UNIVERSITY
OF TECHNOLOGY
SYDNEY

Knowledge-Defined Routing Strategies for Next Generation Wireless Networks

by Sepehr Ashtari Nakhaei

Thesis submitted in fulfilment of the requirements for
the degree of

Doctor of Philosophy

Under the supervision of Professor Mehran Abolhasan and
A/Professor Justin Lipman

University of Technology Sydney
Faculty of Engineering and Information Technology

June 2023

STATEMENT OF ORIGINALITY

I, Sepehr Ashtari Nakhaei, declare that this thesis is submitted in fulfillment of the requirements for the award of doctor of philosophy in the Engineering and Information Technology faculty at the University of Technology Sydney. This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution. This research is supported by the Australian Government Research Training Program.

Signed:

Production Note:
Signature removed prior to publication.

Sepehr Ashtari nakhaei
June 11, 2023

ABSTRACT

The proliferation of smart devices and corresponding applications has resulted in a significant increase in cellular network traffic. To mitigate this, device-to-device (D2D) communication has been proposed as an efficient solution for reducing network congestion and increasing network capacity. D2D was first introduced by the Third Generation Partnership Project (3GPP) as a key enabling technology for the fifth-generation (5G) of cellular networks. However, the limitations of D2D have led to the emergence of multi-hop D2D (MD2D) as a potential replacement and an enabling technology for future generations of cellular networks, such as the sixth generation (6G) and beyond.

The first objective of this thesis is to study the new network architecture known as knowledge-defined networking (KDN) and the benefits of various MD2D routing protocols and frameworks. We conducted a comprehensive literature review that revealed two key findings. First, future network architecture should integrate automated techniques to enhance network performance. Secondly, centralized controllers and virtualization technologies, such as software-defined networking (SDN) and network function virtualization (NFV), are crucial for achieving high performance in MD2D networks. As a result, we targeted to study centralized and automated MD2D routing frameworks and protocols.

The second objective of this thesis was to enhance the energy efficiency of MD2D routing protocols by introducing efficient and intelligent routing methods. We proposed two intelligent routing protocols that leverage network knowledge to construct routing tables. The first method employs a centralized participation mechanism that utilizes a fuzzy logic system to identify eligible nodes for MD2D communication based on multiple network metrics. The second method uses various utility functions to create an automated routing mechanism to alternate paths, thus increasing network lifetime and throughput.

The final objective of this thesis is to enhance MD2D performance with regard to traffic offloading and end-to-end (E2E) delay. To achieve this, we employed virtualization techniques in conjunction with WiFi network slicing to segment various nodes into distinct network slices to optimize throughput and E2E delay. Additionally, we introduced a location-based routing strategy to augment cellular traffic offloading. This MD2D routing protocol utilizes the location of nodes in conjunction with network topology to determine efficient and dependable

routes, thereby reducing E2E delay and increasing throughput, ultimately enabling offloading of a greater number of packets from the BS.

In conclusion, this thesis presents an in-depth examination of MD2D networks and proposes intelligent and high-performance MD2D routing protocols and frameworks.

Keywords: *MD2D, 5G, 6G, KDN, SDN, NFV, routing protocols and frameworks, network automation.*

DECLARATION OF CONTRIBUTION TO THE PUBLICATIONS

I, Sepehr Ashtari Nakhaei declare that the under mentioned publications are the product of my original work done during the Ph.D. candidature. Also, my contribution as the first author in these publications is more than 90%.

Ser	Paper Title	Journal/Conference	Status	Authorship
1	Knowledge-defined networking: Applications, challenges and future work	Elsevier Array	Published	First Author (90%)
2	Performance analysis of multi-hop routing protocols in SDN-based wireless networks	Elsevier Computer & Electrical Engineering	Published	First Author (90%)
3	Joint Mobile Node Participation and Multi-hop Routing for Emerging Open Radio-Based Intelligent Transportation System	IEEE Access	Published	First Author (90%)
4	A Cross-Layer Multi-Hop Device-to-Device Routing Protocol for Future Heterogeneous Wireless Cellular Networks	IEEE Internet of Things Journal	Submitted	First Author (95%)
5	An Adaptive Multi-hop Device-to-Device Routing Framework for Future Cellular Networks	Elsevier Journal of Network and Computer Applications	Under Review	First Author (95%)
6	A Location-based Multi-hop Routing Protocol for Future Wireless Cellular Networks	Elsevier Computer Networks	Submitted	First Author (95%)

Student Signature	Signatures of Co-Authors
Sepehr Ashtari Nakhaei Production Note: Signature removed prior to publication. 05-02-2023	Mehran Abolhasan:--------------------------------- Production Note: Signature removed prior to publication. 05/02/2023 Justin Lipman:-- Production Note: Signature removed prior to publication. 05/02/2023 Wei Ni: -- Production Note: Signature removed prior to publication. 05/02/2023 Negin Shariati:--------------------------------------- Production Note: Signature removed prior to publication. 5/02/2023 Abbas Jamalipour:---------------------------------- Production Note: Signature removed prior to publication. 06/02/2023 Mahrokh Abdollahi:------------------------------------- Production Note: Signature removed prior to publication. 05/02/2023 Ian Zhou: -------------------------------------- Production Note: Signature removed prior to publication. 05/02/2023

Acknowledgements

I am deeply grateful to my supervisor, Professor Mehran Abolhasan, for his invaluable guidance and support throughout this thesis. Without his supervision and guidance, this work would not have been possible. I also extend my sincere thanks to my co-supervisor, Associate Professor Justin Lipman, for his valuable contributions and support.

I am also extremely grateful to Dr. Wei Ni from the Commonwealth Scientific and Industrial Research Organization (CSIRO) for his unwavering support and invaluable guidance throughout this thesis.

I would like to express my sincere gratitude to Dr. Negin Shariati, Professor Abbas Jamalipour, and all others who have helped me throughout my Ph.D. studies.

Finally, and most importantly, I would like to express my deep appreciation and love to my family, without whom none of this would have been possible.

LIST OF PUBLICATIONS

Published Journal Papers

J-1. Ashtari, S., Abdollahi, M., Abolhasan, M., Shariati, N., Lipman, J. (2022). Performance analysis of multi-hop routing protocols in SDN-based wireless networks. Computers Electrical Engineering, 97, 107393.

J-2. Ashtari, S., Zhou, I., Abolhasan, M., Shariati, N., Lipman, J., Ni, W. (2022). Knowledge-defined networking: Applications, challenges and future work. Array, 100136.

J-3. Ashtari, S., Abolhasan, M., Lipman, J., Shariati, N., Ni, W., Jamalipour, A. (2022). Joint Mobile Node Participation and Multihop Routing for Emerging Open Radio-Based Intelligent Transportation System. IEEE Access, 10, 85228-85242.

Under Review Journal Papers

J-1. Ashtari, S., Abolhasan, M., Lipman, J., Ni, W. (2023). A Cross-Layer Multi-Hop Device-to-Device Routing Protocol for Future Heterogeneous Wireless Cellular Networks. IEEE Internet of Things Journal.

J-2. Ashtari, S., Abolhasan, M., Lipman, J., Ni, W., Jamalipour, A. (2023). An Adaptive Multi-hop Device-to-Device Routing Framework for Future Cellular Networks. Elsevier Journal of Network and Computer Applications.

J-3. Ashtari, S., Abolhasan, M., Lipman, J., Shariati, N., Ni, W. (2023). A Location-based Multi-hop Routing Protocol for Future Wireless Cellular Networks. Elsevier Computer Networks.

CONTENTS

List of Figures

List of Tables

LIST OF ACRONYMS

1G	1st generation	DSR	dynamic source routing
2G	2nd generation	DTN	delay tolerant networks
3G	3rd generation	E2E	end-to-end
3GPP	3rd generation partnership project	EI	eligibility index
4G	4th generation	eMBB	enhance mobile broadband
5G	5th generation	FBMC	filter bank multi-carrier
6G	6th generation	FERR	flow error
ACMRP	baseband unit	FPRM	fuzzy-based participation routing protocol MD2D
AI	artificial intelligent	FREP	flow reply
ANN	artificial neural network	GA	genetic algorithm
AoA	angle of arrival	GLSDB	geographical link-state databases
AoD	angle of departure	GPS	global positioning system
AODV	ad hoc on-demand distance vector	HD	high-definition
API	application programming interface	HDR	high definition real-time
BDMA	beam division multiple access	HSAW	hybrid SDN architecture for wireless distributed network
BS	base station	ID	identification
CH	cluster head	IoE	Internet of everything
CNN	convolutional neural network	IoT	Internet of things
C-RAN	cloud/centralized-radio access network	IP	Internet protocol
CRN	cognitive radio network	KDN	knowledge-defined networking
CSI	channel state information	KP	knowledge plane
D2D	device-to-device	LA	learning automata
DNN	deep neural network	LOS	line-of-sight
”DRERR	” data route error	LSDB	link state database
DRL	deep reinforcment learning	LTE	long-term evolution
DSDV	destination sequenced distance vector	M2M	machine-to-machine

MANET	mobile ad hoc network	RIC	RAN intelligent controller
MBS	macro-cell BS	RL	reinforcement learning
MC	mission critical	RREQ	route request
MD2D	multi-hop device-to-device	RSU	roadside unit
MIMO	multiple-input multiple-output	SBS	small base station
ML	machine learning	SDN	software-defined networking
MLSDB	master link state database	SINR	signal to interference noise ratio
mMTC	massive machine type communication	SL	supervised learning
mmWave	millimeter wave	SMDRP	SDN-based MD2D routing protocol
MS	mobile station	SNR	signal-to-noise ratio
NC	non-critical	SON	self-orgonizing network
NFV	network functions virtualization	SPF	shortest path first
NGMN	next generation mobile networks	TC	topology control
NLOS	non-line-of-sight	TCREQ	topology control request
NN	neural network	TDMA	time divisions multiple access
NS-3	network simulator-3	TL	transfer learning
OF	openflow	UAV	unmanned aerial vehicle
OFDM	orthogonal frequency-division multiplexing	UE	user equipment
OFDMA	orthogonal frequency-division multiple access	UL	unsupervised learning
OLSR	optimized link state routing protocol	uRLLC	ultra-reliable and low-latency communication
O-RAN	open-radio access network	V2I	vehicle-to-infrastructure
OSPF	open shortest path first	V2V	vehicle-to-vehicle
P4	programming protocol-independent packet processor	V2X	vehicle-to-everything
PCA	principal component Analysis	VANET	vehicular ad hoc network
PDR	packet delivary ratio	VARP-S	source-based virtual ad hoc routing protocol
PRB	physical resource block	V-RAN	virtualized-radio access network
PSO	particle swarm optimization	VS	virtualized slices
QoE	quality of experience	WiMAX	worldwide interoperability for microwave access
QoS	quality of service	WLAN	wireless local-area network
RAN	radio access netwrok	xRAN	extensible radio access network
RFID	radio frequency identification		

1

Introduction

This thesis aims to present efficient solutions for managing data traffic and increasing network capacity in wireless cellular networks. Our investigation has led us to conclude that one of the most significant challenges facing future wireless networks will be excessive network traffic from heterogeneous users. To address this challenge, we propose using self-managed device-to-device (D2D) communication as a promising solution. Self-managed D2D communication involves utilizing automated techniques to create a self-organizing D2D network, thereby allowing the network to dynamically update and adjust its parameters to optimize performance. This thesis focuses on developing network intelligence and optimization algorithms for WiFi channels to minimize cellular overhead and maximize network performance.

The proliferation of smart devices and corresponding applications has resulted in a significant increase in cellular network traffic [1]. To mitigate this, D2D communication has been proposed as an efficient solution for reducing network congestion and increasing network capacity [2]. However, D2D is limited to one neighboring communication, which decreases the coverage area. To address this limitation and extend the coverage, multi-hop D2D (MD2D) has been proposed as an alternative [3]. MD2D creates an independent secondary infrastructure to improve service coverage and data capacity. To fully leverage the benefits of MD2D networks, new routing protocols and frameworks must be designed and implemented. This thesis comprehensively studies MD2D in a cellular network and proposes new automated routing strategies and frameworks to assist future wireless cellular communication.

1.1 Background

The fifth-generation (5G) wireless network has yet to achieve full automation and intelligence [4], despite offering several advantages over traditional networks, such as improved quality-of-service (QoS), higher data rates, and integration and management of licensed and unlicensed bands [4,5]. However, the current wireless cellular architecture lacks flexibility and intelligence, requiring constant optimization and improvement through software or hardware updates. With the advent of machine learning (ML) and network automation, future cellular networks can provide self-management, self-optimization, and self-adaptation, thus reducing the need for human intervention and updates, thereby decreasing costs and complexity. The sixth-generation (6G) cellular network aims to introduce intelligence and adaptability within the network architecture [6]. Several initiatives, such as Open Networking Foundation (ONF) and Open-Radio Access Network (O-RAN), are currently underway to create intelligent radio networks, focusing on different parts of the network, such as core or edge, with the goal of creating a fully open and intelligent platform for future wireless networks [7].

Software-defined networking (SDN) and network function virtualization (NFV) are foundational technologies to implement new intelligence and self-adaptive cellular network initiatives [8]. SDN is a networking paradigm that decouples the data and control planes [9]. Decoupling these two planes enables SDN to operate as a centralized controller to manage the network. The global view of the SDN controller provides advantages such as network flexibility, programmability, and efficient management over the traditional network. On the other hand, NFV leverages the recent advances in cloud computing to create virtualized services to provide agility and scalability [10]. With the combination of SDN and NFV, a new network architecture was created called knowledge-defined networking (KDN) [11]. KDN adds a layer on top of SDN architecture's control and management layers, called knowledge plane (KP) [12]. The KDN architecture gathers network information and creates network knowledge through optimization and ML techniques. Many of the existing open radio access network ideas come from KDN architecture. KDN was the first network paradigm to integrate intelligence for self-management and self-tuning.

Network knowledge is harvested from user requirements, user applications, network functionality and traffic. However, efficient knowledge is derived based on the outcome of an optimization algorithm or an ML process for a specific application. For instance, a neural network (NN) in conjunction with reinforcement learning (RL) can be used in wireless networks for dynamic power allocation optimization [13]. The output of the RL/NN is taken as adequate knowledge and can be used for future instant decision-making. Knowing about the network's behavior in various environments can be a breakthrough in network performance. First, in resource management, parameters such as bandwidth, QoS, and power can be obtained in different network situations and stored as knowledge for network automation. Second, routing decisions can deploy knowledge in networking for better route discovery while the network is overpopulated. Moreover, user information, including mobility patterns and their velocity, can be used as additional knowledge to improve the accuracy of localization and handover.

In order to accomplish the objectives outlined in this thesis, our focus is on utilizing network knowledge and incorporating it into the design of future cellular network architectures (specifically with MD2D networks) to reduce network congestion and alleviate traffic overload. Network traffic management will continue to be a significant challenge in wireless cellular networks with the constant growth of user devices [14]. Hence, a complementary network communication platform is needed to offload traffic or use it as a new standard for direct data exchange between devices. In recent years, congestion in wireless communication has been growing significantly. It is predicted that the traffic will rise hundreds of thousands of times every year [15]. The emergence of MD2D communication can assist cellular systems in reducing traffic to a secondary infrastructure that can provide similar or even better services [3]. MD2D is a promising technology that can deliver flexibility and scalability to the network. However, new routing algorithms with advanced and intelligent features must emerge to efficiently integrate MD2D in future cellular networks.

The integration of MD2D communication into wireless networks marks a pivotal step in the evolution towards intelligent and self-adaptive cellular systems [16]. To fully realize the potential of MD2D, it is essential to develop new routing protocols and design an intelligent network infrastructure that incorporates knowledge for self-management and self-adaptability. This can be achieved through the application of SDN, NFV, and KDN paradigms. For example, SDN can be used to separate routing decisions from data forwarding, allowing the data plane to solely transmit data instead of undergoing multi-hop flooding. NFV can facilitate advanced optimizations by creating virtualized services as separate entities, while KDN can enable both applications by the creation and implementation of intelligent policies through the gathering of network information, optimization and virtualization.

1.2 Motivations and Objectives

The massive increase in data traffic and excessive load on the base station (BS) is causing performance deterioration in cellular networks in terms of user and application demands. It is important for cellular network providers to address this challenge and invest in new technologies and solutions to upgrade their infrastructure and improve network performance. Moreover, the lack of flexibility and intelligence in cellular networks can hinder their ability to effectively adapt to changing network conditions and user demands. This can result in decreased network performance, reduced reliability, and increased network congestion. The traditional cellular network architecture was designed primarily to support voice communication and was not optimized for the growing demand for data services and Internet of Things (IoT) devices. As a result, the network infrastructure can become overburdened, leading to a decline in network quality and user experience. To address this challenge, cellular network providers and researchers are exploring new technologies and solutions to improve the flexibility and intelligence of cellular networks. This includes the implementation of self-management, self-optimization, and self-adaptation networks. To address this challenge, researchers and network providers are exploring new solutions, including creating optimized MD2D networks and providing automated routing protocols and adaptive routing frameworks and policies. Based on the above motivations, several research questions have arisen. These questions are as follows: how can traffic overheads in future wireless cellular networks be reduced? Second, what are the limitations and advantages of different MD2D routing protocols? Third, how to design efficient MD2D routing protocols and frameworks to improve cellular performance? And finally, How to create adaptive and intelligent MD2D routing protocols?

These questions are the foundation of this thesis and provide direction for its development. As a result, the objectives of this thesis can be summarized as follows:

- Objective 1: To evaluate the possible solutions for network congestion and traffic flow in the cellular network.
- Objective 2: To reduce traffic congestion and increase coverage and capacity in future cellular networks.
- Objective 3: To propose new multi-hop device-to-device (MD2D) routing algorithms and frameworks.
- Objective 4: To make MD2D routing intelligent and adaptive based on user or application requirements.

1.3 Thesis Contributions

The thesis contributions are summarized as follows:

1. A thorough investigation of current limitations and challenges of wireless cellular networks and the future direction is conducted. We explore solutions using KDN architecture for different parts of the network. We investigate ML-based optimization algorithms for various network applications in future cellular networks and summarize the most influential research studies.
2. A detailed study is performed to investigate the performance of proactive and reactive MD2D routing protocols using hybrid SDN architecture for wireless distributed networks (HSAW), SDN-based multi-hop D2D routing protocol (SMDRP), and virtual ad hoc routing protocol (VARP-S). Our results show that the proactive MD2D protocol (HSAW) introduces the highest overhead in the cellular channel while consuming the highest energy compared to the other two reactive protocols (SMDRP and VARP-S). Conversely,

HSAW provides the lowest end-to-end (E2E) delay and packet loss compared to reactive protocols.

3. An intelligent joint topology control and multi-hop routing called fuzzy-based participation and routing protocol for MD2D (FPRM) to increase the network lifetime and packet delivery ratio (PDR) is proposed. In our approach, a fuzzy-based participation mechanism controls and manipulates the network's topology.

4. A new joint utility-based routing protocol called application-driven cross-layer MD2D routing protocol (ACMRP) is presented. ACMRP finds optimal route by incorporating knowledge from the application and network layers. Our protocol creates network knowledge and formulates multiple link-state databases (LSDBs) that can be used to adjust a route by considering the user and application requirements.

5. The concept of virtual network slicing for WiFi channels and a mechanism to enable multiple MD2D routing protocols deployed over each virtual slice was proposed. The network slicing problem is interpreted as a swarm of particles, where UEs/particles represent a candidate solution. As a result, we deploy the improved self-adaptiv particle swarm optimization (IDPSO) algorithm to dynamically identify the network slices and deploy different routing protocols depending on the user application on the network slice.

6. Finally, a new MD2D position-based routing protocol is proposed and compared with a link-state routing protocol. We jointly use the coordinates of nodes with zone-prescribed neighbor discovery to provide fast and flexible routes. To preserve the location of nodes and keep the routing protocol secure and reliable, a privacy-preserving scheme is used in the WiFi channel during MD2D communication.

1.4 Research Methodology

This research endeavors to study the potential benefits of knowledge-defined routing strategies for future wireless networks and to explore and recommend innovative solutions for MD2D communication. To achieve this, we have extensively evaluated various optimization techniques and strategies, selecting those that are computationally efficient and cost-effective. Our methodology incorporates optimization techniques, including fuzzy logic strategies, utility functions, particle swarm optimization (PSO), learning automata (LA), and genetic algorithms (GA). By leveraging a range of optimization techniques and strategies, this research aims to investigate and recommend new solutions for future wireless cellular networks, specifically using MD2D communication. Our proposed solutions have been thoroughly evaluated and validated using open-source network simulation tools such as MATLAB and Network Simulator-3 (NS3). The findings of this research will contribute to the advancement of future cellular network design and deployment.

