SECURE-RPL: APPROACH TO PREVENT RESOURCE-BASED ATTACKS IN WIRELESS SENSOR NETWORKS USING BALANCED CLUSTERING

BASIM AHMAD ABED ALHAMEAD AL ABSI

UNIVERSITI SAINS MALAYSIA

2020

SECURE-RPL: APPROACH TO PREVENT RESOURCE-BASED ATTACKS IN WIRELESS SENSOR NETWORKS USING BALANCED CLUSTERING

by

BASIM AHMAD ABED ALHAMEAD AL ABSI

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

November 2020

ACKNOWLEDGEMENT

With the name of Allah, Most Gracious, Most Merciful.

Before anything, I would like to thank my creator "Allah' for his bless me to complete this work, which I could not do that without his blessings. Moreover

The satisfaction that accompanies the successful completion of any task would be incomplete without a mention of the people who made it possible, whose constant guidance and encouragement crown all efforts with success.

First of all, I wish to express my deep sense of gratitude and sincere thanks to my Research Supervisor and Guide, DR Mohammad FR Anbar, Department of National Advanced IPv6 Centre, Universiti Sains Malaysia (USM), for his valuable guidance, advice and inspiring suggestions at each and every stage of this research project. I thank him for spending his valuable time on this work. I am thankful to Assoc. Prof. Dr Selvakumar Manickam, Co-Supervisor, Department of National Advanced IPv6 Centre, Universiti Sains Malaysia (USM) for his valuable comments and suggestion. I would like also thanks the NAv6 members for their encourage and help, especially Dr Sabri Mohammad who spent his efforts for helping and advising me in my PhD. I am thankful to my colleagues for their moral support extended during my research work.

In this situation I never forget my Father, Ahmad May God have mercy on him, and my mother, Zina May God have mercy on him for their prayers were the most important thing for me to be strong and reach to what I'm now.

Last but not the least, my thank also unplugged to my wife Wafa, daughters Duha and Bushra, sons Mohammad and Ahmad, relatives and my friends whose blessings have given me the much-needed strength to complete my thesis work successfully.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACO	Ant Colony Optimization
AES	Advance Encryption Standard
AMQP	Advanced Message Queuing Protocol
CBC	Coordinative Balanced Clustering
CH	Cluster Head
CoAP	Constrained Application Protocol
DAO	Destination Advertisement Object
DDoS	Distributed Denial of Service
DDS	Data Distribution Service
DIO	DODAG Information Object
DIS	DODAG Information Solicitation
DNS-SD	DNS Service Discovery
DODAG	Destination Oriented Direct Acyclic Graph
DoS	Denial of Service
DTSN	Destination advertisement Trigger Sequence Number
E-ACO	Enriched ACO
EPC	Electronic Product Code
ETX	Expected Transmission Count
FFD	Full Function Device
FSM	Finite State Machine
HMI	Human Machine Interface
ICMP	Internet Control Message Protocol
IDS	Intrusion Detection System
IETF	Internet Engineering Task Force
InReS	Intrusion Detection and Response System
IoT	Internet of Things
LBR	Low Power and Lossy Border Router
LBS	Location Based Service
LLN	Low Power and Lossy Network
LOAD	LoWPAN Ad-hoc On-demand Distance vector routing
LoWPAN	Low Power Wireless Personal Area Network

mDNS	Multicast Domain Name Service
MQTT	Message Queue Telemetry Transport
MST	Minimum Spanning Tree
MTC	Machine Type Communication
NFC	Near Field Communication
OF	Objective Function
PDR	Packet Delivery Ratio
PLC	Programmable Logic Controller
QoS	Quality of Service
RFD	Reduced Function Device
RFID	Radio Frequency Identification
RPL	Routing Protocol for Low Power and Lossy Network
RSSI	Received Signal Strength Indicator
RTU	Remote Telemetry Unit
SIoT	Social Internet of Things
SPRT	Sequential Probability Ratio Test
TM	Trickle Multicast
TRAIL	Trust Anchor Interconnection Loop
UDP	User Datagram Protocol
VeRA	Version number Rank Authentication
WSN	Wireless Sensor Network
XMPP	Extensible Messaging and Presence Protocol

LIST OF APPENDICES

SECURE-RPL: PENDEKATAN UNTUK MENCEGAH SERANGAN-SERANGAN BERASASKAN SUMBER DALAM RANGKAIAN SENSOR WAYARLES MENGGUNAKAN PENGGUGUSAN TERIMBANG

ABSTRAK

Internet benda (IoT) adalah suatu teknologi pengkomputeran yang sedang berkembang yang membolehkan peranti-peranti fizikal saling sambung antara mereka, yang menawarkan banyak kelebihan seperti akses maklumat yang mudah, keberkesanan kos, automasi, penggunaan sumber yang efisien, pengurangan penggunaan daya manusia, dan meningkatkan produktiviti, yang semuanya telah menarik perhatian banyak pemain industri dan para penyelidik. Walau bagaimanapun, penglibatan bilangan peranti dan pengguna IoT dalam kuantiti yang sangat besar menimbulkan banyak isu, termasuk yang berkaitan dengan kualiti perkhidmatan dan keselamatan. Dalam IoT, penghalaan antara peranti-peranti dan nod-nod yang terbatas sumbernya direalisasikan dengan menggunakan protokol penghalaan untuk rangkaian kuasa rendah dan bersifat susut (RPL), yang memilih laluan optimum mengikut fungsi objektif yang khusus. Walau bagaimanapun, RPL berdepan banyak ancaman keselamatan, yang paling ketara adalah serangan berasaskan sumber, termasuk tetapi tidak terhad kepada serangan nafi khidmat teragih (DDoS) dan serangan nombor versi. Oleh itu, keselamatan rangkaian RPL perlu dipertingkatkan. Untuk tujuan ini, kajian ini mencadangkan satu pendekatan yang dinamai sebagai *Secure-RPL* untuk mengesan kewujudan serangan berasaskan sumber dalam rangkaian RPL seperti serangan-serangan DDoS banjir dan nombor versi. Pendekatan ini mempunyai tiga fasa utama, iaitu, (i) algoritma penggugusan berasaskan koordinat (CBC), iaitu algoritma peka tenaga yang memanjangkan jangka hayat rangkaian RPL untuk meminimumkan risiko serangan berasaskan sumber, (ii) pengesanan DDoS berasaskan koloni semut diperkaya, yang bertujuan untuk mengesan serangan DDoS, dan (iii) penghalaan berasaskan algoritma *secure*-RPL, yang bertujuan memilih laluan selamat menurut faktor penskoran bagi nod dan metrik-metrik nod lain yang bererti untuk mencegah serangan nombor versi. Dua fasa terkemudian menggunakan CBC sebagai input. Pendekatan yang dicadangkan dinilai dari segi kelengahan hujung ke hujung, nisbah penghantaran paket, kadar kehilangan paket, bilangan nod mati dan penggunaan tenaga dengan menggunakan simulator NS3. Hasil penilaian menunjukkan bahawa *Secure-RPL* yang dicadangkan mempunyai 7.25%, 6.375% dan 2.625% lebih rendah kelengahan hujung ke hujung berbanding dengan HECRPL, IRPL dan QU-RPL masing-masing. Sementara itu, pendekatan *Secure-RPL* mempunyai 34.96%, 36.76% dan 31.76% nisbah penghantaran paket lebih tinggi berbanding dengan HECRPL, IRPL dan QU-RPL. Dari segi kadar kehilangan paket, pendekatan *Secure-RPL* mempunyai masing-masing 0.115%, 0.0725% dan 0.0825% kadar yang lebih rendah berbanding dengan HECRPL, IRPL dan QU-RPL. Bagi bilangan nod yang mati, pendekatan *Secure-RPL* mempunyai purata sebanyak 16, 14.75 dan 7.75 bilangan nod lebih rendah berbanding dengan HECRPL, IRPL dan QU-RPL. Akhir sekali, pendekatan *Secure-RPL* mempunyai 0.278mW, 0.368mW dan 0.328mW purata penggunaan tenaga lebih rendah berbanding dengan HECRPL, IRPL dan QU-RPL. Keputusan ini mengesahkan bahawa *Secure-RPL* mengatasi prestasi pendekatan sedia ada yang lain.

