networks.

[7] Composite Routing Technique in IoT Application Networks.

[8] Lightweight Load Balancing and Route Minimizing Solution for RPL.

[9] Forwarding Traffic Consciousness Objective Function for RPL Routing Protocol.

[10] QoS-Centric Fault-Resilient Routing Protocol for Mobile-WSN-Based Low-Power and Lossy Networks.

and aimed to improve this fundamental problem. To achieve this, two new objective functions were introduced: FTC-OF based on an evaluation of rank, received signal quality, and node congestion rates, and IRH-OF based on evaluating signal quality (RSSI) and calculating hop counts to the sink node. The most significant advantages of these objective functions include improvements in data exchange quality and successful network receptions. However, these protocols are challenged by momentary negative impacts, the lack of consideration for the dynamic conditions of nodes, and high control overhead.

In [28], Acevedo and et al. introduced a protocol called WRF-RPL.[11] This protocol is designed based on RPL and aims to enhance its performance for use in highly congested networks. To achieve this goal, WRF-RPL evaluates node conditions during the DIO transmission process based on remaining energy, congestion rates, and hop counts to the sink node. Then, by applying weighted criteria, it evaluates node priorities and selects preferred parents. The weighting of criteria and the selection of preferred parents are done in a way that strives to balance traffic and network paths as much as possible. However, this method also lacks measures to mitigate momentary negative impacts and consider the dynamic conditions of nodes in routing decisions. In studies conducted by Wang et al. [29] and Zarzoor et al. in 2021 [30], protocols named RR-RPL[12] and MHOF,[13] respectively, have been introduced. The goal of these protocols is to enhance concurrent services while optimizing energy consumption in the Internet of Things. Both studies focus on optimizing and improving the RPL objective function, to enhance parent selection. In the proposed objective function by RR-RPL, parent selection and routing processes are based on parameters such as ETX, rank, and residual energy. In contrast, MHOF evaluates the signal quality of received signals, examines link quality, and considers energy. The key advantage of these protocols lies in improving concurrent service with energy management. However, on the downside, they lack measures to accommodate the dynamic conditions of nodes, momentary negative impacts, and high control overhead.

In research conducted by Pushpalatha et al. [31] and Shinde et al. [32], protocols named L-RPL[14] and FlAR[15] have been proposed, respectively, to improve routing and data communication in IoT. The objectives of both protocols are to enhance the RPL objective function and improve routing quality. In L-RPL, routing processes are based on ETX, residual energy evaluation, and link error assessment. In contrast, FlAR is designed based on fuzzy logic and selects parents based on energy, link quality, and delay. Simulation results show that L-RPL improves successful data exchanges alongside energy management. However, FlAR lacks empirical results. Limitations of both protocols include high overhead, the absence of measures to

---

[11] Weighted Random Forward RPL for High Traffic and Energy Demanding Scenarios.

[12] Rank Remain Energy RPL.

[13] Enhance-Minimum Rank with Hysteresis Objective Function.

[14] RPL Powered by Laplacian Energy for Stable Path Selection During Link Failures in an Internet of Things Network.

[15] Fuzzy Logic Approach for Routing in Internet of Things Network.

mitigate momentary effects, and interference with the dynamic conditions of nodes in routing decisions. In a proposed study by Kaviani et al. [33] in a protocol called CQARPL,[16] based on optimizing the RPL protocol using the TOPSIS technique, has been suggested. The goal of this protocol is to improve service quality while concurrently controlling congestion. Parent selection in CQARPL is based on multi-criteria evaluation, including energy, hop count, congestion, and ETX, using the TOPSIS technique. The result of this design and performance is an improvement in service quality along with congestion control. However, on the flip side, it suffers from high overhead, lacks measures to mitigate momentary effects, and interferes with the dynamic conditions of nodes in routing decisions.

In studies conducted by Hassani et al., a protocol called E-MRHOF[17] has been introduced based on the development and optimization of the RPL protocol [34]. The objective of this protocol is to improve end-to-end routing by RPL, and it operates based on the cumulative evaluation of the ETX metric. This design and performance lead to improved data exchange reliability and successful receptions in the network. However, the limitations of this protocol include high latency and low service quality. In [35], Charles et al. proposed a mechanism named RLQRPL.[18] This protocol, based on RPL, aims to improve IoT service quality. RLQRPL introduces a new objective function called LQBOF,[19] which is used for selecting preferred parents. This function includes a quality of link metric, based on which nodes with the best link quality are chosen as preferred parents. The primary advantage of this protocol is the improvement in data exchange quality. However, on the downside, it faces challenges related to increased congestion in some nodes, lacks measures to consider the dynamic conditions of nodes, and is susceptible to momentary negative impacts on routing decisions. In [36], Thenmozhi et al. proposed a method called PMQoSR.[20] This method aims to enhance the quality of service, particularly for real-time data transactions. To achieve this, PMQoSR utilizes priority tags and latency patterns during data communication. PMQoSR conducts routing processes by focusing on distance evaluation, successful reception rates, error rates, energy status, and congestion but with the distinct goal of ensuring the quality of real-time data. To fulfill this objective, the transmission of real-time data takes precedence. However, on the downside, this approach leads to increased control overhead and network energy consumption. Additionally, it lacks mechanisms to monitor dynamic node conditions and mitigate momentary negative impacts on routing decisions.

In a proposed project by Darabkh et al. [37], a novel method called FL-HELR-OF[21] is introduced. This method is based on the improvement and optimization of the RPL protocol, to enhance routing quality and data exchanges. To achieve this goal, FL-HELR-OF introduces an objective function based on fuzzy logic, focusing

---

[16] Congestion and QoS Aware RPL for IoT Applications Under Heavy Traffic.

[17] Improvement of Minimum Rank Hysteresis Objective Function.

[18] Reliable Link Quality-Based RPL Routing.

[19] Link Quality-Based Objective Function.

[20] Energy-Efficient Priority-Based Multi-Objective QoS Routing.

[21] Fuzzy Logic Objective Function.

on the evaluation of metrics including hop count, signal quality, energy, and delay. This design and performance result in improved service quality, but this method also lacks measures to consider dynamic node conditions and unforeseen momentary effects on routing decisions. Furthermore, it imposes heavy overhead on the network. Fazli et al. proposed a protocol named V-RPL[22] [38], which is based on optimizing VIKOR[23] and AHP[24] techniques. The objective of this protocol is to improve data exchanges using the RPL protocol. To achieve this, V-RPL is designed based on the evaluation of multiple criteria and operates based on the assessment of connection quality, energy, and ETX. Parent selection is performed using multi-criteria techniques, forming the network graph based on the results of this evaluation. The result of this design and performance is an enhancement in service quality. However, it faces limitations such as increased congestion in some nodes, a lack of measures to mitigate momentary effects, and interference with the dynamic conditions of nodes in routing decisions.

In research conducted by Ms. Kalantar et al. in 2023 [39], a protocol named ELBRP[25] was introduced. This method is also based on the extension of RPL and aims to improve concurrent service quality while maintaining energy consumption and traffic balance among nodes. ELBRP operates based on data traffic distribution. In ELBRP, when network and node traffic is low, routing elections and data communication are focused on quality. Conversely, when node and network traffic is heavy, parent selections are made in a multi-valued manner, with a focus on preserving traffic load balance. Consequently, concurrent quality is maintained while balancing the traffic load. However, this approach faces challenges such as high overhead, a lack of measures to mitigate momentary effects, and interference with dynamic node conditions in routing elections.

In [40], a hybrid energy approach for data routing and transmission on the Internet of Things is presented. In this approach, after the clustering of IoT sensors, the energy consumption of the sensors is predicted using the ML technique and included in the selection of the cluster head, or CH. Then, the meta-heuristic algorithm selects CHs by considering indicators such as centrality and energy status. This technique improves parameters such as energy consumption, end-to-end delay, and load balancing, but on the other hand, measures to consider the dynamic conditions of the network and support parent replacement have not been provided.

In Table 1, some of the most significant research studies conducted in the past two years have been summarized and analyzed:

---

[22] An Effective Routing Algorithm for Low-Power and Lossy Networks Using Multi-Criteria Decision-Making Techniques.

[23] Vlse Kriterijumsk Optimizacija Kompromisno Resenje.

[24] Analytical Hierarchy Process.

[25] Energy and Load Balancing Routing Protocol for IoT.

Table 1 Analysis of the most important research studies conducted in the recent two years

Authors and year	Proposed method	Capability	Limitations
Hassani et al. [34], 2022	Proposed E-MRHOF protocol based on RPL protocol extension with improved end-to-end routing based on cumulative ETX metric evaluation	Improved data reliability and successful receptions	High latency and low service quality
Charles et al. [35], 2022	Introduced RLQRPL mechanism to improve service quality, introduced a new objective function called LQBOF based on a combination of link quality metrics	Enhanced data exchange quality	Increased congestion in some nodes, lack of measures for dynamic node conditions, and momentary impacts on routing selections
Thenmozhi et al. [36], 2022	Proposed the PMQoSR protocol to enhance delay-sensitive data, prioritize labels, and time delay patterns, routing processes focusing on distance, successful reception rate, error rate, energy consumption, and congestion	Improved data exchange	Increased network control overhead and energy consumption, lack of measures for dynamic node conditions
Darabkh et al. [37], 2022	Proposed a new approach called FL–HELR-OF based on improving and optimizing the RPL protocol's objective function, designing a fuzzy logic-based objective function focusing on evaluating metrics like hop count, signal quality, energy, and delay	Quality improvement	Lack of measures for dynamic node conditions and the resolution of momentary effects on routing selections, heavy control overhead
Fazli et al.[38], 2022	Proposed a protocol called V-RPL based on optimizing VIKOR and AHP techniques, aiming to enhance exchanges by the RPL protocol, route decisions based on evaluating connection quality, energy, and ETX using multi-criteria techniques	Improved service quality	Increased congestion in some nodes, lack of measures to mitigate momentary effects, and interference of dynamic node conditions in routing selections

Table 1 (continued)

Authors and year	Proposed method	Capability	Limitations
Kalanter et al.[39], 2023	Proposed ELBRP based on RPL protocol development, designed based on traffic distribution techniques; providing adaptive decision-making to support simultaneous quality with congestion control; capability: traffic management and quality improvement	Supporting simultaneous quality while maintaining traffic load balance	Inefficiency in examining route conditions, lack of measures to control node dynamics, high control overhead

4 Problem statement

While the potential of IoT-based networks and their applicability across various domains is undeniable, there are still ongoing challenges that have not been adequately addressed, despite extensive research. Some of the most important challenges include:

- Momentary negative effects: This issue pertains to the fact that if, at the time of evaluating the objective function, a node is in unfavorable conditions (e.g., experiencing high congestion), this issue continuously affects the ranking of lower-level nodes in the network. Even though the adverse conditions of a node are limited to the current time, they still affect the rankings of the node of interest and other nodes in the network.
- Lack of measures to monitor dynamic node conditions: In RPL and related research, parent selection, network graph formation, and updates are based on the RPL protocol and are transmitted through DIO (DODAG Information Object) messages. This means that the selection of parents and election updates are solely dependent on the evaluation of node conditions at the time of DIO transmission. However, IoT networks are highly dynamic, and conditions can change at any moment. Therefore, evaluating and selecting nodes exclusively during DIO transmission may not be ideal because conditions that are ideal at that moment might change immediately afterward, or vice versa.
- Low quality of communications: As previously mentioned, most IoT applications are critical and of significant importance. Consequently, the data transmitted in these applications is valuable and sensitive. The primary need and necessity for this data are effective coverage of the quality requirements of the transmitted information. This capability can only be well addressed when the quality requirements of the transmitted data are considered during the routing processes. This will be achieved by effectively optimizing the routing objective function based on quality requirements. The lack of such measures leads to increased data loss and data exchange delays, as observed in some prior research.

These issues have been among the most significant challenges in previous research, and the goal of the proposed research is to address and improve them, with a focus on reinforcement learning algorithms. In the upcoming research, these issues will be enhanced through the optimization and improvement of the RPL protocol's objective function using artificial intelligence techniques, specifically Q-learning. By utilizing Q-learning, not only can service quality be effectively improved but also momentary effects on the objective function evaluation can be mitigated, and dynamic node conditions can be considered in the objective function calculations.

For the resolution of the first and second issues, it is worth noting that the Q-learning technique consists of two steps: learning and convergence. By leveraging these two steps, the challenges of momentary negative effects can be

improved, and the dynamic conditions of nodes can be included in the objective function evaluation. With the learning step, the selection of parents is not limited to the moment of DIO (DODAG Information Object) transmission but extends to the after-parent selection, allowing for the evaluation of the objective function and the replacement of parents. This process continues until the algorithm enters the convergence step. In this step, the selected parent will remain stable throughout the process of DIO retransmission. In essence, the learning step provides adaptability and responsiveness to variable network conditions, while the convergence step ensures the efficiency of the proposed method.

The result of this design and operation, in addition to addressing past research issues, will lead to a significant improvement in service quality. The utilization of the concepts of learning and convergence in reinforcement learning not only enhances service quality effectively but also preserves the efficiency of the proposed method.

5 RI-RPL: the proposed method

In this research, considering the significance and value of routing in the Internet of Things and drawing on extensive studies and researcher ideas, a proposed method called RI-RPL[26] is introduced and presented. The primary goal of designing RI-RPL is to improve routing, enhance service quality, and address the challenges of the RPL protocol and the research based on this protocol by focusing on artificial intelligence techniques, particularly Q-learning. To design RI-RPL, key indicators for organizing an effective and efficient method were first obtained, and based on these, the proposed method will be designed. One of the most fundamental objectives of RI-RPL, in addition to addressing the limitations of past research, is to improve reliability and increase the Quality of Service (QoS) of the RPL protocol. The presence of these measures is one of the most outstanding features of the proposed method compared to past research. Additionally, it is worth noting that although past research has been successful in improving the quality and performance of the RPL protocol, some fundamental challenges remain, especially regarding the momentary selection of parents. This is where RI-RPL excels by providing adaptability to variable network conditions through the utilization of learning techniques. Further details of the design of the proposed method and how it operates will be presented in the following sections, demonstrating how the goals envisioned for the performance of RI-RPL have been adequately achieved.

5.1 Energy consumption model

In RI-RPL, energy consumption follows the standard for Internet of things networks. According to this standard, the energy consumption of sensors for sending and

---

[26] Reinforcement Learning-Based RPL Routing Protocol.

receiving data is as per Eqs. (1) and (2). Where $E_S(c, d)$ is the energy consumed for sending c bits of data over a distance of d (the distance between two transmitting and receiving sensors), $E_R(c)$ is the energy consumed for receiving c bits of data by the receiver (c equals the total bits of sent or received packets), E_{elect} represents the energy consumed by electronic circuits to process 1 bit of data, ε_{fs} is the energy required to transmit data to a distance of d^2, and ε_{amp} is the energy required to transmit data to a distance of d^4. Lastly, TR corresponds to the radio range of network members.

$$
\begin{aligned}
E_S(c, d) &= c \times E_{elect} + c \times \varepsilon_{fs} \times d^2 \quad d < \text{TR} \\
E_S(c, d) &= c \times E_{elect} + c \times \varepsilon_{amp} \times d^4 \quad d > \text{TR}
\end{aligned}
\tag{1}
$$

$$
E_R(c) = c \times E_{elect}
\tag{2}
$$

5.2 Communication model

The target IoT network consists of k sensor nodes that are randomly distributed in the environment and are capable of bidirectional communication with each other. In addition to the sensors, the network has a sink node, which is the destination for all data transmitted by network members. For data communications in RI-RPL, it adheres to the RPL protocol and DODAG graph. The DODAG graph is a graph created through the exchange of control data of the RPL protocol and ensures communication between all network nodes and the sink node. Based on this, communications in RI-RPL include:

- P2P[27] Communications: This type of communication involves data exchange between two specific network members (for example, between one of the network nodes and the sink node).
- P2MP[28] Communications: This type of communication pertains to data transmission from one member to multiple network members (for example, sending a packet from the sink node to several network nodes).
- MP2P[29] Communications: This type of communication involves data transmission from multiple network members to one member (for example, nodes sending a packet to the sink node). RI-RPL supports all three of these communication types.

5.3 Assumptions

To design and develop RI-RPL, some general assumptions have been considered in network design:

---

[27] Point-To-Point.

[28] Point-To-Multipoint.

[29] Multipoint-To-Point.