1.5 Thesis Structure

This thesis is organized as follows:

- *Chapter 2*: This chapter presents an overview of cellular networks and provides enabling technologies, challenges, limitations and possible solutions. We also introduce the new RAN architecture and its capabilities. We thoroughly investigate heterogeneous ad hoc networks and a detailed literature review of existing MD2D routing protocols.
- *Chapter 3*: A comprehensive study of proactive and reactive routing protocols for mobile networks based on SDN is presented in this chapter. To this end, previously designed routing protocols are investigated in depth, HSAW as a proactive protocol, VARP-S, and SMDRP as reactive protocols. The routing protocols are compared in distinct environments with different node densities and mobility rates. In each setting, the weakness and strengths of each protocol are analyzed in terms of E2E delay, dropped packets, energy consumption, and routing overhead.
- *Chapter 4*: This chapter presents a fuzzy-based node participation and routing protocol for MD2D networks. A sub-layer at the network layer that can determine nodes with the highest participation probability in routing is utilized by utilizing a fuzzy logic system.
- *Chapter 5*: This chapter proposes a new utility-based routing protocol, known as application-driven cross-layer MD2D routing protocol (ACMRP), to adjust a route by considering the application requirements and packet-level information. Our protocol creates network knowledge and formulates multiple LSDBs that can be used to evaluate different route strategies.
- *Chapter 6*: This chapter introduces an adaptive slicing mechanism with a dynamic MD2D routing protocol selection technique for future wireless cellular networks. A controller is responsible for collecting the network data and utilizing the improved self-adaptive particle swarm optimization (IDPSO) algorithm to virtually slice WiFi channels into various virtual sub-layers based on network traffic. In particular, our algorithm uses data request from users to create virtual slices where users with similar content can share and download data to or from other users. Moreover, the controller prescribes the most efficient routing protocol using the genetic algorithm (GA) in conjunction with learning automata (LA) for any virtual slices based on the content, mobility rate, throughput, number of neighbors, and network density.
- *Chapter 7*: This chapter presents a new MD2D position-based routing protocol in conjunction with zone-prescribed neighbor discovery. Our protocol acquires the position of the nodes and identifies the the zones that need to perform neighbor discovery. The BS prescribes reliable routes and the next hop based on the distance and number of hops to the destination. The BS utilizes node positions to create geographical link-state databases

(GLSDB). The GLSDB evaluates the paths to the destination by deploying a privacy-preserving algorithm and an efficient cost function.

- *Chapter 8*: The final chapter concludes our research work. We summarize all the chapters in the thesis and draw a conclusion based on our study. We further exploit the research and provide insights into the future direction of this thesis and how one can continue and contribute further in this field.

2

Literature Survey

2.1 Overview

Part of the upcoming chapter is published in Elsevier Array [1]. The main objective of this chapter is to provide an overview of the current challenges faced by wireless cellular networks and potential solutions. The challenges discussed include traffic congestion and future user demands. The focus is then placed on device-to-device (D2D) communication as a promising solution to address future cellular traffic congestion.

2.2 Introduction

The substantial increase in user and application demands has necessitated cellular network operators to explore innovative solutions for managing the exponential growth in data traffic. There are several potential strategies that can provide relief, such as the deployment of smaller cells to accommodate more users, but this approach reaches a point of diminishing returns and incurs significant costs. Another alternative is to develop new algorithms to facilitate shared communication channels, however, there are limits to the amount of bandwidth that can be shared among users. Hence, it is imperative for new technological advancements to emerge, offering more efficient and effective solutions.

[1] Ashtari, S., Zhou, I., Abolhasan, M., Shariati, N., Lipman, J., Ni, W. (2022). Knowledge-defined networking: Applications, challenges and future work. Array, 100136.

The future of wireless cellular systems demands innovative solutions to address the growing demands of users and applications. To meet these demands, two promising approaches have emerged. The first is the utilization of D2D communication, which enables adjacent nodes to exchange data without the involvement of a base station (BS). The second approach is to enhance the intelligence and self-management capabilities of the network architecture. Despite the advancements of 5G technology, it falls short in providing sufficient flexibility and intelligence to handle the increasing demands for massive machine-type communication (mMTC), low-latency, and enhanced mobile broadband (eMBB) services [17]. The next-generation, sixth-generation (6G) cellular network holds the potential to overcome these challenges and improve performance through terahertz frequencies, ultra-reliable low latency communications (URLLC) and the implementation of advanced intelligence within the network architecture [18]. Thus, a transformation is necessary from the current 5G architecture to a more intelligent cellular network.

For an architectural transformation, we need to look at the missing components in the current cellular architecture. One critical component that is missing is intelligence. More specifically, the intelligence that comes from knowledge. As shown in Figure 2.1 the knowledge resides in the controller and uses network information to create intelligence. Knowledge is used for recommendation and automation across wired and wireless network applications. For instance, in resource management problems, parameters such as bandwidth, quality of service (QoS), and power can be obtained and processed using an machine learning (ML) algorithm in different network situations [19]. The output of ML can be referred to as intelligence over the network, which is used for automation purposes. Moreover, routing decisions can benefit from knowledge in the networking applications for better route discovery while the network is congested [20].

The remainder of this chapter delves into the examination of network architecture and the latest developments in D2D technology. We begin with an exploration of the evolution of wireless cellular communication, delving into the underlying enabling technologies and highlighting the prevailing challenges faced by the current cellular network. Subsequently, we examine the architecture of wireless networks and the advancements in D2D communication. This thesis endeavors to present a novel network architecture that incorporates advanced network topologies and protocols, leveraging the capabilities of D2D communication to enhance the overall network performance.

2.3 Wireless Cellular Network

In wireless cellular communication, users request and receive information from the BS. Nowadays, the majority of cellular traffic is consumed by smart devices. The growth in demand for information and data among users has resulted in a rapid increase in the use of wireless cellular communication. In recent years, the widespread adoption of smart devices has revolu-

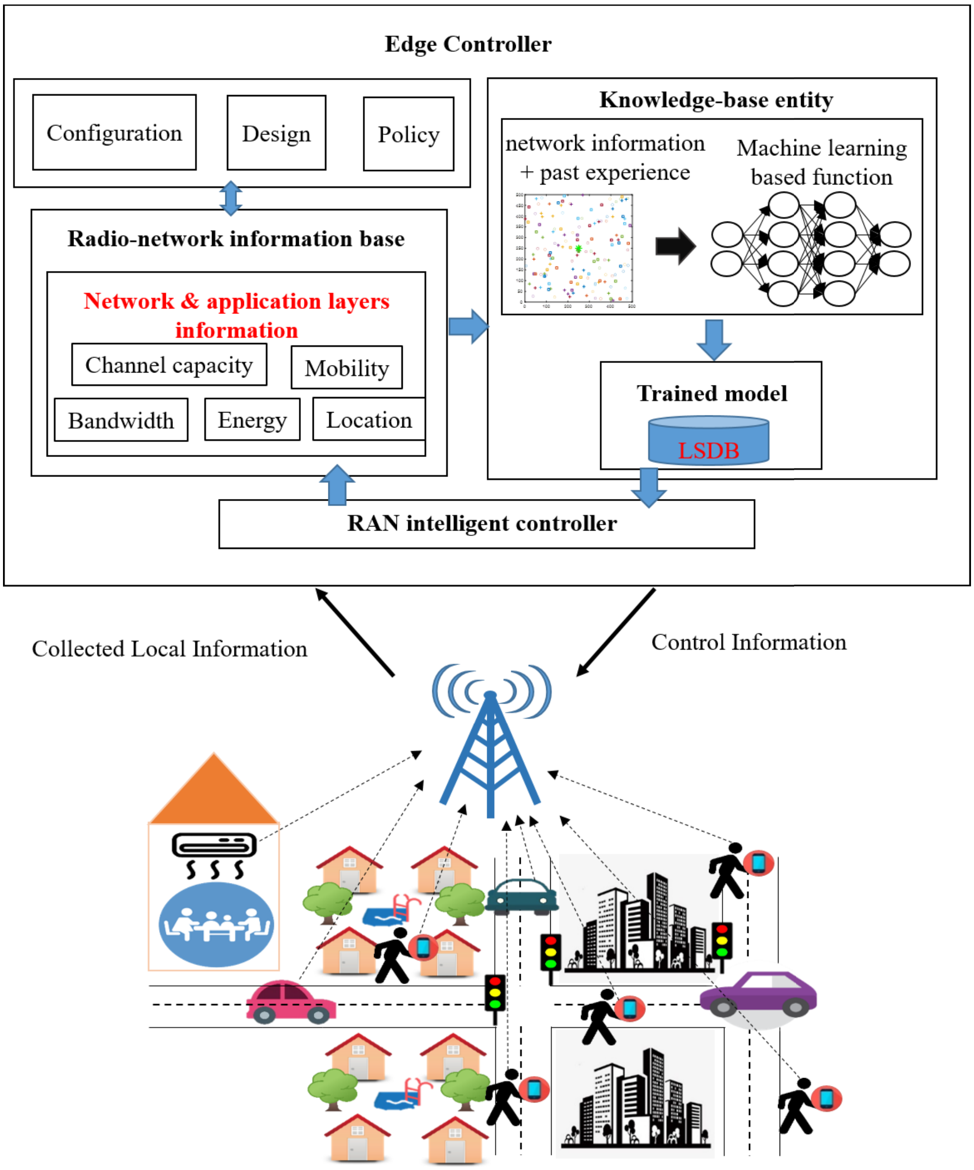


Figure 2.1: Integration of KDN into cellular architecture.

tionized the way people communicate and access information. Today, three out of four people worldwide own smartphones and rely on cellular communication as their primary means of communication. The increasing popularity of new applications such as internet-of-things (IoT) and autonomous vehicles has also contributed to the growth of cellular traffic. To keep pace with the growing demand, ongoing research and development efforts are focused on enhancing the capabilities of radio access networks (RANs) and cellular networks. In the following sub-sections, we provide a brief overview of the recent cellular network innovations and the enabling technologies. Finally, we present the direction of this thesis using the current challenges and possible solutions.

2.3.1 Evolution of Wireless Cellular Networks

Since the first generation (1G) of cellular communication, increasing adoption, user expectations and new bandwidth-hungry applications have led to the steady advancement in the evolution of cellular networks [21]. For instance, the upgrade from 1G to second generation (2G) provided faster connectivity by enabling time-division multiple access (TDMA) and expanding the communication bandwidth. Later advancements in modulation techniques and RAN enabled third-generation (3G) wireless networks, where low bit rate web traffic type applications were enabled. There has always been a significant change in every evolution. However, the fourth generation (4G) of wireless cellular networks was a significant leap forward in terms of providing higher data rates and increasing network capacity.

The evolution of cellular networks has been driven by a combination of increasing demands, rising user expectations, and bandwidth-intensive applications. Since the first generation (1G) of cellular communication, steady advancements have been made, leading to faster connectivity and initiation of new applications. For example, the transition from 1G to the second generation (2G) involved the implementation of time-division multiple access (TDMA) and an increase in communication bandwidth. Subsequent advancements in modulation techniques and radio access networks (RANs) led to the introduction of third-generation (3G) wireless networks, which enabled low-bitrate web traffic type applications. The fourth generation (4G) of wireless cellular networks marked a significant leap forward in terms of providing higher data rates and increased network capacity, and every subsequent evolution has brought about significant changes as well.

Summary of current and future cellular technologies:

- **4G**: 4G is still in use in various parts of the world. 4G provides fast connectivity and high-resolution data content. 4G was introduced in the late 2000s but upgraded rapidly throughout the 2010s [22]. New technological innovations such as multiple-input multiple-output (MIMO) antennas helped 4G to attain a significant difference from the other generation of cellular networks. 4G was the main evolution in wireless cellular communication in Internet connectivity, multimedia service, voice, and online streaming. The key advantages of 4G compared to its predecessors were faster connectivity, capacity, security, and low cost of services. However, all around the world number of smart devices is exponentially increasing. Consequently, more devices and applications will require cellular connectivity, causing massive data traffic. Therefore, due to higher user demands and the emergence of new applications, 4G was replaced by 5G [23].

- **5G**: The fifth generation (5G) of wireless cellular networks was established in December 2017 by the 3rd generation partnership project (3GPP). 5G was in operation as of 2022 in most developed countries, and it's now officially the latest cellular network technology. 5G provides 100 times faster connectivity, 10 to 100 times increase in the number of connected devices, 5 times lower E2E delay, and 10 times extended battery life compared to 4G. The enabling technologies in 5G are advanced access technology named beam division multiple access (BDMA), non- and quasi-orthogonal or filter bank multi-carrier (FBMC) multiple access, and millimeter waves (mmWaves). Millimeter waves provide a high-band spectrum for high speed and low latency, enabling up to 10 Gbps data rate. Ideally, a two-hour movie would take less than four seconds to be downloaded by 5G. Although 5G provides massive advantages compared to the previous generation, the existing 5G architecture lacks the flexibility and intelligence to enable self-management, self-optimization, and self-adaptation [6].

- **6G**: The study on the sixth generation (6G) of wireless cellular networks started in October 2020 by next generation mobile networks (NGMN) [24]. They proposed a new "6G Vision and Drivers" project to provide early direction for global 6G activities. Many countries and universities across the world have started their research focusing on several challenging areas, including reliable near-instant unlimited wireless connectivity, distributed computing, intelligence, and technological advancements in antennas, circuits, and devices. Among the research community, many are concentrating on AI-enabled technologies to facilitate ML/AI to make the network intelligent and self-drivable. 6G is intended to make architectural changes to the existing cellular network and create an open environment to provide flexibility and intelligent functionalities. Future networks must be prepared for the massive number of services and applications that require ultra-reliable and low-latency communication (uRLLC), massive machine type communication (mMTC), and enhanced mobile broadband (eMBB) [25].

The evolution of cellular networks has always concentrated on data rate, user connectivity, and E2E delay. However, the massive growth of applications and user demands for ultra-reliable services with the highest QoE leads us to 6G and beyond.

2.3.2 Enabling Technologies for Future Wireless Networks

Several enabling technologies can assist the emergence of the next generation of wireless cellular networks, where some of them are explained as follows:

- **Network Function Virtualization (NFV)**: The growth and complexity of traditional networks have resulted in inefficiency in management and service provision. To address these issues, network function virtualization (NFV) was introduced as a solution that decouples network functions from proprietary hardware and transforms them into software. This decoupling improves service provision and reduces operational and capital expenditures. NFV is a crucial enabler for improving network performance in mobile networks, where it enables on-demand agile provisioning of mobile functions. When combined with Software-Defined Networking (SDN), NFV offers a scalable and elastic ecosystem for automated network management and orchestration. The integration of NFV and SDN is considered a key breakthrough in 5G and future wireless networks, such as cellular, IoT, vehicular, and machine-type communication. The incorporation of NFV with wireless networks offers high flexibility in network orchestration and enables the optimization of physical resources, reduction of costs, and decrease in network energy consumption. In 5G, most of the core network functions are virtualized, providing adaptability in various network scenarios [26].

- **Open-Radio Access Network**: Open-radio access network (O-RAN) is an emerging technology that enables service heterogeneity, on-demand service deployment, and simultaneous coordination of heterogeneous devices [27]. O-RAN was introduced in 2018 by merging the xRAN Forum and C-RAN alliance to create a new openness functionality that would support the evolution beyond 5G and 6G wireless. Over the past few years, various efforts from the research industry have been made to enhance the radio access network (RAN) [27, 28]. Among them, cloud-RAN or centralized-RAN (C-RAN) is a promising RAN architecture that helps to reduce baseband expenses. Further, a virtualized radio access network (V-RAN) brings the benefit of cloud and virtualization to increase the network's agility, scalability, and flexibility. This architecture has helped RAN with new opportunities for virtualization and cost reduction. V-RAN simplified the management of RAN devices and deployment. Although C-RAN and V-RAN are both cost-effective and readily available for any changes to service requirements, they lack the openness to maximize the benefit of virtualization. To overcome the limitations associated with both C-RAN and V-RAN, O-RAN accommodates the baseband unit (BBU) and remote radio unit (RRU) software and hardware from different vendors [29]. O-RAN is open hardware with an operator-defined RAN architecture that provides intelligent radio control for future cellular networks. The O-RAN intelligent controller (RIC) disaggregates the control, and data planes of the RAN to provide flexibility and programmability [30]. O-RAN standardizes the control plane using open infrastructure and provides programmability by exploiting SDN and NFV principles. Intelligence is the fundamental building block of wireless networks beyond 5G [31, 32]. The main objective of O-RAN is to incorporate open interfaces and intelligence in RAN through virtualized network elements to enhance the RAN performance. O-RAN has influenced the market by presenting new opportunities for operators to achieve core network independence by manipulating network protocols and topology [33]. This will allow operators to leverage their core services by integrating the latest technological supports and next-generation communication protocols and frameworks.

- **Network Slicing**: As a result of network softwarization (NFV/SDN), network slicing is considered an ultimate solution to transform the existing wireless networks into software-based network slicing that can operate on a physical network infrastructure [34]. Network slicing provides the dynamic creation of logically isolated networking and service-customized solutions using a shared infrastructure. 3GPP defined the network slicing for basic principles concerning RAN operations in [35]. SDN and NFV help in slicing the RAN architecture, which is an emerging research direction towards virtualization, cloudification and centralization of RAN resources beyond 5G mobile networks. Network slicing is part of the 5G networks that allows multiple instances of pre-defined slices for services, including eMBB, mMTC, vehicle-to-everything (V2X), etc. Currently, the most popular network slicing in 5G cellular networks is scheduling the physical resource blocks (PRBs) at RAN. Mobile operators can efficiently slice the entire infrastructure based on user requirements and industry demands. Network slicing in future wireless networks is essential, especially in open radio networks, where intelligence exists. The network slicing in the intelligent controller enforces policy-based resource management with near-real-time functions to achieve the desired performance gain. Therefore, network slicing is one of the necessary enablers and performance indicators in future networks that can monitor and manage the QoS of different services, including mission-critical, safety, etc. [36].

- **Artificial Intelligence or Machine Learning**: Artificial intelligence (AI) is a progressive branch of computer science that deals with automation across various fields, and ML is an application of AI. ML is applied to developing systems to learn from patterns and data without explicitly being programmed [37]. The 6G architecture requires adapting an ML technique to optimize and create intelligence for the network. ML applications have been successfully utilized for network analysis, online customer support, search engines, computer vision, and signal processing applications [38]. As a result, research studies based on ML on various aspects enable the network paradigm to access and adapt a suitable ML algorithm for the appropriate task. Future wireless networks depend on AI to simultaneously generate policies, detect network anomalies, traffic management, etc. This new concept of an AI-enabled network is known as a self-driving network, which helps adjust and tune the system. ML empowers machines to learn patterns and optimize them as much as possible, which is an integral part of self-driving networks. ML is the most critical enabler of 5G and beyond networks to allow intelligence into the network architecture.

The enabling technologies provide insights and show the capability to create an intelligent network. These technologies are geared towards the goal of constructing a highly advanced and intelligent cellular network infrastructure. With their programmable and self-tunable capabilities, these technologies provide a robust platform to achieve the desired network performance.

2.3.3 Network Challenges

Despite the numerous advantages of SDN in 5G, several challenges need to be addressed to fully leverage its potential in future generations of cellular communication networks. These challenges include optimal placement of the SDN controller [39, 40], efficient retrieval of network topology and link-state information [41], traffic engineering and orchestration [42], ensuring quality of service (QoS) requirements in heterogeneous networks [43], and exploring AI/ML techniques for network automation and management [44].

Along with the many benefits NFV offers, network operators face several common challenges when deploying network virtualization applications. Virtualization has shown abnormal latency and significant throughput instability [45]. Therefore, making sure the NFV performance will at least be as good as the network without virtualization is one of the main problems. Another significant problem associated with network virtualization is how to smoothly migrate from the existing network infrastructure to virtualized-based network functions. Moreover, NFV faces some critical challenges in practice, including an increase in the number of virtual networks, resource orchestration in shared infrastructure, management complexity, several requirements from different tenants, and many more [46].

One of the main building blocks of 6G is intelligence, and as discussed before, intelligence comes from AI/ML-enabled techniques. However, the integration of AI with edge networks faces many challenges because the application services running in 6G must operate in real-time. AI embedded systems in 6G must meet services and applications' requirements. 6G is expected to support self-driving vehicles, deployment of the fourth industrial revolution (Industry 4.0), smart city/building services, augmented reality/virtual reality (AR/VR) services, and more [47, 48]. For 6G to support all these services, new technological advancements and software updates are needed (e.g., O-RAN, SDN, NFV, ML/AI, etc.).