SECURE-RPL: APPROACH TO PREVENT RESOURCE-BASED ATTACKS IN WIRELESS SENSOR NETWORKS USING BALANCED CLUSTERING

ABSTRACT

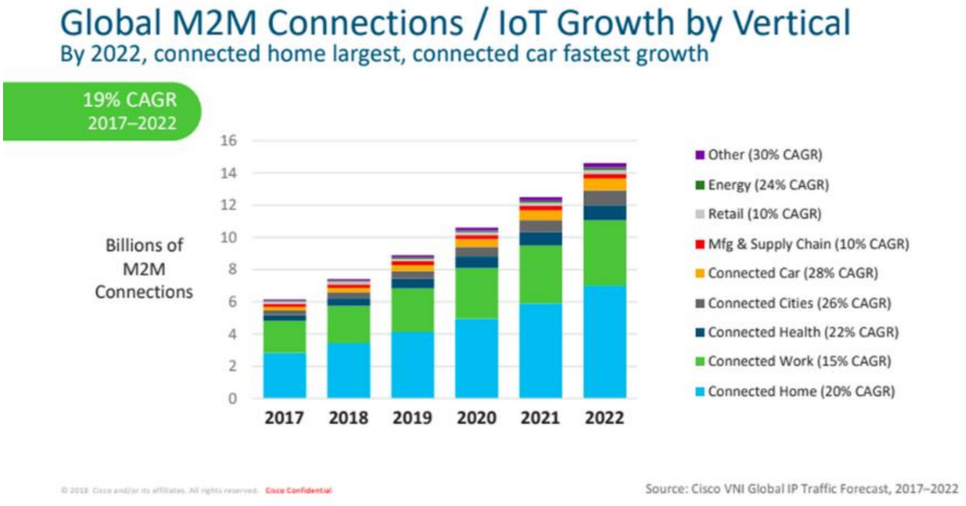
Internet of Things (IoT) is an evolving computing technology that enables an interconnection amongst physical devices, which offers many advantages, such as easy access to information, cost effectiveness, automation, efficient resource utilisation, reduced human effort and high productivity, all of which have attracted many industry players and researchers. However, the involvement of a vast number of devices and IoT users introduces many issues, including those related to quality of service and security. In IoT, routing amongst resource-constrained devices and nodes is realised by using the routing protocol for a low-power and lossy network (RPL), which selects an optimal route according to the specific objective function. However, RPL is not energy-aware protocol which make it faces many security threats, the most significant of which are resource-based attacks, which include but are not limited to distributed denial-of-service (DDoS) and version number attacks. Therefore, the security of the RPL network needs to be improved. To this end, this research proposes an approach named Secure-RPL for preventing resource-based attacks in an RPL such as DDoS flooding and version number attacks using balanced clustering. This approach has three main phases, namely, (i) coordinative-based clustering algorithm (CBC), which is an energy-aware mechanism that extends the RPL network lifetime to minimise the risk of resource-based attacks, (ii) enriched-ant-colony-based DDoS detection, which aims to detect DDoS attacks and (iii) Sec-RPL mechanism prevent the data transmission from version number attacks in accordance to the scoring factor of the node and other significant node metrics. The latter two phases utilise CBC as an input. The proposed approach is evaluated with the presence of resource-based attacks in term of end to end delay, packet delivery ratio, packet loss rate, number of dead nodes and energy consumption using the NS3 simulator. The evaluation results reveal that the proposed secure-RPL has 7.25%, 6.375 % and 2.625% lesser end to end delay compared with HECRPL, IRPL and QU-RPL, respectively. Meanwhile, Secure-RPL has 34.96%, 36.76 % and 31.76% higher packet delivery ratio compared with HECRPL, IRPL and QU-RPL, respectively. In term of packet loss rate, Secure-RPL approach has 0.115%, 0.0725 % and 0.0825% lesser packet loss rate compared with HECRPL, IRPL and QU-RPL, respectively. As for number of dead nodes, Secure-RPL approach has 16, 14.75 and 7.75 lower average of dead nodes compared with HECRPL, IRPL and QU-RPL, respectively. Lastly, Secure-RPL approach has 0.278mW, 0.368 mW and 0.328 mW lower average of energy consumption compared with HECRPL, IRPL and QU-RPL respectively. These results confirmed that secure-RPL outperforms other existing proposed approaches.

CHAPTER 1

INTRODUCTION

1.1 Overview

The recent growth of the Internet of Things (IoT) represents the arrival of an entirely new technology. The common communication technologies in IoT include radio-frequency identification and wireless sensor network (WSN) technologies (Zou et al , 2014). IoT has been applied in six application areas, including the social domain, industrial domain, transportation domain, health domain, smart city and smart environments, all of which partially overlap with one another as they share some applications (Al-Fuqaha et al, 2015). Recent studies reveal that the majority of the companies worldwide will use IoT applications by the end of 2019 (Wortmann & Flüchter, 2015), whereas others predict that 20 to 50 billion IoT objects will be connected by 2020. Meanwhile, Cisco predicts that the global adaptation rate of IoT will increase by 2022 as shown in Figure 1.1 (Evans, 2011).



Source: Cisco VNI Global IP Traffic Forecast, 2017–2022

Figure 1.1 IoT to drive growth in the number of connected devices through 2022: Cisco | ZDNet(Systems, n.d.)

IoT consists of smart things, including smart buildings, vehicles and devices (e.g. mobile phones and sensor devices). Sensors are deployed in WSNs for data acquisition, collection and analysis (Atzori et al, 2010). WSN nodes have resource constraints in battery, process capability and range transmission. To allow a limited constrained node in WSN to be compatible with IoT, the Internet Engineering Task Force (IETF) introduced the Internet Protocol version 6 (IPv6) over low-power personal area network (6LoWPAN) standard to enable a feasible communication in WSNs via IPv6 (Kim & Gomez, 2012).

The conventional routing protocol that operates on the IPv6/IPv4 network is unable to route packets in WSNs equipped with a tiny device. Therefore, IETF introduced a standard routing protocol for low power and lossy network (RPL) to enable communication via IPv6 in WSNs (Winter et al, 2012). IPv6 networks use the RPL protocol in IoT to overcome certain problems, including address space exhaustion, security issues, complex configuration and routing table enlargement (Clausen et al , 2011; Razali et al, 2017)). However, RPL has many variabilities that expose this network to different types of attacks. Resource-based attacks, such as distributed denial-of-service (DDoS) and version number attacks, are the most common attacks that threaten an RPL network. According to the Arbor report, around 86% of DOS/DDoS attacks target IoT-based networks as shown in Figure 1.2 (Arbor, 2016).

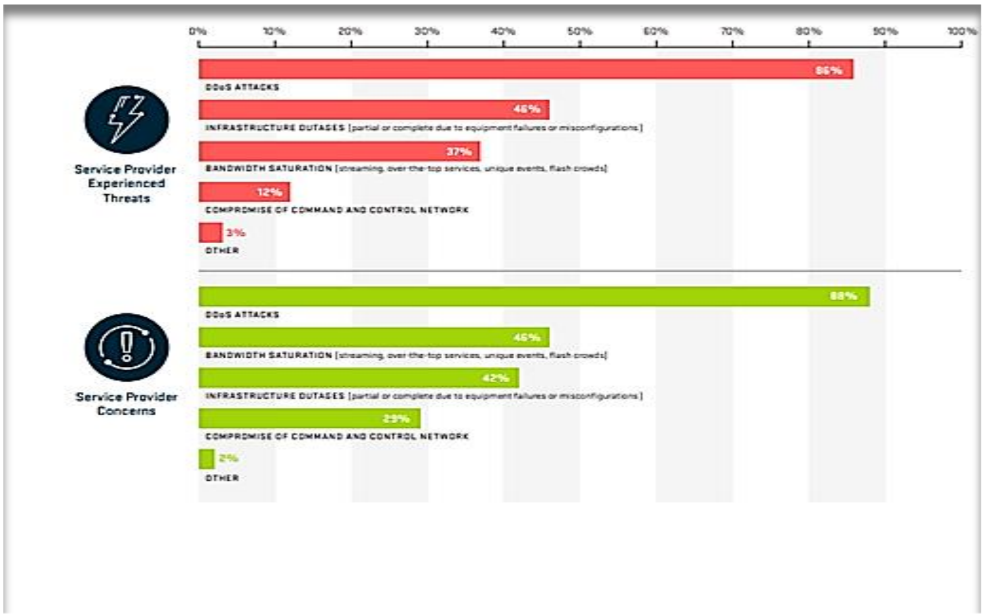


Figure 1.2 DDoS Attacks Reported in 2016 (Arbor, 2016)

1.2 Background

This section presents an introduction to WSNs, low-power and lossy networks (LLNs) and their general characteristics, RPL, the related challenges and resource-based attacks.

1.2.1 Wireless Sensor Network (WSN) and Low Power and Lossy Network (LLN)

LLN is a typical WSN in which all sensor nodes are operated in low power (Garcia et al, 2017). In other words, LLN is a WSN that is constructed with large resource-constrained devices (i.e. devices with limited power and memory). To build an LLN, nodes are interconnected by lossy links given their support for low data and packet delivery rates and their instability. Numerous constrained nodes are deployed in LLN to handle a small amount of data. Many applications in IoT, including smart homes, smart environments and forest monitoring, are realised through LLN. Table 1.1 compares LLN with an ad hoc network.

Table 1.1 Comparison between LLN and Wireless Ad Hoc Networks

Parameter		**LLN**	**Wireless Ad Hoc Network**
Number of Nodes		100 to 1000	10 to 100
Microprocessor/ microcontroller	Specification	16-bit microcontrollers	32-bit microprocessor
	Operational Frequency	In the order of MHz	In the order of GHz
Memory		In the order of KB	In the order of GB
Energy		1500 mAh to 2500 mAh	~3000 mAh
Maximum Radio Output		0 dBm	10 dBm to 16 dBm
Data Rate		Maximum of 250 kbps	In the order of Mbps

Table 1.1 compares the specifications of LLN with those of wireless ad hoc networks and reveals that compared with wireless ad hoc networks, LLN networks support more nodes and have lower power requirements. The characteristics of an LLN include large number of low-end devices, unreliable lossy links with low data rates, multi-hop communications, and small frame size and energy scarcity.

Given these characteristics, LLN offers several advantages, including scalability, flexibility, accuracy, simplicity and easy deployment. However, LLN is subjected to several challenges such as security as described in (Piste et al, 2013); Dohler, 2009)).

1.2.2 Routing Protocol for Low Power and Lossy Network (RPL)

Routing protocol over LLN (ROLL) that follows the working group in / IETF). LLN has a limited power, battery and memory capacity. RPL, which serves as the IPv6 routing protocol for LLN, creates a topology by using a destination-oriented directed acrylic graph (DODAG) (Ko et al., 2011). The objective function of an RPL is to select an optimal route following several significant constraints. RPL is a distance-vector routing network that effectively supports the traffic and maintenance of nodes. The occurrence of events in an RPL is detected in consideration of the quality of information being exchanged in messages (Tian et al, 2017). In RPL, the traffic flows are either multipoint-to-point (i.e. sensor-to-root) or point-to-multipoint (i.e. sensor-to-sensor) and these two traffic flows prefer upward or downward routes. Traffic is considered to be the most common occurrence due to this the arrival of traffic needs higher attention. RPL is also used in healthcare, such as within hospitals, or in home building environments (Gara et al, 2015).

RPL is based on a directed acyclic graph (DAG) with a tree-like structure to specify the default routes between nodes. RPL also builds a DODAG to enable an efficient route selection. In DODAG, the most popular destination node is most probably the sink node or those providing default route to the internet probably gateway is acting as the root node in the DAG graph. The topological concept in RPL is created based on DAG (Gaddour & Koub, n.d.). RPL usually has a DODAG with a root node that can generate a new DODAG, serve as the sink and act as the final destination. DODAG is constructed by exchanging ICMPv6 control messages.