Table 2 Acronym table

Variable	Explanation	Variable	Explanation
$EInt_i$	Initial energy of node i	EC_P	Energy consumption of processing
EC_S	Energy consumption of transmissions	EC_L	Energy consumption of low-energy mode
EC_R	Energy consumption of receptions	SP_i	Probability of success in transmissions by node i
SSP_i	Packets correctly sent by node i	PR_i	Packets received by node i
HC_i	Number of hops from node i to the sink node	BU_i	Node i's buffer
BS_i	Total buffer capacity of node i	BB_i	Buffer occupancy rate of node i
R_i	Node i's rank	RA_{Father}	Rank of the node i's parent
RA_{inc}	Constant rank 255	ETX_i	Expected transmission rate by node i
$Rw[a_t, b_t]$	Reward value obtained from performing action "a" by a network node	$Q[a_t, b_t]$	Value of the Q parameter at the current time (moment "t")
$Q[a_{t+1}, b_{t+1}]$	Value of the Q parameter in the future (moment "$t+1$")	CC_v	Constant value (1) associated with the cost incurred by nodes for data transmission
$CV(a_{t+1})$	Variable cost of data transmission	SDR	Successful reception rate of the network
ED	Average end-to-end delay	OH	Quantity of control overheads
PEC	Energy consumption rate for each packet	PDR	Network data loss rate
TG	Network throughput		

- The network is distributed and lacks any central infrastructure or controller.
- RI-RPL is based on the RPL protocol, and in terms of performance, adheres to RPL routing standards.
- Network members (nodes except the sink node) are homogeneous in terms of energy and radio range limitations.
- Each network node has a unique and distinct digital identity.
- Interactions between nodes are symmetrical and bidirectional.
- Network connectivity is dynamic and variable.
- Network members communicate with the sink node using single/multi-hop based on their location in the network.
- The common media access standard is CSMA-CA [41].
- Nodes do not possess the capability to self-locate in the network or any additional tools for positioning.

5.4 Details of the proposed method

In order to better understanding, Table 2 depicts all used symbols in the proposed method.

Based on the aforementioned fundamentals, we will introduce the design details and performance of RI-RPL. RI-RPL is based on the RPL protocol and focuses on using learning algorithms. The goal of this technique is to improve routing and enhance service quality, which is achieved through the following three operational steps:

1. Routing operations are conducted by the optimization of the RPL protocol based on RI-RPL's proposed measures (with a focus on the Q-learning algorithm).
2. Support for parent node changes in various network conditions (based on learning).
3. Control and management of changes in network topology.

It is worth noting that each of the introduced steps plays a crucial role in advancing the objectives of the proposed method. In the following sections, a functional flowchart of the method is presented in Fig. 1, followed by a discussion of the design and operation of each of its stages.

5.4.1 Phase one: routing aligned with RI-RPL proposed measures

This phase aims to establish the network's communication graph in line with RI-RPL objectives. The RPL protocol operates by utilizing the exchange and sharing of three control packets: DIO, DAO, and DIS. RI-RPL has been designed based on these three processes, with this phase focusing on optimizing the transmission of DIO and DAO packets. In essence, this phase seeks to harmonize the delivery of DIO packets (by RPL) with the proposed measures for facilitating high-quality routing in RI-RPL. Figure 2 illustrates the flowchart for this phase of RI-RPL. It

is noteworthy that in the provided flowcharts, the color-coded boxes represent the main components of the proposed method and RI-RPL innovation.

This phase of RI-RPL consists of two fundamental pillars:

1. Optimized and Targeted Development: In addition to providing the necessary prerequisites for performing the desired routing, measures are taken to prevent overburdening the network for the advancement of RI-RPL's performance. It is important to note that the proposed protocol, aligned with RPL, has evolved without adding a specific procedure to the network.
2. Facilitating the Analysis of RI-RPL's Routing Requirements: This pillar focuses on providing sufficient provisions for analyzing the routing requirements of the proposed protocol to effectively support QoS.

In RPL, the creation and updating of the DODAG graph (routing operations) occur through the periodic transmission and reception of DIO and DAO packets. This means that DIO packets are generated by the sink node at specified intervals and disseminated throughout the network. Nodes, by receiving these packets, become aware of their new conditions and parent nodes, and then, based on their new conditions and parent nodes, make selections and inform them by sending DAO packets. This process continues repeatedly until the end of the network's lifetime. According to the RPL protocol, the routing process in RI-RPL also consists of two main sub-steps as follows:

• DIO transmission process

In RPL, DIO packets are generated by the sink node and disseminated among nodes to the extent that all network members receive the transmitted packets. Thus, all sensors and parent nodes identify each other and become aware of their positions in the network. The DIO transmission process in RI-RPL has some differences from the RPL protocol, as outlined in Table 3. As shown in the flowchart in Fig. 2, after the network starts, the sink node creates a DIO packet and sends it to the nodes in the network. Each receiving node first checks whether the received packet is new or not (the first condition in the flowchart). Based on the result of this assessment, two different scenarios are presented, as described in Table 4.

As mentioned in Table 2, one of the fundamental differences between RI-RPL and RPL is the addition of proposed routing metrics to the DIO packet. RI-RPL updates the DIO packet according to its requirements. To achieve this, in addition to the node's rank, remaining energy status, probability of successful transmission, congestion, and the number of hops to the sink node are added to the DIO packet and shared along with the shared packet. Equations (3)–(10) provide details on the calculation of RI-RPL's routing metrics along with relevant discussions.

As mentioned, one of these metrics is the remaining energy status of nodes. Equation (3) provides details on evaluating the remaining energy of nodes. The use
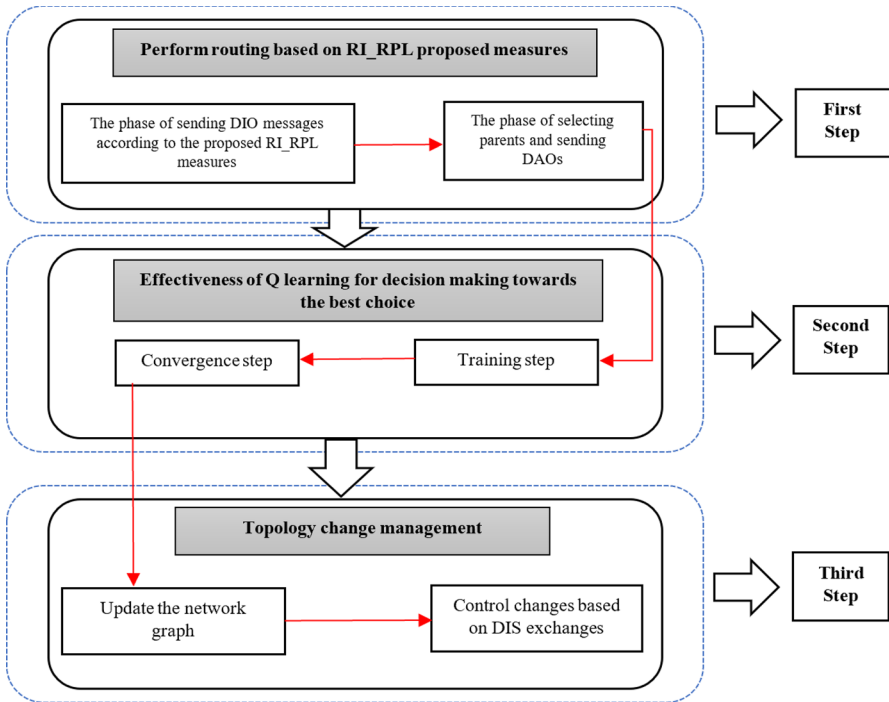
Fig. 1 Block diagram of the proposed method

of this metric leads to the selection of nodes with higher energy on one hand, and the other hand, prevents early energy depletion of some nodes.

$$ER_i = \frac{EInt_i - (EC_S + EC_R + EC_P + EC_L)}{EI_i} \qquad (3)$$

In Eq. (3), $eInt_I$ is equivalent to the initial energy of node i, and EC_S, EC_R, EC_P, and EC_L, respectively, represent the energy consumption related to transmissions, receptions, processing, and low-energy mode (sleep mode).

The next parameter is the probability of successful transmission. Based on this metric, the higher the parent's probability of success in transmissions, the higher their priority for selection. Using this parameter improves the reliability of exchanges and enhances successful receptions. Equation (4) presents the evaluation of this metric. In this equation, SP_i represents the probability of success in transmissions by node i, $\sum No.of SSP_i$ is the sum of packets correctly sent by node i during previous interactions (a random mode will be used to check the success of transmissions), and $\sum No.of PR_i$ is equal to the sum of packets received by node i. In this case, if the sending node receives its sent message again, it means that the packet has been sent correctly.
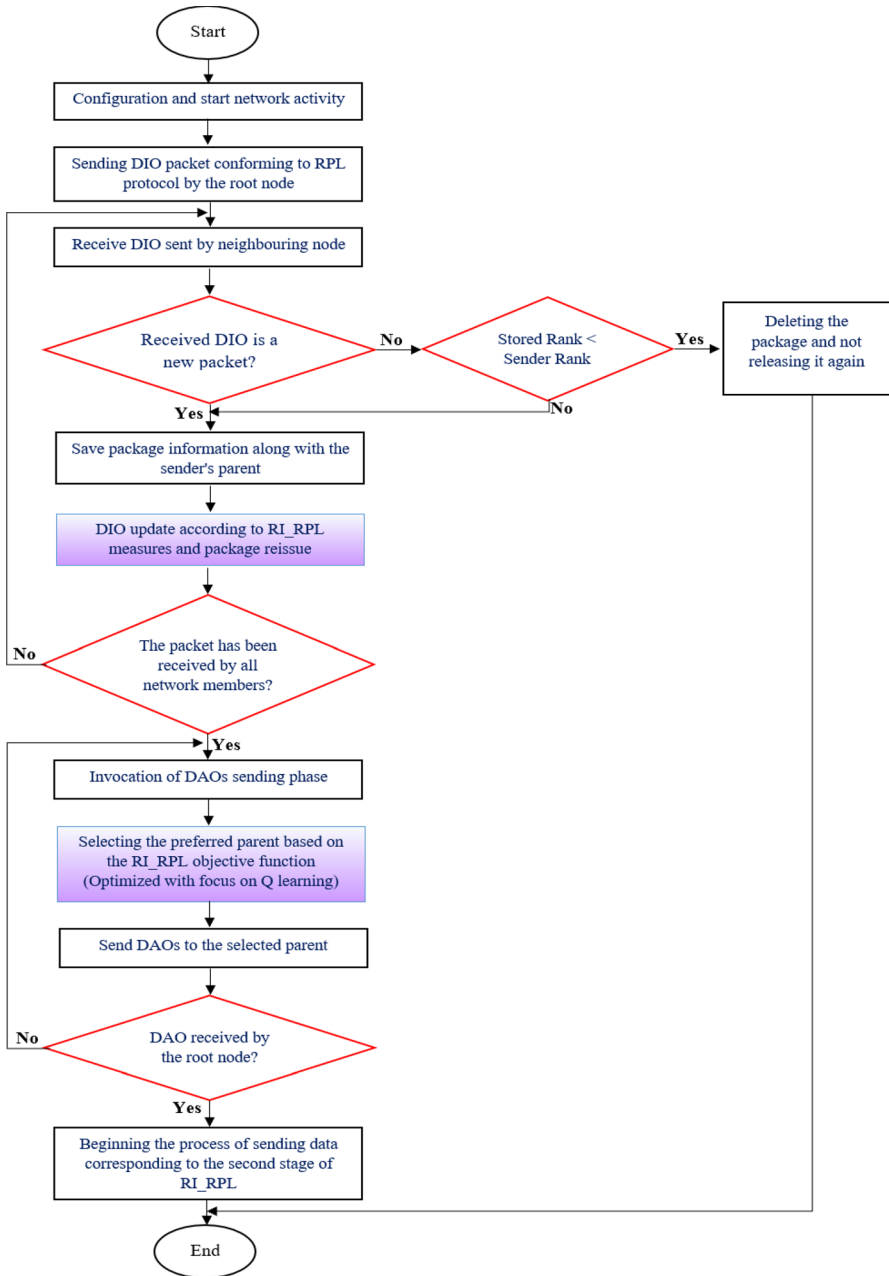
Fig. 2 Flowchart related to the first stage of RI-RPL

$$SP_i = \frac{\sum \text{No. of SSP}_i}{\sum \text{No. of PR}_i} \tag{4}$$

Table 3 Differences in the DIO transmission process between RI-RPL and RPL	Row	Difference
	1	The first difference is that in RI-RPL, in addition to performing the DIO transmission-related procedures (including rank updates, adding addresses, etc.), nodes also calculate their other conditions such as congestion, success probability in transmission, remaining energy, and the number of hops to the sink node. The results of these assessments are then added to the DIO packet. This is done to provide awareness of the nodes' status and conditions relative to each other in RI-RPL. Therefore, each node, in addition to identifying the positions of its parents relative to the sink node, will also be aware of their other conditions
	2	The second difference lies In the calculation of the objective function for the nodes. In RPL, the objective function is calculated based on rank and/or the ETX metric. This is in contrast to RI-RPL, where the objective function is based on a multi-criteria concept and is evaluated using Q-learning concepts

The next parameter is the number of hops to the sink node. Examining this parameter will play a significant role in optimizing the communication graph and data exchanges. Using this metric leads to the selection of parents with fewer hops to the sinkhole. Equation (5) provides the evaluation of this parameter, where No.ofhoproot$_i$ is equal to the number of hops from node i to the sink node. This parameter is increased by one for each hop that the DIO packet is sent.

$$HC_i = No. \ ofhop \ root_i \tag{5}$$

The next parameter is the nodes' congestion rate. Examining this parameter results in nodes with lower congestion being selected as the main parents. Using this metric prevents some parents from becoming overloaded, mitigates the problem of child explosion, and improves communication delay significantly. Equation (6) presents the evaluation of this parameter, where BU$_i$ equals the efficiency of the node i's buffer, BS$_i$ equals the total buffer capacity of node i, and BB$_i$ equals the buffer occupancy rate of node i [5].

Table 4 Various scenarios related to investigating the redundancy of DIO packets in RI-RPL

Scenario	Explanation
First	The received DIO packet is new. In this scenario, the node stores the parent's rank (rank included in the packet), adds its identifier along with other proposed metrics to the packet and re-broadcasts it within the network
Second	Receiving Duplicate DIO Packets (Second Condition in the Flowchart)
	If the rank in the new DIO packet (rank of the parent sending the DIO) is lower than the previously stored rank of the previous parent (rank of the previous parent sending the DIO), the node stores the packet, updates it according to the first scenario, and re-broadcasts it within the network
	If the rank in the new DIO packet (rank of the parent sending the DIO) is higher than the previously stored rank of the previous parent (rank of the previous parent sending the DIO), the node simply discards the received packet without any specific action
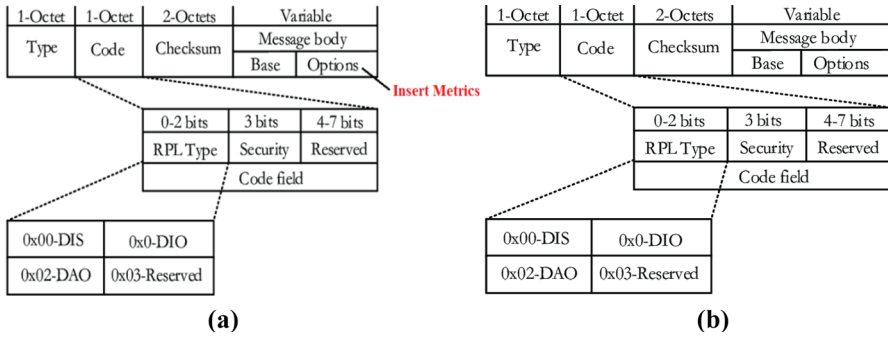
Fig 3 DIO package in **a** RPL and **b** proposed RI-RPL

$$BU_i = 1 - \frac{BB_i}{BS_i} \tag{6}$$

In addition to the mentioned parameters, each node evaluates its rank and replaces the rank field in the DIO message. The rank refers to the position of nodes from the sink node. Using this parameter leads to the selection of parents closer to the sink node, which has significant effects on improving communication and exchange delays. Equation (7) presents the evaluation of this parameter, where R_i equals the node i's rank, RA_{Father} equals the rank of the node i's parent, and RA_{inc} equals the constant rank increment value for nodes. The value of this parameter in the RPL protocol is considered as 255 [14].

$$R_i = R_{Father} + R_{inc} \tag{7}$$

After evaluating the introduced parameters, the results of these calculations are added to the DIO message and the packet is sent back into the network in a flooding manner. This process continues until all members receive it. Figure 3a shows the DIO message in RPL, and in Fig. 3b, it is represented in RI-RPL. As shown in the figure, the desired parameters are added to the option section of the packet.

Following the dissemination and sharing of DIO messages based on what has been presented, the second sub-step of the first phase, based on the provided information, is invoked and executed.