2.4 Wireless Network Architecture

Traditional cellular architecture suffer from complexity, lack of flexibility, proprietary and expensive equipment. However, the idea of running network applications in a logically centralized controller shows the potential to address these challenges. SDN architecture enables programmability, more straightforward configurations, and network management. To go one step further in software-defined cellular networks, KDN introduced intelligence in addition to programmability and centralized control of SDN [49]. It is essential to understand the KDN paradigm because it illustrates the main building blocks of future wireless cellular networks.

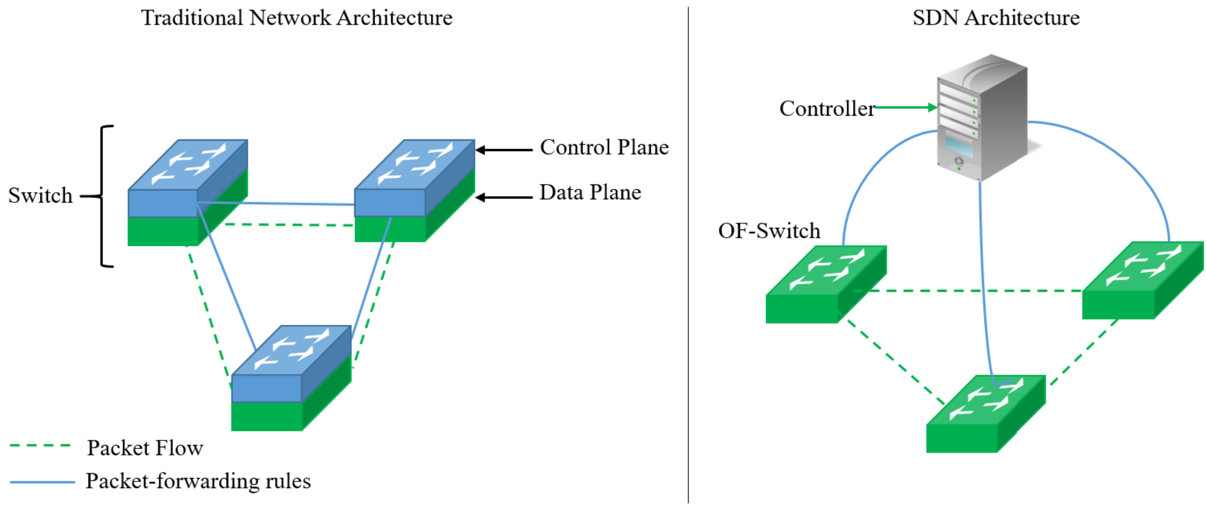


Figure 2.2: Software-defined networking versus the traditional network architecture.

2.4.1 Software-Defined Networking

Traditionally, network upgrades or configurations were performed manually as packet forwarding and routing decisions were executed by network devices [50]. In contrast to traditional networks shown in Figure 2.2, which have tightly coupled control and data plane equipment, SDN logically separates the control functions from network devices and simplifies them as forwarding elements. The forwarding instructions are determined by the centralized control plane, which has a global view of the network. This enables the dynamic allocation of network resources based on the application requirements. The control plane installs flow rules for the forwarding devices, and its main applications include the management of routing algorithms, load balancing, and access control. SDN is considered a critical enabler for current and future generations of wireless cellular networks, and its paradigm was developed in 2008 to address the problems associated with traditional networks. The programmability functionality of SDN in the network simplifies network management and enables innovation [51]. SDN played an essential role in 5G functionality and performance [52]. Based on its unique characteristics, it will be indispensable for the evolution of beyond 5G and 6G mobile networks. The SDN controller interacts between the switches via the application programming interface (API) protocol. The two most recognized protocols are OpenFlow (OF) and the programming protocol-independent packet processor language abbreviated as P4. SDN paradigm enables innovation and deployment of new services by simple software updates. SDN paradigm provided many advantages across different networks, including wireless local-area networks (WLANs), mobile clouds, vehicular ad hoc networks (VANETs), air-space-ground integration networks, satellite networks, and unmanned aerial vehicle (UAV) networks [50, 53, 54].

The SDN platform has two main southbound protocols to communicate between network planes, OF and P4. With the arrival of OF, the term SDN was officially operational and used by the research community as early as 2009 [55]. However, it did not have much impact on networking vendors until 2011, when OF eliminated the configuration complexity and automated network management. The development of OF started in 2011, and the latest version was released in 2016. SDN operates with OF standards to manage large-scale networks to provide more straightforward configuration options. OF applies a fixed set of protocols to populate the rules in the data plane using the control plane. However, OF is understandable by a fraction of available hardware routers and switches [56]. Moreover, the OF does not control the switch's behavior of the supported protocols. It only provides a way to populate the tables in the switch. The current OF has specific and fixed protocol headers for forwarding a packet. This increased the complexity of the specifications without providing any flexibility for adding new headers. To solve the problems with OF, the programming protocol-independent packet processor P4 language was developed. P4 is a tool that reduces the complexity of OF. The necessity of P4 alongside (or operating separately without OF) OF for improving network functionality is promising [57, 58]. P4 was first introduced in 2014 to address the limitations of the data plane by providing flexibility in programming the data plane in network switches that support OF standards [59, 60]. The P4 high-level language for programmable protocol-independent packet processors was developed through the collaboration of Barefoot Networks, Intel, Stanford University, Princeton University, Google, and Microsoft [58]. P4 enables programmability of the data plane and allows switches to process the packet. Hence, vendors and enterprises will be able to develop their application-oriented software for a programmable switch chip, resulting in several benefits to the network, such as reducing the packet processing time, modifiable packet headers, and switch protocol independence. These programmable switch chips are based on a protocol-independent switch architecture (PISA) [61].

2.4.2 Knowledge-Defined Networking

The concept of knowledge was introduced by Clark *et al.* [12], where the authors proposed a new idea called knowledge plane (KP). As shown in Figure 2.3 KP is an additional plane over a network with inbuilt ML capabilities. The incorporation of KP in the network architecture can be referred to as knowledge-defined networking (KDN) [11], where knowledge is the processed network information using an ML algorithm. The fundamental building blocks in KDN are network telemetry, SDN, and ML. Network telemetry is network information, such as NetFlow data, sFlow data, queue occupancy, policy rules, and processing time. However, to fully automate the network, information such as hop latency, link utilization, packet drop, and queue congestion states are also required. This temporary data is available through in-band network telemetry (INT) or packet-level network states. The packet-level network state information can be collected using a new southbound domain-specific language called P4 [58], which allows the collection of information directly from the data plane.

With the new advancements in data plane elements, routers and switches are capable of computation and storage, which makes network monitoring and network telemetry accessible [62]. Network telemetry provides flow information, real-time packet information, and other critical packet-level data, as well as network state monitoring and organization with centralized network analytics. Hence, network telemetry and network analytics present a richer view of network performance metrics, providing an extra advantage over conventional network management techniques. The incorporation of SDN and network analytics provides essential elements required by the KP. However, the last piece of the puzzle to make the KDN fully functional is to integrate ML. ML uses network telemetry and historical data to process and finds valuable information about the network, where this information is stored as knowledge to improve network performance.

ML algorithms are generally classified as supervised learning (SL), unsupervised learning (UL), and reinforcement learning (RL) [63]. In SL, the learning agent learns using a dataset as an input vector and maps the inputs to the outputs based on the previous inputs with their provided outputs. The dataset is a collection of labeled samples, where each element is called a feature vector. In UL, the dataset collects unlabeled samples, where the learning agent tries to categorize the input. Finally, in RL, the machine continuously observes the environment to improve decision-making. This technique constantly provides updated policies based on environmental feedback. Each ML algorithm can assist in the different applications of wireless networks. For instance, in SL, the KP learns the behavior of the network variables, such as network configuration and traffic load, which will enable the system to increase the network performance once the features are fed to the algorithm. UL assists the network operator by following the correlations in the data. For example, ML may predict the user's mobility effect on a communication link. Moreover, in RL, the learning algorithm will discover the best action that leads to an optimal configuration in a network. RL determines the target policy and adapts to the environment. Based on the target strategy, RL can be integrated into the KDN architecture to perform optimal actions [64]. As a result of network softwarization, network telemetry, and integration of ML in the KP, intivitatives such as O-RAN have emerged. KDN can overcome the drawbacks associated with conventional resource management, mobility management, networking, and localization [11].

2.4.3 Knowledge-Defined Networking Architecture

The concept of KDN is to add one more plane to the traditional two planes of SDN, which incorporates SDN, data analytics, and ML. The KDN paradigm has several advantages: first, it has a global view of the network, and second, it enables telemetry data to be collected by the management plane to transform the data into knowledge via ML. The knowledge will later turn into decisions by nodes to achieve efficient network operations [65, 66]. The benefit of having the KDN over traditional networks is that it automatically operates based on the knowledge obtained from the network. Figure 2.3 illustrates the KDN architecture, including the data

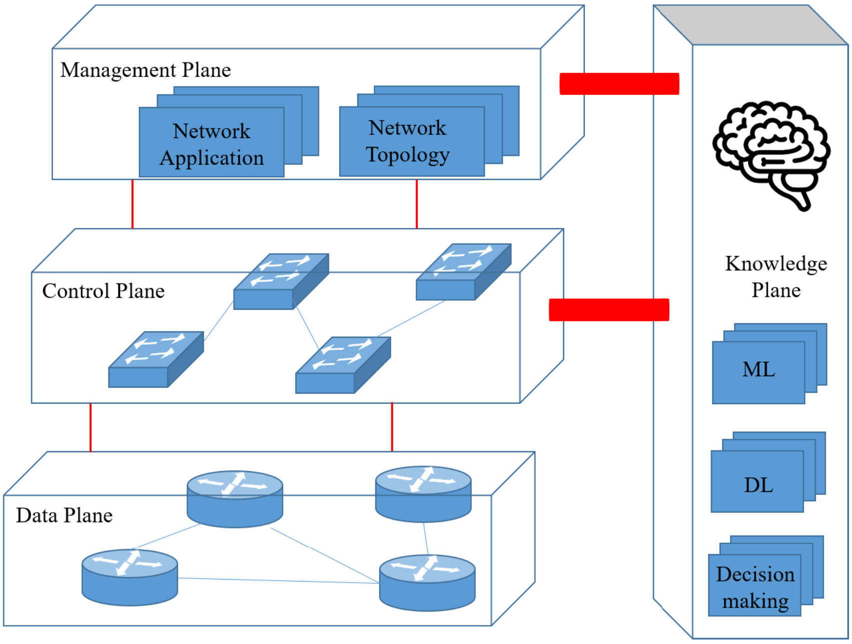


Figure 2.3: Knowledge-defined networking architecture.

plane, control plane, management plane, and knowledge plane.

The data plane in KDN is responsible for forwarding, dropping, processing, and packet modification. This layer works precisely like the data plane in SDN, where it consists of physical and virtual device elements. This layer operates unaware of the rest of the network and relies on the instructions and control rules coming from other planes.

The control plane exchanges information and updates the data plane processing and matching strategy rules. The logically centralized controller exchanges data and updates policies using a southbound application programming interface (API). The controller gathers the data and network state from the data plane and updates the flow tables to perform actions. In KDN, the data are also utilized to allow the KP to know which appropriate action is required. Then, the controller receives an action from the KP and updates the flow tables accordingly. These actions are usually used for forwarding and routing packets while the data plane is populated.

The management plane facilitates network topologies, support services, and configuration of the network devices. This layer must ensure that the network operates fully with maximum performance. This functionality of the network and the responsibility of monitoring the data plane and network analytics is also the job of the management plane. KP can influence management plane decisions when a new strategy or policy is available.

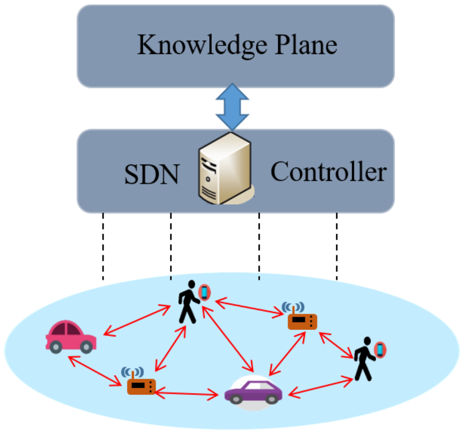
KP is the brain of the architecture and is responsible for modeling network behavior and decision-making. The decisions are made for different network applications, including resource management, networking configurations, mobility management, and localization. ML algorithms create knowledge in this layer, and new policies are established. For instance, data layer information, such as routing policies and parameters, can be obtained and fed to an RL algorithm for further processing to find the optimum Hello packet interval. ”Hello packet” - a special packet transmitted and used to determine network adjacency. Therefore, using an ML algorithm at the KP enables the operation of network protocols to be enhanced.

A number of studies have introduced the KDN architecture. For instance, the authors of [67] restated the concept of the KP in the context of SDN architecture in addition to the two planes of the SDN paradigm. As can be seen in their network architecture, the KP is located on top of the control and management planes. The integration of the KP generates behavioral models and reasoning processes for decision-making. This architecture enables the KP to have a full view and control of the network through the control and management plane. Other research studies [19, 68, 69] have a similar architecture for KDN. In [68], the same KP is utilized on top of all the layers, but it uses a cross-layer management and monitoring plane with ML algorithms to manage the rest of the planes. The proposed method utilizes an ML-based algorithm in both separate orchestration layers and embedded in the management plane. Therefore, it is important to investigate the architecture of KDN in wireless networks to identify the most suitable architecture in terms of flexibility and performance.

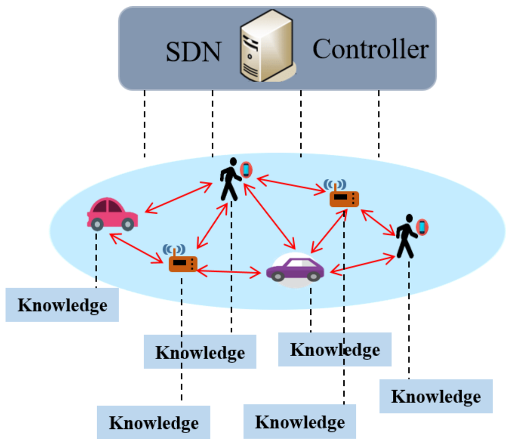
The architecture of the KDN in wireless networks can be centralized, distributed, or hybrid. In a centralized architecture, as shown in Figure 2.4a, the controller is located at the center of the network and collects information from nodes. The information is processed through knowledge plane, and then instructions and rules are transmitted to nodes. The new rules are updated using both direct and indirect approaches. The direct approach uses the previously processed information and sends new strategies immediately back to the user equipment (UE). The indirect approach uses an ML algorithm to determine new rules before being sent to UEs. In the distributed architecture illustrated in Figure 2.4b, the individual devices maintain local knowledge. Each node collects data from the environment and its surroundings and then independently applies a greedy-based ML strategy to acquire knowledge. The greedy strategy may be determined from prior knowledge (such as transfer learning (TL)) or using new ML-based optimization algorithms. For instance, in a routing scheme, a node can collect information from other nodes and use ML-based approaches to find the best route. In the hybrid architecture, as depicted in Figure 2.4c, knowledge is maintained at both the extreme edge and core with updated or synchronized knowledge. Both the controller and devices act intelligently together based on the information they collect. This information is processed by ML algorithms to acquire knowledge and injects new rules into the system. The hybrid approach combines the greedy strategy and centralized knowledge to increase the network performance. Further, there can also be a switching strategy to switch between centralized and distributed according to the application.

2.5 Heterogeneous Device-to-Device networks

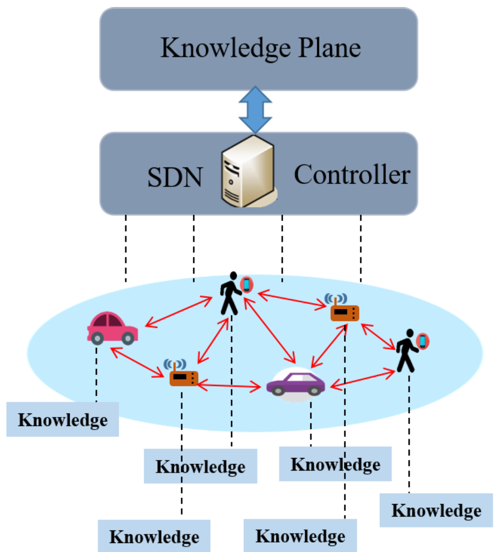
Due to the popularity of different applications, such as self-driving cars, IoT devices, smartphones, and recently unmanned aerial vehicles (UAVs), researchers and network providers began introducing new technique and solutions for network traffic management [70]. Network



(a) Centralized architecture of KDN.



(b) Distributed architecture of KDN.



(c) Hybrid architecture of KDN.

Figure 2.4: Proposed KDN architectures.

providers started to increase the number of BS by deploying small cells (BSs with short coverage areas) [71]. However, this solution is costly and makes network management more difficult. Consequently, a new feature was presented to distinguish various applications into different networks. This feature was called a wireless ad hoc network, a decentralized type of wireless network [72]. In ad hoc networks, nodes request and receive information from neighboring nodes. Although ad hoc networks can decrease cellular network congestion and traffic, they lack flexibility and coverage. Moreover, they were not designed to be integrated into cellular networks. For this reason, D2D communication was introduced [73]. In a D2D network, devices may communicate directly with or without the aid of the cellular network.

2.5.1 Device-to-Device (D2D) Communication

Device-to-device (D2D) communication was first proposed to extend the coverage at the edge network where nodes cannot communicate to the BS due to low signal strength [74]. In general, D2D can expand the network coverage and capacity and offers less E2E delay, energy consumption, and cellular overhead [75]. D2D communication is claimed to be an essential part of 5G networks [76]. The 3GPP standardized D2D in public safety, emergency services, V2X communication, autonomous ships, and IoT [77]. In these applications, the data must be exchanged directly between nodes using D2D to avoid data loss and late data delivery. However, various challenges, such as security protocols, mobility management, and resource management, need to be addressed. One of the main challenges in D2D communication is identifying an optimal route. Without an efficient routing framework or protocol, vital information can be delayed or, in worst-case scenarios, lost. Several research studies in D2D communication have proposed new practical algorithms to improve routing performance. Nevertheless, many believe that future routing protocols must facilitate ML for adaptation and security reasons.

2.5.2 Multi-hop Device-to-Device (MD2D) Communication

D2D communication lacks flexibility and scalability because it can only communicate with nodes in one hop neighbor. Therefore, to increase the network throughput and expand the D2D coverage [78], MD2D communication was introduced [79]. MD2D is an ongoing investigation by 3GPP because it carries many challenges [80]. Among the challenges, an adaptive and intelligent routing protocol is essential for MD2D. In current wireless networks, devices work with high-frequency signals to increase the data rate. However, the consequence of using high signal frequency is the growth of signal attenuation. In the case of static networks, we can predict the attenuation and prescribe a routing technique, but in mobile networks, attenuation can cause a significant deterioration in performance. Therefore, an efficient routing protocol is crucial for deploying MD2D in future cellular networks. If the routing protocol in MD2D is designed carefully, it can have significant benefits in IoT, mobile, and vehicular applications. For instance, in agriculture farms where IoT devices are scattered across a large area, instead of

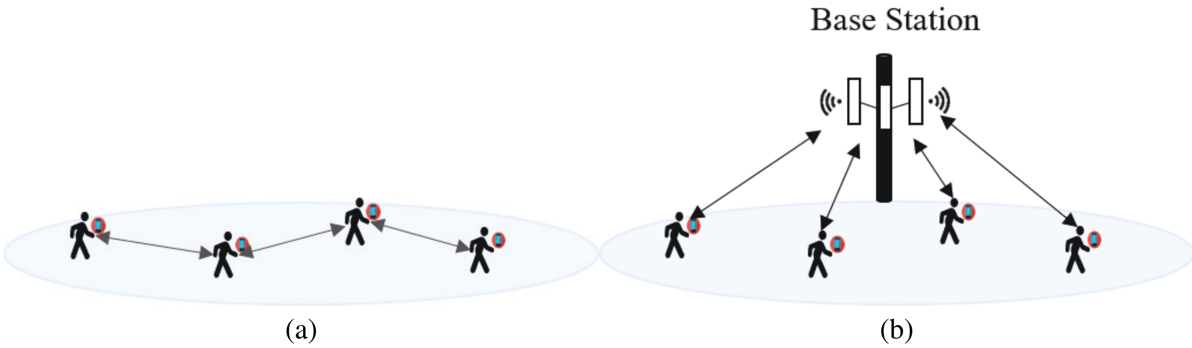


Figure 2.5: a) Illustrate the decentralized MD2D framework, and b) shows the centralized framework.

deploying relay centers to collect data, MD2D can use devices as relay nodes to collect the data. In mobile networks, where social media and the Internet are the primary cellular network traffic, MD2D can be used to share social media content. Moreover, self-driving cars can receive road information from other vehicles and roadside units to make safety and traffic decisions.