1.2.3 Security Issues and Resource-Based Attacks on RPL

One of the main challenges in RPL is security issues given that RPL employs different applications, with each application involving the participation of a substantial number of legitimate and illegitimate users. Consequently, a secure communication amongst users in the network must be ensured. The key security challenges in RPL include bootstrapping, trust management, mobility monitoring, interoperability, resource provisioning, legacy systems, computation complexity, scalability and time maintenance (Bekara, 2014). As far as the sensitivity of data sharing is concerned, the security level needs to be improved. Attacks directed towards RPL are classified according to different aspects, the most significant of which include resource-based attacks (Mayzaud et al, 2016);(Sharma, et al, 2017).

Resource-based attacks aim to exhaust network resources. Specifically, resource-based attacks drive legitimate nodes to perform unnecessary processing that leads to additional resource consumption. These attacks also consume the energy and memory of legitimate nodes, which in turn leads to link unavailability in the network. The major attacks classified under this category are described as follows:

Flooding attack: Flooding attacks generate a large amount of unwanted traffic in the network to make the nodes and links unavailable. In RPL, flooding attacks are carried out by (i) broadcasting a DODAG information solicitation (DIS) message to neighbours to reset the trickle timer and (ii) by unicasting a DIS message to a node that needs to respond with a DODAG information object (DIO) message. Both of these transmissions lead to traffic congestion and saturation of RPL nodes. A flooding attack is a direct attack that is initiated by either an external or internal attacker.

***Version number attack**:* Version number is an important component of a DIO message in an RPL-based network. This number should be increased by the DODAG root only when the topology is rebuilt. However, when attacker nodes are present in the network, these nodes change the version number, thereby leading to an unnecessary rebuilding of the DODAG graph and consuming additional time and energy.

1.3 Problem Statement

The RPL protocol is applied over WSNs. Therefore, RPL must achieve efficient routing, balance the nodes and guarantee security. The routing process in RPL is based on the object function (OF), which includes a set of metrics/constraints to select the optimal parent set (optimal root set) and to balance the nodes in the RPL network. An inappropriate selection of the parent set can negatively affect networking balance, maximise energy consumption and minimise the RPL network lifetime. Moreover, the OF can be misused by an attacker to perform different types of attacks with a deliberate intent to minimise the RPL network lifetime.

The abovementioned problems have motivated researchers to propose approaches for improving network balance, maximising network lifetime and protecting the network from attacks. Network lifetime has an inverse relationship with the presence of attacks in an RPL network, which minimises the network lifetime.

The existing approaches for maximising network lifetime are categorised into (i) OF-based approaches (Iova, et al.) and (ii) clustering approaches based on RPL operation, such HECRPL (Zhao et al, 2017), EECPK-means (Ray & De, 2016a) and IRPL (Zhang et al, 2018). Clustering is one of the important methods for prolonging the network lifetime in wireless sensor networks (WSNs). It involves grouping of sensor nodes into clusters and electing cluster heads (CHs) for all the clusters. The CH performs aggregation of the packets received from all the nodes present in their cluster. Also, all the nodes get a chance to become the CH to balance the overall energy consumption across the network. Although these clustering approaches are considered most efficient in maximising network lifetime, they face three main limitations. Firstly, they require an additional control message that increases the control message overhead, such in the case of HECRPL. Secondly, CH selection is based on inefficient metrics, such as residual energy and distance in the case of EECPK-means. Thirdly, these approaches are inappropriate for parent node selection, such in the case of IRPL.

Meanwhile, the existing approaches for protecting the network against several types of attacks, such as those proposed by (Nancy & Chrisment, 2016), Mehare and Bhosale (2017), Chen et al, (2016), Dvir (2011) and Ahmed and Ko (2016), cannot accurately detect the presence of attacks for three reasons. Firstly, these approaches use simple heuristics to detect the presence of DDoS attacks by counting the number of messages within a time window. Secondly, these approaches are unable to detect different attack scenarios, such as two DDoS attackers operating beside each other to produce unreliable information. Thirdly, these approaches identify only a single attack in each round, and restarting the process leads to the misdetection of some attackers.

Therefore, secure RPL approach that is aware of attacks must be proposed to maximise the network lifetime and protect the RPL network from attacks.

1.4 Research Motivation

WSNs and LLNs are key networks that realise IoT in many real-time applications (Iova et al., 2017). Routing in these low-power resource-constrained networks is challenging due to their energy constraints, multi-hop topologies, frequent topology changes and mobility of nodes. These issues are addressed by designing the RPL protocol, which considers a backbone protocol for LLN. RPL is a special protocol that enables routing amongst energy- and resource-constrained devices, supports a variety of link layers (e.g. constrained and potentially lossy link layers) and can be utilised in conjunction with host or router devices. In RPL, the nodes are connected to root nodes through multi-hop paths. RPL utilises link costs, node attributes, node status information and an OF for route selection and distributes route knowledge amongst the neighbour nodes in the network. RPL also provides IoT-based networks with the following advantages: (i) communication between devices and machines, (ii) reduced costs and complexity, (iii) environmental monitoring, (iv) relatively fast and timely output and (vi) automation of regular tasks.

Despite its advantages, RPL faces several challenges, the most significant of which is related to security. RPL-based IoT networks are vulnerable to many attacks, including DDoS, rank, version number and spoofing attacks. The presence of these attacks in RPL-based networks can downgrade the performance of the RPL network and minimise its lifetime.

1.5 Research Objectives

The main goal of this thesis is to propose a secure RPL approach based on an efficient clustering mechanism to maximise RPL network lifetime and prevent resource-based attacks. The following objectives are set to achieve this goal:

- To propose a mechanism that utilises the clustering process to maximise network lifetime and minimise the risk of resource-based attacks,
- To propose a mechanism based on the ant colony algorithm (ACO) to prevent DDoS attacks in the RPL network and

- To propose a mechanism to prevent the data transmission from version number attacks.

1.6 Scope and Limitations

The proposed Secure-RPL approach considers a WSN that consists of randomly deployed nodes. The sensing and transmission processes in this approach are the same as those in an RPL network. The other scopes of this research are listed as follows:

- The router is fixed and located inside the network,
- The router has no energy limitations,
- All sensors in the RPL network are homogeneous and energy constrained,
- All nodes have the same sensing nodes,
- The nodes are deployed randomly in the area,
- An application that requires the building of multiple DODAGs, such as wildlife monitoring, is considered and
- The proposed approach focuses on preventing two types of source-based attacks, namely, (i) DIO flooding DDoS and (ii) version number attacks.

1.7 Research Contributions

The main contribution of this research is the proposed secure RPL approach that maximizes the RPL network lifetime based on an energy-efficient clustering algorithm and prevents DDoS and version number attacks. This approach is called Secure-RPL. The contributions of this research are summarized as follows:

- A mechanism that utilizes the clustering process to maximize the lifetime and minimize the risk of resource-based attacks,
- A mechanism that detects DDoS attacks by utilizing the ant colony pheromone, and
- A mechanism that prevents data transmission from version number attacks.

1.8 Research Steps

The Secure-RPL approach is developed after several stages of theoretical and experimental analyses to maximise network lifetime and minimise the risk of resource-based attacks, DDoS attack detection and secure route identification. To achieve the research objectives, this research is divided into the following stages: (1) reviewing the related literature, (2) formulating the research problem and conducting the (3) proposed work, (4) experimental work and (5) evaluation work. Figure 1.3 illustrates these research stages.

In the first stage, several studies on RPL performance and security are reviewed to understand the RPL protocol. This protocol is selected for this research because RPL is becoming the routing protocol for IoT and is used in many application scenarios. The drawbacks and gaps of previous studies are then reviewed, and the output is to identify the problem statement and main objective of this work.

In the second stage, the challenges faced by RPL when operating in harsh environments, changing the nodes, recharging batteries and guarding itself against DDoS and version number attacks are discussed. This stage also examines the proposed approach for improving the energy consumption and maximising the lifetime of RPL.

In the third stage, a coordinative-based clustering (CBC) mechanism is developed to optimise network lifetime. This mechanism imitates the cluster head selection process, where the cluster head root node is selected by applying distance based on the Radio Signal Strength Indicator (RSSI) equation algorithm. This stage also develops an enriched ACO (E-ACO) mechanism to detect the presence of DDoS attackers in five sequential steps, namely, initialisation, evaluation, fitness calculation, probability detection and DDoS attack identification. This stage also introduces a secure route mechanism for selecting the best parent set to avoid version number attacks.

The fourth stage performs an RPL protocol secure environment simulation by using the NS3 simulator, a discrete-event network simulator that strongly supports the RPL routing module. The simulation involves the flow-mentor and WSN modules, which play important roles in demonstrating the performance of the proposed security mechanisms.

The fifth stage evaluates and compares the results of this work with those of previous studies to underscore the improvements realised by the proposed mechanism.