• The DAO sending process

According to the presented flowchart for this stage of RI-RPL, after completing the DIO transmission process, the DAO transmission process begins. In this process, if any node in the network has only one parent, it informs that parent of its choice by sending a DAO packet. In other cases, according to the RI-RPL objective function, the best parent is selected from the candidates and this choice is reported to it by sending a DIO packet. The RI-RPL objective function and its evaluation details are designed based on the Q-learning algorithm. In this algorithm, decision outcomes are quantified under an index called "Q." RI-RPL also
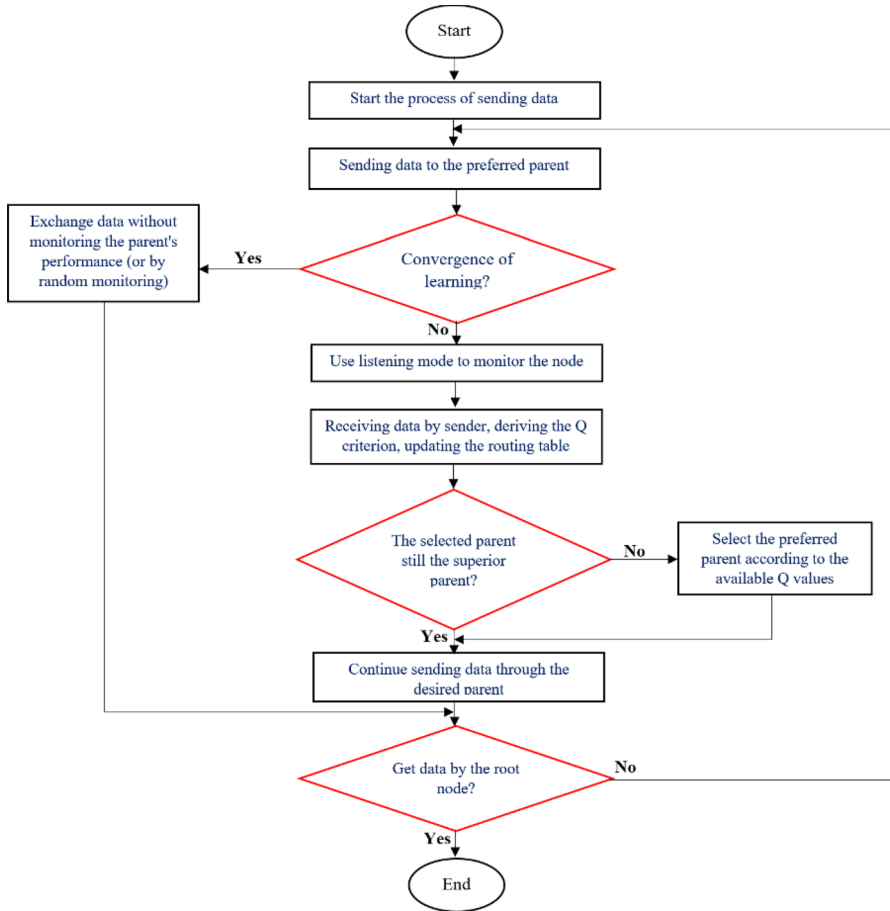
Fig. 4 Flowchart related to the second stage of RI-RPL

adheres to this concept and calculates and evaluates the objective function based on discussions related to Q-learning. The details of this design and evaluation are provided below.

Before discussing the details of calculating the Q-learning-based objective function, it should be noted that in RI-RPL, each network member has a learning table (routing table) that stores a pair (a, b) with a value of Q for each of its parents. During network initialization, the values of Q assigned to these pairs are set to zero, and they are updated during each round of the DIO transmission process. Since Q represents a learning index, given that there is no learning during network initialization, its value is zero.

Based on this pair, whenever a node intends to make a decision, it examines the current conditions of "a" and acts similarly to "b." After performing the desired
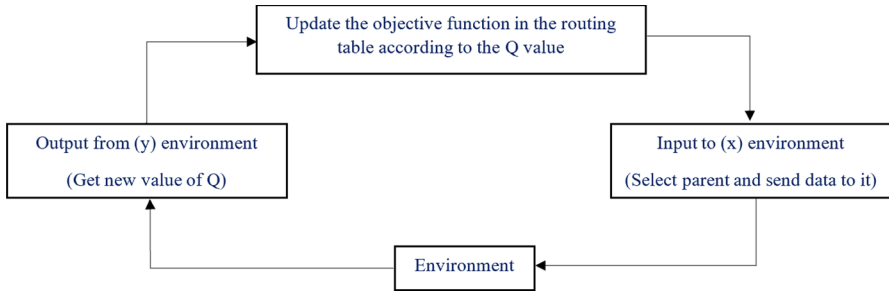
Fig. 5 Learning cycle in the proposed RI-RPL method

action, a reward value is obtained (represented by *Rw* (*a*, *b*)), based on which the Q parameter is updated. The calculation of the Q-learning parameter is presented in Eq. (8). In this equation, $Q[a_t, b_t]$ represents the value of the Q parameter at the current time (moment "*t*"), $Q[a_{t+1}, b_{t+1}]$ corresponds to the value of the Q parameter in the future (moment "*t*+1"), $Rw[a_t, b_t]$ equals the reward value obtained from performing action "*a*" by a network node, $\theta[u, v]$ is the coefficient determining how the Q parameter learning is carried out, and ϑ is equivalent to the coefficient determining the importance of future rewards in reinforcing learning.

$$Q[a_t, b_t] = Q[a_t, b_t] + \theta[a, b] \cdot \left(Rw[a_t, b_t] + \vartheta \cdot \max Q[a_{t+1}, b_{t+1}]\right) \tag{8}$$

As discussed, $\theta[u, v]$ is a coefficient that determines how the Q parameter is learned. In the equation, this coefficient specifies to what extent the conditions obtained in the current state are preferred and prioritized over past conditions. Therefore, if the value of this symbol is set to zero, nodes will not be able to learn against their current activities, and if it is set to one, nodes will learn to the maximum extent in longer time intervals for reward. If Q is initially set to zero for all network members, $\theta[a, b]$ can be considered as one to accelerate the reinforcement of learning. If we take Q as zero at the beginning and $\theta[a, b]$ as one, then Eq. (8) becomes Eq. (9).

$$Q[a_t, b_t] = \left(Rw[a_t, b_t] + \vartheta \cdot \max Q[a_{t+1}, b_{t+1}]\right) \tag{9}$$

Considering what was mentioned about ϑ, we set the importance of future rewards for maximizing greediness in current learning actions to zero ($\vartheta = 0$) in RI-RPL. If the value of ϑ is set to zero, then Eq. (9) becomes Eq. (10). It should be noted that in RI-RPL, $Q[a_t, b_t]$ is the same as the objective function for selecting the primary parent. According to Eq. (10), OF$_{(RI-RPL)}$ represents the RI-RPL objective function, which determines the decision outcomes.

$$Q[a_t, b_t] = Rw[a_t, b_t] \rightarrow OF_{RI-RPL} = Q[a_t, b_t] = Rw[a_t, b_t] \tag{10}$$

Based on Eq. (10), the evaluation and update of the Q parameter directly depend on the value and magnitude of the reward $Rw[a_t, b_t]$, and Q is updated accordingly.
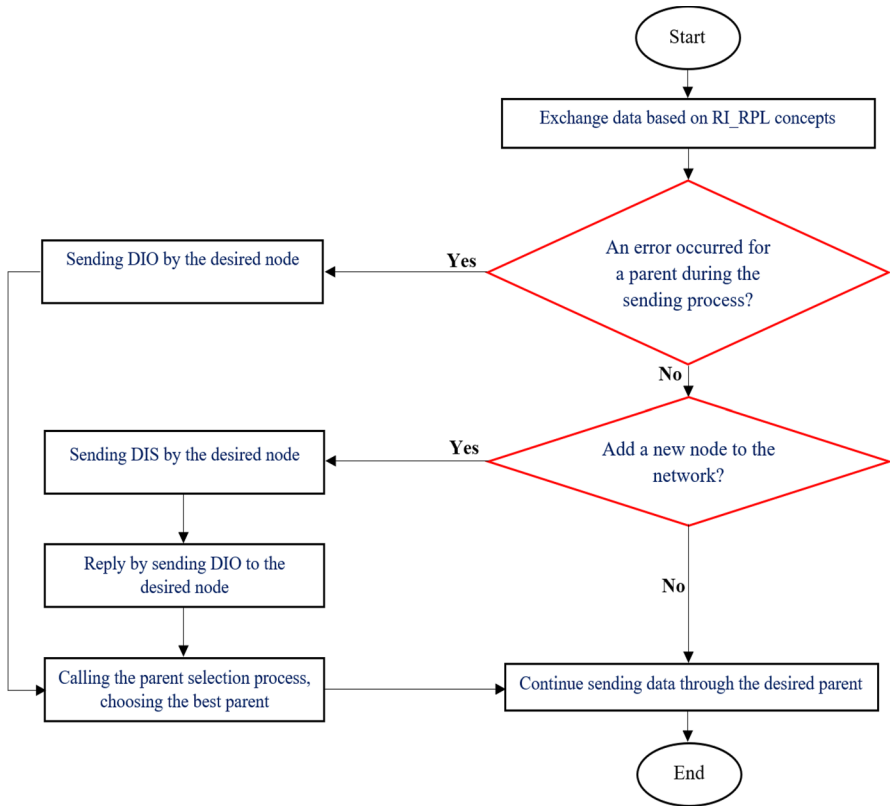
Fig. 6 Flowchart related to the last stage of RI-RPL

To calculate and evaluate the reward $Rw[a_t, b_t]$ in RI-RPL, Eq. (11) is used. $Rw[a_t, b_t]$ represents the reward obtained from acting "*a*" in the current state "*b*." CC_v is a constant value associated with the cost incurred by nodes for data transmission (in RI-RPL, considering the step-by-step data transmission, this constant value is assumed to be one), and $CV(a_{t+1})$ is equivalent to the variable cost of data transmission by candidates (candidate parents), or in other words, the data exchange cost through candidate parents.

$$Rw[a_t, b_t] = \left(CC_v + \left[\frac{1}{(CV(a_{t+1}) + 1)} \right] \right) \tag{11}$$

Depending on the discussed concepts, the $CV(a_{t+1})$ factor represents the cost or price of communicating data through candidate parents. The data transmission cost in the proposed RI-RPL protocol is calculated based on Eq. (12) and takes into account energy efficiency, transmission success probability, congestion status,

Table 5 Simulation parameters

Parameter	Values
OS	Contiki/Cooja 2.7
Communication protocols	CSMA, IEEE 802.15.4, IPv6, UDP
Number of nodes	20,40,60,80,100
Sensor type	Tmote sky
Deployment area	200*200 m
Transmission range	50 m
Data rate (Kbits/s)	250 Kbps
Routing protocols	RI-RPL, RLQRPL, ELBRP, RPL
Data message length	512 bit
Simulation time	Until the first node dies
Result	15 round
Link failure model	UDGM with distance
Traffic rate	60-120-180-240 packet per minutes (PPM)
Adaptation	6LoWPAN
Radio model	Unit disk graph medium (UDGM)
Queue size	10 packet
Queuing model	M/M/k/1
TX current consumption	21 mW
RX current consumption	23 mW
Low-power mode (LPM) current consumption	1.2 mW
CPU current consumption	2.4 mW

expected transmission count (ETX), the number of hops, and parent ranks. ETX_i is equivalent to the expected transmission rate by node i, which is equal to $\frac{1}{\text{df} \times \text{dr}}$, where df corresponds to the success probability of receiving the transmitted data packet and dr corresponds to the probability of successfully receiving the sent message acknowledgment. ω_E, ω_{SP}, ω_R, ω_B, and ω_{ETX} are coefficients assigned to the evaluation criteria ER_i, SP_i, HC_i, R_i, Bu_i, and ETX_i, such that $1 \geq \omega_E, \omega_{\text{sp}}, \omega_B, \omega_R, \omega_{\text{ETX}} \geq 0$, and $\omega_E + \omega_{\text{SP}} + \omega_B + \omega_R + \omega_{\text{ETX}} = 1$. Other criteria have been introduced earlier.

$$\text{CV}(a_{t+1}) = \left(\omega_E \times (1 - \text{ER}_i) + \omega_{\text{SP}} \times (1 - \text{SP}_i) + \omega_B \times (1 - \text{BU}_i) \right.$$
$$\left. + \omega_R \cdot \left(\left(1 - \frac{1}{\text{HC}_i} \right) \times \left(1 - \frac{1}{R_i} \right) \right) + \omega_{\text{ETX}} \cdot \left(1 - \frac{1}{\text{ETX}_i} \right) \right) + \alpha \times \text{FC}(u_{t+1})$$

(12)

Based on the discussed concepts, the Q parameter is calculated, and its result is recorded as the objective function and decision criterion for parent nodes in the routing table. The result and output of this stage of the proposed method are the selection of the best parent as the preferred parent.

Table 6 Simulation scenarios

Traffic rate	Number of nodes	Scenario	Traffic rate	Number of nodes	Scenario
180	60	11	60	20	1
240	60	12	120	20	2
60	80	13	180	20	3
120	80	14	240	20	4
180	80	15	60	40	5
240	80	16	120	40	6
60	100	17	180	40	7
120	100	18	240	40	8
180	100	19	60	60	9
240	100	20	120	60	10

5.4.2 Second stage - utilizing Q-learning for optimal selection

This stage, serving as the second phase of RI-RPL, is designed to ensure optimal performance under varying and dynamic network conditions based on the principles of Q-learning. This step comprises two sub-phases: training and convergence. Figure 4 illustrates the flowchart associated with this phase of RI-RPL, and we will proceed to analyze and elaborate on the details of this phase, referring to the relevant flowchart.

The input to this stage is the output of the previous stage, namely, the selected parents.

The design principles and performance of this stage in RI-RPL consist of two fundamental pillars as follows:

1. Ensuring optimal performance under various network conditions: One of the key objectives of this stage is addressing issues stemming from momentary selections. The training and learning sub-phase facilitate this capability for the proposed protocol.
2. Preserving performance efficiency: It should be noted that a protocol can only be deemed successful if it not only ensures optimal performance under different conditions but also remains efficient. The learning convergence sub-phase provides this capability for RI-RPL.

The foundation of the design for this stage is the provision of measures to ensure optimal performance. For this purpose, RI-RPL leverages reinforcement learning concepts. Accordingly, this phase of RI-RPL consists of two key sub-phases known as training learning and convergence learning. The details of these two stages are explained below.
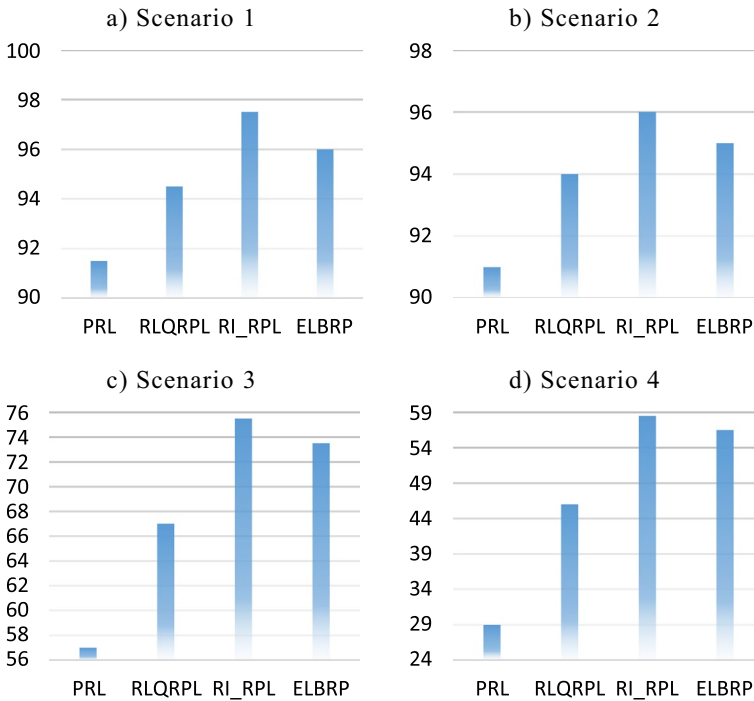
Fig. 7 The results of the success delivery ratio parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 20 nodes with traffic rates: **a** 60 PPM, **b** 120 PPM, **c** 180 PPM, and **d** 240 PPM

- Training learning

According to the provided flowchart, after completing the first stage of RI-RPL, the preferred parent selection, and data exchange process commences through the selected parent. In this scenario, the child node, after sending data to the selected parent, monitors the learning status of the chosen parent. Based on the results of this monitoring, two general scenarios arise. The first scenario involves learning during the training phase, and the related process follows the training learning sub-phase (the process associated with this sub-phase). The second scenario involves learning during the stabilization phase, and the process related to the convergence learning sub-phase follows (the process associated with the next sub-phase). Typically, after selecting the preferred parent, learning occurs during the training phase. Before discussing the details of training learning, it is worth mentioning that during network initialization or any change in network topology, learning is initiated, and after completing the learning process, the convergence step is executed.