2.6 MD2D Routing Frameworks and Protocols

Routing protocols are used in heterogeneous ad hoc networks to enable communication between two devices. Routing protocols discover the appropriate path from the source node to the destination and maintain the route in the network. The performance of routing protocols depends on various factors, including E2E delay, packet delivery ratio (PDR), power consumption, and throughput. MD2D network is an integral part of future wireless networks to enable devices in close proximity to communicate and acquire information without the intervention of a centralized controller. To benefit from the MD2D networks, the first thing that needs to be designed carefully is the framework. MD2D frameworks are the architectural infrastructure or the spine of the structure. Routing protocols are built on top of the framework. Three frameworks exists for MD2D networks, as stated below:

- **Decentralized Frameworks**: As shown in Figure 2.5a, decentralized frameworks act without the intervention of a BS or a centralized controller. In this framework, nodes act distributedly and must manage the routing procedure and resource allocation by themselves. The advantage of decentralized frameworks is very low cellular overhead. The disadvantages are lack of agility, difficulty of management, security and scalability.
- **Centralized Frameworks**: As illustrated in Figure 2.5b, a centralized controller like the BS is responsible for managing the network. In this framework, the BS must provide appropriate network information to the nodes. Nodes receive rules and policies from the BS, and they act accordingly. The disadvantage of this framework is the additional cellular overhead. The advantages are ease of management, optimal resource utilization, improved security and adaptability of the framework to the network environment. In this framework, the controller has up-to-date knowledge of the entire network and can make

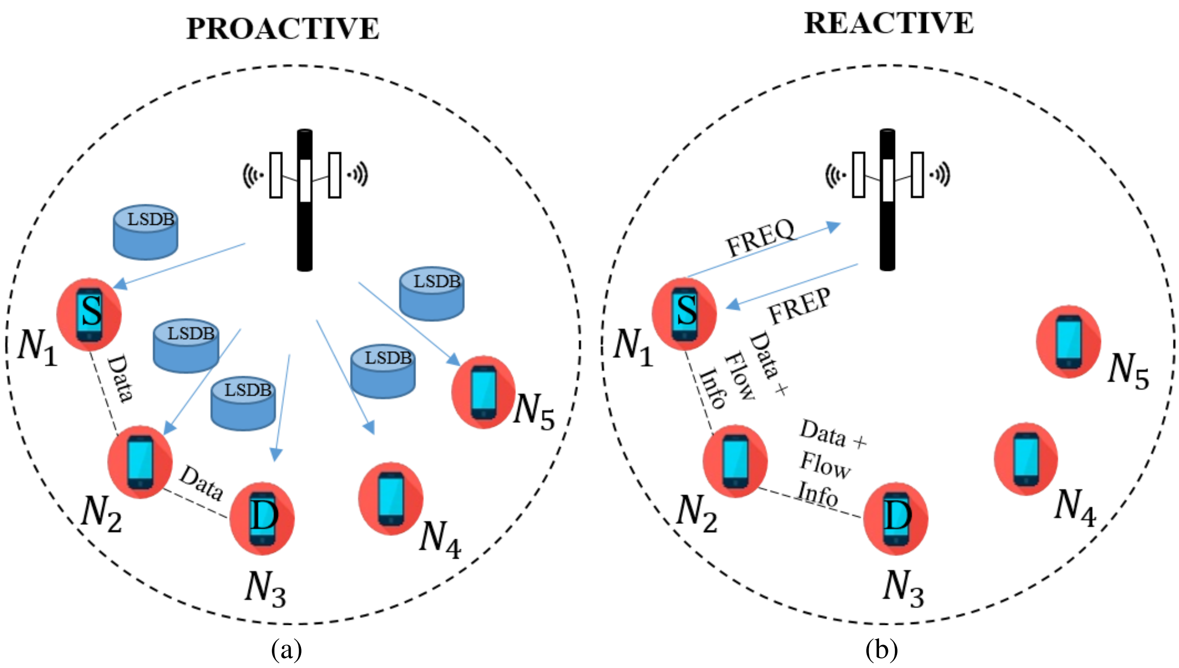


Figure 2.6: a) Proactive routing protocol, and b) reactive routing protocol.

the right decisions in highly dynamic environments. Centralized frameworks proved to perform better than decentralized [3].

- **Hybrid Framework**: The combination of the decentralized and centralized architecture is a hybrid framework. In the hybrid framework, the controller provides the necessary information to the nodes, where nodes are responsible for actively processing the data in a decentralized manner and evaluating a route. For instance, frequent observation of the controller is not needed in static and semi-static IoT networks. Therefore, nodes can obtain new information and update their policies using decentralized protocols.

In order to develop an effective routing protocol, it is crucial to first design a comprehensive framework. This framework serves as a guide for the architectural regulations of the network. Upon the establishment of the framework, the next step is to introduce a routing protocol to determine the optimal route. Routing protocols can be broadly categorized into proactive, reactive, and hybrid. Detailed descriptions of these categories will be provided in the following paragraphs. This thesis will concentrate on centralized and hybrid frameworks. There have been several proposed routing protocols for MD2D communication, each with its own advantages and disadvantages, as reported in the following studies [81, 82]. Some of the prominent routing protocols for MD2D include:

- **Proactive Routing Protocols**: These protocols use a table-based approach, where a routing table is maintained at each device that contains information about the routes to other devices. As shown in Figure 2.6a, nodes receive the complete network information in proactive routing protocols [83]. Where this information helps them to find routes in a distributive manner, there are various proactive routing protocols proposed for mobile networks [84, 85]. These protocols use the proactive routing strategy but different methods and policies to collect data and update routing information. The advantages of the proactive routing protocol are the fast response to the network dynamics (low E2E delay) and low packet failure. The disadvantages are high cellular channel overhead and energy consumption.

- **Reactive Routing Protocols**: These protocols use a request-reply approach, where a device sends a request for a route when it needs to communicate with another device. In contrast to proactive, where the centralized controller provides the whole network topology, here, nodes request a route on-demand [86]. Therefore, if a node has a packet to transmit, it demands a route to the destination from the controller. Then, based on the network criteria and optimization algorithms, the controller obtains the path to the destination and provides it to the network. As shown in Figure 2.6b, the source node sends a flow request (FREQ), and the controller forwards a flow reply (FREP) to the source node with routing information. The advantages are low cellular overhead and lower energy consumption. Compared to the proactive routing protocols, the disadvantages are higher chance of packet failure and E2E delay.

- **Hybrid Routing Protocols**: These protocols combine elements of proactive and reactive routing protocols. The best feature of both proactive and reactive routing protocols is integrated to achieve high levels of scalability [82]. The hybrid approaches are cluster-based, zone-based, or tree-based. In cluster-based methods, nodes are divided into clusters with a cluster head that collects the data proactively, and nodes in that cluster use reactive techniques to acquire route information from the cluster head. In zone-based approaches, the network is physically separated into multiple zones, where nodes in each zone are connected proactively, and to access nodes in other zones, they request a route from the BS. In tree-based techniques, the network is defined as a tree where nodes are connected with branches to the trunk (the gateway). Each branch represents a zone, and nodes can communicate in a proactive or reactive approach, depending on the application.

- **Geographic Routing Protocols**: Geographic routing protocols are a class of routing protocols that use the geographic location of devices to determine routes for MD2D communication [87]. These protocols are based on the assumption that devices have some form of location awareness, such as GPS or a location service provided by the network. It is important to note that Geographic Routing Protocols can suffer from limited reliability in dynamic environments, where mobility and changes in network topology can affect the accuracy of location information. However, geographic routing protocols can handle large numbers of devices in a network, making them well-suited for dense deployment scenarios. Moreover, geographic routing protocols can reduce energy consumption by using the location of devices to minimize the number of control packets exchanged and the distance that packets need to be transmitted.

- **Utility-based Routing Protocols**: These protocols are routing protocols that use utility functions to determine the best path for data transmission [88]. Utility-based routing protocols use different metrics such as energy consumption, throughput, delay, and reliability to evaluate the quality of the different paths and choose the one that maximizes the overall utility. Utility-based MD2D routing protocols can be adapted to different network scenarios and changing conditions by adjusting the weighting of different metrics. These routing protocols can optimize multiple objectives at the same time, such as energy efficiency, E2E delay, PDR and throughput.

We focus on improving various routing protocols to optimize MD2D networks. The approach and framework selected for the implementation of the routing protocol will be contingent on the specific application and performance requirements. The objective is to create an efficient MD2D routing protocol for the next generation of wireless cellular networks. Subsequently, this thesis delves into a thorough investigation of routing protocols in MD2D networks.

2.6.1 Review of MD2D Routing Protocols

The research study on routing protocols in MD2D networks has increased exponentially with the emergence of 5G. It started with upgrading the conventional D2D routing protocols to MD2D with higher performance. For instance, authors in [89] proposed a reactive MD2D routing protocol for 5G networks to enable nodes to exchange data at the edge network. They modify the conventional dynamic source routing (DSR) algorithm and use a hybrid framework with 5G features to make faster routing decisions. The proposed protocol reduces the WiFi overhead by decreasing the number of control messages exchanged between the devices. Although the proposed protocol shows better PDR and energy consumption, DSR lacks scalability in highly dynamic networks due to long route maintenance time. Similarly, other authors tried to improve and extend the existing routing protocols into multi-hop routing [90, 91]. For instance, authors in [90] used the quality of the links to improve the performance of optimized link state routing (OLSR), and in [91], destination sequenced distance vector (DSDV) is modified to allow particular nodes create routing tables. Furthermore, authors in [92] proposed a new restricted broadcasting mechanism of route requests for ad hoc on-demand distance vector (AODV) to improve the routing performance. They specifically restrict the problematic nodes and links during the path identification process. However, modifying conventional routing protocols for D2D and ad hoc networks cannot advance beyond a certain point. Therefore, new routing frameworks with advanced routing mechanisms must be developed to fulfill the requirements of future wireless networks.

One of the main advantages of MD2D is caching strategies in 5G, where the user content, such as popular videos and viral content, can be stored at the edge network and shared among several neighboring devices. Mobile networks can fully use the catching strategies with appropriate routing protocols to enable mobile nodes to share information and communicate in a decentralized manner. There are various routing protocols and frameworks proposed for mobile nodes. For instance, authors in [86] proposed a centralized reactive routing protocol called virtual ad hoc routing protocol (VARP) for mobile devices. An SDN controller is used to extend cellular services by adapting MD2D routing. If a node has a packet to transmit, it requests a route from the SDN controller. The SDN controller collects all the adequate network information before the node's request to process and obtain the most efficient path. Once the flow request (FREQ) is received, the controller checks the packet criteria to see if the MD2D routing can deliver sufficient QoS. Then it will authorize MD2D communication by giving the source node the route information. The proposed protocol increases the network lifetime, PDR, and throughput. Authors in [93] used a similar approach, but instead of only providing the source node with the routing information, the involved relay nodes will also receive the routing information. They present a new SDN-based MD2D routing protocol (SMDRP) to reduce the E2E delay and increase the PDR in highly dynamic networks. For a routing protocol to be reliable, it must show high performance even in situations with high mobility rates. In [83] used a hybrid routing protocol with a centralized SDN controller. The SDN controller collects all the network information and creates a link state database (LSDB). The network information consists of Hello packets established by nodes to confirm adjacency relationships, including the list of neighboring nodes with their link quality. The SDN controller uses this data to create a table that includes all nodes with their neighbors and associated link qualities. This table is transmitted to all nodes, where nodes independently find a route to any destination (by using the shortest path first (SPF) or Dijkstra's algorithm). The advantage of this table is when a relay node can not relay the information, the source node can immediately find a new path to the destination. Therefore, the proposed protocol provides low E2E delay and high PDR.

Another application of MD2D protocols can be in VANETs. VANETs are generally classified into two types of communication, vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I). Vehicular communication to roadside units (RSUs) and other vehicles enables diverse applications related to safety, traffic efficiency and management, autonomous driving, and emergency services. To have these application work efficiently, routing protocols for VANETs must be designed carefully to provide the highest performance. Authors in [94] proposed an intelligent greedy position-based multi-hop routing using fuzzy logic techniques. The overall goal of the proposed method is to provide efficient and reliable data transmission links in V2V communication. The best next-hop node routing of a packet is calculated using the fuzzy logic system in a centralized framework. Simulation results show improvement in PDR compared to similar algorithms. In [95], the authors presented an inter-vehicle distance-based location-aware MD2D routing to enhance the vehicle's connectivity. The proposed protocol uses a decentralized framework with a reactive-based routing protocol. Their algorithm predicts the next location of the nodes to identify the next best hop for forwarding a packet. A geometry-based localization technique is used to obtain the inter-vehicle distance (the distance to the relay node) if the location information of the nodes is not received from the global positioning system (GPS). This location-based system can handle the location information of vehicles even in highly congested areas. The simulation results show the significance of the proposed protocol compared to the existing VANET routing protocols. Authors in [96] used a centralized position-based routing protocol to effectively analyze the geographical positions of the vehicles in the network and provide updated routing information. A centralized RSU acts proactively and provides vehicles with the next hop vehicle. For security preserving purposes, a secure identity validation is used in messages for V2V and V2I communications. The simulation result shows that the introduced routing protocol has better safety features and reduces the possibility of attacks in the network.

IoT networks are another application of MD2D protocols. IoT networks present a huge portion of future networks due to the constant improvement of their applications. This application can be seen in industry, agriculture, and smart cities. IoT enabled the emergence of smart sensors, Internet technologies, radio-frequency identification (RFID), and communication protocols. Current IoT sensor nodes are treated as smart devices which are widely used to gather and forward sensed information. As a result of the current advancements and improvements in this field, there has been significant effort to find the most efficient routing protocol. In [97], an energy-efficient multi-hop routing protocol is proposed. The proposed protocol shares the body-sensing data from various sensors deployed on a human body to a sink node using distributed multi-hop flooding. They use a fuzzy logic system to choose the node with the highest residual energy in the route. The advantages of the proposed protocol are shown in the simulation results using the throughput and overall network energy consumption. In [98], the authors used an energy-aware clustering-based multi-hop routing algorithm to guarantee balanced energy consumption throughout the network. They have proposed a dynamic clustering algorithm to re-form clusters if required. The k-mean and open source development model algorithm (ODMA) are utilized for clustering. A genetic algorithm is also employed for multi-hop route identification. The proposed protocol shows significant improvement in PDR and energy consumption. Similarly, in [99], the network is split into clusters, and each cluster has nodes with different responsibilities. In every iteration of the network, every node undergoes a procedure to decide on its responsibility. The responsibilities are being the cluster head, independent nodes, and member nodes. Therefore, the energy load will be distributed among all the network nodes and members. In IoT networks, energy consumption contributes significantly to network performance, and most researchers concentrate on this aspect. However, security issues are as important as energy. IoT sensors sense the data and share it with nearby sensors or transmit it to a central hub. New secure routing protocols must be designed to preserve the data from any threats. In [100], a new energy-aware and secure multi-hop routing (ESMR) protocol were proposed. The algorithm uses a secret sharing technique to improve the security against malicious actions while keeping energy consumption low. To create a secure routing protocol, they have clustered the network into zones based on the node location. In each zone, a cluster head is responsible for forwarding the data to the sink node using a secure sharing scheme. The simulation presents a significant increase in network lifetime with a very low E2E delay.

2.6.2 Challenges of MD2D Routing Protocols

There are number of challenges in the current routing protocols and we intended to answer some of them in this thesis to assist in the future deployment of MD2D networks. These challenges are summarized as follows:

1. **Lack of Intelligence**: Networks typically use a routing protocol without considering current or historical networking conditions. However, to fully capture the network dynamics, recording general networking data during network operation hours is essential to create an intelligent processing unit. Then, an ML algorithm can process the data and provide optimized information. For instance, for every particular time (rush hours, night, etc.), an ML algorithm can record and identify the most efficient routing protocol and keep track of the network. Hence, different routing protocols can be prescribed based on the network intelligence for MD2D networks to maximize performance. To solve this problem, we proposed a new routing protocol using various LSDBs. The controller uses the application and network layers information and creates a master LSDB (MLSDB) to keep the network's intelligence. Based on the user application, the network prescribes different routing information.

2. **Network Lifetime**: In most of the proposed MD2D routing protocols, all nodes in the network are actively participating in the network. Therefore, if a node has a low battery or doesn't want to participate in MD2D communication, it is forced to relay other nodes' traffic. As a result, the network lifetime dramatically decreases, damaging the network coverage and PDR. To solve this problem, we proposed joint mobile node participation and the MD2D routing protocol to identify nodes in the network with low battery levels and exclude them from routing.

3. **Lack of Adaptation**: The current MD2D routing protocols lack self-organization and self-tuning. For instance, most routing protocols find a route using one particular network metric, and that metric is usually the shortest distance. Suppose a node is regularly used as a relay node due to the shortest path algorithm. This node might quickly deplete its battery and be removed from the network. Therefore, the routing protocol unintentionally causes battery depletion leading to packet drop and deterioration of network lifetime. As a result, routing protocol should be designed in an adaptive manner to adjust to the network. Another example could be the route parameter adjustment based on the user application and requirements. To provide insights and deliver a solution, we have proposed a dynamic virtual slicing mechanism to adapt the route based on the user application.

2.6.3 Integration of Machine Learning with MD2D

Owing to the exponential growth of data traffic, wireless networks will require advanced technical solutions in the near future. In particular, challenges, including the imbalanced distribution of traffic loads among BSs and wireless channel dynamics, need to be addressed. To overcome some of the issues, the knowledge acquired by ML algorithms can assist networks in building intelligence and automation. A summary of the automated routing techniques that can potentially be used in MD2D routing protocols is presented. Table 2.1 summarizes the studies surveyed in this section.

i) *Knowledge derived from supervised learning*: Sharma *et al.* [101] proposed a routing protocol for efficient routing in opportunistic networks called MLProph, which uses a decision tree and neural network (NN) algorithms to compute the node's successful delivery probability. The algorithm uses past network routing data to calculate the probability of whether the data will be delivered to the destination by the relay nodes. The probability value helps to decide on the next-hop selection. The ML method trains itself based on network characteristics, such as node density, buffer capacity, hop count, node energy, mobility rate, and the number of successful deliveries. Here, the NN is trained to determine whether the forwarded message has been delivered; it can have two outputs p=1 for successful deliveries and p=0 otherwise. The NN is trained iteratively by setting initial values, and subsequently, it provides optimal predictions for successful and unsuccessful deliveries. The simulation results indicate that the proposed algorithm performs much better than previous works in terms of overhead ratio, average latency, and packet delivery ratio. They also compared the two ML models and found that the NN performs better than the decision tree in terms of overhead, delivery probabilities, and packet loss.

   To improve the traditional routing strategies and increase the performance of the wireless backbone, the authors of [102] proposed an intelligent routing scheme using a deep convolutional neural networks (CNN). The proposed method learns from the previous experience based on congestion and uses this information to train a two-phase procedure, namely a cold start period and an intelligent running period. The cold start period is the initialization of the training set, where the algorithm only defines a route with a minimum hop path. After this period, the algorithm switches to the intelligent running period, which performs real-time updating and routing judgments. More importantly, a CNN is constructed for each routing decision, which takes the collected information based on traffic patterns from routers, including traffic generation rate, to predict whether the selected routing strategy can cause congestion in the network. This process is periodically updated until it is predicted that the chosen route will cause no congestion. Simulation results prove that the proposed algorithm performs much better in terms of E2E delay and packet loss ratio than conventional routing strategies, where there is no intelligence. The proposed method can provide real-time intelligence for traffic control.

   In [103], a supervised deep neural network was proposed for routing optimization in heterogeneous networks to predict the path from the source to the destination node. Each router in the network uses a deep neural network (DNN) to predict the next hop; the

DNN takes traffic patterns as inputs, and based on these inputs, it generates the desired output. The output of the deep learning structure significantly improved the network traffic management. There are three phases to obtaining a fully functional ML. The first phase is the initial phase, where the traditional routing protocols, such as OSPF, provide the network route, and the network starts to operate. At the same time the second phase is the training phase to train the deep learning system from the collected information based on the traditional operating system. Finally, the running phase is the stage in which the machine is thoroughly trained and can provide real-time routing strategies. This method has been proven to have higher throughput and less overhead than OSPF. The proposed study suggests a greedy-based distributed architecture over a knowledge-based network to increase the throughput.

ii) *Knowledge derived from unsupervised learning*: In [104], the authors focused on load balance routing based on principal component analysis (PCA) and NN for dimension reduction and prediction of the network load status. The algorithm integrates ML and data analytics into the SDN architecture to obtain intelligence from the network. The use of these algorithms has led to efficient and intelligent routing decisions. This article aims to address the shortcomings associated with the next generation of wireless mobile networks, such as video streaming and online gaming, to mitigate the delay caused by traffic. The proposed routing strategy is based on an ML scheme, where PCA was used to reduce the dimension of the vector-matrix by applying it to the original adjacency matrix of the network topology. Based on the normalization, they designed a queue-utilization routing algorithm for routing prediction. Moreover, routers were continuously updated based on neighbors' information to select the routers with more resources. In this vein, authors explored the current SDN architecture and represented an ML algorithm to predict routes.