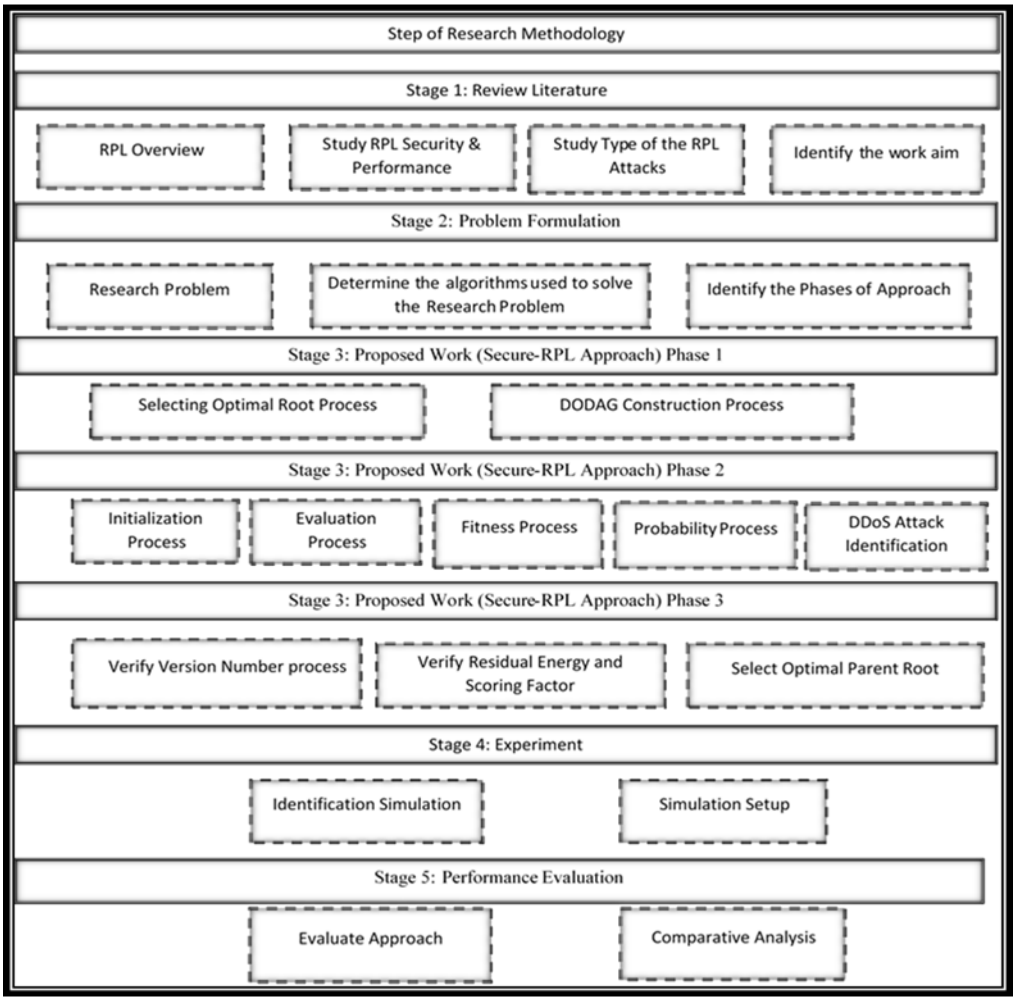

Step of Research Methodology
Stage 1: Review Literature
RPL Overview
Study RPL Security & Performance
Study Type of the RPL Attacks
Identify the work aim
Stage 2: Problem Formulation
Research Problem
Determine the algorithms used to solve the Research Problem
Identify the Phases of Approach
Stage 3: Proposed Work (Secure-RPL Approach) Phase 1
Selecting Optimal Root Process
DODAG Construction Process
Stage 3: Proposed Work (Secure-RPL Approach) Phase 2
Initialization Process
Evaluation Process
Fitness Process
Probability Process
DDoS Attack Identification
Stage 3: Proposed Work (Secure-RPL Approach) Phase 3
Verify Version Number process
Verify Residual Energy and Scoring Factor
Select Optimal Parent Root
Stage 4: Experiment
Identification Simulation
Simulation Setup
Stage 5: Performance Evaluation
Evaluate Approach
Comparative Analysis

Figure 1.3 Research Stages

1.9 Thesis Organization

The rest of this thesis is organised as follows:

Chapter 2 reviews the related literature that addresses the challenges in RPL and WSNs and the adopted countermeasures.

Chapter 3 discusses the proposed approach to optimise network lifetime and dealing with DDoS and version number attacks in WSNs and describes the procedure for analysing such methodology.

Chapter 4 presents design and implementation of the proposed approach and in addition to the proposed attacks detection mechanisms, with the verification of models of the proposed mechanisms.

Chapter 5 discusses the evaluation result of the proposed approach by using a simulation tool and compares the simulation results with those obtained by the existing protocol.

Chapter 6 concludes the research work and presents some suggestions for future work.

CHAPTER 2

LITERATURE REVIEW

This chapter presents a background of the RPL protocol in WSNs, reviews the literature on enhancing this protocol and highlights the limitations of these studies that serve as the motivation of this work. The reviewed literature also includes those studies that have attempted to deal with security issues in RPL-based IoT and quality of service (QoS) issues in RPL-based networks.

This chapter is organised as follows. Section 2.1 presents an overview of RPL and related terminologies, control messages, DDOAG construction and challenges. Section 2.2 discusses the existing approaches for maximising RPL network lifetime Section 2.3 discusses the related works that have attempted to detect the presence of attacks in an RPL network. Section 2.4 highlights the research gaps. Section 2.5 summarises the chapter.

2.1 Background

RPL has been extensively used in IoT for routing over WSNs. However, RPL faces several security issues. Specifically, source-based attacks, such as DDoS and version number attacks, can disrupt the data transmission amongst nodes and downgrade the performance of WSNs. This section presents an overview of RPL and its terminologies, control message, DODAG construction and security challenges.

2.1.1 Overview of RPL

RPL is a distance-vector routing protocol designed and developed by the ROLL working group (Vasseur et al., 2011) that operates on top of several link-layer mechanisms. RPL supports three types of communication, namely, point-to-point,

point-to-multipoint and multipoint-to-point communications. The basic concept of RPL is based on a DAG with a tree-like structure that specifies the default routes amongst nodes. RPL also builds a DODAG to enable an efficient route selection. The most popular destination node in DODAG is most probably the sink node or those providing default route to the internet probably gateway is acting as the root node in the DAG graph. The major features of RPL are listed in Table 2.1 (Gaddour & Koubâa, 2012).

Table 2.1 Major Features of RPL

Features of RPL	**Description**
Auto-configuration	The involvement of neighbour discovery mechanisms in RPL realises the auto-configuration of new paths and destinations. The dynamic discovery of new routes and destinations can improve network performance through auto-configuration.
Self-healing	With the involvement of RPL, the logical network topology changes and node failures can be adapted. Given that the links and nodes in an LLN are dynamic and vary frequently, self-healing is important for the network. The risk of failure is addressed by selecting more than one parent node per node in DAG.
Loop avoidance and detection	The RPL protocol can include a reactive mechanism for loop detection in case of a topology change. A node in DAG must be ranked higher than its parents given that

Features of RPL	Description
	DAG is acyclic in nature. RPL also utilises global and local recovery mechanisms to resolve the loop occurring in the network.
Independence and transparency	The main advantage of RPL is its independence from data-link layer technologies. RPL is designed to operate over multiple link layers as in the IP architecture where resource-constrained nodes are present.
Multiple edge routers	In an RPL-based LLN, multiple DAGs can be constructed in which each DAG has a root. Therefore, if a node belongs to more than one DAG, then this node plays different roles in each DAG. This property of RPL enables high availability and load balancing in the network.

2.1.2 RPL Terminologies

As shown in Figure 2.1, the DAG graph comprises a source node, destination node, RPL nodes and RPL router. The basic terminologies used in RPL are defined as follows (Nathan & Scobell, 2012; Pavkovic et al, 2014; Umamaheswari & Negi, 2017):

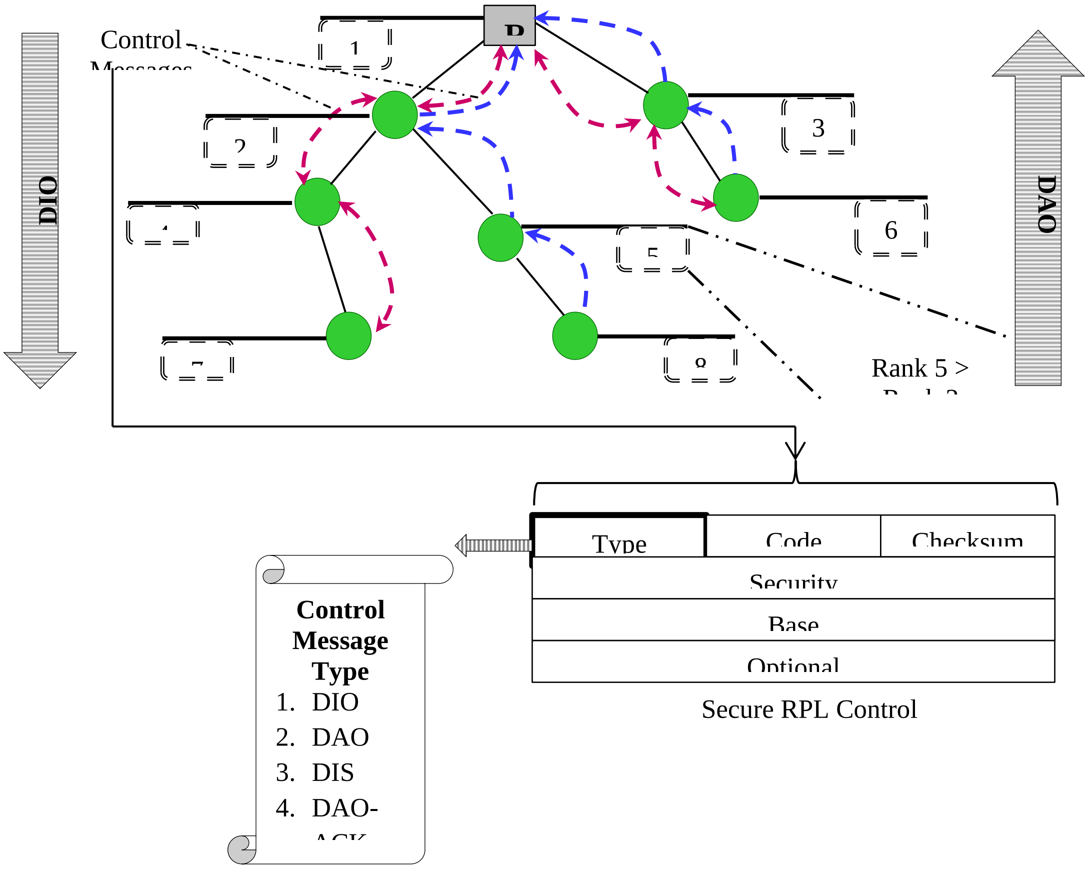


Figure 2.1 Basic Terminologies Used in RPL

DAG: All edges in a DAG are oriented such that no cycles exist in the graph. In each path, all edges are oriented towards and are terminated at one or more root nodes.