During the training learning phase, as presented in the flowchart in Fig. 4, the node, after sending data to the selected parent, enters a listening state and begins monitoring the chosen parent. Here, after initiating data transmission, the sending
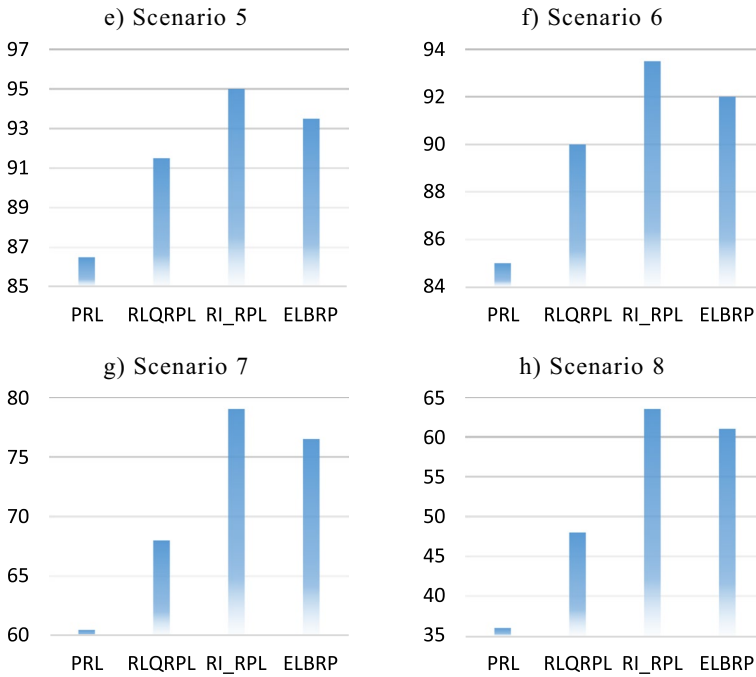
Fig. 8 The results of the success delivery ratio parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 40 nodes with traffic rates: **e** 60 PPM, **f** 120 PPM, **g** 180 PPM, and **h** 240 PPM

node (child node) enters a listening state and listens to the parent node. The preferred parent, after receiving the packet from the child, evaluates the Q parameter for itself, in addition to performing the transmission process (sending data to its selected parent in the path toward the sink node), and includes the Q value along with the data packet. After sending the data by the parent, the child node (the sending node in the previous step, which was in the listening state) receives the sent packet and extracts the Q value from it. Subsequently, the new Q value replaces the previous value in the routing table for the target node. This process continues iteratively until the data is received by the sink node. In fact, during this learning sub-phase, learning is continuously updated and reinforced with each listening and the dynamic conditions of the nodes.

It should be noted that the goal of using the learning concept in RI-RPL is to manage the challenge of momentary parent selections and consider the dynamic conditions of nodes, which are addressed through learning. Furthermore, in subsequent transmissions, if the preferred parent still has the best Q value, this node will continue to be chosen as the preferred parent. Otherwise, the node with the best Q value will replace it as the preferred parent, and data transmission will continue through this node. Importantly, the process of sending DIOs and updating the Q value during the transmission of these packets is also part of the training process, as it involves learning and updating
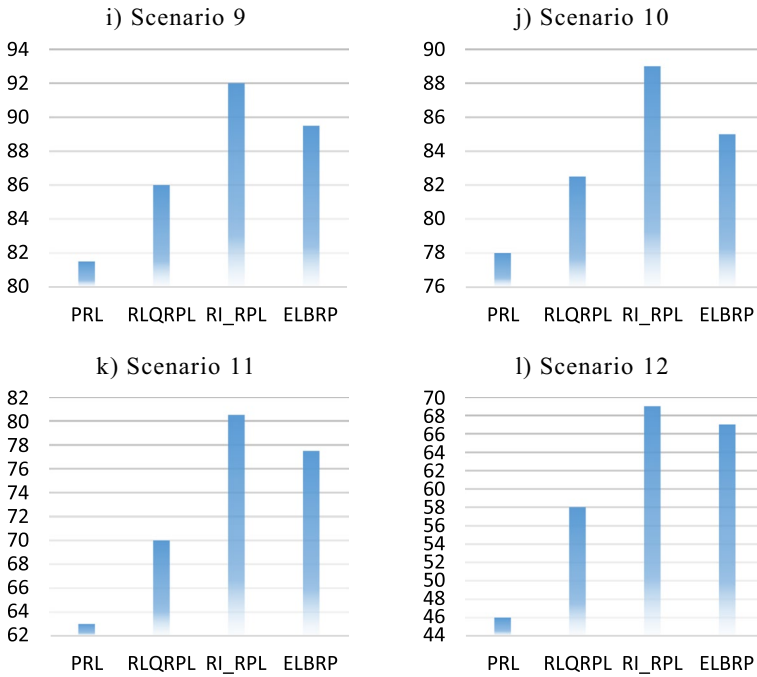
Fig. 9 The results of the success delivery ratio parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 60 nodes with traffic rates: **i** 60 PPM, **j** 120 PPM, **k** 180 PPM, and **l** 240 PPM

the Q value. Figure 5 depicts the learning cycle in the proposed RI-RPL method. In this figure, x represents the action, and y represents the reinforcement signal.

After completing the training process, the convergence learning process is invoked.

- Convergence stage

The primary goal of the Convergence stage is to maintain the efficiency of the protocol's performance. Since frequent listening and consecutive Q value updates impose additional costs on the network and can lead to a decrease in protocol efficiency, it is essential to control and optimize these side costs. To support this, RI-RPL utilizes the concept of learning convergence. Learning convergence ensures the efficiency of the proposed protocol's operation. Learning convergence in RI-RPL is presented in two ways, depending on the network conditions:

1. The network is intended for use in typical applications. After repeating the learning step "*n*" times, the algorithm enters the convergence phase. In the convergence phase, there is no need for listening, monitoring the parent, or updating the Q value. It is worth noting that increasing the "*n*" parameter can enhance learning but may
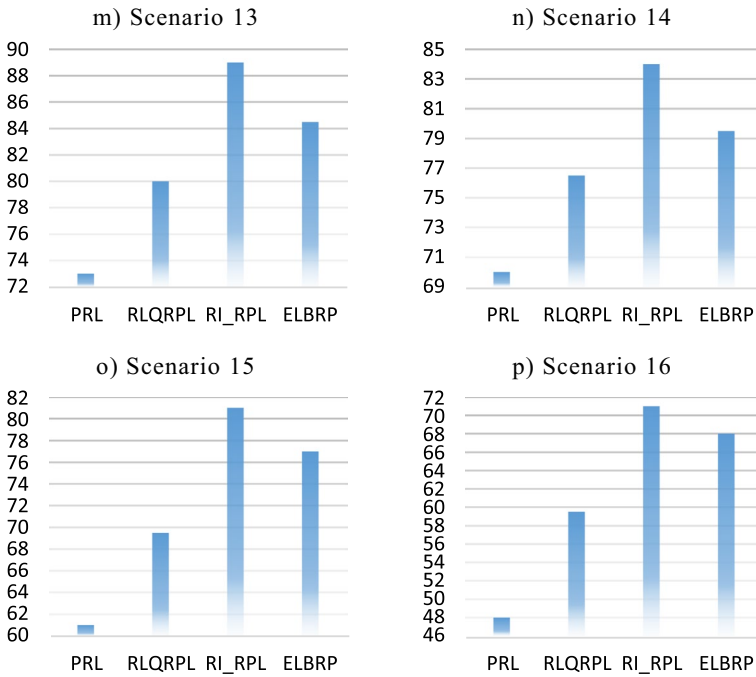
Fig. 10 The results of the success delivery ratio parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 80 nodes with traffic rates: **m** 60 PPM, **n** 120 PPM, **o** 180 PPM, and **p** 240 PPM

also lead to more effective stability assurance. The value of "*n*" depends on the application and will be determined accordingly.

2. The network is intended for use in highly dynamic applications where network traffic changes significantly. In such cases, after repeating the learning step "*n*" times, the algorithm enters the convergence phase. In the convergence phase, after every "*k*" packet, one packet is randomly monitored by the child node, and the Q value is updated based on the result of this monitoring.

In RI-RPL, convergence is based on the first execution state.

5.4.3　Third stage - network topology changes management

This stage, as the third and final stage of RI-RPL, is introduced to control network topology changes and maintain stability under different conditions. Figure 6 illustrates the flowchart related to this stage of RI-RPL.

By the depicted flowchart for this stage, if an error occurs during data transmission to a parent node, the concerned node informs other nodes by sending a DIO message. Upon receiving the sent DIO, nodes invoke the parent selection process, choose a new parent, and continue their exchanges through this new
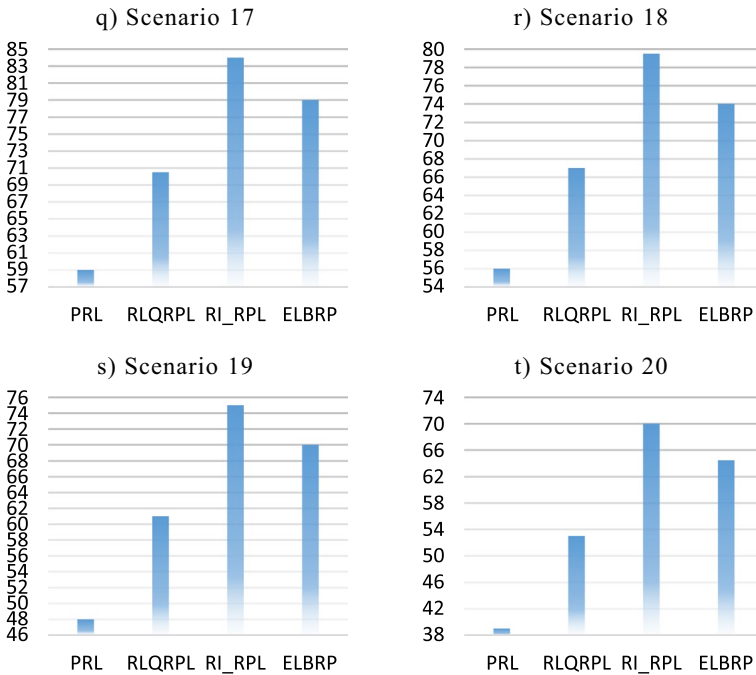
Fig. 11 The results of the success delivery ratio parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 100 nodes with traffic rates: **q** 60 PPM, **r** 120 PPM, **s** 180 PPM, and **t** 240 PPM

parent. On the other hand, if a new node is added to the network or if the position of one of the nodes in the network changes for any reason, the concerned node notifies other nodes by sending a DIS message. The receiving nodes respond to the concerned node by sending a DIO message to it. Then, the node selects the best parent by receiving DIOs and becomes a member of the network graph by sending a DAO. As presented, RI-RPL is designed based on three general stages to achieve its objectives. It starts with the performance of the first stage, where routing (aligned with the proposed RI-RPL measures) is implemented. Subsequently, in the second stage, learning convergence is supported to ensure the best selection in various network conditions. Finally, in the third stage, network topology changes are controlled and managed.

6 Analyzing the results

Implementation and evaluation of RI-RPL were carried out in the Cooja simulator software. Cooja is one of the most widely used and efficient simulators for IoT. This software provides all the components, standards, and patterns required for implementing and simulating various IoT applications and is readily available by default. Cooja can be considered a dedicated simulator for IoT applications. Its prominent
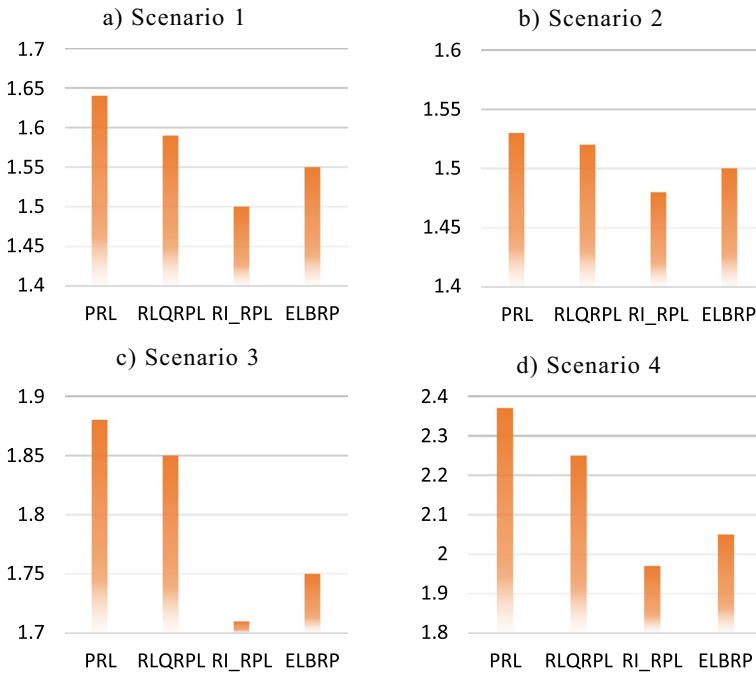
Fig. 12 The results of the end-to-end delay parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 20 nodes with traffic rates: **a** 60 PPM, **b** 120 PPM, **c** 180 PPM, and **d** 240 PPM

features, compared to other software, include accurate simulation of low-power and lossy networks (especially IoT applications), providing a desirable graphical simulation environment, access to various components and tools related to IoT applications, the presence of the RPL protocol by default, and the provision of extensibility and optimization capabilities as needed. Cooja simulator is based on the Contiki operating system.

6.1 Simulation scenarios

To conduct accurate experiments, simulation scenarios for all four protocols: Proposed RI-RPL, RLQRPL [35], ELBRP [39], and RPL [14], have the same conditions. Table 5 lists the simulation parameters and their assigned values. The selection and assignment of network communication protocol components, including the carrier sense multiple access (CSMA) medium access technique, IEEE-standardized communication techniques (IEEE 802.15.4), the IP version used (IPv6), and the transport layer protocol (UDP) are based on IoT standards and prior research. Additionally, simulation scenarios are ranked according to Table 6. The important and influential parameters for the performance of the method are the number of nodes
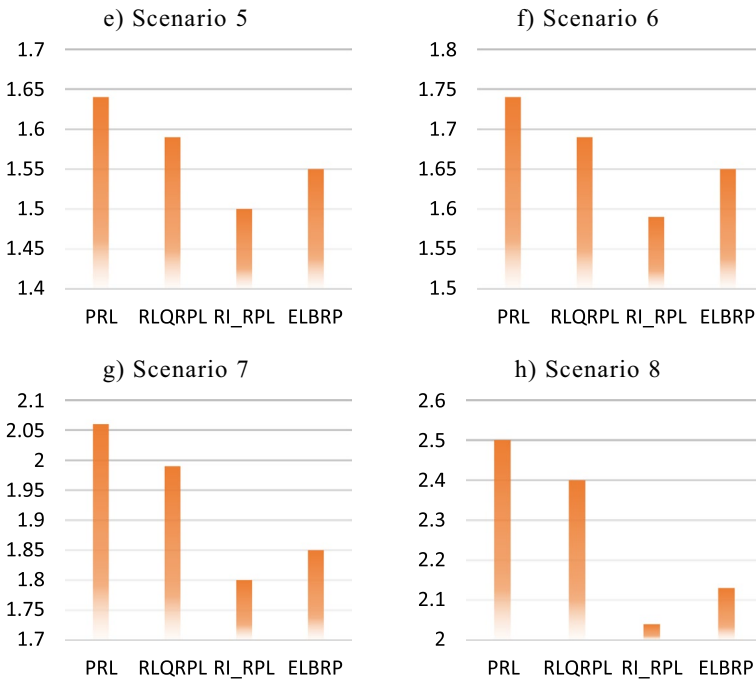
Fig. 13 The results of the end-to-end delay parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 40 nodes with traffic rates: **e** 60 PPM, **f** 120 PPM, **g** 180 PPM, and **h** 240 PPM

and the traffic rate. In addition, these two parameters are the most important parameters in evaluating the efficiency of methods in the network, especially IoT-based networks. In this regard, we have proposed various simulation scenarios based on different values of the two parameters to simulate all possible situations.

6.2 Evaluated metrics and results

In this section, the results obtained from the simulations are presented, and the relevant analyses are discussed. The parameters used to evaluate the performance of the protocols are among the most critical parameters related to the routing domain and IoT exchanges, significantly affecting the performance of the compared protocols. These parameters and details are presented below.