Owing to the fixed network architecture of some routing protocols and the massive volume of data traffic exchanged between devices, the authors of [105] introduced a context-aware routing protocol named KROp. This protocol uses several network features to make routing decisions based on the network conditions. KROp uses the K-mean clustering algorithm and exploits network features to select the best next hop. This algorithm is based on the knowledge acquired from the node's behavior to identify a cluster of the best forwarders. The numerical results show superior performance in KROp in terms of dropped packets, overhead, and average hop count compared to other routing strategies, such as history-based prediction routing (HBPR) [106], and probabilistic routing protocol using the history of encounters and transitivity (PRoPHET) [107].

Tang *et al.* [108] proposed a centralized routing scheme with mobility prediction (CRS-MP) for VANETs. The proposed method utilizes an SDN controller with an artificial neural network (ANN) to gather information and predict the user's arrival rate. Based on the arrival rate of each vehicle, RSUs or BSs can model statistical traffic patterns and estimate traffic mobility. Intelligence was also used in this study by integrating the CRS-MP model at the RSU/BS to predict the mobility patterns of vehicles and find vehicle connections. The ANN takes input according to the number of arrival vehicles at different instances. Based on the initial random weights, predicts the vehicle arrival rates. The numerical results of the CRS-MP scheme outperform other vehicular routing protocols, such as V2I and V2V communication, in terms of overall vehicular service delay. Furthermore, the proposed algorithm is independent of the mobility rate, making it more robust to high mobility rates. The proposed routing protocol utilizes multi-hop routing in vehicular systems using an SDN controller that solves the overload on the BS.

iii) *Knowledge derived from reinforcement learning:* With the massive growth of IoT devices connected to the edge network, the design of routing strategies is complicated. In particular, in smart cities, routing is significantly more difficult owing to the distribution of the crowd and network congestion. The authors of [109] designed a deep reinforcement learning (DRL) algorithm for smart routing decisions for load balancing and mitigating network congestion when massive crowds are moving around the city for daily activities. They adopted a DRL agent to use the NN and generate Q-values directly. First, the SDN controller collects the network state information. Then, the DRL agent makes an action (routing decisions) based on the current state, and finally, the agent receives a reward. The objective of the reward function is to maximize the successful service access rate, minimize the data transmission delay, and balance the network load. The algorithm performance was better than the OSPF and enhanced- OSPF (EOSPF).

Cognitive radio network (CRN) has attracted considerable attention owing to their importance in future wireless communication systems. This technology overcomes the channel spectrum's scarcity by allowing secondary or unlicensed users to benefit from underutilized licensed channels. However, the dynamic nature of CRNs makes routing a complicated task. The authors of [110] proposed a clustering mechanism or cluster-based routing to boost network scalability and functionality. Once the cluster heads are identified in the network, each cluster head estimates the Q-value of each neighboring node. The routing table is constructed based on the Q-values, and the largest Q-value is the next chosen node for the next hop. During the learning procedure, the state of the network represents the destination node, and the decision to select the next hop is the action. Finally, the system's reward is the throughput resulting from the chosen hop. In this study, the knowledge is derived from each state and action pair, which provides an appropriate action for the next instant.

In [111], the authors studied three route selection schemes in a real testbed environment to improve the performance of multi-CR networks. One of the schemes is based on spectrum leasing, and the other two are based on RL. Spectrum leasing is a new term used for communication between unlicensed and licensed users in CR networks. The two RL algorithms are based on Q-learning to predict the next-hop neighbor. Similar to other studies [110], the next hop is selected based on the highest Q-value. The state action is the destination node and the selected next-hop node for the source node to transmit the data. The reward is the channel-state information. The proposed routing scheme was compared with the highest-channel protocol in a multi-hop network and has shown better performance.

To manage the network overheads in highly mobile scenarios, the authors of [112] proposed the mobility, residual energy, and link quality-aware multipath (MRLAM) for routing decisions. To do this, they used a Q-learning algorithm to select the optimal route with energy-efficient nodes. The proposed routing scheme aims to maintain a reliable and stable network during a particular timeframe. They have successfully improved several metrics, including energy cost, end-to-end latency, convergence time, and packet loss ratio, when compared with other routing techniques, including the multipath optimized link state routing (MP-OLSR) protocol [113] and the extension version of MP-OLSR known as MP-OLSRv2 [114]. Both the proposed protocols in [112, 115] act as a distributed network, where each node decides on the next hop.

The authors of [116] added intelligence to the network to mitigate the complexity of network topologies. They integrated both centralized and distributed network functionality to guarantee high QoS. The hybrid approach uses AI routers for distributed intelligence and a processing unit for centralized intelligence. AI routers are responsible for hop-by-hop Internet protocol (IP) routing to ease network overhead. The processing unit acts as a global controller for connection-oriented tunneling-based routing protocols, including segment routing and multiprotocol labeled switching. A DRL-based routing strategy is deployed in the processing unit for tunneling-based routing to ensure QoS. The state of the RL algorithm is represented by the network traffic characteristic information and device information, and the action is the forwarding path. The reward it acquires is the effectiveness of action to the delay requirements and optimization targets. The proposed RL method converges to the global minimum, and the routing strategy performs better than other routing strategies in congested areas. The proposed scheme has a centralized architecture with an intelligent control plane. This plane controls and installs the rules in the network. Stampa *et al.* proposed a DRL algorithm for optimizing routing in a centralized knowledge plane [117]. The actor-critic learning method is used, where the state of the learning algorithm is calculated by the traffic matrix (defined by the bandwidth request between source and destination pairs), and the action is the path taken to transmit data (obtained using link weights). Finally, the reward of the algorithm is based on the average network delay. The proposed method provides operational advantages compared to traditional optimization algorithms for routing strategies.

2.7 Thesis Aim and Focus

While 5G cellular networks currently offer fast and reliable communication for network users, the rapid proliferation of smart devices and user applications has led to the need for an upgrade to 6G. This next-generation network promises to enable intelligent and adaptive architecture with capabilities for network application development. This means that new policies and protocols can be specifically designed to address the challenges presented by the high volume of traffic and resultant network congestion. The high volume of traffic will cause the performance to degrade. Integrating MD2D with 6G is a potential solution to decrease the network congestion on the BS and provide satisfactory performance.

MD2D communication involves direct or indirect communication between two devices. When two devices are within transmission range, a direct link is established to exchange data, while relay nodes are used to relay traffic when devices are out of range. MD2D enables ad-hoc communication without the need for a fixed infrastructure, creating a secondary or WiFi interface for devices to communicate. However, ad-hoc networks can lack flexibility and scalability due to distributed policy management and route establishment, which can lead to increased power consumption and slower packet delivery. To address these limitations, integrating a centralized controller is essential for MD2D networks. This controller can acquire network information and prescribe routing strategies, while also maintaining the ad-hoc nature of MD2D during disasters when the controller is unavailable. The controller's role is to create the most efficient policies and routing strategies when it is online, ensuring that MD2D performance is similar to or better than cellular communication. This thesis aims to provide adaptable and efficient routing frameworks and protocols for future wireless networks to assist in constructing the most efficient routing protocol for MD2D networks.

In this thesis, we identified the key enabling technologies in future wireless cellular networks and studied them carefully to introduce new routing protocols that suit the technologies. This thesis explores the benefits of integrating MD2D communication with future wireless networks and investigates how MD2D routing protocols could be developed to enable traffic offloading. A routing protocol should be designed very carefully to maximize the efficiency of the MD2D networks. We used SDN and NFV concepts to propose centralized routing protocols for MD2D networks. Then, we exploited knowledge-based algorithms and proposed a routing protocol. Later, we introduce new automated network slicing techniques to help self-manage and self-tune the routing protocol in MD2D networks. The future open RAN-based networks allow us to create a secondary infrastructure such as MD2D and deploy automated routing protocols to enhance coverage and capacity.

2.8 Summary

All the generations of cellular networks have faced some common problems, including the growing demands of user applications and the exponential increase of user devices. The latest generation of cellular network, 5G lacks flexible architecture due to standardization and proprietary hardware, protocols, and interfaces. As a result, dynamic network changes due to the growth of user applications require network changes and the development of new services. The literature has proposed to transfer the current network architecture to an open infrastructure where MD2D can be a solution to increase the network performance. One of the main advantages of MD2D is the ability to offload network traffic from the cellular BSs, which can help to reduce congestion and improve overall network performance. Additionally, MD2D can enable more efficient use of spectrum resources and support the development of new network applications, such as enhanced mobile broadband and the Internet of Things. When it comes to MD2D protocols, there are several different approaches that have been proposed for enabling efficient and effective MD2D communication in beyond 5G and 6G networks. Some examples include geographic routing, utility-based and proactive routing protocols.

In conclusion, our literature review has highlighted the importance of MD2D routing protocols in future wireless cellular networks. Section 2.3 presented an overview of future cellular network generations and their enabling technologies. The focus of this thesis is then introduced, which centers on addressing the challenges and opportunities presented by these technologies. Section 2.4 introduced a new network paradigm that aims to bring intelligence and self-management to the network architecture. In Section 2.5, we discussed heterogeneous D2D networks and introduce MD2D as a potential solution to their limitations, and Section 2.6 demonstrated different routing protocols for MD2D networks. As such, the primary objective of this thesis is to investigate and propose new routing frameworks and protocols for MD2D networks in order to enhance the performance and capabilities of future cellular networks.

Table 2.1: Knowledge-based strategies for MD2D routing.

Article	Knowledge Objective	Architecture	ML Algorithm	Deliverable	Limitations	Conclusion
[101]	Enabling efficient route selection	Centralized	Decision tree and neural networks	Decreasing average latency, improving overhead and packet delivery ratio	Overfitting of the dataset	Highly efficient route selection based on various network parameters
[102]	Real-time updating and routing judgments in heterogeneous networks	Centralized	Supervised learning and deep CNN	Minimize the average delay and improves packet loss ratio	Dependency on labeled data	Intelligent traffic control
[103]	Self-driving networks by learning traffic control mechanism in heterogeneous networks	Distributed	Supervised deep neural networks	Better signaling overhead, delay, and throughput compared to OSPF	Need for large training sets	Capable of learning complex patterns and functions to predict the least cost path
[104]	Efficient and intelligent route decisions in wireless mobile networks based on KDN architecture	Centralized	Principle component analysis and neural network	Lower packet loss ratio and acceptable throughput and E2E delay	Lack of continuity	Provides a dimension-reduction vector matrix to reduce the algorithm response time but it must be verified over larger networks
[105]	Trained algorithm can select next hop with a least total average hop count and successful delivery probability in wireless networks	Centralized	K-mean clustering	Less dropped packets and network overhead	highly sensitive to the initial selection of cluster centroids	Simple, efficient routing protocol but needs improvement in average message latency and should involve energy consumption as a node feature
[108]	Optimal route selection and prediction capability using global information for VANETs	Centralized	Unsupervised learning-based algorithm	Better transmission delay compared to existing VANET routing protocols	Difficulty handling high-dimensional data	The proposed scheme is robust to varying mobility rates
[109]	Smart routing decision in IoT-based smart cities	Centralized	Deep reinforcement learning	Mitigating the network congestion and load balancing	Challenging to balance between exploring new actions and exploiting the knowledge	Simultaneous QoE satisfaction and crowd management
[110]	Identifying stable routes in cognitive radio networks	Distributed	Q-learning	Minimizing the interference between SUs and PUs and less frequent route discovery	Time and memory requirements	Boosting network scalability and functionality
[111]	Route selection for multi-hop cognitive radio networks	Hybrid	Reinforcement learning	Improves the QoS	Difficult to deal with high-dimensional state and action spaces	Selecting the best possible route in terms of throughput and packet delivery ratio
[112]	Intelligent QoS-aware route selection	Distributed	Q-learning	Energy efficient, QoS-aware and mobility tolerance	Sensitive to sparse and delayed rewards	Reliable, stable and extended lifetime network
[116]	Intelligent traffic routing control with KDN approach for next generation of wireless networks	Centralized and hybrid	Deep reinforcement learning	Minimizing the overhead	Requires frequent interactions with the environment	The proposed paradigm combines distributed and centralized intelligence to achieve highest performance
[117]	Automatic adaptation based on the traffic conditions via KDN paradigm	Centralized, distributed, or hybrid	Deep reinforcement leaning	Traffic engineering	Designing appropriate reward functions is crucial	Near optimal solution after one single training step

3

Performance Analysis of Multi-hop Routing Protocols in SDN-based Wireless Networks

3.1 Overview

The upcoming research is published in the Elsevier Journal of Computers and Electrical Engineering [1]. This chapter's main objective is to thoroughly investigate three different centralized MD2D routing protocols. This will help us realize the advantages of each routing protocol in environments with varying mobility rates and densities. Therefore, it provides insights into how to efficiently design a routing framework and protocol for different networks with distinct characteristics.

3.2 Introduction

Current 4G networks are no longer capable of providing users demands for better quality-of-service (QoS) and high-speed access to various multi-media applications. This is due to constraints, such as limited spectrum, limited scalability, proprietary interfaces, device-centric architecture, complex protocols, expensive infrastructure and existence of heterogeneous networks [75]. 5G wireless networks have been proposed to address some of the above chal-

[1] Ashtari, S., Abdollahi, M., Abolhasan, M., Shariati, N., Lipman, J. (2022). Performance analysis of multi-hop routing protocols in SDN-based wireless networks. Computers Electrical Engineering, 97, 107393.

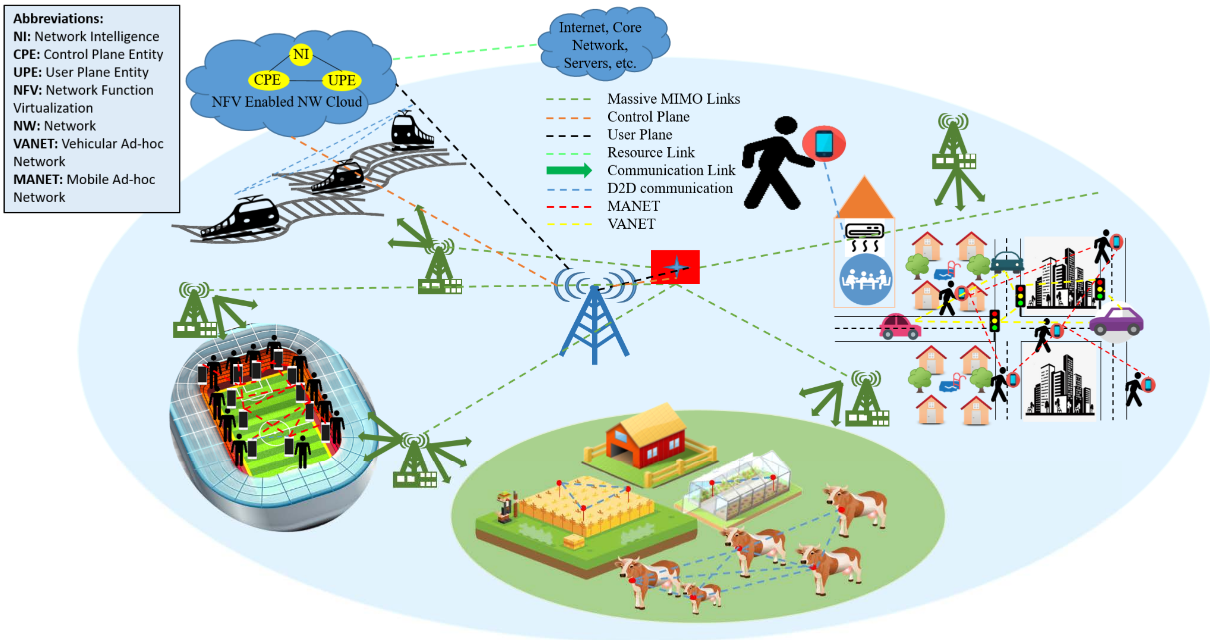


Figure 3.1: Multi-hop routing underlaying SDN-based cellular network.

lenges [118,119]. 5G networks are expected to provide anytime-anywhere network connectivity with significant improvement in network performance compared to current 4G networks, such as 1000 times more network capacity, 10-100 times higher data rates, more than 90% energy saving, decreased end-to-end delay to less than 1ms and 99.999% service availability [120,121]. To meet the expectations and to fulfill the presumptions and challenges of 5G networks, new technologies and solutions are being developed [122]. For instance, massive antenna configurations and millimeter wave technologies are two promising features of 5G. Utilizing massive antenna or multiple-input multiple-output (MIMO) technology provides robust, secure and energy efficient communications. Additionally, millimeter waves (mmWaves) and higher frequency range (3-300GHz) provide higher data rate to satisfy the 5G requirements. Although these two technologies increase the capacity of 5G and quality-of-experience (QoE), there are however other promising technologies, such as MD2D could be added to 5G and beyond to meet the ambitious goals set for future wireless networks and growing demands from mobile users [123].

Integrating MD2D communications with cellular networks has a number of benefits, such as cellular coverage expansion, increased QoS and QoE, reduced communication delay, improved load balancing by offloading cellular traffic, enhanced resource and power allocation, and improved spectral efficiency [124]. Further, it is essential for disaster recovery applications when cellular infrastructure is not available. To integrate MD2D communications with cellular networks, a number of different solutions have been proposed. One of the solutions is to establish mobile ad-hoc networks (MANETs) in cell edges or congested areas to allow devices to form one-hop or multi-hop connections to reach to the base station (BS) or other nodes in close proximity [76].

MANETs consist of a number of nodes or devices that are connected wirelessly in a self-configured and self-organized manner without the help of any fixed infrastructure. Devices in MANETs move freely in the network and form multi-hop links to provide network connectivity. Each node acts as a router to forward traffic to one or more specific targets. However, due to a lack of central management in MANETs, these networks are not scalable as the number of hops and participating nodes in the MANET grow. Hence, integrating the best features of MANETs and cellular devices can improve the total performance and will eliminate their limitations in the new hybrid structure.

Regardless of achieved improvements, MANETs under cellular networks are faced with a number of challenges: peer discovery techniques, selecting the communication mode (device-to-device (D2D) or cellular) to get the maximum throughput, efficient radio resource allocations, interference management in in-band communications where D2D and cellular communications are within the same cellular spectrum, security concerns and pricing policies for relaying devices [125]. Several techniques [126, 127] are proposed to minimize interference, optimize resource allocation and spectral usage in D2D underlaid cellular networks. Despite the proposed solutions and techniques for cellular networks, there are still a number of constraints which are worth mentioning, such as absence of virtualization, high cost of network upgrade, lack of flexibility, complexity in service deployment, tight coupling between the control plane and data plane, lack of fine-grained control over resources in cellular network and scalability issues due to centralized data plane functions in the long-term evolution (LTE) core entity [128, 129]. One approach to address the above shortcomings is through the integration of software-defined networking (SDN) concept in cellular networks. Hence, SDN aims to address such challenges by introducing a new programmable framework with open interfaces wherein control plane functions are removed from forwarding devices and are logically centralized in one or more control entities. Network functions are defined as applications run on top of the controller that make the network upgrade and network innovation faster and more flexible via software update rather than hardware or infrastructure upgrade that is costly and time consuming.

In SDN framework, control functions are taken from the forwarding devices and are logically centralized in one or more control units called SDN controller. The network devices are simple forwarding elements that receive instructions from the controller. Unlike traditional networks that have distributed network management, in SDN, the controller has a global and real-time knowledge of the network via periodic network status collection. Network functions such as load balancing, resource management, traffic monitoring, security, routing and topology discovery are applications running on top of the controller. Further, SDN provides open application programming interfaces (APIs), namely southbound and northbound interfaces that enable the controller to communicate with the forwarding devices and application layer, respectively [130, 131].

Employing SDN in cellular LTE networks has been studied by many researches [132–134]. The research covers two main components of cellular architecture, namely core and edge. These studies aim to make the core or edge parts programmable by using the SDN paradigm. Our study focuses specifically on the edge part of the network and investigates different SDN-based solutions for multi-hop routing under cellular networks. Figure 3.1 depicts different applications of MD2D in wireless networks.