DAG Root: A DAG root is a node present in DAG without any outgoing edge. In a DAG, each path is terminated with a root given that the graph is acyclic in nature.

DODAG: A DODAG is a DAG where all edges are rooted with a single destination (i.e. at a single DAG root) without an outgoing edge.

DODAG Root: The DODAG root acts as a border router for DODAG. This root aggregates the routes in the DODAG and redistributes them to other routing protocols.

Virtual DODAG Root: A virtual DODAG root comprises a combination of two or more RPL routers. The border routers in the network can be coordinated to synchronise the DODAG state, and the coordinated root is often called a virtual DODAG root.

Up: In a DODAG tree, up refers to the data transmission from the leaf nodes to the DODAG roots.

Down: In a DODAG tree, down refers to the data transmission from the DODAG root to the leaf nodes.

Rank: Rank defines the position of a node relative to other nodes in the DODAG root. In general, the rank of a node is computed based on object function (OF). The rank increases in the down direction and decreases in the up direction.

OF: An OF is a criterion for parent selection in which routing metrics, optimisation objectives and related functions are involved. The rank for each node is computed based on OF and the routing metrics.

RPL Instance ID: Each RPL instance within a network is distinguished by an RPL instance ID. DODAGs with different RPL instance IDs indicate that each instance has a different OF. However, DODAGs with the same RPL instance ID shares the same OF.

RPL Instance: An RPL instance is a set of DODAGs that share the same RPL instance ID. The RPL node is allowed to participate in only one DODAG in an RPL instance, and each RPL instance is independent of the other RPL instances.

Figure 2.2 presents an RPL instance with three DODAGs. In this figure, R1, R2 and R3 denote the root or DODAG IDs of different DODAGs.

DODAG ID: As mentioned above, an RPL instance comprises multiple DODAGs, in which each DODAG is differentiated through a DODAG ID, which serves as an identifier of a DODAG root. In a network, a DODAG is identified by both RPL instance ID and DODAG ID.

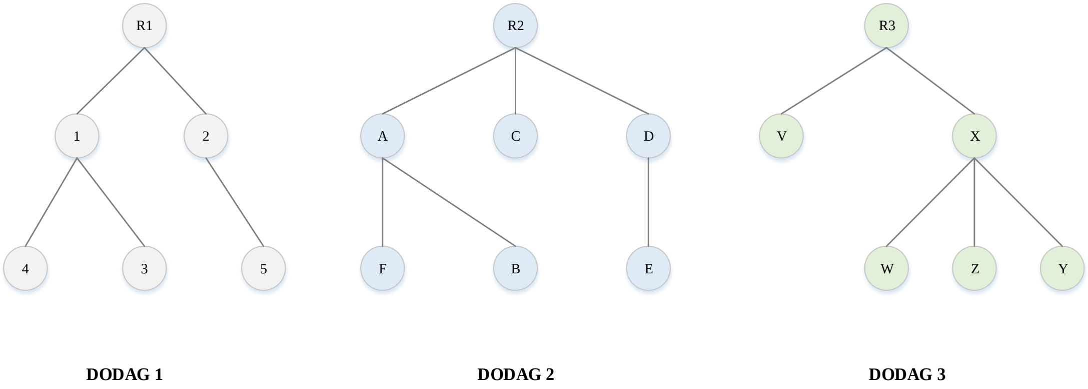


Figure 2.2 RPL Instance

DODAG version: The DODAG version specifies the iteration, that is, the version of a DODAG with a given DODAG ID in the network.

DODAG version number: A DODAG version number is a sequential counter that is incremented by the root to form a new DODAG version. A DODAG version is identified by the RPL instance ID, DODAG ID and DODAG version number. The version number changes along with the topology. Whenever the topology of DODAG changes, the version number of this DODAG increases by 1.

Figure 2.3 illustrates the version number change in DODAG. In this figure, the version number of a DODAG is denoted by ‘β’. In DODAG, the RPL node ‘4’ changes its parent node from ‘5’ to ‘2’, which then leads to the overall topology change of the

DODAG. Therefore, the version number of DODAG increases by 1 and becomes '$\beta + 1$'.

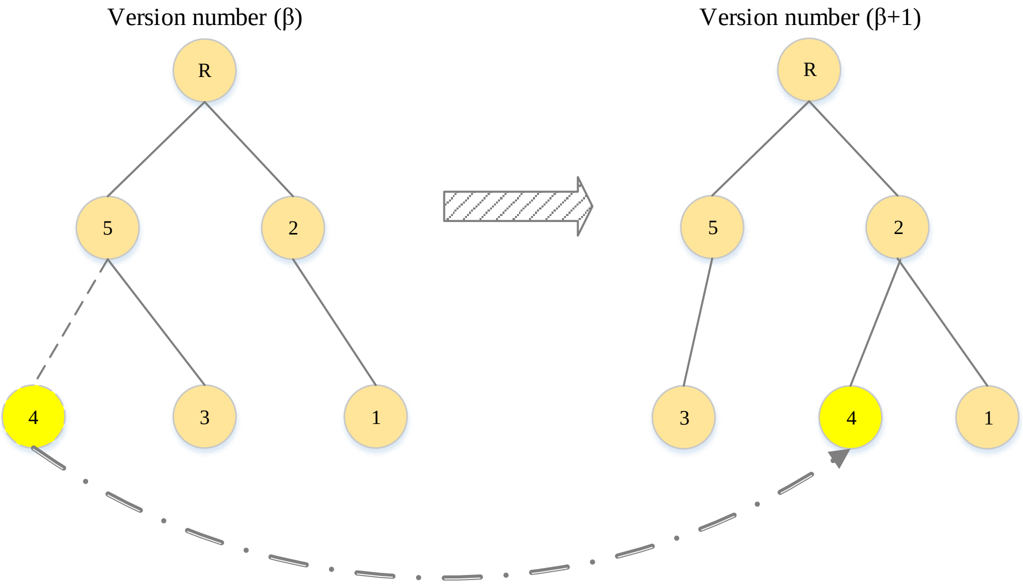


Figure 2.3 DODAG Version Number

DODAG parent: In a DODAG, the parent node is an intermediate successor of a node on the path towards the DODAG root. A parent node has a lower rank than the other nodes (i.e. satisfies OF).

Sub-DODAG: The sub-DODAG of a node is a set of other nodes whose routes to the root node pass through that node. In other words, the sub-DODAG of a node is a set of other nodes that are ranked below that node.

Local DODAG: In a local DODAG, a single root node is present, and the single root can allocate and manage the RPL instance that is identified by a local RPL instance ID without coordinating with the other nodes.

Global DODAG: A global DODAG uses a global DODAG ID that is coordinated amongst several nodes in the network.

2.1.3 RPL Control Messages

The Internet control message protocol for IPv6 (ICMPv6) introduces a new type of control message in RPL (Conta et al, 2006). The control message of an RPL has two main fields, namely, the header and the message body. The RPL control message format is depicted in Figure 2.4, where the header comprises three main fields, namely, the type, code and checksum, the message body is divided into base and options and the code in the header consists of RPL type, security and reserved fields.

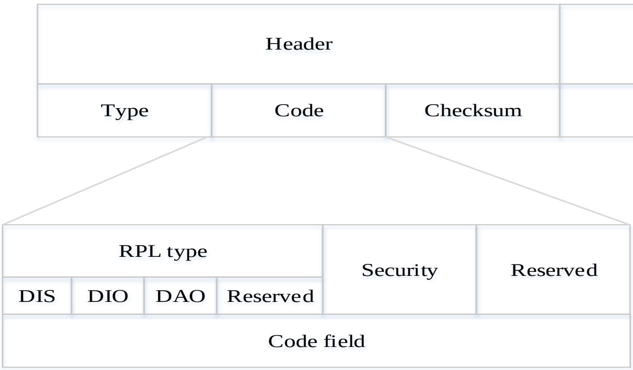


Figure 2.4 RPL Control Message

RPL uses four types of control messages as described below:

DODAG Information Solicitation (DIS) Message: A DIS message is initiated to request for a DIO message from the root node in a DODAG. In the RPL type field, 0*00 indicates the DIS message. This message is also used in the neighbour discovery process to probing its neighbour nodes to nearby DODAG.

DODAG Information Object (DIO) Message: The DIO message in a DODAG is multicast and initiated by the root node. This message, which is represented by 0*01 in the type field, is initiated to construct a new DAG. A DIO message includes the network information for discovering an RPL instance, learning its configuration parameters and selecting a parent set.

Destination Advertisement Object (DAO) Message: A DAO message is initiated by each node involved in propagating reverse route information to record the visited nodes along the upward path. The unicast message is initiated by a child node to its parent node or root node based on the operating mode.

DAO Acknowledgement (DAO-ACK) Message: The unicast message is initiated by the DAO recipient (i.e. DODAG root node or parent node) to the sender node that initiates the DAO message. This message is sent in response to the DAO message.

Table 2.2 presents an overview of RPL control messages and their unique purposes. These control messages are further involved in DODAG construction and RPL routing.

Table 2.2 Overview of RPL Control Messages

Control Message	**Purpose**
DIO	• Multicasts an RPL instance downward. • Allows other nodes to discover an RPL instance to join.
DIS	• Enables neighbour discovery. • Enables link-to-local multicast.
DAO	• Enables unicast from the child to the parent node. • Requests to join DODAG.
DAO-ACK	• Replies to a DAO message.