6.2.1　Success delivery ratio of the sink node

The graphs presented in Figs. 7, 8, 9, 10 and 11 depict the evaluation results of the network's successful reception rate under varying numbers of nodes and variable traffic rates according to the scenarios defined in Table 5 for the proposed RI-RPL
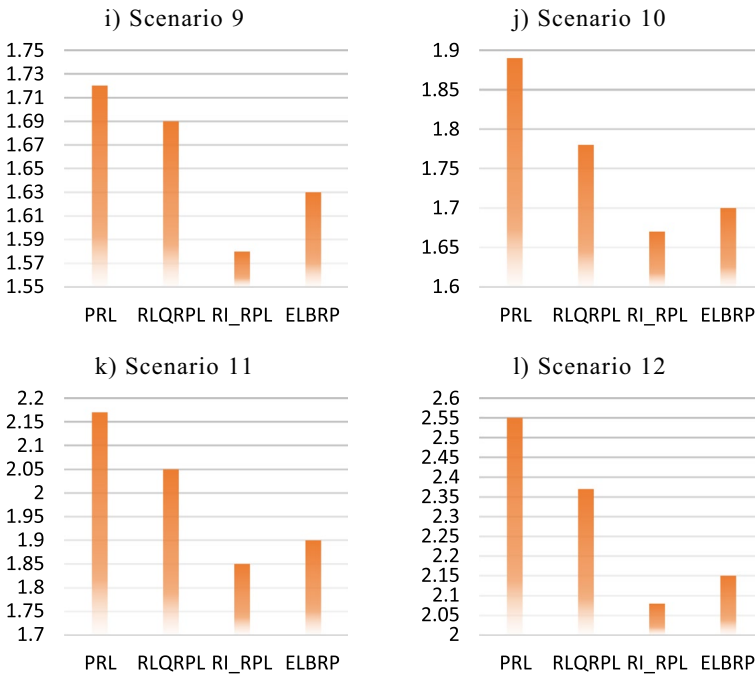
Fig. 14 The results of the end-to-end delay parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 60 nodes with traffic rates: **i** 60 PPM, **j** 120 PPM, **k** 180 PPM, and **l** 240 PPM

protocol, in comparison with three other protocols: RLQRPL, ELBRP, and RPL. This parameter is calculated based on the total successfully received packets by the sink node concerning all transmissions and is calculated using Eq. (13), where SDR represents the successful reception rate of the network (sink node).

$$SDR = \frac{\sum No.\ of\ Data\ packet\ Recevied}{\sum No.\ of\ Data\ packet\ Send} \times 100 \qquad (13)$$

Please note that in the Internet of Things, the network's successful reception rate is evaluated based on the ratio of packets successfully received by the sink node to the total sent packets. Therefore, in the conducted experiments, only packets that were received correctly and successfully by the sink node were considered in the calculations for this parameter. The data transmission rate is the same for all four protocols. However, the successful reception results vary, such that the protocol that performs better in routing and data exchanges, supporting QoS more effectively, will have better successful reception. Overall, this parameter has changed based on two main factors and results. The first factor relates to the discussion of QoS (Quality of Service). This refers to the protocols' ability to provide quality routing and support for QoS in a way that the better the protocol performs in routing and data exchanges,
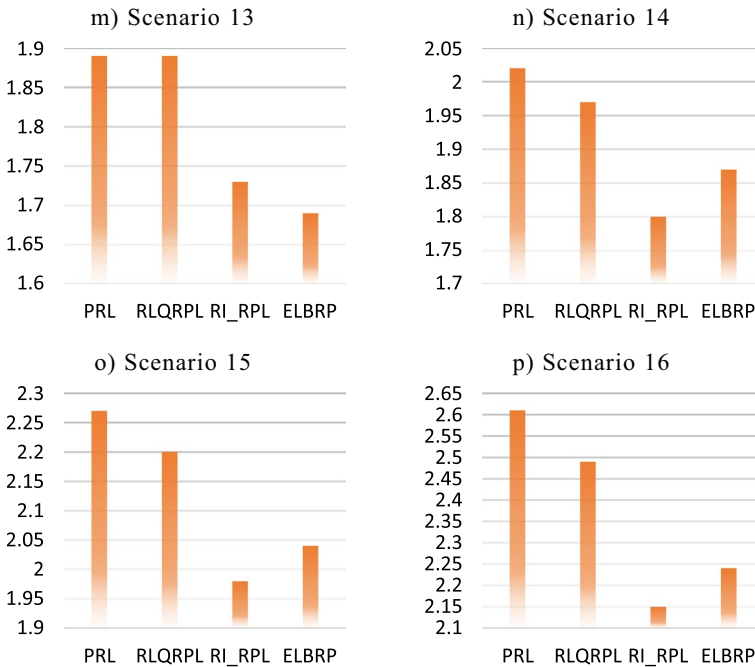
Fig. 15 The results of the end-to-end delay parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 80 nodes with traffic rates: **m** 60 PPM, **n** 120 PPM, **o** 180 PPM, and **p** 240 PPM

the better the successful receptions will be. Therefore, RI-RPL and ELBRP protocols effectively ensure service quality. These two protocols evaluate multiple metrics during routing, ensuring that the highest quality and most suitable intermediate routes are selected. RLQRPL also performs well in this regard but focuses only on evaluating ETX and signal quality, making it less effective than the other two protocols in this aspect. RPL does not provide measures for this purpose and is inefficient in supporting QoS. The second factor relates to the reliability and stability of the communication graph. In this regard, only RI-RPL is effective and efficient. The proposed protocol is the only one that has been successful in evaluating the reliability of routes and ensuring the adaptability of the graph in the face of parent dynamics, based on the concept of learning. Based on the results, with an increase in the number of nodes, the successful reception rate for all four protocols decreased, as this is due to the increased interference and issues caused by the increasing number of nodes. However, the reduction rate in successful receptions for RI-RPL compared to the other three protocols was lower. The reason for this is that RI-RPL's practical measures are more tangible in scenarios with a larger number of nodes, making the proposed protocol more successful in these scenarios. Figures 7, 8, 9, and 10 present the results of the success delivery ratio parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 20, 40, 60, and 80 nodes, respectively.
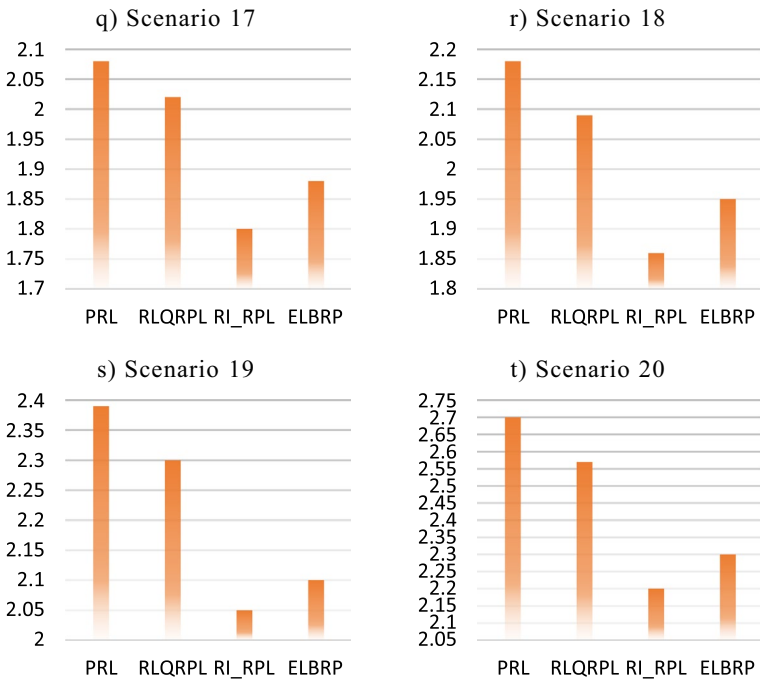
q) Scenario 17

r) Scenario 18

s) Scenario 19

t) Scenario 20

Fig. 16 The results of the end-to-end delay parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 100 nodes with traffic rates: **q** 60 PPM, **r** 120 PPM, **s** 180 PPM, and **t** 240 PPM

Figure 11 presents the results of the success delivery ratio parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 100 nodes with variable traffic rates. According to the obtained results, with an increase in traffic rate, the successful reception rate showed a decreasing trend. However, it is important to note that RI-RPL was more efficient in scenarios with heavy traffic compared to scenarios with light traffic. RI-RPL effectively managed congestion events and unexpected errors, making it more successful in traffic management. ELBRP focuses on traffic management and congestion control, and in this regard, its performance is better as network traffic increases.

6.2.2 End-to-end delay

The graphs presented in Figs. 12, 13, 14, 15 and 16 depict the evaluation results of the end-to-end delay rate under varying numbers of nodes and variable traffic rates according to the scenarios defined for the Proposed RI-RPL protocol, in comparison to three other protocols: RLQRPL, ELBRP, and RPL. This parameter is based on the average delay of received data, as calculated by Eq. (14), where ED represents the average end-to-end delay, Delay$_i$ is the delay of the received packet, and No.ofthepacketreceived is the total number of received packets.
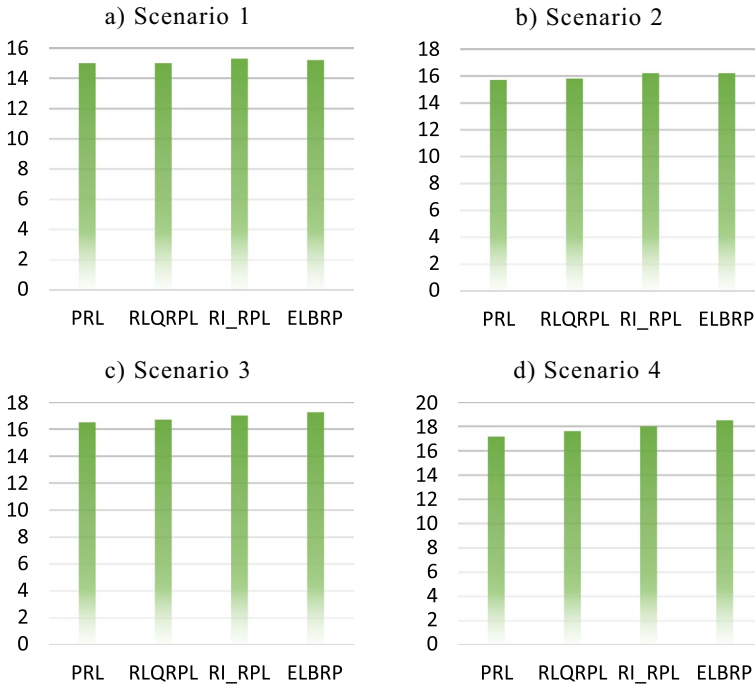
Fig. 17 The results of the overheads parameter for the Proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 20 nodes with traffic rates: **a** 60 PPM, **b** 120 PPM, **c** 180 PPM, and **d** 240 PPM

$$ED = \frac{\sum_{i=1}^{\text{No. of packet received}} \text{Delay}_i}{\text{No. of packet received}} \tag{14}$$

Delay is a qualitative parameter and one of the crucial factors, especially about the Quality of Service (QoS) of transmitted data. It's worth noting that this parameter has been evaluated in experiments based only on the data that have been successfully received. Essentially, the average delay presented in the results is the delay of packets that have been successfully received by the sink node. This parameter has a direct connection with the performance of protocols in supporting QoS. However, maintaining the stability of the network topology and the convergence of the network also significantly affects the reduction of unwanted factors, increasing delay. According to the results, with an increase in both the number of nodes and the traffic rate, the delay has increased. The reason for this is that an increase in the number of nodes leads to longer intermediate routes, which is a key factor in increasing delay. Furthermore, an increase in traffic rate leads to network congestion, which directly affects an increase in the delay of exchanges. However, RI-RPL has performed more successfully in scenarios with increasing factors. RI-RPL, based on its capabilities in ensuring and supporting QoS, minimizes and controls the delay of
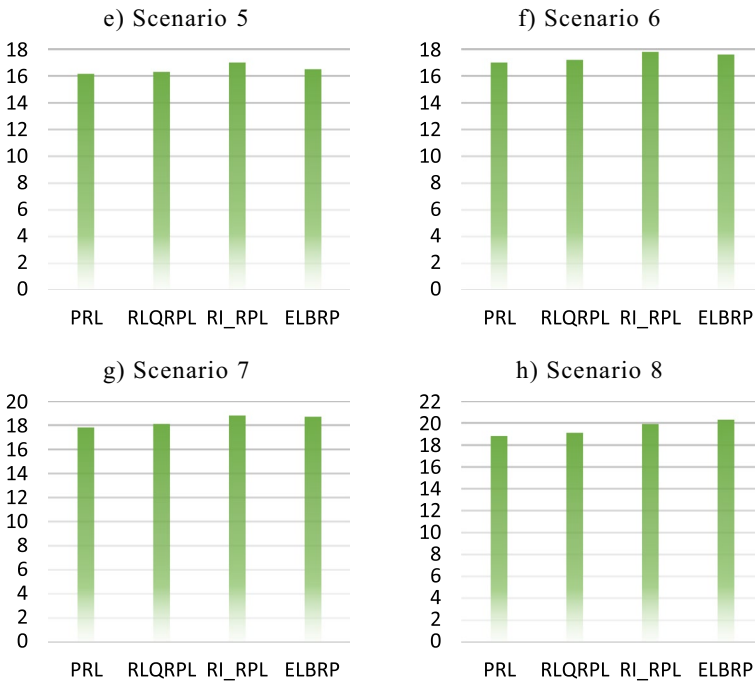
Fig. 18 The results of the overheads parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 40 nodes with traffic rates: **e** 60 PPM, **f** 120 PPM, **g** 180 PPM, and **h** 240 PPM

exchanges as much as possible. These provisions have had more significant effects in scenarios with a higher number of nodes, resulting in a noticeable improvement in delay in these scenarios. Additionally, RI-RPL effectively manages issues arising from increased congestion and inefficiency of parents by ensuring the adaptability and dynamism of parents. This has had significant effects in scenarios with higher traffic rates (due to traffic intensity and fluctuations), resulting in more successful performance in these scenarios.

ELBRP is efficient in traffic management and congestion control, and in this regard, its performance is much more successful in scenarios with heavier traffic. The other two protocols, especially RPL, do not provide significant measures for controlling and optimizing delay, and as a result, exchanges in these methods are accompanied by an increase in delay compared to the other two protocols. Figure 12 presents the results of the end-to-end delay parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 20 nodes with variable traffic rates. Figure 13 presents the results of the end-to-end delay parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 40 nodes with variable traffic rates. Figure 14 presents the results of the end-to-end delay parameter for the proposed RI-RPL protocol compared to
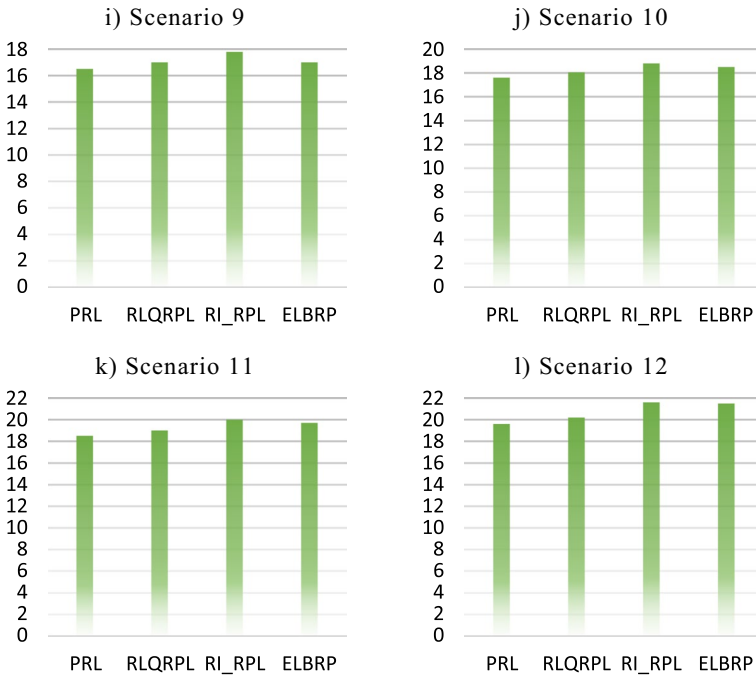
Fig. 19 The results of the overheads parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 60 nodes with traffic rates: **i** 60 PPM, **j** 120 PPM, **k** 180 PPM, and **l** 240 PPM

RPL, ELBRP, and RLQRPL in the presence of 60 nodes with variable traffic rates. Figure 15 presents the results of the end-to-end delay parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 80 nodes with variable traffic rates. Figure 16 presents the results of the end-to-end delay parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 100 nodes with variable traffic rates.