Authors in [135] propose a fast and low-overhead routing protocol called low-overhead D2D (LODR) for MD2D under SDN-based cellular system. Their architecture consists of a cellular network with SDN capabilities and MD2D communication, which utilizes an OpenFlow controller to manage the forwarding strategies. For each packet transmission between the nodes, the controller installs flow rules on the associated nodes. The proposed technique decreases control overhead and improves communication performance in MD2D communication. In [136], a hierarchical D2D communication architecture to offload traffic from the BS is presented. In the architecture proposed in [136], there is an SDN controller that manages MD2D communications. The authors believe that their proposed architecture could improve the overall energy and spectral efficiency of the network that is suitable for disaster recovery scenarios. In [137], authors introduce a communication layer for hybrid wireless networks called WASP[2] to manage mobile devices and to inform each device to handle their own traffic in heterogeneous networks. In [137], the proposed architecture consists of a centralized controller with OpenFlow protocol that decides on communication protocols between the mobile nodes. The controller calculates the route and updates each mobile device as required for data transmission between devices. Furthermore, the controller is used for content transmission between the Internet and mobile devices, which reduces overhead and offload traffic in cellular communication. This study shows scalable and energy efficient communications compared to the existing ad-hoc protocols, such as ad-hoc on-demand distance vector (AODV) and optimized link state routing protocol (OLSR). Authors in [138], designed a framework for mobile devices in disaster zones called FINDER (Finding Isolated Nodes using D2D for Emergency Response). In the proposed framework, there is a global SDN controller that detects the failed BS and instructs other active nearby BSs to extend the cellular coverage. In the case of disaster, the isolated nodes change their operation mode from cellular to D2D and form a multi-hop connection to the mobile nodes in the edge of nearby BSs or WiFi access points (APs). Ant Colony Optimization-based routing is used for energy efficient routing in D2D communication. Simulation results indicate that the proposed framework improves the network lifetime.

The conducted survey studies confirm that in SDN-based cellular networks, the scalability, energy and spectral efficiency are improved while network management is simplified. Further, routing strategies compared to the traditional ad-hoc routing protocols achieve better performance due to global knowledge of the controller and its immediate response to the network

[2]WASP is a prototype designed by authors in [137], which consists of Wi-Fi, Ad-Hoc, SDN, and Personal-Mobile technologies.

topology changes.

In our previous work, we designed an SDN-based cellular framework and proposed three different protocols, namely HSAW [83], SMDRP [93] and VARP-S [86]. The proposed protocols use proactive and reactive approaches to discover the routes in the network and are detailed in Section 3.3. The simulation results indicate that the proposed protocols significantly improve the network overhead compared to the traditional ad-hoc networks. Further, the reactive approach performs much better in densely populated areas compared to the proactive approach in terms of routing overhead and energy consumption. However, the performance of the proposed protocols have not been thoroughly investigated for different routing metrics and various network applications. To this end, a detailed performance comparison is performed by this study and also network parameters, such as routing overhead, number of dropped packets, average end-to-end delay, and overall power consumption are evaluated for these three routing protocols. This performance analysis enables a comprehensive investigation to realize which routing protocol is appropriate for which network condition. Therefore, the performance analysis provides us with oversight of possible options for the future integration of multi-hop routing in 5G and beyond wireless networks. The contributions of this chapter are as follows:

- Two different SDN-based routing approaches are compared, namely reactive and proactive. To this end, our previously designed routing protocols are thoroughly investigated, HSAW as a proactive protocol, VARP-S and SMDRP as reactive protocols.
- A system model compatible with 5G networks is designed. The mmWave and MIMO systems are deployed to model the cellular channel, to increase the channel capacity, and to enable faster transmission over the network. Our channel model uses directional antenna with beamforming approach to estimate and tune the channel for each user.
- The routing protocols are compared in distinct environments with different node densities and mobility rates in the designed system model. In each environment, the weakness and strength of each protocol is analyzed in terms of E2E delay, dropped packets, energy consumption and routing overhead.

The rest of this chapter is organized as follows. Section 3.3 gives a brief overview of the designed routing protocols. Section 3.4 describes the simulation tools and parameters modeled for our analysis. Section 3.5 presents the results of our simulations. Our conclusion and future research directions are drawn in Section 3.6.

3.3 Detailed Illustration of Routing Protocols

In this chapter, three SDN-based multi-hop routing protocols are analyzed and compared. In the routing framework, a SDN-based BS manages multi-hop communications between the mobile nodes under its coverage area. Two different frequency bands are utilized in the framework;

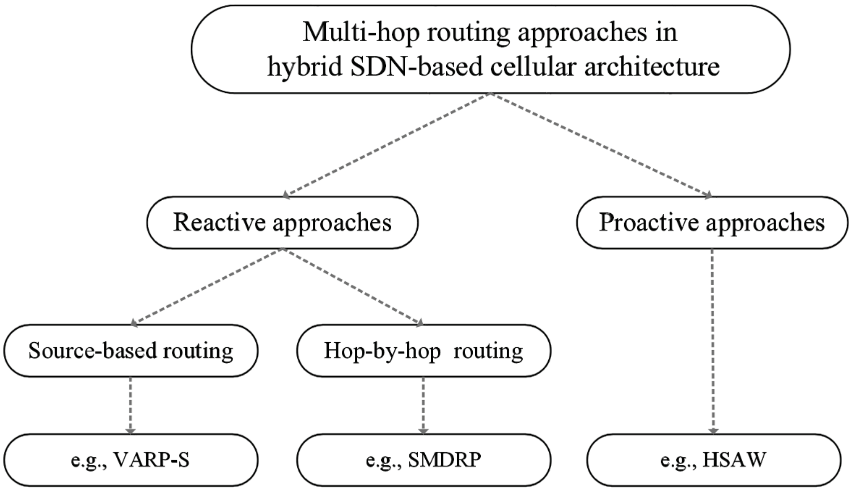


Figure 3.2: Multi-hop routing approaches in hybrid SDN-based cellular framework.

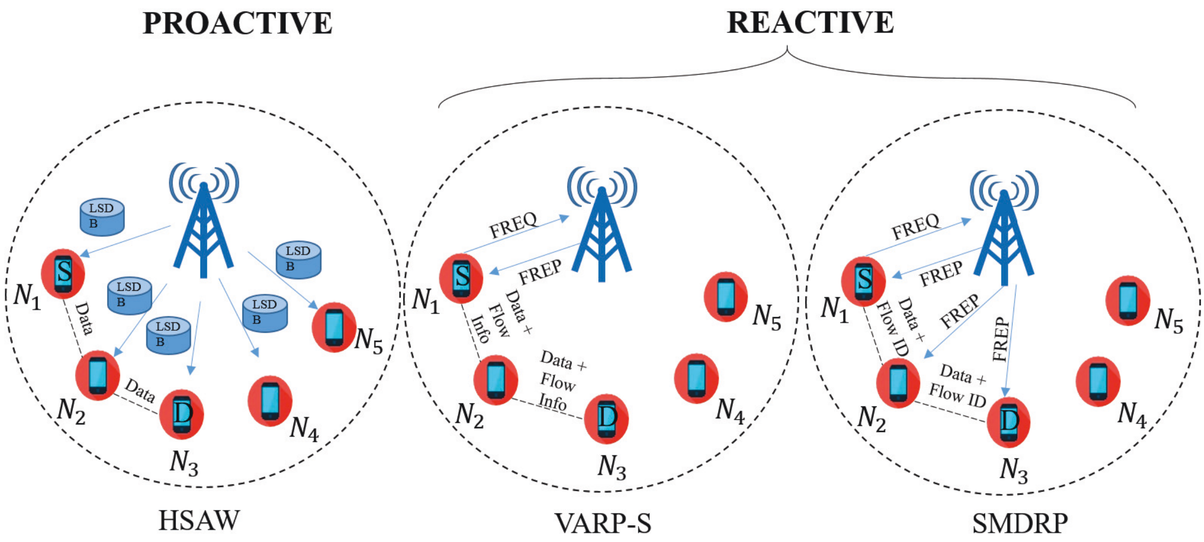


Figure 3.3: Demonstration of reactive and proactive routing approaches in hybrid SDN-based cellular framework.

licensed frequency for cellular and unlicensed frequency for D2D communication. The controller employs two different approaches to manage data forwarding as highlighted in Figure 3.2, namely proactive and reactive. In the proactive approach, the controller provides each node with the link state data base (LSDB) of the whole network, while in the reactive approach, the controller only provides forwarding information to active flows. The routing protocols for the mentioned approaches are HSAW for proactive, and VARP-S and SMDRP for the reactive approach. In all the mentioned protocols, the controller participates in the multi-hop routing. However, the level of participation is different from one to another as explained in the following subsections and also depicted in Figure 3.3.

3.3.1 Virtual Ad-hoc Routing Protocol-Source Based (VARP-S)

VARP-S is a source-based routing protocol in which the source node of the data packet requests for a route from the controller to obtain the forwarding information. Once a source node tries to send a packet to a specific destination node, if it cannot find any route in its flow table, then it sends a flow request to the controller. The controller decides which mode of communication should be selected (cellular or D2D) and sends back the forwarding information to the source node. The source node forwards data packets to the BS if the controller acknowledges the cellular transmission is a better option for that data traffic. Otherwise, the source node attaches the received multi-hop forwarding information, which is a complete source to destination address, and sends the packets to the next hop. Relay nodes forward the received packets based on the attached routing information in the packet header. In case of any link failure, the upstream node of the broken link sends an error message to the controller. Subsequently, the controller informs the source node about the failed link and updates routing information.

3.3.2 SDN-based Multi-hop Device-to-Device Routing Protocol (SMDRP)

SMDRP is hop-by-hop routing protocol, where the source node of a data packet forwards a flow request to the controller to find a route to the target. If the controller decides that MD2D is a better option compared to the cellular transmission for this data traffic, then the controller assigns a unique flow ID to the flow and informs the source node and selected relay nodes about the flow ID and the next hop address. Following that, the relay nodes store the flow ID and next hop address in their flow table. Unlike VARP-S, in SMDRP, the source node only attaches the flow ID to the data packets. Based on the attached flow ID, the relay nodes forward the packet to the next hop. In case of any link failure, the upstream node of the broken link sends an error message including the flow-ID to the controller. Then, the controller broadcasts the flow-ID to inform about the error that occurred in the flow. Any relay node that has this flow-ID in their flow table will change the status of the flow to invalid and wait for the updated routing information.

3.3.3 Hybrid SDN Architecture for Wireless Distributed Networks (HSAW)

In HSAW, each mobile node receives the LSDB of the whole network from the controller. Consequently, each node can decide on the packet forwarding by building its own routing table as they have a global knowledge of the network. If any link failure occurs in the flow, the upstream node of the broken link runs the Dijkstra's algorithm and finds the next available hop to the target. It also informs the controller about the error, and following that, the controller broadcasts the information of the failed links. Accordingly, all nodes in the network will update their LSDB.

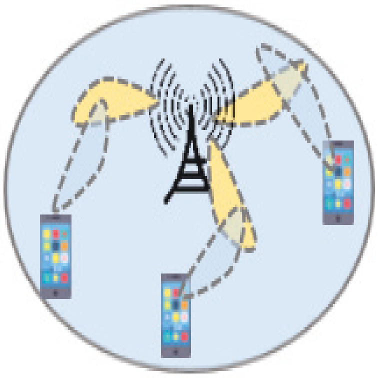


Figure 3.4: A mmWave and MIMO cellular system model, where BSs and MSs communicate via directional beamforming.

3.4 System Model

The main components used to build our system model is detailed in the following sections, including channel model, time delay model, energy consumption model, and network model.

3.4.1 Channel Model

In our system model, directional beamforming with a large array of antennas is adapted at both transmitter and receiver. Due to the exponential increase of users and their demands for higher data rate, utilizing directional precoding with large antenna arrays will increase the data rate that result in sufficient received signal strength. Figure 3.4 represents the MIMO directional beamforming in the cellular networks.

The BS and mobile nodes in the network are equipped with a MIMO-orthogonal frequency-division multiplexing (OFDM) cellular system as depicted in Figure 3.5, where TX and RX represent transmitter and receiver, respectively. For any single connection between the BS and a mobile node, N number of antennas and M number of RF chains are utilized. In our system model, the number of RF chains (N_{RF}) are less than the number of receiving (N_{RX}) and transmitting (N_{TX}) antennas [139–141]. Moreover, the number of antennas in mobile station (MS) (N_{MS}) are less than or equal to the number of antennas in BS (N_{BS}). Considering the fact that the use of RF chains will lead to power consumption and hardware constraints in mobile nodes, two different precoding algorithms is used, namely analog and digital, to address this issue.

To establish a link between the BS and mobile nodes, the proposed algorithm in [142] is employed. Based on the algorithm, the precoding matrix at the transmitter, referred to as (F_T), is made by the baseband precoder (F_{BP}) and analog phase shifters (F_{RF}). The transmitted signal by BS is calculated as below:

$$F_T = F_{RF} F_{BP}, \tag{3.1}$$

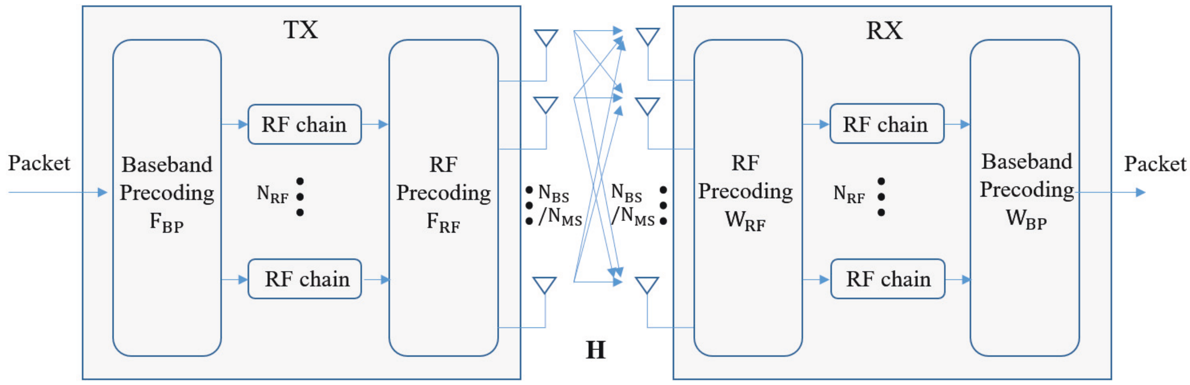


Figure 3.5: System diagram of BS-MS transceiver that uses baseband precoders, RF chains and precoders for both ends.

$$x = F_T p, \tag{3.2}$$

where x is the combined BS or MS precoding matrix (depending on downlink or uplink) with transmitted symbols, namely the discrete-time transmitted signal from a BS or MS. p is the $N_p \times 1$ vector of transmitted packet symbols, and F_T is the $N_{BS} \times N_p$ precoding matrix.

Further, a narrowband block-fading channel model is adapted to our system model in which the received signal r at the BS/MS will be observed as:

$$r = \mathbf{H}x + n, \tag{3.3}$$

where $\mathbf{H}$ is the channel matrix between the BS and mobile node and its size depend on the number of antennas at the transmitter and the receiver (e.g., in downlink, the channel matrix size is a dimension of N_{BS} $\times$N$_{MS}$), n is the noise attenuating the signal and is calculated from Gaussian model as $n \sim N(0, \delta^2)$.

Upon receiving the signal in the receiver antenna, there is a combiner, referred to as W_R, that acts as a precoding matrix at the receiver to reduce the channel effect. The combiner consists of two components, namely a phase shifter (W_{RF}) and a baseband combiner (W_{BP}) as shown below:

$$W_R = W_{RF} W_{BP}, \tag{3.4}$$

The process of combining the received signal in the combiner is shown with y and indicated as follows, where the notations A^H, A^T, A^{-1}, I are Hermitian (conjugate transpose), transpose, inverse and identity matrix throughout this chapter, respectively:

$$y = \mathbf{w_R^H} \mathbf{H} x + \mathbf{w_R^H} n, \tag{3.5}$$

Keep in mind that for uplink (signals from MS to BS), the precoding and combiner matrices will switch.

While it is assumed there are no interfering BSs, then, channel model can be formulated with L

scattered signals, where each of them represents a single propagation path between the TX and RX. Therefore, under these assumptions the channel $\mathbf{H}$ is expressed as:

$$\mathbf{H} = \sqrt{\frac{N_{TX}N_{RX}}{\rho}} \sum_{l=1}^{L} \alpha_l k_{RX}(\theta_l) k_{TX}^{H}(\phi_l) \mathbf{w_R^H H F_T} p + \mathbf{w_R^H} n, \tag{3.6}$$

where ρ is the average path loss between the transceiver and receiver, and α_l is the complex gain of the l^{th} scattered path, where each path amplitude is Rayleigh distributed. The variables k_{RX} and k_{TX} are antenna array response vectors for the receiver and transmitter, respectively. Moreover, for phase shifters, the azimuth angles of arrival or departure (AoAs/AoDs) for the l^{th} path varies between $\theta_l, \phi_l \sim [0, 2\pi]$. In this study, only the azimuth angles are examined which implies that the BS and mobile nodes only use horizontal (2D) beamforming. Considering only azimuth and eliminating elevation means that all the scattering takes place in azimuth. Finally, assuming uniform array of antennas at both sides, antenna array response vectors is written as:

$$k_{RX}(\theta_l) = \frac{1}{\sqrt{N_{RX}}}[1, \exp^{j\frac{2\pi}{\lambda}dsin(\theta_l)}, ..., \exp^{j(N_{MS}-1)\frac{2\pi}{\lambda}dsin(\theta_l)}]^T, \tag{3.7}$$

in the same fashion, the array response vector for TX is obtained, where λ is the wavelength of the transmitted signal, and d is the distance between antenna elements of BS and mobile nodes.

If we assume that both BS and MS have prior knowledge of the channel, it is possible to obtain the capacity of our channel by using the singular value decomposition (SVD) of the MIMO channel, as shown below:

$$\mathbf{H} = \mathbf{USV}^H, \tag{3.8}$$

where U with dimensions of $N_{RX} \times N_{RX}$ and V with dimensions of $N_{TX} \times N_{TX}$ are the unitary matrices, and S is the diagonal matrix with non-negative singular values of $\mathbf{H}$. The resulted matrices from SVD equation, V and U, are equivalent to the optimal values of precoding (F_T) and combiner matrices (W_R), respectively. Using SVD enables us to obtain the channel capacity or the channel rate R by:

$$R = \log_2 |I_{N_p} + \frac{P}{N_p} \mathbf{R_n^{-1} W_R^H H F_T F_T^H H^H W_R}|, \tag{3.9}$$

where I_{N_p} is the unitary matrix, P is the total transmission power, N_p is number of transmitted symbols and P/N_p is the normalized total transmitted power, and R_n is the noise covariance matrix.

The data rate of MIMO channel between the transmitter and receiver can be obtained by considering the achieved channel rates from (3.9), as follows:

$$DataRate = N_{TX} BWR, \tag{3.10}$$

where BW is the channel bandwidth in Hz.

Consequently, for each single node in the network, the transmission delay can be approximated by using the obtained channel data rate from (3.10).

$$TransmissionDelay(s) = \frac{DataSize(b)}{DataRate(b/s)}, \tag{3.11}$$

where the $DataSize$ is the overall packet size that should be transmitted.

3.4.2 Time Delay Model

In this chapter, there are two types of packets being transmitted across the network, namely data packets and control packets. Data packets consist of the information that a node requires to send to the destination. The control packets include the Hello packets for neighbor discovery and other routing information used by the routing protocols for their route discovery and route maintenance. In our routing framework, the size of control packets is different from one routing protocol to another. In HSAW, the control packets are overall LSDB of the network that is sent by the controller to each node. In VARP-S and SMDRP, the control packets include: 1) the route queries sent by the source node of the data packet to the controller 2) flow information sent by the controller to the source node in VARP-S and to the relay nodes in SMDRP 3) the routing information attached to the data packets, flow ID in SMDRP and full path in VARP-S. For route maintenance procedure when a link failure occurs in the network, in HSAW, all the nodes will be informed about the failure, while, in the two other protocols, only the source node (VARP-S and SMDRP) and relay nodes (SMDRP) will be notified.