2.1.4 DODAG Construction

As mentioned earlier, an RPL instance comprises one or more DODAGs depending on the application requirements (Dhumane et al., 2015);(Kim et al, 2017);(Zhao et al, 2017)$^{)}$. For instance, in the application automation scenario, an RPL instance is constructed with a single DODAG and DODAG root. However, in the case of urban data collection, an RPL is constructed with multiple DODAGS to improve connectivity. In any RPL-based application, DODAG plays a vital role, and DODAG construction serves as an initial process. The following steps are involved in DODAG construction:

- The DODAG construction is initiated by the root node by disseminating a DIO message. This message is received by all nodes present within the communication range of the root node.
- The DIS message is initiated by the nodes to select optimal parent nodes. This message contains new information about the DODAG structure.
- In the upward direction, the DODAG is constructed with the help of a multicast DAO message. All nodes send the DAO message to the root node (i.e. multipoint-to-point transmission). A node belonging to a DODAG has to send a DAO message to its parent nodes within the same DODAG.
- The parent nodes send a DAO-ACK message to the nodes from which the DAO message is sent.
- Each node in the DODAG is ranked according to its position in the DODAG and OF.

Figure 2.5 illustrates the entire DODAG construction process. This process starts by disseminating the DIO message in the network by root node R. Afterwards, the DIS message is initiated by all RPL nodes to discover the neighbour nodes in the network. The child nodes then send the DAO message to the parent node as a request to join the DODAG. The DODAG construction is completed by sending a DAO-ACK message to the child nodes from the parent nodes.

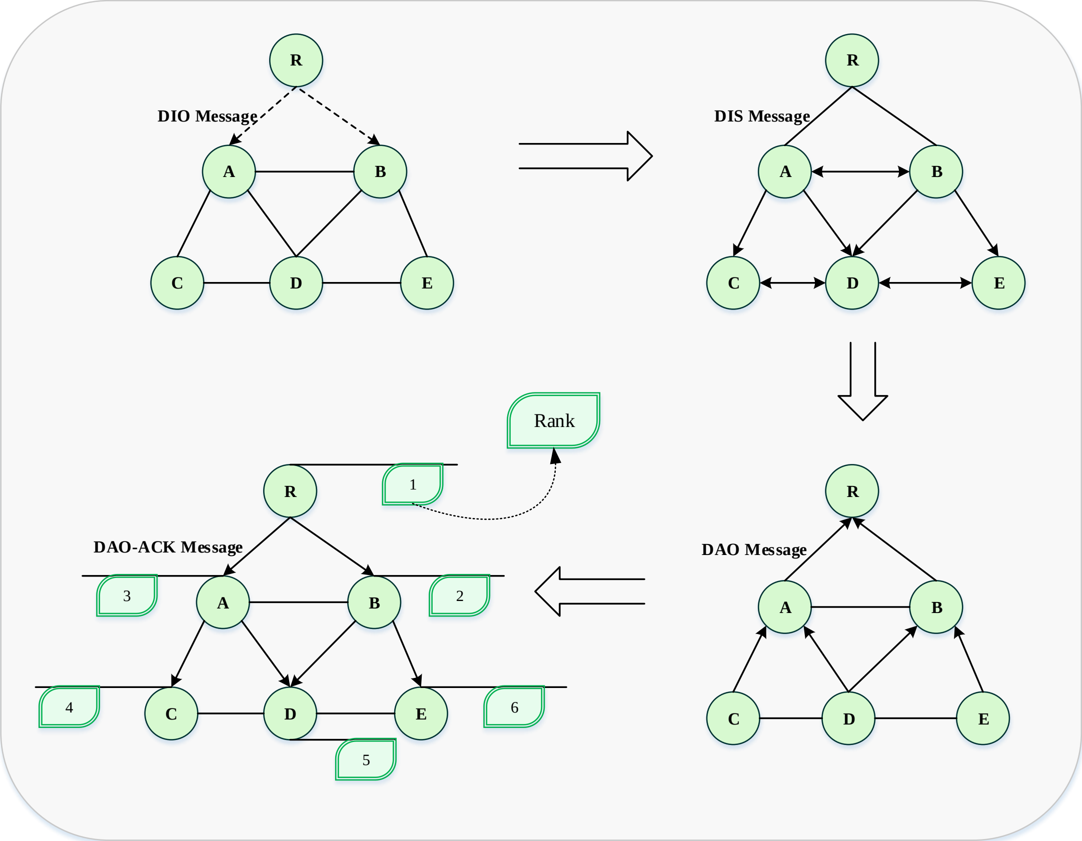


Figure 2.5 DODAG Construction

In the constructed DODAG, RPL selects an optimal parent node for data transmission based on OF, which is formulated by routing metrics. The major routing metrics involved in OF are listed in Table 2.3.

Table 2.3 Major Routing Metrics Involved in OF

Routing Metrics	Purpose
Node state and attribute object	Provides information about the node characteristics
Node energy	Prevents the selection of a node with a low residual energy
Hop count	Reports the number of nodes visited along the path
Link throughput	Reports the range of throughput that the links can handle
Link latency	Serves as a constraint or path metric
Link reliability	Can be degraded for several reasons, including signal attenuation and interferences of various forms
Link colour	Avoids or attracts specific links for certain traffic types
Expected retransmissions	Represents the number of transmissions that a node expects to make to a destination to successfully deliver a packet
Received signal strength	Measures the power level received by the receiver node from a source node

Based on the routing metrics, the OF is constructed and the data transmission is performed. Despite its many advantages, RPL faces some challenges, such as in DODAG construction, OF formulation and security (Lamaazi et al, 2018; Zhao, et al, 2017). Amongst these issues, security poses a major problem that affects the entire network.

2.1.5 Challenges in RPL

RPL faces several challenges that can be categorised into Security Related and Non-Security-Related challenges.

2.1.5(a) Security Related Challenges in RPL

Security poses a major concern in RPL given that this network is involved in different applications with a huge number of legitimate and illegitimate users (Pongle et al, 2015; Mayzaud et al, 2016; Grgić et al, 2016; Kamble et al , 2017; Medjek et al). The challenges identified in RPL include trust management, bootstrapping, interoperability, mobility monitoring, legacy systems, resource provisioning, scalability, computational complexity and timely maintenance. Before sharing sensitive data, the security level in the network must be improved. However, RPL is vulnerable to many security threats. Figure 2.6 classifies the attacks in RPL into resource-, traffic and network-topology-based attacks.

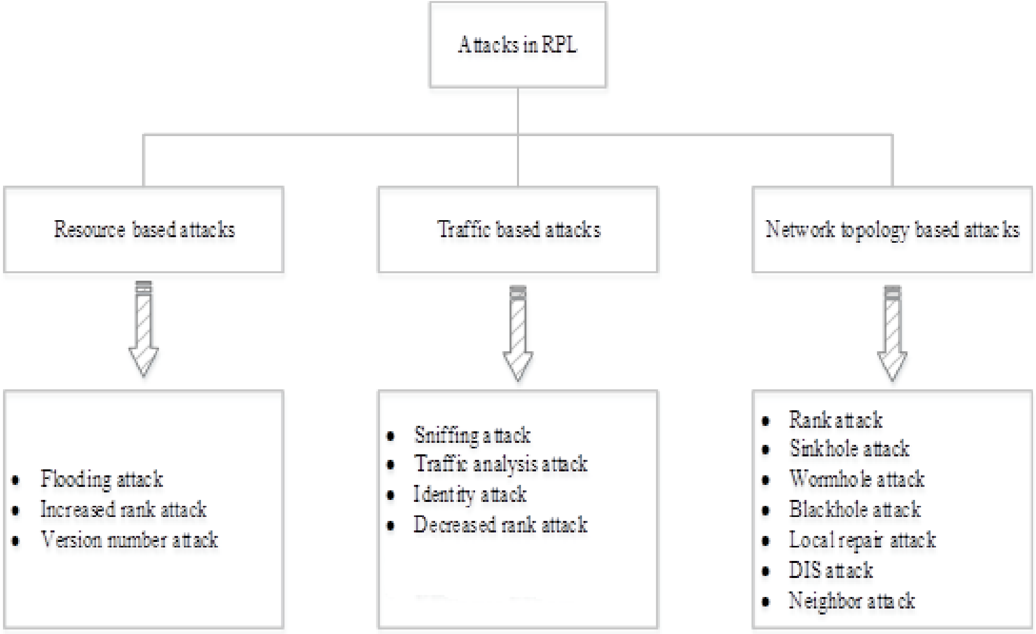


Figure 2.6 Classification of Attacks in RPL (Mayzaud et al, 2016; Alabsi et al, 2018)

A. Resource-Based Attacks

Resource-based attacks aim to exhaust network resources. In these attacks, the legitimate nodes have to perform unnecessary processing that leads to additional resource consumption. These attacks also aim to consume the energy and memory of legitimate nodes and promote congestion, which in turn leads to link unavailability in the network. The major attacks falling under this category include flooding, version number, and increase rank attacks (Rghioui et al, 2014). This thesis focuses on resource-based attacks (DIO flooding attack and version number attacks) as these attacks has a destructive impact on network performance which lead to reduce Packet Delivery ratio (PDR), Packet Loss Rate (PLR) and energy consumption.

B. Traffic-Based Attacks

Traffic-based attacks attempt to modify or overhead the network traffic. These attacks are described as follows:

Sniffing attacks: Sniffing attacks overheard the network traffic or eavesdrop in the network. These attacks are launched through a compromised device that captures packets from a shared transmission medium. Various types of information, such as partial topology, routing information and data content, can be obtained from the sniffed packets. In the RPL network, if the attacker sniffs the control message, then s/he can access information regarding DODAG in the network.

Identity attack: Identity attacks include spoofing and Sybil attacks, which attempt to modify the identity of the RPL node. In a clone ID attack, the attacker pretends to be a legitimate existing node. This attack produces a significant impact if the attacker spoofs the address of the root node in the DODAG graph. Root node plays a vital role in the DODAG graph by building and maintaining the DODAG topology. The root node can be identified from the control messages sent over the network. This type of identity attack is called a spoofing attack. In the case of a Sybil attack, a single malicious node uses several identities in the same physical node.

Decreased rank attack: In RPL, having a lower rank indicates that an RPL node is located closer to the root node and that all other legitimate RPL nodes connect to the root through this node. If a compromised or attacker node illegitimately advertises a lower rank value, then all other legitimate nodes transmit their packets through this attacker node, thereby leading to a huge PLR and network performance degradation.