6.2.3 Control overheads

The presented graphs in Figs. 17, 18, 19, 20 and 21 provide the evaluation results of the control overhead rate with varying numbers of nodes and variable traffic rates for the proposed RI-RPL protocol in comparison to three other protocols: RLQRPL, ELBRP, and RPL. This parameter is calculated based on the quantity of control data transmitted in the network, as detailed in Eq. (15), where OH represents the quantity of control overheads in the network.
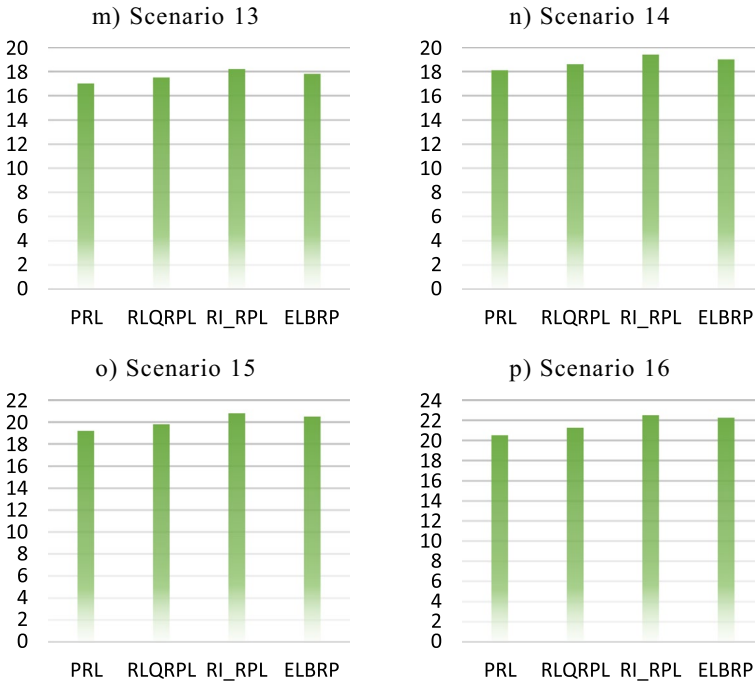
m) Scenario 13

n) Scenario 14

o) Scenario 15

p) Scenario 16

Fig. 20 The results of the overheads parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 80 nodes with traffic rates: **m** 60 PPM, **n** 120 PPM, **o** 180 PPM, and **p** 240 PPM

$$OH = \frac{\sum No. \text{ of control bit send}}{Time \text{ (s)}} \tag{15}$$

Control overheads refer to non-data packets that are exchanged in the network to perform auxiliary tasks, especially routing. In general, the data sent in IoT networks can be divided into two categories. The first category is the primary network data (such as pollution measurement results in environmental applications). The second category includes non-primary network data, which is transmitted to advance the objectives of routing processes in the communication network. These supplementary data are called control overheads. The existence of these overheads is necessary for achieving the goals of the methods and implementing their measures, but if not controlled, they can lead to increased network traffic and energy consumption, ultimately resulting in reduced network performance. Therefore, the performance of proposed protocols should be optimized as much as possible.

In line with what was mentioned regarding the performance of RI-RPL, the proposed protocol uses the extension of the option field in DIO packets to share the
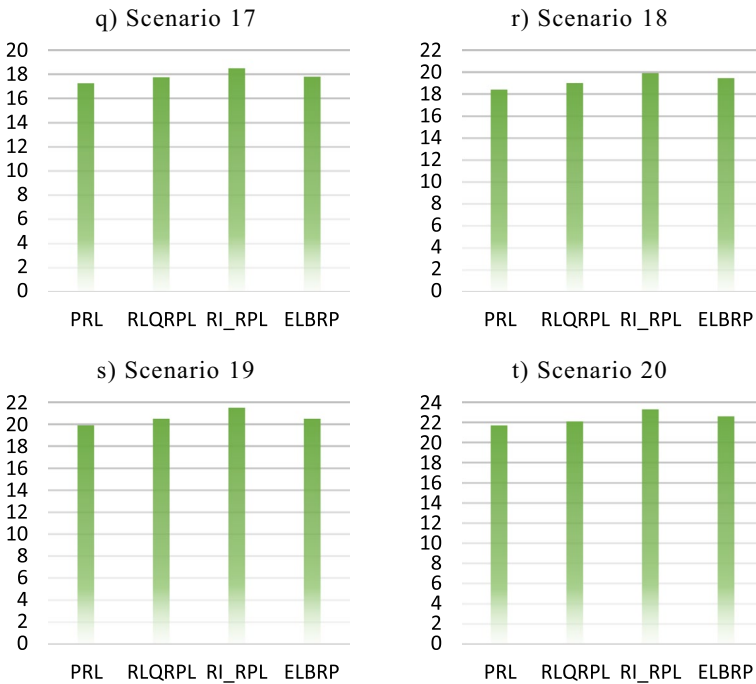
Fig. 21 The results of the overheads parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 100 nodes with traffic rates: **q** 60 PPM, **r** 120 PPM, **s** 180 PPM, and **t** 240 PPM

quality conditions of nodes, resulting in a negligible increase in control overheads. Additionally, in the proposed protocol, the concepts of Q-learning are used to control the conditions of parents and ensure graph dynamism, which results in parent changes when necessary.

While this strategy ensures the preservation of graph superiority under various conditions, it does impose a minor control overhead on the network. Among other methods, ELBRP, due to its routing development based on multi-metric evaluation and the use of data dissemination techniques (selecting multiple parents to maintain load balance), has imposed higher overheads on the network. This increase in overheads has been more noticeable in scenarios with denser traffic due to the increased need for load distribution.

RPL, on the other hand, does not implement any additional measures, resulting in lower control overheads compared to the other protocols. With an increase in the number of nodes, the complexity of the routing process and graph formation has increased, leading to increased control overheads in all four protocols. Similarly, with an increase in traffic rates, control overheads have also seen a limited
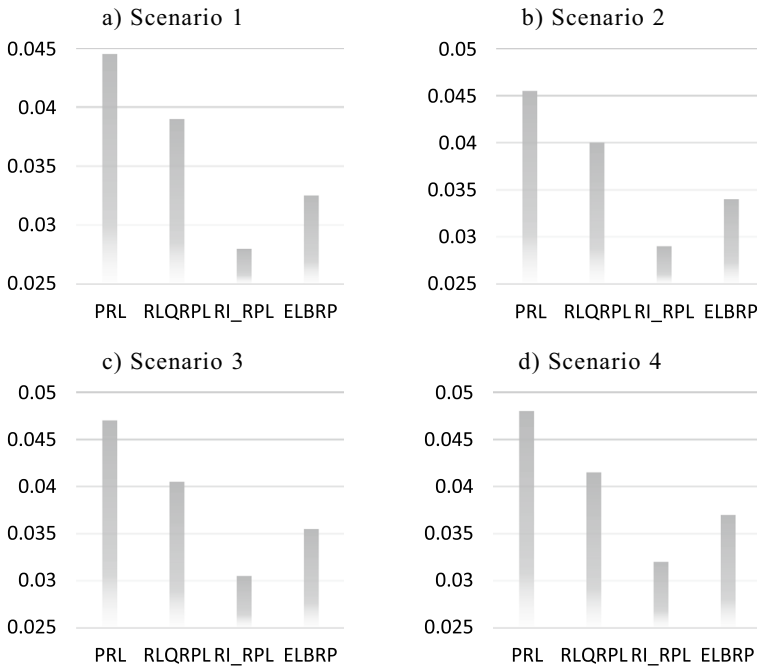
Fig. 22 The results of the energy consumption parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 20 nodes with traffic rates: **a** 60 PPM, **b** 120 PPM, **c** 180 PPM, and **d** 240 PPM

increase. However, it should be noted that in these scenarios, due to the increase in node rates and network congestion, parent changes in RI-RPL have been more frequent, resulting in higher overheads. Nevertheless, considering the substantial and notable improvements in the proposed method in five critical and influential network parameters, namely, successful network receptions, end-to-end delay, energy consumption, throughput, and the insignificance of these overhead rates, this overhead rate becomes relatively inconsequential.

Figure 17 presents the results of the overheads parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 20 nodes with variable traffic rates.

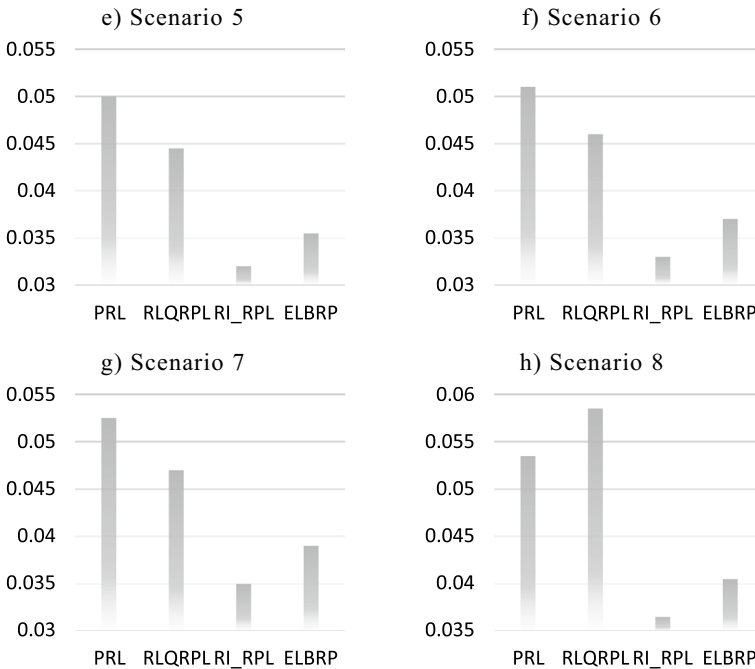Figure 18 presents the results of the overheads parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 40 nodes with variable traffic rates. Figure 19 presents the results of the overheads parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 60 nodes with variable traffic rates. Figure 20 presents the results of the overheads parameter for the proposed RI-RPL protocol compared to RPL, ELBRP,
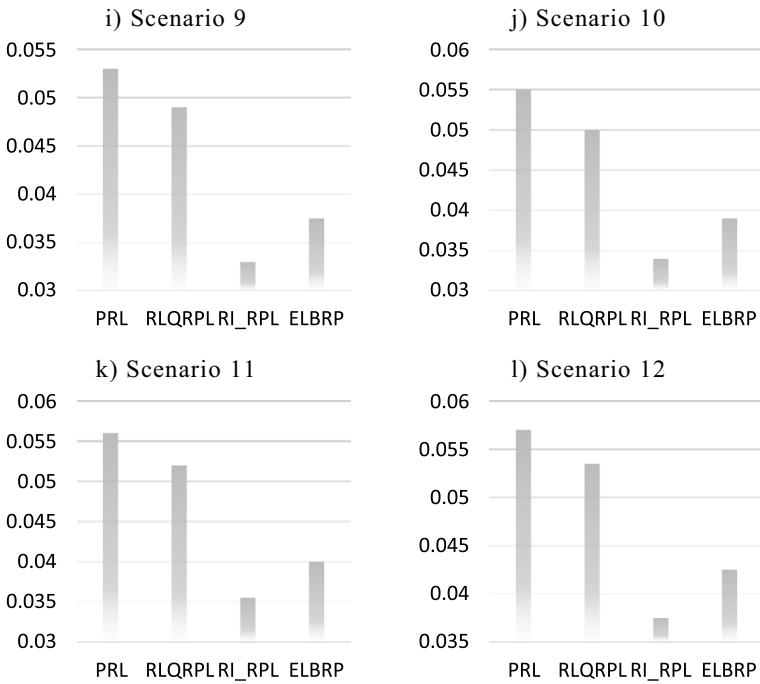
Fig. 23 The results of the energy consumption parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 40 nodes with traffic rates: **e** 60 PPM, **f** 120 PPM, **g** 180 PPM, and **h** 240 PPM

and RLQRPL in the presence of 80 nodes with variable traffic rates. Figure 21 presents the results of the overheads parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 100 nodes with variable traffic rates.

6.2.4 Energy consumption

The presented graphs in Figs. 22, 23, 24, 25 and 26 display the results of energy consumption with varying numbers of nodes and variable traffic rates for the proposed RI-RPL protocol, in comparison with three other protocols: RLQRPL, ELBRP, and RPL. This parameter is evaluated based on the energy consumption rate for network exchanges, as detailed in Eq. (16), where PEC represents the energy consumption rate for each packet.

$$PEC = \text{Energy consumption/packet } (j) \qquad (16)$$

Fig. 24 The results of the energy consumption parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 60 nodes with traffic rates: **i** 60 PPM, **j** 120 PPM, **k** 180 PPM, and **l** 240 PPM

According to the results of experiments, energy consumption has increased with an increase in the number of network members and traffic rates. The reason for this is that an increase in the number of nodes leads to longer intermediate routes and a higher rate of exchanges, resulting in increased network energy consumption. Additionally, an increase in traffic rates has also led to increased network traffic volume, directly impacting increased energy consumption. However, RI-RPL has performed more successfully in scenarios with increasing factors. In scenarios with increasing disturbances, network instability, data loss, and more retransmissions, measures provided by RI-RPL effectively support service quality and graph reliability, minimizing energy consumption. RPL had the highest energy consumption, attributed to its inefficiency in preventing unintended factors and disruptions to quality and reliability. The presence of these issues was more pronounced in increasing scenarios,
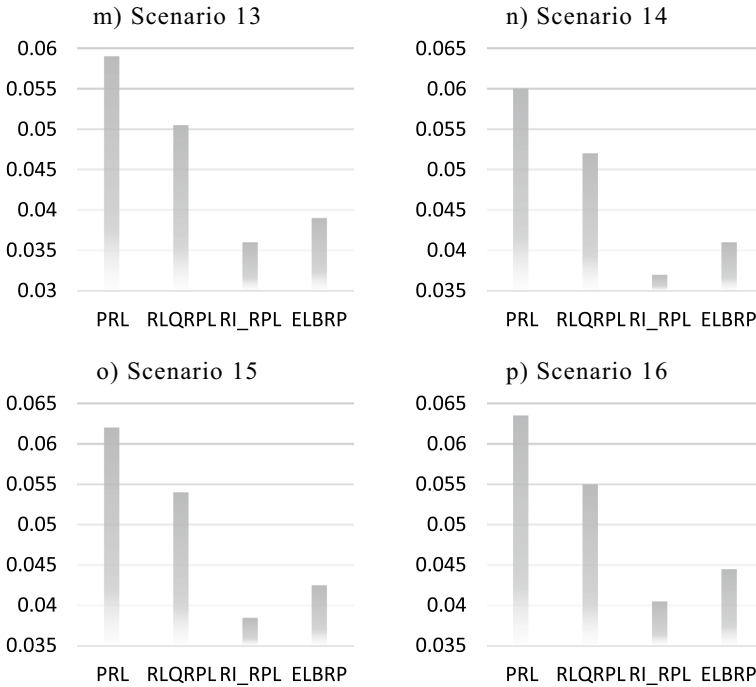
Fig. 25 The results of the energy consumption parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 80 nodes with traffic rates: **m** 60 PPM, **n** 120 PPM, **o** 180 PPM, and **p** 240 PPM

thus leading to higher energy consumption for RPL. ELBRP exhibited better performance than RLQRPL, but both methods were inefficient in maintaining stability and continuity in exchanges, resulting in reduced performance in certain conditions. Figure 22 presents the results of the energy consumption parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 20 nodes with variable traffic rates. Figure 23 presents the results of th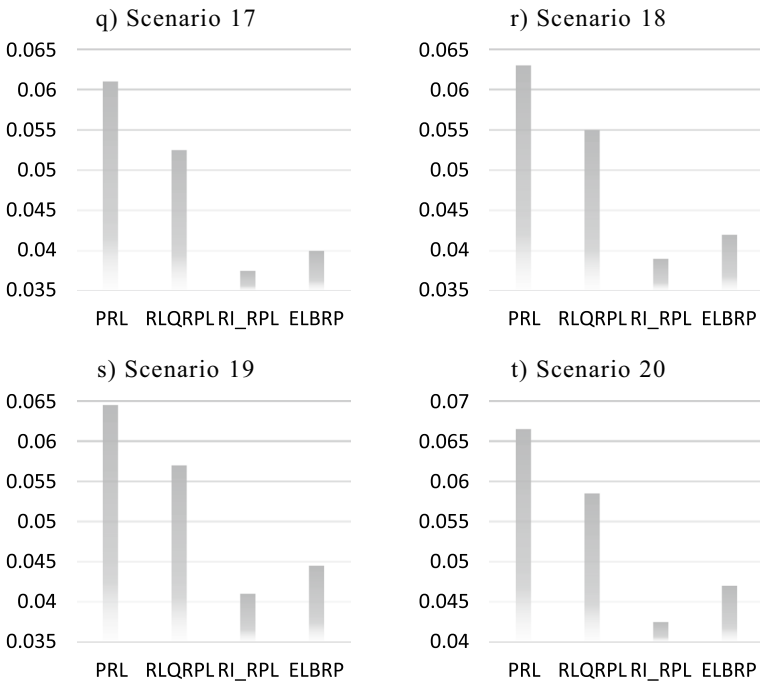e energy consumption parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 40 nodes with variable traffic rates.

Figures 24 presents the results of the energy consumption parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 60 nodes with variable traffic rates. Figure 25 presents the results of the energy consumption parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 80 nodes with variable traffic rates. Figure 26 presents the results of the energy consumption parameter for the proposed

Fig. 26 The results of the energy consumption parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 100 nodes with traffic rates: **q** 60 PPM, **r** 120 PPM, **s** 180 PPM, and **t** 240 PPM

RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 100 nodes with variable traffic rates.

6.2.5 Network data loss rate

The graphs presented in Figs. 27, 28, 29, 30 and 31, respectively, show the results of the network data loss rate with varying numbers of nodes and variable traffic rates for the proposed RI-RPL protocol, in comparison with three other protocols: RLQRPL, ELBRP, and RPL. This parameter is evaluated based on the percentage of lost data compared to the total transmissions, as detailed in Eq. (17), where PDR represents the network data loss rate, calculated as the sum of lost packets over the sum of transmissions.