For our study, it is assumed that each node in the network is represented by a unique ID, referred to as N_{ID}, and each request sent by source node of the flow to the controller is identified by a unique ID denoted by Req_{ID}. Consequently, in both reactive protocols, the size of route request (S_{RREQ}) is calculated as follows:

$$S_{RREQ} = Req_{ID} + N_{ID}, \tag{3.12}$$

After receiving the flow request, the BS will process the request and send back a flow reply containing the information of relay nodes to the requested target. The time taken by the controller to calculate the requested route is affected by the network traffic. This traffic depends on the data packets that are being processed by the BS and also the number of nodes in the network. Hence, for higher traffic, the process time of the request will be increased as each request has to wait in the queue for a longer time to be processed by the controller. This will lead to extra delays, unintentional energy usage and packet loss in the network. In our study, all the effecting parameters are taken into account. In VARP-S, the BS sends only the entire route to the source node via 5G radio. If the length of the route is L_R, then the size of flow reply in VARP-S, referred to as S_{VARP-S}, is calculated by:

$$S_{VARP-S} = L_R \times N_{ID}, \tag{3.13}$$

For SMDRP protocol, when a node requests for a route, the controller calculates the path, assigns a unique flow ID to the flow and forwards the forwarding information including the next-hop address + flow ID (F_{ID}) to the relays in the active path (source node and all the relay nodes between the source and target). Consequently, the overall size of flow reply transmitted by the BS in SMDRP, referred to as S_{SMDRP}, is calculated by:

$$S_{SMDRP} = (L_R - 1) \times (N_{ID} + F_{ID}), \tag{3.14}$$

It should be noted that MIMO technology and directional beamforming is used to send the forwarding information from the BS to the relay nodes. This will increase the data rate and reduce the network congestion.

For HSAW, nodes find the route to the packet's destination node by a different approach. In this approach, the BS broadcasts the whole network's LSDB to all nodes in the network via the 5G radio. Subsequently, each node builds its own LSDB. If a node has a packet to send, then it checks its routing table to find a route. If no route is found, then it runs Dijkstra's algorithm to find the least-cost path to the target. The size of the LSDB packet, referred to as S_{LSDB}, depends on the number of nodes in the network and the neighboring nodes in their proximity which is calculated by:

$$S_{LSDB} = \sum_{j=1}^{N} \sum_{i=1}^{N-1} (2 \times N_{ID}) \times L_{i,j}, \tag{3.15}$$

$L_{i,j}$ indicates the existence of a link between $node_i$ and $node_j$ if the following condition is met:

$$L_{i,j} = \begin{cases} 0, & if \;\; d_{ij} \geq min(R_i, R_j) \; or \;\; i == j; \\ 1, & if \;\; d_{ij} < min(R_i, R_j); \end{cases} \tag{3.16}$$

where $d_{i,j}$ is distance between node$_i$ and node$_j$, and R is the transmission range of nodes.

3.4.3 Energy Consumption Model

The initial power of nodes is represented as p_0, and the power consumption model described in [143] is used, where the power loss is assumed based on the free space power loss and multipath propagation loss. Therefore, the consumed power for node$_i$ to transmit or receive a packet P^{TX} is expressed below. The reason for choosing this energy model is to solely showcase the impact of energy consumption in the WiFi channel while considering various packet transmissions across different routing protocols. This approach enables us to conduct a thorough analysis of the differences between routing protocols, which is the primary objective of this chapter.

$$P^{TX}_{i\in N} = \begin{cases} P_e \times S_i + P_{fm} \times S_i \times d^4, & if\ d > a; \\ P_e \times S_i + P_{fs} \times S_i \times d^2, & if\ d < a; \end{cases} \tag{3.17}$$

whereas $a = \sqrt{\frac{p_{fs}}{P_{fm}}}$ is a threshold distance that determines whether multipath or free space model should be used to calculate transmit power, P_e is the power consumed by the electronic devices, P_{fm} and P_{fs} are the power loss by multipath fading model and free space, respectively. S_i is the packet size that node$_i$ ($i \in 1, 2, ..., N$) tries to send or receive and d is the distance between the sender and receiver. Moreover, the consumed power by electronic devices can be expressed as:

$$P_e = P_s + P_{da}, \tag{3.18}$$

where P_s and P_{da} are the sender and data aggregation power, respectively. The power consumption of each node for receiving a signal P^{RX} is calculated as follows:

$$P^{RX}_{i\in N} = S_i \times (P_e + P_{da}), \tag{3.19}$$

The total consumed power of N number of nodes in the network, referred to as P_T, is defined as below:

$$P_T = [\sum_{i=1}^{N} P_i^{TX} + \sum_{i=1}^{N} P_i^{RX}], \tag{3.20}$$

It is worth mentioning that the types of devices connecting to cellular networks has been changing to include more machine-to-machine (M2M) and the internet of things (IoT) based sensor devices and would be more heterogeneous than the homogeneous nodes used in this chapter. Consequently, the initial power and power specifications of those devices would be different from one another.

Table 3.1: Simulation parameters.

Parameters	Value
Simulation tool	MATLAB
Seeds	1000 times
Simulation area	500m $\times 500m$
Packet size	50Mbits
Protocols	HSAW, VARP-S and SMDRP
Number of nodes	100, 600, 1000
Network specification	sparse, semi-dense, dense
Mobile node Transmission range	75m
Mobile node speed	3 m/s and 20 m/s
Mobile node movement model	Random waypoint mobility
Max bandwidth	80MHz
Carrier frequency	30GHz
Pathloss constant	3
Max MS antennas	4
Max BS antennas	64
Wireless standard	IEEE802.11g

3.4.4 Network Model

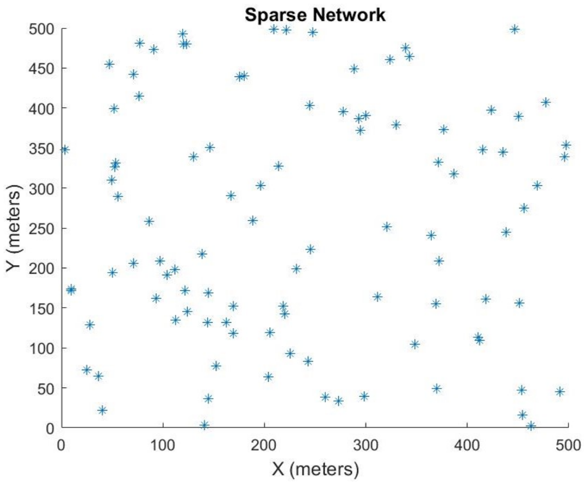
For our network model, two frequency bands are utilized, licensed (e.g., 5G) and unlicensed (e.g., WiFi) frequencies. The former is used to exchange control packets and also data packets between the BS and mobile nodes, downloading and uploading from or to the BS to connect to networks that are not reachable by MD2D or MD2D cannot provide the required QoS. The latter is used for multi-hop communications to exchange data traffic between devices and also for uploading or downloading when one-hop connection to BS does not provide the quality requirements.

The type of control packets exchanged between mobile devices differs from one protocol to another. However, in all the proposed protocols, the mobile nodes inform their existence to their one-hop neighbours by exchanging Hello packets. Further, a centralized controller, connected to the BS, manages the whole network by collecting each mobile's node information.

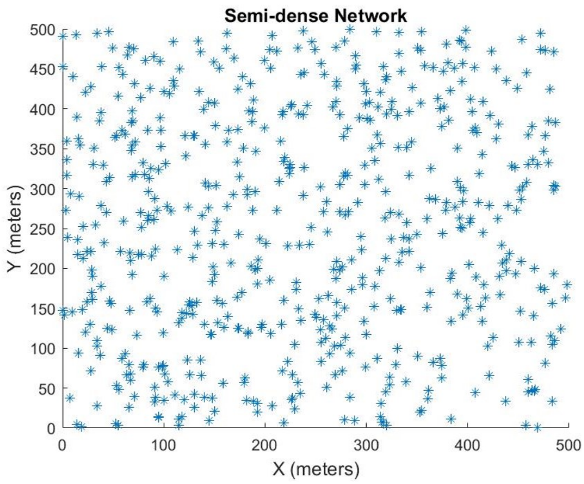
In this study, our previous three designed routing protocols are analyzed in three scenarios – 500×500 (m^2) – with various densities; sparse (<400 nodes), semi-dense (400∼800), and dense (>800) as illustrated in Figure 3.6. Node density gives us a very good insight into the protocols scalability as the number of nodes change. The performance of each protocol is evaluated in each scenarios separately considering two different average mobility rates for mobile nodes in each area, 3 m/s and 20 m/s, to model pedestrian and vehicular networks respectively. Therefore, six different scenarios are evaluated in our simulation. For each scenario, the performance of routing protocols are estimated in terms of the dropped packets, energy consumption, E2E delay, cellular (vertical link) overhead and WiFi overhead (horizontal link). The main parameters of our system model and implementation are illustrated in Table 3.1.

Table 3.2: Simulation scenarios.

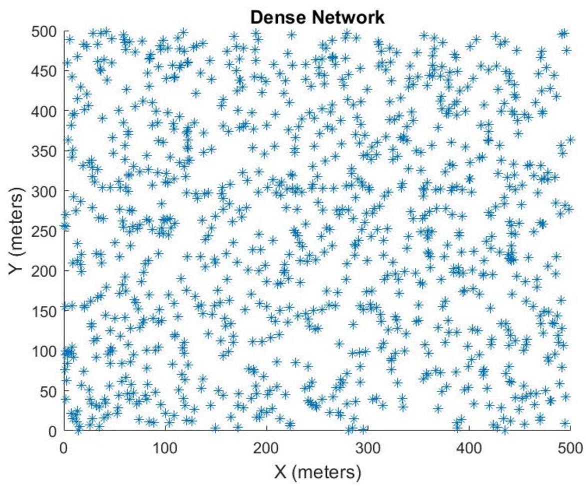
Scenarios	Density	No. of Nodes	No. of Active Flows	Mobility (m/s)
1	sparse	100	25,50,100,150,200	3
2	sparse	100	25,50,100,150,200	20
3	semi-dense	600	150,300,600,900,1200	3
4	semi-dense	600	150,300,600,900,1200	20
5	dense	1000	250,500,1000,1500,2000	3
6	dense	1000	250,500,1000,1500,2000	20



(a) node's density = 100



(b) node's density = 600



(c) node's density = 1000

Figure 3.6: Representation of the network with three different node density.

3.5 Simulation Results and Performance Analysis

All lifetime results are averaged over 1000 independent runs and reported with their 95% confidence intervals. In this section, simulation results and performance analysis of three routing protocols, namely HSAW, VARP-S and SMDRP, with different network densities and mobility rates are evaluated. The SDN-based BS is located in the center of the network and all the nodes are placed under coverage of the BS. Each mobile device is randomly generated from the continuous uniform distributions and moves in the network based on Random Waypoint mobility model. Further, each mobile node is equipped with two wireless interfaces; 5G interface for BS communications and WiFi interface with transmission range of $75m$ for D2D communications (WiFi communication) over IEEE 802.11g[3] standards. The considered propagation models for 5G and WiFi interfaces are Log-distance path loss and Friis free space [144], respectively. In the conducted simulations, different numbers of active flows are defined with randomly selected source and targets. In our simulation, for all the defined scenarios illustrated in Table 3.2, the performance of each routing protocol is studied.

Here is a brief explanation of the network performance metrics used by this study to evaluate each routing protocol:

- Packet loss: When the network starts to operate, for reactive protocols, a random number of nodes will request a route. Then, this request will be processed by the BS and a reply will be sent in response. For proactive protocol, the controller broadcasts the LSDB of the whole network. Due to the mobility of the nodes, the link between the relay nodes in the active flows maybe be broken, which could lead to packet loss. In all the protocols, the upstream node of the failed link will inform the controller about the failure and following that, the controller broadcasts the information of the failed link to the network. Subsequently, each node updates its routing table. For proactive protocols, each node run the Dijkstra's algorithm to find the new route, while in reactive protocols, the controller informs the involved nodes on the updated forwarding information.

- E2E delay: To calculate the end-to-end delay, delays resulted from the route discovery and route maintenance in both cellular and WiFi links are computed. In cellular communication, to calculate the time taken for a control packet to travel between the BS and the mobile nodes, firstly, the data rate (3.10) for each user is calculated based on the channel capacity (3.9) of each link. The channel capacity depends on the channel estimation (3.6) that the BS performs in the network. Secondly, the transmission delay (3.11) for any control packet is calculated based on the obtained data rates. Delay in WiFi links is the time taken by a data packet to travel from the source to destination. This delay time mostly depends on the number of link failures that could occur in the path of data packets. This failure will grow in accordance with the nodes' mobility. Delay also depends on other factors, such as number of nodes and data traffic density. The reason is that the load on the BS for processing data packets will increase when the number of nodes grows in the network. As a result, the response time of the controller to network changes, and also route replies would be slower.

- Energy consumption: To calculate the energy consumption, nodes will start with an ini-

[3]The proposed architecture is not limited to using IEEE802.11g, as any other ISM/802.11-based radio could be used. However, for the purpose of this study the 802.11g is chosen as a potential ISM-based radio technology.

tial energy. As soon as they transmit or receive a packet, the energy will start to reduce based on (3.20). The amount of consumed energy for sending or receiving the packets depends on the packet size; larger packets require more energy to be transmitted or received. Further, based on (3.17) and (3.19), mobile nodes consume more energy for packet transmission compared to packet reception.

- Overhead: For overhead, the cellular and WiFi overheads are calculated separately. Cellular overhead is related to the control packets that are exchanged between the BS and mobile nodes. WiFi overhead is related to the control packets that are attached to the data packets in reactive protocols. For example, in VARP-S and SMDRP, the full path to the destination and the F_{ID} will be attached to the data packets, respectively, which introduces overheads on the network. Further, in all the protocols the amount of data that must be resent due to link failure also will result in additional overhead in both cellular and WiFi. To further illustrate the performance of proactive and reactive approach in terms of cellular overhead, a theoretical cost calculation of the routing protocols is conducted in Figure 3.7. If there are n number of nodes in the network, then the amount of information required to differentiate one node from other nodes in the network is $\log_2 n$, denoted by N_{ID}. Additionally, the amount of information needed to identify one flow from the other flows is represented by F_{ID}. The size of F_{ID} depends on the number of flows in the network. For this theoretical analysis, (3.13), (3.14) and (3.15) are used to compare the cellular channel overhead of the proposed protocols.

Figures 3.8-3.17 represent the simulation results for mobility rate of $3m/s$ and $20m/s$, which are explained in the following subsections. Moreover, the representation of simulation results of main performance parameters are shown in tabular form in Tables 3.3-3.7, where m, μ, K, M, G, represent Milli, Micro, Kilo, Mega and Giga, respectively.

3.5.1 Simulation Analysis for Dropped Packets

Figure 3.8 and Table 3.3 depict the number of dropped packets for different number of flows with various network densities. As it can be seen, HSAW has the least number of dropped packets compared to the two other protocols when the number of active flows are low. However, once the number of nodes increases in the network and the number of active flows rises, HSAW performance degrades, because nodes run out of energy (i.e., in semi-dense networks when number of flows is more than 600). In sparse networks, VARP-S experiences more dropped packets compared to SMDRP, while the response time of the controller in SMDRP is quicker as the size of the flow reply is smaller compared to VARP-S. Although, both protocols operate almost the same in semi-dense and dense networks. This is because, in those networks the load on the BS is much higher and packets must wait in the queue for a longer time resulting in slower response time. In networks with mobility rate of $20m/s$, the number of dropped packets is significantly higher compared to $3m/s$ scenarios because nodes in the high mobile networks

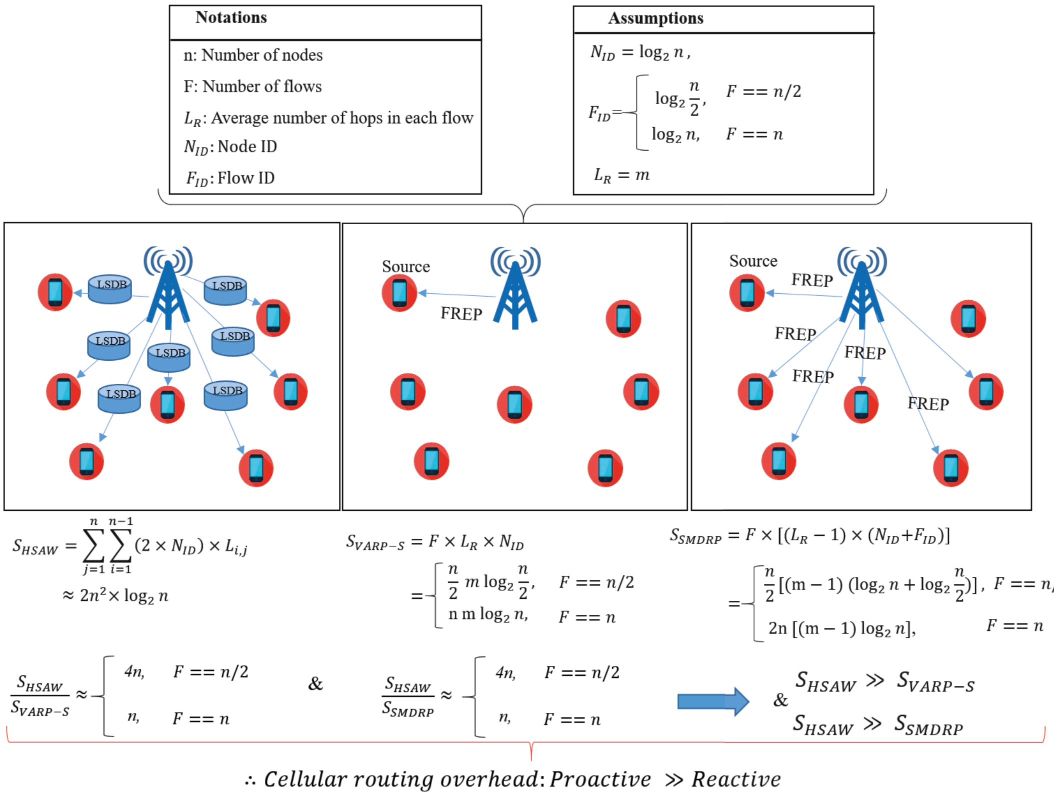


Figure 3.7: Theoretical performance analysis of the proposed routing protocols in terms of cellular overhead.

Table 3.3: Number of dropped packet in Kbits.

Network Size	Number of Flows	Protocols					
		HSAW	VARP-S	SMDRP	HSAW	VARP-S	SMDRP
		Mobility 3 m/s			Mobility 20 m/s		
		Dropped Packets Kbits					
Sparse	0	0	0	0	0	0	0
	25	0.0341	0.7448	0.2591	0.3035	6.6571	2.0302
	50	0.0807	1.5842	0.4535	0.6792	11.7701	3.8335
	100	0.1428	2.7732	0.8169	1.4173	23.0505	7.8406
	150	0.3152	4.3184	1.3642	1.8742	36.0307	12.4482
	200	0.5739	6.1731	1.7121	3.2168	48.4477	16.223
Semi-dense	0	0	0	0	0	0	0
	150	0.912	1.5845	1.4577	1.6582	26.292	20.501
	300	1.792	2.9294	2.689	3.206	48.942	40.347
	600	3.713	6.0155	5.8953	6.479	98.757	97.536
	900	45	9.2205	8.5693	450.8	147.44	149.91
	1200	60	11.623	11.418	600.6	200.83	199.14
Dense	0	0	0	0	0	0	0
	250	5.7215	9.6478	9.1274	14.334	178.03	161.45
	500	11.937	18.933	17.336	28.472	318.22	308.02
	1000	50	36.751	35.631	807	645.22	628.921
	1500	75	54.854	53.487	1200	976.13	956.86
	2000	105	73.979	73.696	1957	1297.8	1125.539

Table 3.4: Energy consumption in %.

Network Size	Number of Flows	Protocols					
		HSAW	VARP-S	SMDRP	HSAW	VARP-S	SMDRP
		Mobility 3 m/s			Mobility 20 m/s		
		Energy Consumption in %					
Sparse	0	0	0	0	0	0	0
	25	1.379	1.245	0.5251	1.427	1.4	0.917
	50	1.897	1.573	1.23	1.913	1.601	1.51
	100	3.34	2.146	1.444	3.37	2.223	1.895
	150	3.754	2.78	2.74	3.924	2.858	2.53
	200	4.78	3.426	3.4	4.837	3.615	3.445
Semi-dense	0	0	0	0	0	0	0
	150	14.721	3.137	0.231	15.04	3.164	2.798
	300	24.068	5.347	1.351	28.98	5.483	4.533
	600	58.51	9.788	8.11	77.099	10.005	8.373
	900	100	14.227	14.168	100	14.582	12.896
	1200	100	18.683	18.664	100	19.162	19.133
Dense	0	0	0	0	0	0	0
	250	38.715	6.244	3.482	40.622	6.32	5.122
	500	79.233	11.577	10.015	87.174	11.643	10.346
	1000	100	22.28	20.016	100	22.8	20.474
	1500	100	37.98	37.566	100	40.844	40.899
	2000	100	79.38	79.146	100	83.478	83.13

Table 3.5: End-to-end delay in s.