Traffic analysis attacks: Traffic analysis attacks obtain routing information by analysing network traffic and utilise the traffic patterns and characteristics of the link to obtain routing information. Therefore, these attacks can also be launched against encrypted packets and be combined with rank attacks, which can greatly affect network performance.

C. Topology-Based Attacks

Topology-based attacks target the network topology and its information, including routing and topology information. Topology-based attacks can be classified as follows:

Rank attacks**:** In RPL, rank attacks increase the rank value from the root node to the child node to affect network performance (Le et al., 2013). By modifying the rank value, an attacker can attract a huge amount of network traffic. This type of attack produces the following consequences:

- Generation of a non-optimal path,
- Undetected formation of unnecessary loops,
- Presenting an unusable optimal path in the topology,
- Decreasing PDR and increasing PLR and
- Subjects the network to a topology change that increases the control overhead.

Sinkhole attack: Sinkhole attacks build a sinkhole to increase the PLR in the network. These attacks are launched in two phases. Firstly, a malicious node attracts a vast amount of packets by advertising falsified information, such as information regarding the link quality and rank value. Secondly, the malicious nodes are dropped, and the received packets from other legitimate nodes in the network are modified. Sinkhole attacks significantly affect the network topology by modifying its information, thereby severely degrading network performance.

Wormhole attacks: Wormhole attacks distort the routing path with the support of two RPL attacker nodes (Patel, 2016). Each packet received by an attacker is forwarded to another attacker, which makes the replay later. The attacker can also transmit the routing information from one part of the network to another, thereby changing the routing paths. This attack leads to the unavailability of optimal routing paths depending on the OF.

Blackhole attack: In blackhole attacks, the malicious node drops the packets that are supposed to be forwarded. These attacks are either combined with sinkhole or DoS attacks. When combined with sinkhole attacks, blackhole attacks greatly damage the network by dropping a large amount of packets.

Local repair attack: In local repair attacks, the attacker sends local repair messages continuously even when there are no problems in link quality. After receiving the local repair message, all surrounding nodes attempt to perform a local repair, thereby leading to unwanted processes. These attacks greatly affect the delivery ratio and delays by increasing the control message overhead.

DIS attack: In RPL, the DIS message is initiated by a new node to obtain DODAG topology information before joining the network. In DIS attacks, an attacker node periodically sends a large amount of DIS messages to its neighbour nodes. Upon receiving this message, the neighbour node rests the DIO timer assuming that there is a problem with the topology. Therefore, DIS attacks promote network congestion and energy consumption given that broadcasting DIS messages involves a large number of nodes.

Neighbour attack: In neighbour attacks, a malicious node broadcasts a DIO message without including information that seems to be requested from a new node. Upon receiving the DIS message, the other nodes assume that the new node is joining the network. Therefore, the legitimate node attempts to select the node that is not a neighbour node as a parent node. The major aim of these attacks is to affect the QoS in the network.

An efficient security scheme must be designed to protect RPL-based IoT or LLN networks from various types of attacks.

2.1.5(b) Non-Security -Related Challenges

A. Selected Objective Function

RPL routing is implemented with a certain OF that is selected for routing a packet between nodes (Liu et al, 2010). The most commonly used OFs include TX, hop count, stability, signal-to-noise ratio, energy, distance, and connectivity. To obtain better routing results, the best OF is selected, but the selection presents a challenge in RPL routing.

B. Battery-Assisted Node

The other significant challenges in RPL routing have been discussed in (H. S. Kim et al. ,2017). Given the participation of the battery-assisted node in the network, the load must be balanced, especially when the level of the traffic is very high, which occurs when thousands of nodes are involved in data transmission. This work discusses load balancing under a heavy traffic scenario.

C. Multicast Routing

Multicast routing is a critical process in RPL where data are disseminated and broadcasted in various networks. The bi-directional multicast RPL forwarding , trickle multicast (TM) and stateless multicast RPL forwarding (SMRF) algorithms are traditionally used in multicast routing (Oikonomou et al, 2012; Oikonomou, Phillips, & Tryfonas, 2013b; Gastón et al, 2016). These algorithms suppress the re-broadcasted packets, and re-broadcasting produces overhead in the network. Therefore, broadcasting data by using conventional algorithms presents a challenge in RPL.

D. Mobility

The dynamic node in the RPL directly influences the networking balance to maximise the energy consumption and minimise network lifetime.

2.2 Related Works on Maximizing the Lifetime of RPL-Based Networks

The presence of attacks is correlated with the lifetime of RPL-based networks. Therefore, minimising the lifetime of these networks may indicate the presence of attacks. An RPL network has an energy-consuming design. Therefore, the presence of attacks, such as resource-based attacks, can rapidly drain the energy of nodes and minimise the lifetime of RPL-based networks. Accordingly, many researchers have proposed approaches for minimising the energy consumption of these networks, maximising their lifetime, reliable routing and performance and Load balancing, and maintaining their sustainability in the presence of attacks.

Zhao et al ,2016 proposed an energy-efficient region-based RPL (ER-RPL) to solve the key issue of energy consumption. ER-RPL has two stages, namely, the network initialisation stage and route discovery stage. In the former, the relative distance and hop counts were estimated by the reference nodes. Afterwards, the distributed self-regional strategy was used to segment nodes into different region numbers. The routes were selected according to their reliability and energy conservation to minimise routing overhead. However, ER-RPL does not perform well in dynamic environments where many real-time applications are realised.

Zhang et al ,(2017) proposed energy-efficient heterogeneous ring clustering (E2HRC) routing to address the energy problem in sensor networks. Ring domain communication was enabled by determining the domain grade in terms of RSSI. A cluster construction was then performed based on the cluster probability threshold, and a cluster head rotation mechanism was designed to balance the energy consumption in the network. Route selection was performed afterwards in consideration of the optimal direction angle, node residual energy and hop difference. However, E2HRC increases the amount of time consumed in detecting the ring and location of a node given that the nodes present in the rings are clustered.

Alamelumangai and Nachiappan (2015) proposed a hybrid routing protocol and load balancing technique to improve the performance metrics, including PDR, residual energy, delay and packet drop. Both proactive- and reactive-based routing were involved in this approach. If the source node was a DAG member, then this node would use a proactive approach. Otherwise, this node would use a reactive approach. The data mule with the shortest ID was selected as the leader node, which major responsibility was to divide the nodes into sharable and non-sharable nodes. The load could be balanced by estimating the load on a mule. However, this approach cannot efficiently achieve low balancing.

Zhao et al, (2017) proposed a Hybrid Energy efficient Cluster parent based RPL (HECRPL). Optimal selection of Cluster Parent Set (CPS) was a top-down approach for reducing the energy depletion. DODAG is constructed and CPS is selected based on residual energy, cost and node's priority. DODAG requires repeated updating if the nodes are dynamic, this makes the system complex to handle.

Yang and Ping (2016) proposed cognitive-receiver-based RPL (CRB-RPL), a receiver-based routing protocol that improves the delay and energy efficiency in radio-enabled smart grids. This protocol was designed to support routing in real-time smart grid applications with low latency and routing in green smart grids with low energy consumption. In CRB-RPL, the packet from the sender node was received by all neighbour nodes instead of a single receiver node to improve link success probability. The transmission quality in this approach was defined by cognitive transmission quality (CTQ), which was used to describe the trade-off between transmission quality and interference. Moreover, the energy efficiency was quantified by hop energy efficiency (HEE). However, transmitting packets to all neighbours instead of a single receiver would introduce congestions in the network and increase the energy consumption for all neighbour nodes.

Kamgueu et al ,(2013) achieved an energy-aware route selection by considering the residual energy of the node in an RPL-based network. The path cost between the source and sink was computed as

$$Cost_i = \min\,[\max(Cost_j, E_i)] \tag{2.1}$$

The path cost of node i ($Cost_i$) was computed by using the path cost of node j ($Cost_j$) and the remaining energy of node i (E_i). After computing the path cost, the path with the minimum cost was selected as the optimal path for transmission. Nevertheless, considering energy alone in the path selection would increase the number of retransmission and energy consumption.

Li et al , ,(2015) improved the network lifetime of the RPL routing protocol by using an energy balancing scheme, where each node has three objects, namely, the INSTANCE object that contains the OF, the PARENT object that contains information on the parent node and the DAG object. The parent node selection involved routing metrics, including rank, link quality and energy consumption. The routing metric was computed as

$$Metric(E, Cost) = erg \times W_e + (rank + link) \times W_{cost} \tag{2.2}$$

Where E denotes energy, W_e denotes the weight of energy and W_{cost} denotes the weight of path cost. Based on these metrics, the quality of the parent node was determined, and route selection was performed. However, this scheme only considers the parent information for route selection and is therefore inefficient. Table 2.4 summarises the RPL energy consumption improvement approaches proposed in the literature.

Sankar et al ,2018 proposed a multi-layer cluster-based energy aware routing protocol for RPL (MCEA-RPL) to enhance network lifetime based on dividing area into rings. MCEA-RPL has three process, namely, the ring creation process, intra ring clustering process and interclassing routing process. In the former, the intra-ring clustering process performs two operations, namely cluster formation and CH selection. The cluster formation is based on the energy consumption of nodes in each ring Afterwards, the inter-cluster routing applies the fuzzy logic over ETX and RER to select the best CH parent node, for data transfer from participant node to DODAG root. However, MCEA-RPL does not perform well in dynamic environments where many real-time applications are realised, and this approach is inappropriate for parent node selection which increase the packet loss.