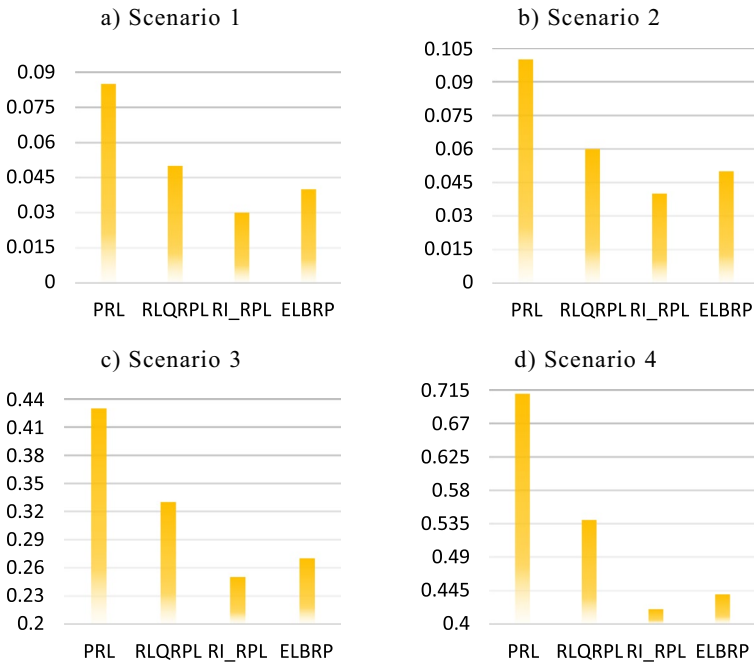$$PDR = \frac{\sum \text{No. of Data Message Drop}}{\sum \text{No. of Data Message Send}} \qquad (17)$$

Fig. 27 The results of the data loss rate parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 20 nodes with traffic rates: **a** 60 PPM, **b** 120 PPM, **c** 180 PPM, and **d** 240 PPM

Regarding this parameter, it is essential to note that while the performance of protocols has a direct and reciprocal impact on data loss, several other issues, including wireless medium interference, node performance degradation, and other unintended factors, also contribute to data loss. However, because the configured scenarios for the protocols were uniform and similar, the impacts of these factors on the results of the protocols were consistent, with the only difference being the performance of the compared protocols. Data loss has a reciprocal relationship with routing confidence and network graph stability. In other words, the more effectively a protocol operates in these aspects, the less data loss occurs, and vice versa.

RI-RPL, based on multi-metric evaluation graph formation (with a focus on energy, congestion, successful reception probability, and link reliability), along with providing the ability to examine route conditions, forms the communication graph through the most reliable nodes and intermediate paths. The presence of this capability, along with support for adaptability and parent changes, ensures the stability and continuity of exchanges, minimizing data loss. ELBRP has also been successful in this area, especially in scenarios with heavy traffic, and has minimized data loss. However, this protocol introduces the ability to assess route conditions and the dynamics of the graph, which has led to data loss
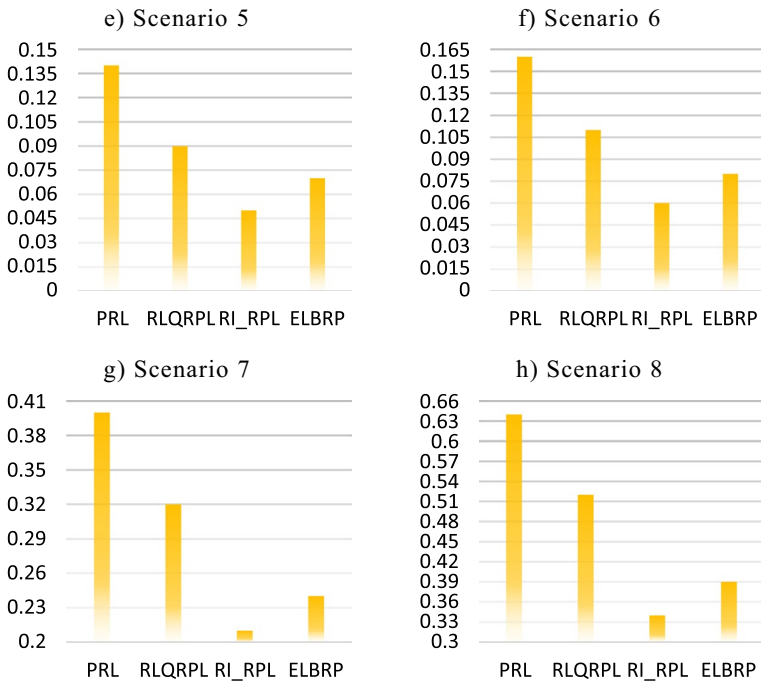
Fig. 28 The results of the data loss rate parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 40 nodes with traffic rates: **e** 60 PPM, **f** 120 PPM, **g** 180 PPM, and **h** 240 PPM

in certain situations. RLQRPL, compared to the two mentioned protocols, has performed weaker but has been more successful than RPL due to its support for quality of service (QoS). RPL has been ineffective in supporting reliability and quality, resulting in increased data loss rates for this method. With an increase in the number of nodes, disruptions and consequently data loss have increased, a trend observed in all four protocols. With an increase in traffic, congestion, and its associated effects, such as data loss, have also increased. However, this issue has been less pronounced for RI-RPL and ELBRP compared to the other two protocols.

Figure 27 presents the results of the data loss rate parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 20 nodes with variable traffic rates. Figure 28 presents the results of the data loss rate parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 40 nodes with variable traff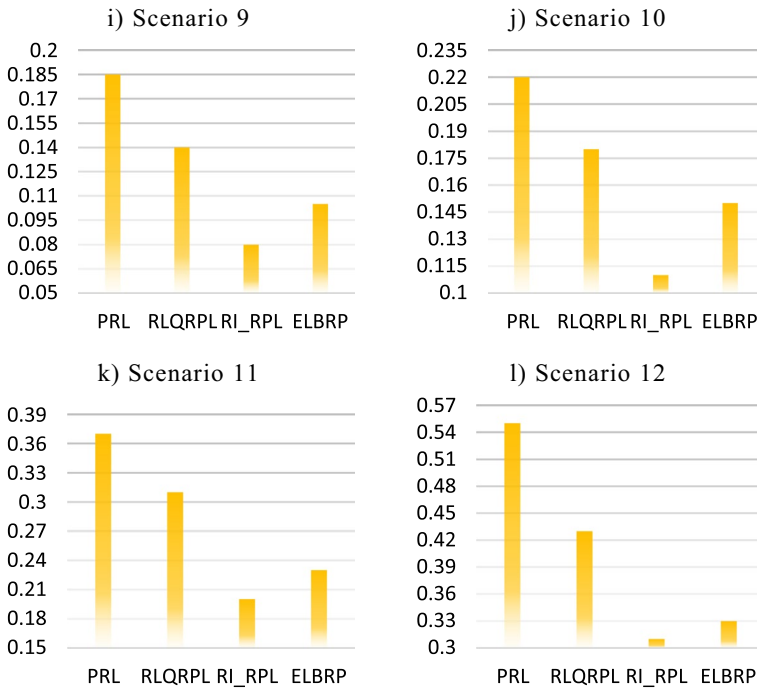ic rates. Figure 29 presents the results of the data loss rate parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 60 nodes with variable traffic rates. Figure 30 presents the results of the data loss rate parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 80 nodes with variable traffic rates.

Fig. 29 The results of the data loss rate parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 60 nodes with traffic rates: **i** 60 PPM, **j** 120 PPM, **k** 180 PPM, and **l** 240 PPM

Figure 31 presents the results of the data loss rate parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 100 nodes with variable traffic rates.

6.2.6 Network throughput

The graphs presented in Figs. 32, 33, 34, 35 and 36 depict the network throughput results with varying numbers of nodes and variable traffic rates for the proposed RI-RPL protocol, in comparison to three other protocols: RLQRPL, ELBRP, and RPL. This parameter is evaluated based on the total received interactions over time, as detailed in Eq. (18), where TG represents the network throughput rate.

$$TG = \frac{\sum \text{No. of all byte recivied} * 8}{\text{Time (s)}} \tag{18}$$

Network throughput is the ideal network transmission rate and has a direct relationship with the performance of protocols, especially their capabilities in supporting quality and reliability. However, other factors have also played a role in changes in this parameter, with their impacts being consistent across the compared protocols.
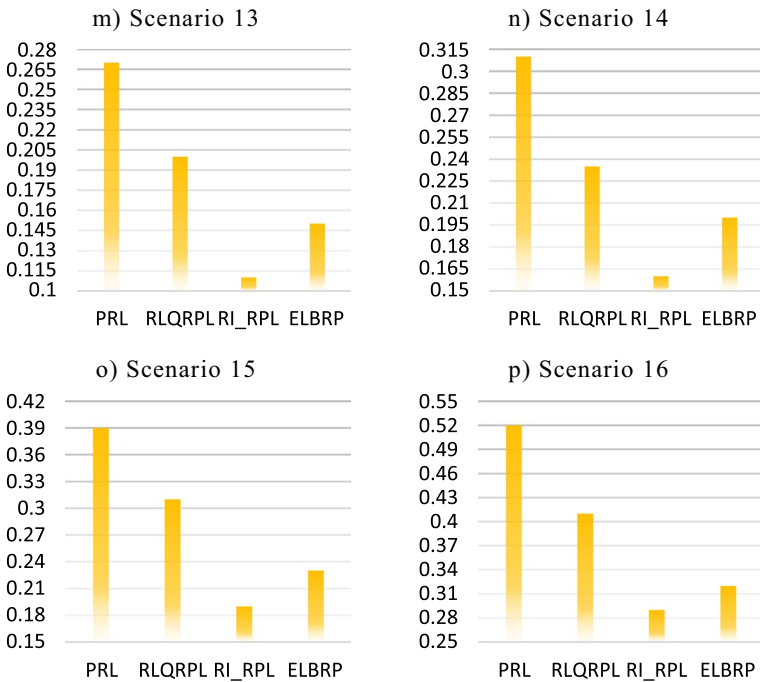
Fig. 30 The results of the data loss rate parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 80 nodes with traffic rates: **m** 60 PPM, **n** 120 PPM, **o** 180 PPM, and **p** 240 PPM

Therefore, what has been highlighted in the conducted experiments is the variations in network throughput about protocol performance. Based on what has been presented in the previous sections and the results shown, the proposed protocol has operated in a way that prioritizes QoS first and then improves reliability to the best extent possible. To achieve this, RI-RPL initially creates the communication graph through paths with maximum quality, subsequently supports parent dynamics based on learning to prevent performance degradation in temporary network conditions, and finally manages topology changes. The result of this targeted design and adaptive performance has been an increase in network throughput for RI-RPL. Among the other three protocols, ELBRP has been more successful due to the effective measures it has taken in supporting QoS and traffic management compared to RLQRPL and RPL. However, in terms of routing aspects, especially ensuring graph dynamism and supporting adaptive routing, it has not performed as well and has shown lower improvements in these aspects compared to RI-RPL. RLQRPL has been more successful than RPL in evaluating connection quality and ETX for routing selections.

However, this protocol does not provide measures to support other needs and requirements, resulting in weaker results compared to RI-RPL and ELBRP in this regard. RPL focuses solely on routing and does not provide the capability to ensure quality and reliability. With an increase in the number of nodes and traffic rates,
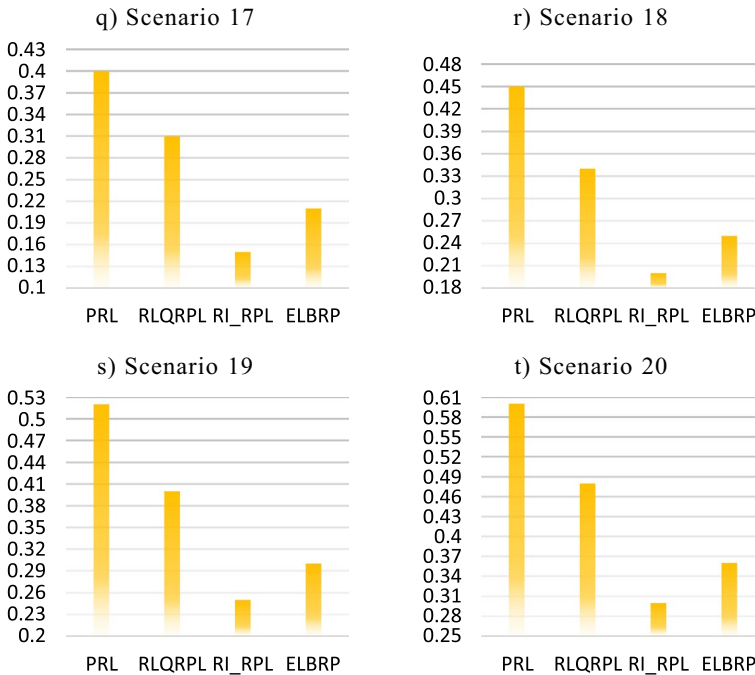
q) Scenario 17

r) Scenario 18

s) Scenario 19

t) Scenario 20

Fig. 31 The results of the data loss rate parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 100 nodes with traffic rates: **q** 60 PPM, **r** 120 PPM, **s** 180 PPM, and **t** 240 PPM

the throughput has shown an increasing trend, and after a certain number of nodes and specific traffic levels, throughput becomes stable and even exhibits a decreasing trend. This issue is a result of the negative impacts of increased traffic and the number of nodes on protocol performance.

Figure 32 presents the results of the throughput parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 20 nodes with variable traffic rates. Figure 33 presents the results of the throughput parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 40 nodes with variable traffic rates. Figure 34 presents the results of the throughput parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 60 nodes with variable traffic rates. Figure 35 presents the results of the throughput parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 80 nodes with variable traffic rates. Figure 36 presents the results of the throughput parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 100 nodes with variable traffic rates.
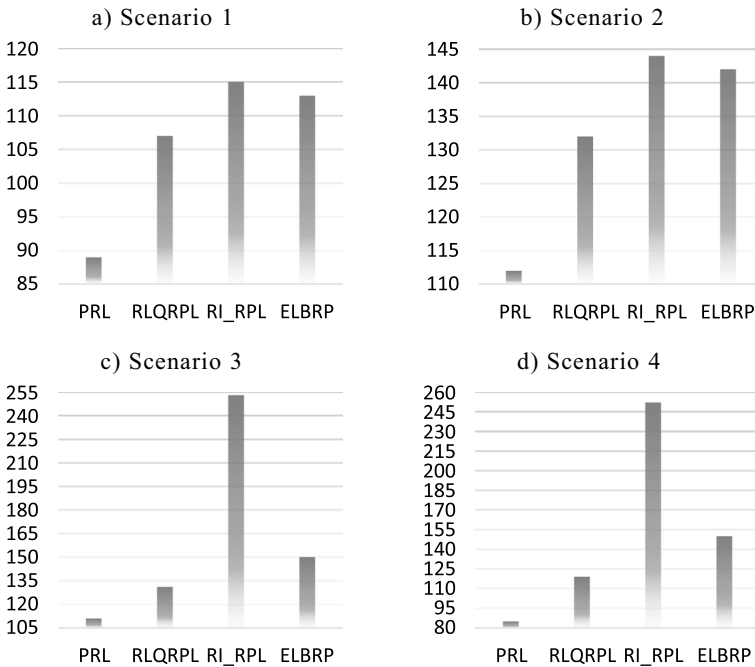
Fig. 32 The results of the throughput parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 20 nodes with traffic rates: **a** 60 PPM, **b** 120 PPM, **c** 180 PPM, and **d** 240 PPM

6.3 Analysis of results

The proposed protocol, RI-RPL, has been designed to enhance the Quality of Service (QoS) and ensure the reliability of Internet of Things (IoT) communications. The results obtained from RI-RPL experiments, across various parameters and factors, indicate a significant improvement and acceptable performance compared to prior research in the field. Table 7 provides an overall evaluation of the performance of the proposed RI-RPL protocol compared to three other protocols under consideration. The parameters under our review are among the most important parameters that increase the efficiency of the network. By improving these parameters, the network has become much more efficient and resistant. Also, these parameters are all in the same direction, which means that the improvement of each one will have a positive effect on the improvement of the rest of the parameters and ultimately the stability of the network.
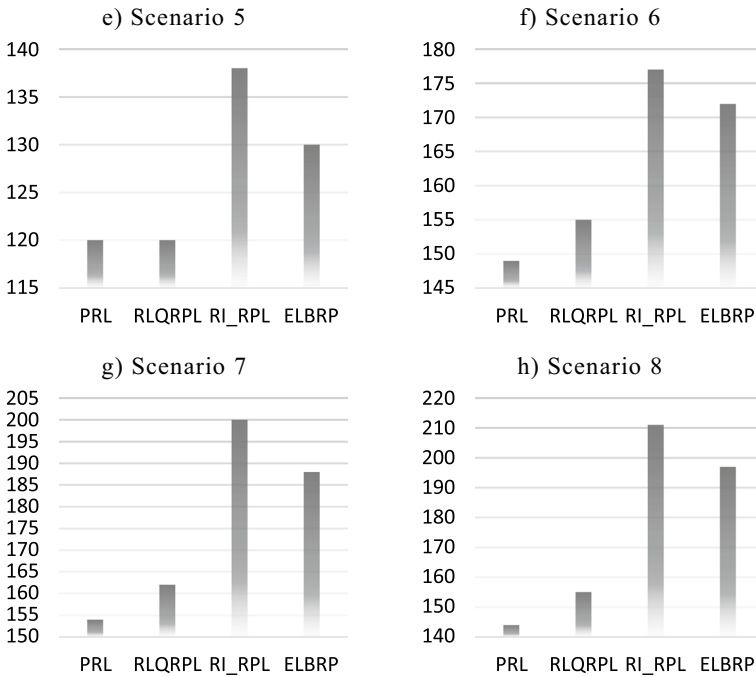
Fig. 33 The results of the throughput parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 40 nodes with traffic rates: **e** 60 PPM, **f** 120 PPM, **g** 180 PPM, and **h** 240 PPM