Table 2.4 Summarises Existing Approaches for RPL Energy Consumption Improvement

Previous Work	**Purpose**	**Metrics**	**Drawback**
ER-RPL	Energy efficiency	Distance, hop count, reliability and energy conservation	Not suitable for dynamic network environments
E2HRC	Balancing energy consumption	Direction angle, residual energy and hop difference	Increases time consumption
Hybrid routing	Improves PDR, residual energy, delay and packet drop	Load	Inefficient load balancing
HECRPL	Improves energy efficiency	residual energy , cost and node’s priority	increases time consumption
CRB-RPL	Improves delay and energy efficiency	CTQ and HEE	Introduces severe congestions High energy consumption

Previous Work	Purpose	Metrics	Drawback
Energy-aware route selection	Improves energy efficiency	Path cost and remaining energy	Increases number of retransmissions
Energy balancing scheme	Extends network lifetime	Energy and path cost	Inefficient route selection
MCEA-RPL	Extends network lifetime	ETX and RER	Increases number of retransmissions

Barcelo et al ,(2016) proposed Kalman Positioning-RPL (KP-RPL) to achieve a reliable routing in WSNs. In KP-RPL, the confidence region of a node was determined based on RSSI measurements, and the location of each node was predicted by setting a higher probability value within its confidence region. Kalman filter was used along with velocity measurements for refining. A possible route was then identified by following the estimated end-to-end ETX. The ETX performance metric gradually increased but did not exceed the positioning RPL routing.

Pavkovi et al ,(2011) modified the MAC layer of RPL-based IEEE 802.15.4 by adapting a cluster-tree topology to enable opportunistic routing. The nodes in the modified cluster-tree were allowed to associate with multiple parent nodes through an adequate organisation of superframes in the MAC layer. The opportunistic forwarding scheme was built over a modified MAC layer with a cluster-tree topology. The nodes were allowed to transmit their packets through multiple parents in an opportunistic manner to meet the transmission budget, and the transmission budget for each node was computed based on the deadline and hop count metrics. Collusions were avoided by scheduling superframes. Despite showing improvements in multipath routing, this method has an unreliable data transmission. Moreover, under conditions with a large traffic load in the network, this method increases the frequency of collisions.

Zhao et al ,2015 proposed an opportunistic coordination forwarding scheme over a cluster-parent-based RPL protocol. The end-to-end cost for each node was minimised by using a top-down approach and an optimal cluster parent set selection. The end-to-end cost in this approach was defined by the number of transmissions required by each node to achieve a successful packet transmission. Each node was provided with a cluster parent set and assigned a cost value. Afterwards, the optimal parent node was selected based on the link quality and cost value of a node. However, this method increases the number of retransmissions if the parent node fails to overhear the transmission of the other nodes.

Gonizzi, Monica and Ferrari (2013) minimised end-to-end delay in RPL routing by designing a delay metric that uses forward packet delay. The minimum forwarding time (MFT) was computed by adding the following time components:

1. time interval of a partial reception of the packet,
2. time interval of a complete reception,
3. time interval for the reception of additional packets,
4. time spent in internal processing,
5. waiting time until the node wakes up,
6. backoff time to check channel availability and
7. time interval to repeat the packet transmission until the receiver wakes up.

The forward delay was computed as

$$Delay = \frac{C_{T,C}}{2+MFT} \tag{2.3}$$

Where $C_{T,\,C}$ denotes the duty cycle time. Afterwards, the cost for each path was computed based on the average delay announced by the parent node, the forwarding delay of the parent node and the maximum delay threshold. Afterwards, the route that minimises the cost was selected as the optimal route for transmission. Although this method minimises the delay metric, reliability and energy efficiency still pose major concerns. Moreover, computing all delay metrics increases the time consumption and complexity.

Guo and Orlik (2016) jointly achieved a mixed mode of operation (MOP) and resource adaption in IoT by using the resource-aware hierarchical RPL (H-RPL) protocol. They also used requiring routing memory (RRM) and expected routing lifetime (ERL) to detect the mode in the network. RRM was computed as

$$M_L = N_P \times (|P_{ID}| + |P_{MOP}| + |DL| + |LU|) + |HR_{ID}| + |HD_{ID}| + |HD_{VN}| + |N_{MOP}| + OL \tag{2.4}$$

where N_p denotes the number of parents, P_{ID} denotes the parent ID, P_{MOP} denotes the MOP of the parent, DL represents the default lifetime, LU denotes the lifetime unit, HR_{ID} represents the H-RPL instance ID, HD_{ID} represents the H-DODAG ID, HD_{VN} represents the H-DODAG version number, N_{MOP} represents the MOP of the node and OL represents the memory required by the leaf.

ERL was defined as the period during which the node acts as a router and was computed based on the battery level of the node and leaf lifetime of the parent set nodes. However, this method increases the computational complexity and energy consumption in the network.

Omer et al ,2017 formulated an OF by considering several metrics, including available bandwidth, buffer occupancy and ETX, to improve the performance of RPL. The available bandwidth represents the capacity of the network and was computed as

$$\omega_n = \rho - \left(\frac{\sum_{\mu=1}^{\theta} \beta_\mu + \gamma_\mu}{\theta}\right) \tag{2.5}$$

where ω_n represents the average available bandwidth at any node, θ represents the current size of the averaging window, β_μ represents the total generation rate, γ_μ represents the total overhead at the MAC layer and μ represents the index number. Meanwhile, the buffer occupancy metric was considered to prevent the node from selecting a parent node with a high congestion. These metrics were used to improve network performance. However, this OF can only be implemented in upward routing and is only suitable for networks with a small number of nodes.

Kamgueu et al. (2013) achieved an energy-aware route selection by considering the residual energy of the node in an RPL-based network. The path cost between the source and sink was computed as

$$Cost_i = \min\,[\max(Cost_j, E_i)] \tag{2.6}$$

The path cost of node i ($Cost_i$) was computed by using the path cost of node j ($Cost_j$) and the remaining energy of node i (E_i). After computing the path cost, the path with the minimum cost was selected as the optimal path for transmission. Nevertheless, considering energy alone in the path selection would increase the number of retransmission and energy consumption.

Table 2.5 Summarises Existing Approaches for RPL Reliable Routing Improvement

Previous Work	Purpose	Metrics	Drawback
KP-RPL	Reliable routing	End-to-end ETX	Not efficient in route selection
MAC-RPL	Allows opportunistic routing	Deadline and hop count	Cannot guarantee reliable transmission, Introduces collisions in the network
Cluster-parent RPL	Minimises end-to-end cost	Link quality and cost value	Increases number of retransmissions
MFT-based RPL	Minimises end-to-end delay	MFT, average delay and duty cycle time	Major issues in reliability and energy efficiency, Increases complexity
H-RPL	Achieves mixed MOP and resource adaption	RRM and ERL	Increases computational complexity , High energy consumption
Multipath RPL	Improves QoS metrics	Buffer occupancy, hop count, PDR,	Large PLR Increases control message overhead

Previous Work	**Purpose**	**Metrics**	**Drawback**
		packet arrival rate and packet service time	

Lodhi et al, 2017 concentrated QoS metrics, such as fault tolerance, reliability, congestion mitigation and hole avoidance, via the multipath extension of the RPL protocol. The key idea behind this protocol was to enable multipath routing over a single routing path to avoid congestion. The node that was free from congestion was selected as the optimal parent node, and transmission was performed through this node. Node congestion was determined by using buffer occupancy, hop count, PDR, packet arrival rate and packet service time. Congestion detection and mitigation were performed by using two control messages, namely, emergency DIO and congestion notification messages. A parent list that contains details on the potential parent nodes was maintained at each node. However, congestion mitigation through multipath transmission leads to a large PLR, and the involvement of additional control messages increases the amount of overhead in the network.

Oikonomou et al ,(2013) performed multicasting in RPL-based networks by using the TM and SMRF algorithms. TM was enabled by exchanging frequency of periodic information without leading to control message flooding. In this approach, each packet was allowed to carry multiple options, such as sequence number, single bit M parameter and unique identifier. However, this method also has certain limitations, including delays, complexity and multicast problems. The multicasting and arrival order problems in TM were addressed by SMRF by considering the topology information. However, duplicate data propagation presents a major problem in SMRF that increases time and energy consumption.

Qorany and Fadeel (2015) proposed the enhanced SMRF (ESMRF) algorithm to address the problems in the SMRF algorithm. ESMRF initially constructs a multi-hop tree to enable multicasting in both the up and down directions. The multicast packet of the source node was encapsulated into the ICMPv6 delegation packet in the root node. In this way, the packet of the source node was multicast from the root node instead of the root node. All nodes in the network would send their multicasting packets to the root, and then the root would verify whether these packets already exist. If these packets were already transmitted by the root, then they were dropped by the root to minimise flooding in the network. However, this method increases overhead at the root node and is not suitable for large networks.

Table 2.6 Summarises Existing Approaches for RPL Network Lifetime Multipath Approach Improvement

Previous Work	**Purpose**	**Metrics**	**Drawback**
Multipath RPL	Network lifetime improvement	ELT (ETX and energy)	High energy and time consumption
TM and SMRF	Efficient multicasting	Topology information	Introduces longs delays and high complexity, Duplicate data propagation increases time and energy consumption

Previous Work	Purpose	Metrics	Drawback
ESMRF	Improves SMRF	Multi-hop tree	Increases overhead at the root node, Unsuitable for large networks

Kim et al ,(2016) proposed queue-utilisation-based RPL (QU-RPL), which eliminates the congested nodes to achieve a best parent node selection. QU-RPL attempted to improve the end-to-end packet delivery performance of the network through load balancing. The route selection process considered the queue utilisation (QU) factor, which was computed as

$$QU = \frac{Number\ of\ packets\ in\ the\ queue\ of\ node}{Total\ queue\ size\ of\ the\ node} \tag{2.7}$$

The optimal parent node was then selected based on the QU, ETX and hop count metrics. Despite addressing the congestion in the network and selecting the node with the minimum congestion, this approach cannot avoid congestion given that the major reason for network congestion is the presence of an attacker. Moreover, parent selection based on QU, ETX, and hop count limits packet transmission efficiency.

Lee et al ,(2014) improved transmission performance in RPL-based 6LoWPAN by considering RSSI-based IPv6 routing metrics. The RSSI metric was associated with the link-oriented metric ETX. The nodes would periodically update the ETX_RSSI value to enable an efficient neighbour selection. After updating the ETX_RSSI value, the node would select one- and two-hop neighbour nodes for the data transmission. The payload utilisation was increased after the data transmission.