7 Conclusion

Considering the studies related to the present research in the realm of the Internet of Things (IoT), which is one of the most significant and widely used communication technologies in scientific research, it becomes evident that if QoS is not adequately provided and supported, or if challenges related to routing and data exchanges are not effectively managed, the data communication process will fail, rendering the network non-functional. Given this, and in line with the research focus on ensuring and guaranteeing the QoS of IoT data exchanges, the proposed protocol, RI-RPL (based on the extension of the RPL routing protocol), has been introduced and its details have been elucidated. RI-RPL, to achieve its objectives, implements its performance in three operational phases: performing routing by the optimization of the RPL protocol based on RI-RPL's proposed measures (with a focus on the Q-learning algorithm), ensuring the best selection under various network conditions (based on learning and convergence),
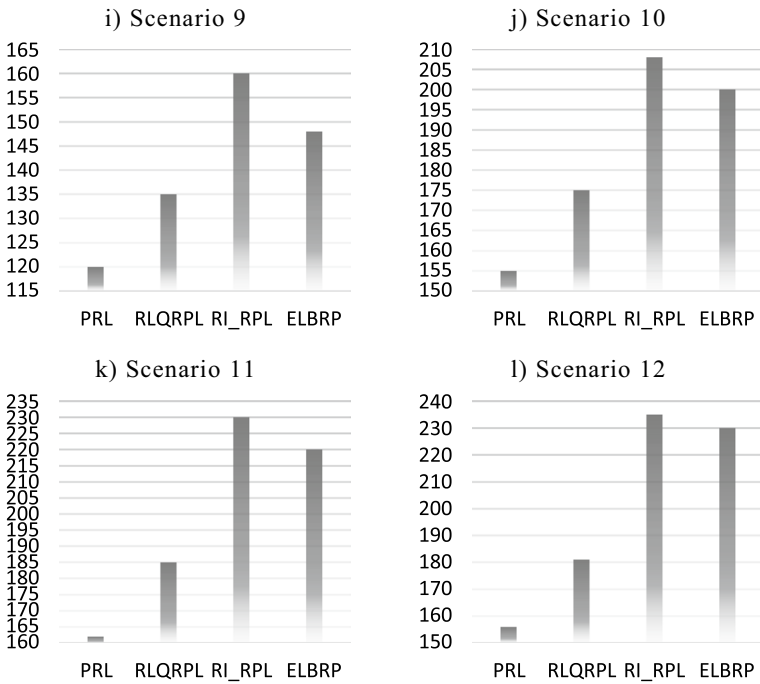
Fig. 34 The results of the throughput parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 60 nodes with traffic rates: **i** 60 PPM, **j** 120 PPM, **k** 180 PPM, and **l** 240 PPM

and controlling and managing changes in network topology. The proposed performance design of RI-RPL has been executed in a way that, in addition to pursuing and achieving research objectives, the proposed protocol is compatible with the characteristics and conditions governing IoT and is well-suited for implementation in various applications. Simulation results using the Cooja software demonstrate improvements in key parameters, including successful receptions, end-to-end delays, energy consumption, network throughput, and data loss rates compared to similar recent methods such as ELBRP, RLQRPL, and RPL. By improving these parameters, the proposed method stabilizes the topology and DODAG graph. Also, despite some works that have problems with traffic or a higher number of nodes, our method has often shown better and more optimal behavior with an increase in the number of nodes and traffic rate. The positive results obtained from the simulation in 20 different scenarios indicate that our method has good scalability and adaptability to different network conditions. The proposed method, in terms of the investigated parameters, has only a small amount of weakness
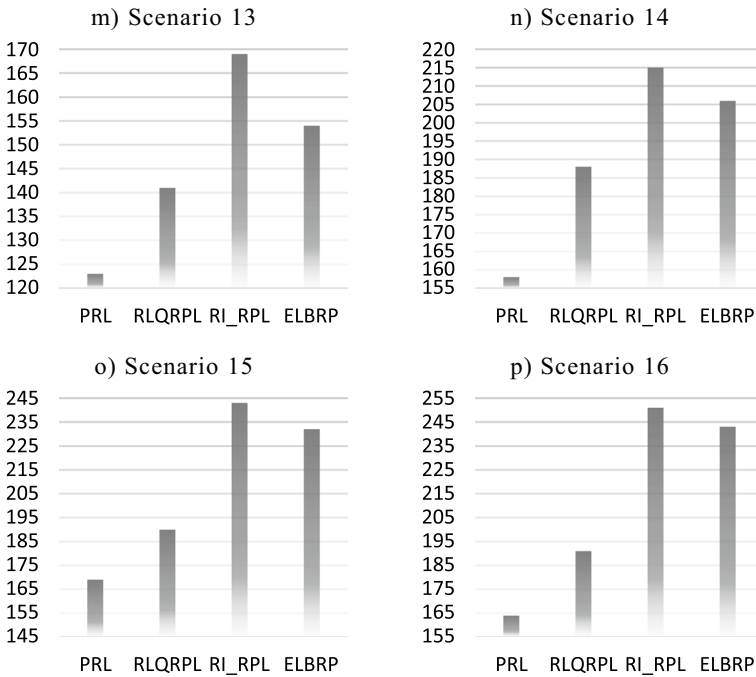
Fig. 35 The results of the throughput parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 80 nodes with traffic rates: **m** 60 PPM, **n** 120 PPM, **o** 180 PPM, and **p** 240 PPM

in the amount of control overhead. which, in exchange for the many improvements it has had, is very insignificant and can be ignored. As future work for the development of the proposed method, expanding the RI-RPL protocol based on location-based routing techniques to ensure QoS for specific IoT applications (including latency-sensitive applications such as medical and industrial uses), equipping the proposed protocol with trust models to support routing trust and data exchanges, in addition to ensuring service quality and extending RI-RPL based on data distribution techniques to improve fault tolerance and increase communication reliability, is proposed.
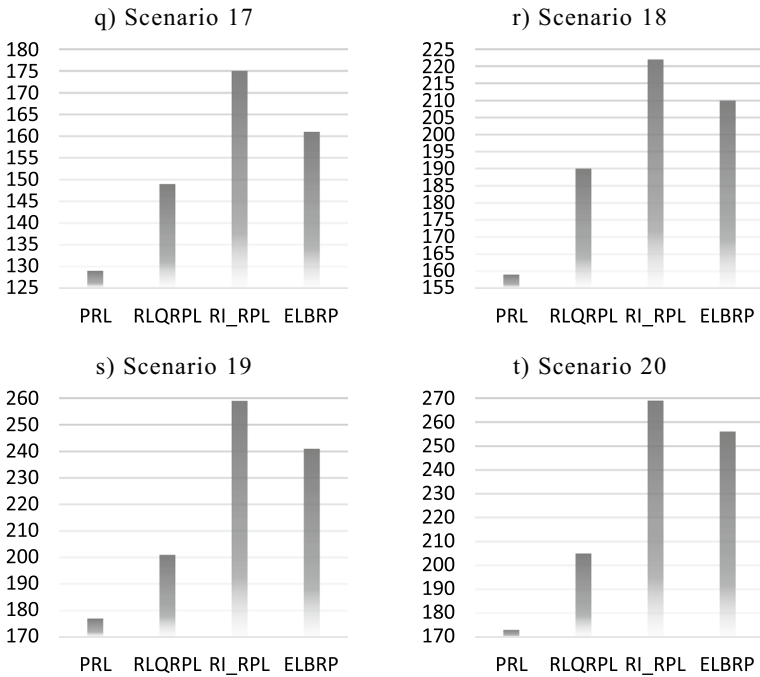
q) Scenario 17

r) Scenario 18

s) Scenario 19

t) Scenario 20

Fig. 36 The results of the throughput parameter for the proposed RI-RPL protocol compared to RPL, ELBRP, and RLQRPL in the presence of 100 nodes with traffic rates: **q** 60 PPM, **r** 120 PPM, **s** 180 PPM, and **t** 240 PPM

Table 7 Review and general conclusion of the simulation results for different scenarios

Criterion	ELBRP	RLQRPL	RPL
Successful receipts	4.03% improvement	13.26% improvement	28.87% improvement
End-to-end delay	3.04% improvement	9.82% improvement	13.12% improvement
Energy consumption	10.43% improvement	28.91% improvement	36.35% improvement
Throughput	10.23% improvement	28.45% improvement	46.88% improvement
Data loss	15.06% improvement	34.95% improvement	49.66% improvement
Control overheads	1.44% weakening	4.29% weakening	6.68% weakening

Author contributions The first two authors worked on all aspects of the paper and the third one checked the simulation process as well as the idea.

Declarations

Competing interests The authors declare no competing interests.

References

1. Osorio A, Calle M, Soto JD, Candelo-Becerra JE (2020) Routing in LoRaWAN: overview and challenges. IEEE Commun Mag 58(6):72–76
2. Bhuiyan MN, Rahman MM, Billah MM, Saha D (2021) Internet of things (IoT): a review of its enabling technologies in healthcare applications, standards protocols, security, and market opportunities. IEEE Internet Things J 8(13):10474–10498
3. Tightiz L, Yang H (2020) A comprehensive review on IoT protocols' features in smart grid communication. Energies (Basel) 13(11):2762
4. Oleiwi HW, Al-Raweshidy H (2023) Cooperative Hybrid-NOMA/Dynamic SWIPT-Pairing Mechanism for 6G THz Communications, in 2023 5th Global Power, Energy and Communication Conference (GPECOM), IEEE, pp 524–529
5. Hassan R, Qamar F, Hasan MK, Aman AHM, Ahmed AS (2020) Internet of Things and its applications: a comprehensive survey. Symmetry (Basel) 12(10):1674
6. Aslam S, Michaelides MP, Herodotou H (2020) Internet of ships: a survey on architectures, emerging applications, and challenges. IEEE Internet Things J 7(10):9714–9727
7. Souri A, Hussien A, Hoseyninezhad M, Norouzi M (2022) A systematic review of IoT communication strategies for an efficient smart environment. Trans Emerg Telecommun Technol 33(3):e3736
8. Chandnani N, Khairnar CN (2020) A comprehensive review and performance evaluation of recent trends for data aggregation and routing techniques in IoT networks. Social Networking and Computational Intelligence: Proceedings of SCI-2018, pp 467–484
9. Oleiwi HW, Al-Raweshidy H (2022) SWIPT-Pairing mechanism for channel-aware cooperative H-NOMA in 6G Terahertz communications. Sensors 22(16):6200
10. Marietta J, Chandra Mohan B (2020) A review on routing in internet of things. Wirel Pers Commun 111:209–233
11. Gopika D, Panjanathan R (2020) Withdrawn: energy efficient routing protocols for WSN based IoT applications: a review. Elsevier
12. Sari RF, Rosyidi L, Susilo B, Asvial M (2021) A comprehensive review on network protocol design for autonomic internet of things. Information 12(8):292
13. Dey AJ, Sarma HKD (2020) Routing techniques in internet of things: a review. Trends in Communication, Cloud, and Big Data: Proceedings of 3rd National Conference on CCB, 2018, Springer, pp 41–50
14. Gaddour O, Koubâa A (2012) RPL in a nutshell: a survey. Comput Netw 56(14):3163–3178
15. Kushalnagar N, Montenegro G, Schumacher C (2007) IPv6 over low-power wireless personal area networks (6LoWPANs): overview, assumptions, problem statement, and goals
16. Shah Z, Levula A, Khurshid K, Ahmed J, Ullah I, Singh S (2021) Routing protocols for mobile Internet of things (IoT): a survey on challenges and solutions. Electronics (Basel) 10(19):2320
17. Lamaazi H, Benamar N (2020) A comprehensive survey on enhancements and limitations of the RPL protocol: a focus on the objective function. Ad Hoc Netw 96:102001
18. Thubert P (2012) Objective function zero for the routing protocol for low-power and lossy networks (RPL)
19. Gnawali O, Levis P (2012) The minimum rank with hysteresis objective function
20. Pancaroglu D, Sen S (2021) Load balancing for RPL-based internet of things: a review. Ad Hoc Netw 116:102491
21. Sankar S, Srinivasan P, Luhach AK, Somula R, Chilamkurti N (2020) Energy-aware grid-based data aggregation scheme in routing protocol for agricultural internet of things. Sustain Comput Inform Syst 28:100422
22. Touzene A, Al Kalbani A, Day K, Al Zidi N (2020) Performance analysis of a new energy-aware RPL routing objective function for internet of things. 2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), IEEE, pp 1–6
23. Rana PJ, Bhandari KS, Zhang K, Cho G (2020) EBOF: a new load balancing objective function for low-power and lossy networks. IEIE Trans Smart Process Comput 9(3):244–251
24. Mishra SN, Elappila M, Chinara S (2020) Eha-rpl: a composite routing technique in iot application. First International Conference on Sustainable Technologies for Computational Intelligence: Proceedings of ICTSCI 2019, Springer, pp 645–657
25. Seyfollahi A, Ghaffari A (2020) A lightweight load balancing and route minimizing solution for routing protocol for low-power and lossy networks. Comput Netw 179:107368

26. Hassani AE, Sahel A, Badri A (2021) FTC-OF: forwarding traffic consciousness objective function for RPL routing protocol. Int J Electr Electron Eng Telecommun 10:168–175
27. Eloudrhiri Hassani A, Sahel A, Badri A (2021) IRH-OF: a new objective function for RPL routing protocol in IoT applications. Wirel Pers Commun 119:673–689
28. Acevedo PD, Jabba D, Sanmartín P, Valle S, Nino-Ruiz ED (2021) WRF-RPL: weighted random forward RPL for high traffic and energy demanding scenarios. IEEE Access 9:60163–60174
29. Wang H, Fan Z, He X, Li P, Zhang C (2021) Improvement of RPL routing strategy based on 6LoW-PAN. Sensor Networks and Signal Processing: Proceedings of the 2nd Sensor Networks and Signal Processing (SNSP 2019), 19–22 November 2019, Hualien, Taiwan, Springer, pp 21–35
30. Zarzoor AR (2021) Optimizing RPL performance based on the selection of best route between child and root node using E-MHOF method. Int J Electr Comput Eng 11(1):224–231
31. Pushpalatha M, Anusha T, Rao TR, Venkataraman R (2021) L-RPL: RPL powered by laplacian energy for stable path selection during link failures in an internet of things network. Comput Netw 184:107697
32. Pingale R, Shinde SN (2021) fuzzy logic approach for routing in internet of things network. Tehnički glasnik 15(1):18–24
33. Kaviani F, Soltanaghaei M (2022) CQARPL: congestion and QoS-aware RPL for IoT applications under heavy traffic. J Supercomput 78(14):16136–16166
34. Hassani AE, Sahel A, Badri A (2022) Towards an enhanced minimum rank hysteresis objective function for RPL IoT routing protocol. WITS 2020: Proceedings of the 6th International Conference on Wireless Technologies, Embedded, and Intelligent Systems, Springer, pp 483–493
35. Charles ASJ, Kalavathi P (2022) A reliable link quality-based RPL routing for internet of things. Soft comput 26(1):123–135
36. Thenmozhi R, Sakthivel P, Kulothungan K (2022) Hybrid multi-objective-optimization algorithm for energy efficient priority-based QoS routing in IoT networks. Wirel Netw. https://doi.org/10.1007/s11276-021-02848-z
37. Darabkh KA, Al-Akhras M, Ala'F K, Jafar IF, Jubair F (2022) An innovative RPL objective function for broad range of IoT domains utilizing fuzzy logic and multiple metrics. Expert Syst Appl 205:117593
38. Fazli F, Mansubbassiri M (2022) V-RPL: an effective routing algorithm for low power and lossy networks using multi-criteria decision-making techniques. Ad Hoc Netw 132:102868
39. Kalantar S, Jafari M, Hashemipour M (2023) Energy and load balancing routing protocol for IoT. Int J Commun Syst 36(2):e5371
40. Seyfollahi A, Taami T, Ghaffari A (2023) Towards developing a machine learning-metaheuristic-enhanced energy-sensitive routing framework for the internet of things. Microprocess Microsyst 96:104747
41. Sutton RS, Barto AG (1998) Introduction to reinforcement learning, vol 135. MIT press, Cambridge