Network Size	Number of Flows	Protocols					
		HSAW	VARP-S	SMDRP	HSAW	VARP-S	SMDRP
		Mobility 3 m/s			Mobility 20 m/s		
		E2E delay (s)					
Sparse	0	0	0	0	0	0	0
	25	83.4	0.441m	0.158m	0.3313m	0.0013	0.436m
	50	85.4	0.404m	0.146m	0.3528m	0.0012	0.411m
	100	86.9	0.415m	0.137m	0.3458m	0.0012	0.425m
	150	84.9	0.4m	0.147m	0.3181m	0.0012	0.44m
	200	86.7	0.455m	0.141m	0.3327m	0.0012	0.429m
Semi-dense	0	0	0	0	0	0	0
	150	0.01578	0.1897	0.1705	0.0619	0.8583	0.8591
	300	0.1589	0.1817	0.1724	0.63	0.8214	0.8391
	600	0.1537	0.1837	0.1712	0.6281	0.8318	0.8189
	900	inf	0.1861	0.1766	inf	0.8287	0.8368
	1200	inf	0.188	0.1771	inf	0.841	0.834
Dense	0	0	0	0	0	0	0
	250	0.07281	0.7069	0.6978	0.2893	3.1934	3.1289
	500	0.6211	0.6989	0.6791	2.909	3.1988	3.1007
	1000	inf	0.6815	0.6767	inf	3.214	3.1523
	1500	inf	0.6925	0.6758	inf	3.4905	3.1726
	2000	inf	0.6798	0.6339	inf	3.5258	3.423

Table 3.6: Cellular overhead in bits/s.

Network Size	Number of Flows	Protocols					
		HSAW	VARP-S	SMDRP	HSAW	VARP-S	SMDRP
		Mobility 3 m/s			Mobility 20 m/s		
		Cellular Overhead bits/s					
Sparse	0	0	0	0	0	0	0
	25	23K	6.5K	6.9K	113K	60K	69.9K
	50	96K	13K	14K	282K	134K	147K
	100	115K	27K	28K	824K	269K	299K
	150	149K	40K	41K	1.1M	390K	404K
	200	205K	53K	55K	3.2M	528K	535K
Semi-dense	0	0	0	0	0	0	0
	150	4.4M	1.3M	1.4M	284M	13.5M	14.69M
	300	35M	2.6M	2.7M	462M	27M	28.99M
	600	69M	5.3M	5.4M	685M	54M	55.81M
	900	104M	8M	8.1M	814M	81M	82.32M
	1200	129M	10M	11.79M	899M	108M	119.44M
Dense	0	0	0	0	0	0	0
	250	20M	6.19M	8M	834M	62M	69.3M
	500	40M	12M	13M	1.3G	124M	145.61M
	1000	120M	24M	25M	4.2G	249M	269.92M
	1500	518M	36M	38M	6.5G	372M	391.56M
	2000	1.9G	43M	44M	8.6G	447M	448.71M

Table 3.7: WiFi overhead in bits/s.

Network Size	Number of Flows	Protocols					
		HSAW	VARP-S	SMDRP	HSAW	VARP-S	SMDRP
		Mobility 3 m/s			Mobility 20 m/s		
		WiFi Overhead bits/s					
Sparse	0	0	0	0	0	0	0
	25	0	19136	3920	0	24864	9800
	50	0	40352	6000	0	109952	17400
	100	0	76288	13520	0	148608	35560
	150	0	100544	21040	0	254848	44720
	200	0	127616	27320	0	325792	70920
Semi-dense	0	0	0	0	0	0	0
	150	0	108128	20640	0	396576	54896
	300	0	181120	39960	0	946560	194160
	600	0	372064	79080	0	1723040	382040
	900	0	563232	127640	0	2686304	560400
	1200	0	746304	166720	0	3621920	743240
Dense	0	0	0	0	0	0	0
	250	0	155840	69720	0	693952	147360
	500	0	314080	66800	0	1360864	299840
	1000	0	625600	132120	0	2866944	615120
	1500	0	3390816	830256	0	5752448	1612600
	2000	0	11024864	2181250	0	14774336	4852330

experience more link failure. The same trend for $20m/s$ is observed for three protocols as illustrated in Figure 3.13 and Table 3.3. Moreover, for all the scenarios, when the number of flows increases, the number of dropped packets will increase accordingly.

3.5.2 Simulation Analysis for Energy Consumption

Figure 3.9 and Table 3.4 illustrate the energy consumption of nodes for various number of flows in sparse, semi-dense and dense networks. As shown, nodes consume the most energy and the least energy in HSAW and SMDRP, respectively. The consumed energy of nodes in HSAW protocol is much higher in semi-dense and dense networks compared to sparse networks. This is due to the large size of LSDB that must be received by each node in the network which leads to more rapid depletion of their battery when number of flows are more than 600 and 500 for semi-dense and dense networks, respectively. In SMDRP, nodes use slightly less energy in contrast with VARP-S. Because in VARP-S, the information of relay nodes are attached to the data packets that will be received by each hop between source and target, while in SMDRP, only flow ID is attached to the data packets. Consequently, nodes consume less energy as they receive smaller packets. The consumed energy for three protocols will increase in accordance with the number of flows and mobility rate in the network as depicted in Figure 3.9, Figure 3.14 and Table 3.4.

3.5.3 Simulation Analysis of E2E Delay

E2E delay between source and target of the flows in the network are indicated in Figure 3.10 and Figure 3.15 with Table 3.5. As it can be seen, nodes in HSAW experience the least delay as response time to the link failures is much faster. The reason is that in HSAW, nodes do not wait to receive the forwarding information from the controller. They simply run the Dijkstra's algorithm and find the new path to the target. In contrast, nodes in the other two protocols have to wait for the controller updates that leads to more end-to-end delay. SMDRP has less delay compared to VARP-S because of the smaller size of updated information sent by the controller, since in BS, smaller packets are processed faster. Therefore, the updated information takes less time to get to the relay nodes. However, when number of flows is more than 600 and 500 in semi-dense and dense networks, HSAW is not functional as nodes will run out of battery and relay nodes are no longer available to deliver packets. Regardless of higher delay time in mobility rate of $20m/s$ (Figure 3.15) compared to $3m/s$ (Figure 3.10), the trend of delay is the same for three protocols. Further, E2E delay increases when the number of flows increases.

3.5.4 Simulation Analysis of Overhead

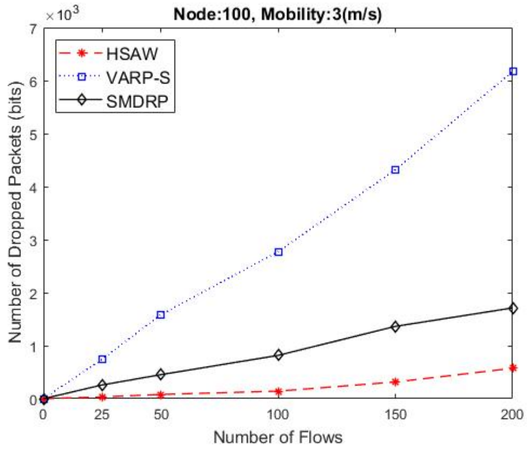
Figure 3.11 and Figure 3.16 with Table 3.6 represent the cellular overhead, when the mobility rate of nodes is $3m/s$ and $20m/s$, respectively. As expected, HSAW puts significant cellular overhead on the network compared with the two other protocols. This is due to the consumed bandwidth by the controller for forwarding LSDB updates to the whole network. This overhead will increase when the size of the network grows. In SMDRP and VARP-S, cellular channel is utilized for sending and receiving forwarding information to and from the controller. However, in SMDRP, more cellular bandwidth is required as all the relay nodes in one flow should receive the forwarding information, while in VARP-S, only the source node of the flow will receive the forwarding information.

The WiFi overhead for mobility rates $3m/s$ and $20m/s$ is analyzed in Figures 3.12 and 3.17, and Tables 3.7 respectively. As shown, network experiences more overhead in WiFi channel in VARP-S and SMDRP protocols, since the control packets attached by those protocols to data packets cause the extra overhead on WiFi. On the other hand, HSAW has no overhead on WiFi channel because it does not attach any control header to the data packets.

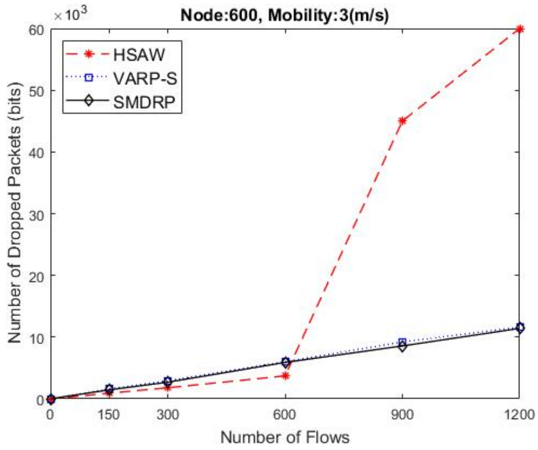
In all the defined scenarios, the total vertical and horizontal overhead will increase when the number of nodes, number of flows and mobility rate increase in the network as depicted in Figures 3.11, 3.12, 3.16, and 3.17.

3.6 Conclusion

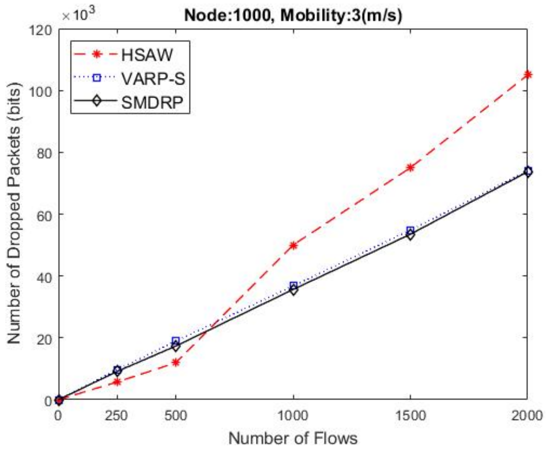
A detailed study is performed to investigate the performance of proactive and reactive MD2D routing frameworks, using HSAW, SMDRP and VARP-S protocols. Three network scenarios: sparse, semi-dense and dense networks were modelled over different mobility rates. Our results show that the proactive MD2D protocol, namely HSAW, introduces the highest overhead in the cellular channel and while consuming the highest energy compared to the other two reactive protocols. Conversely, HSAW provides the least E2E delay and packet loss compared to reactive protocols for lower network and traffic densities. Since, in HSAW when node and traffic density is high, nodes will eventually run out of battery, leading to larger packet loss compared to reactive protocols. Our results suggest further research is required to develop an adaptable routing protocol, which can optimize its core routing functionality and parameters according to networking conditions. This is the focus of the following chapters in this PhD.



(a) Sparse

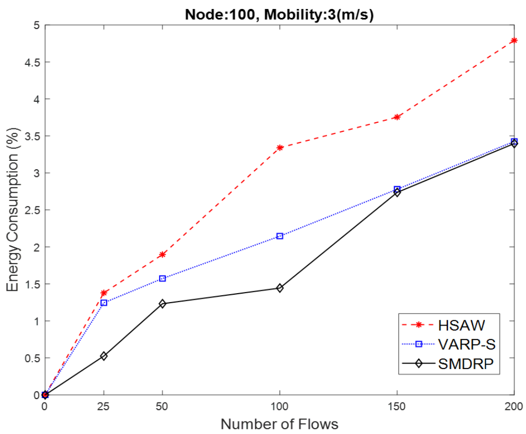


(b) Semi-dense

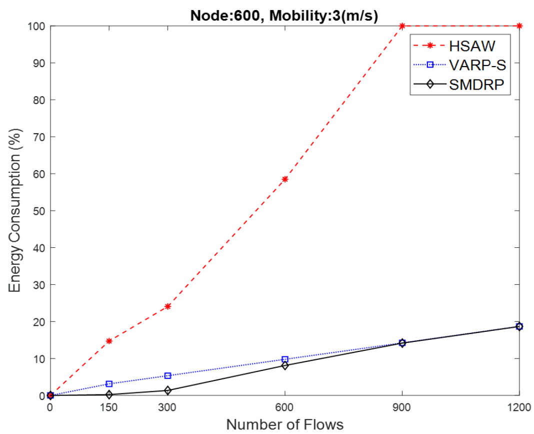


(c) Dense

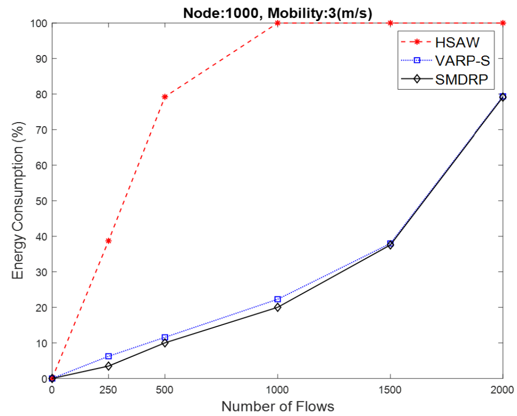
Figure 3.8: Number of dropped packets (bits) with node velocity 3m/s.



(a) Sparse

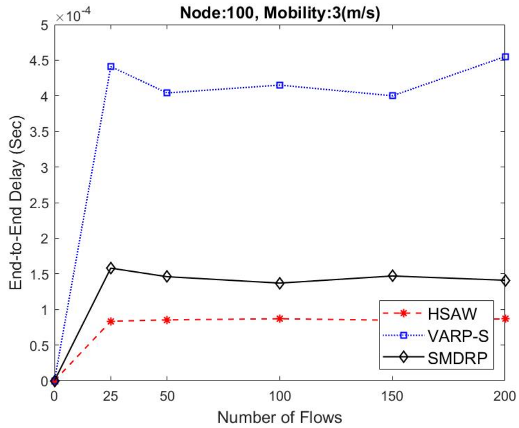


(b) Semi-dense

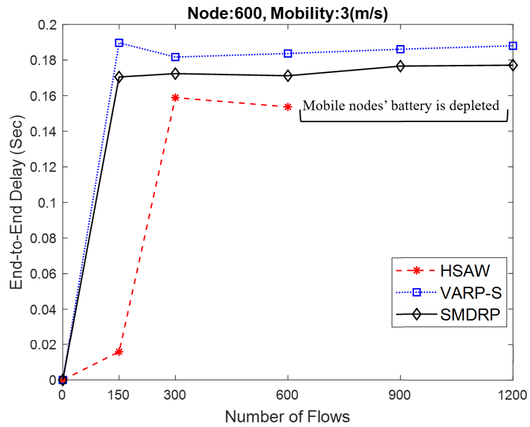


(c) Dense

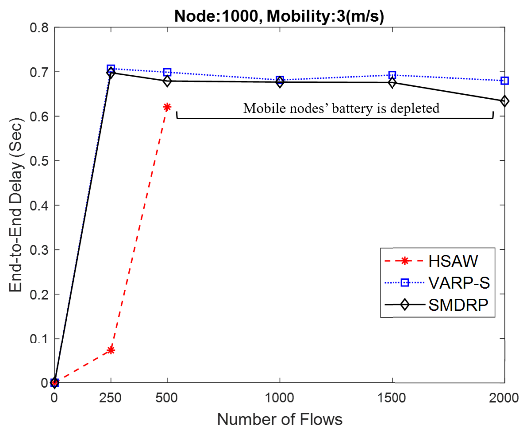
Figure 3.9: Energy consumption of nodes (%) with node velocity 3m/s.



(a) Sparse

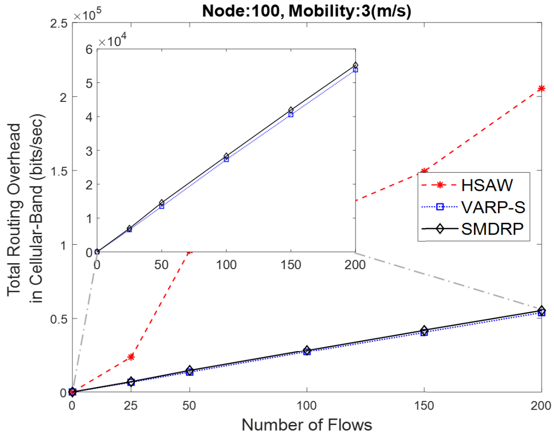


(b) Semi-dense

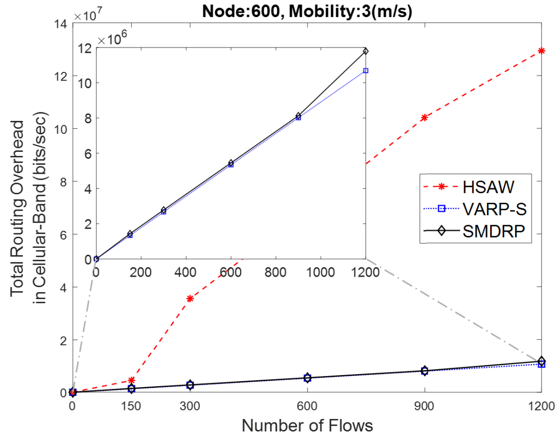


(c) Dense

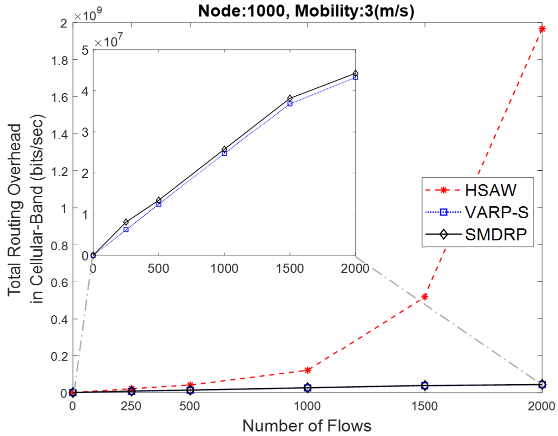
Figure 3.10: End-to-end delay (s) with node velocity 3m/s.



(a) Sparse

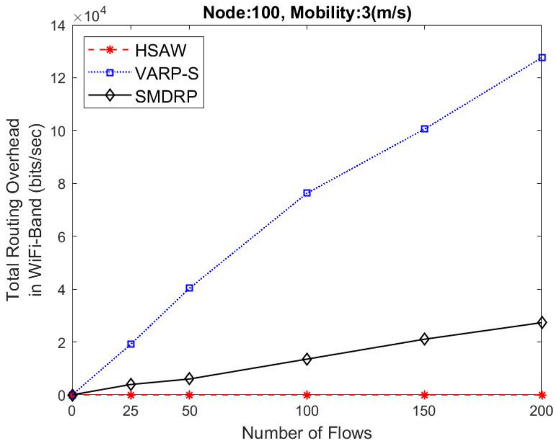


(b) Semi-dense

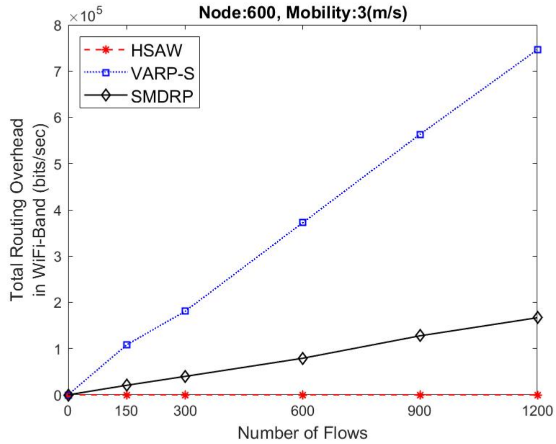


(c) Dense

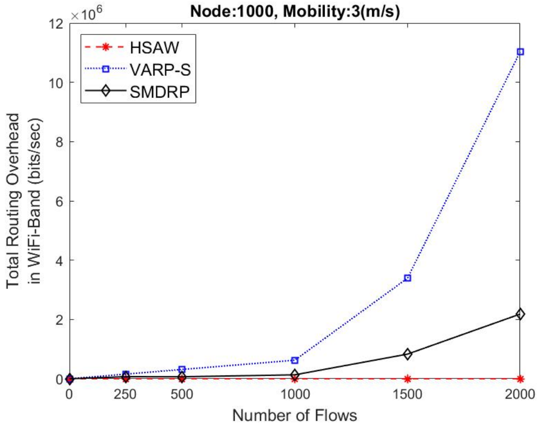
Figure 3.11: Total routing overhead in cellular-band (bits/s) with node velocity 3m/s.



(a) Sparse

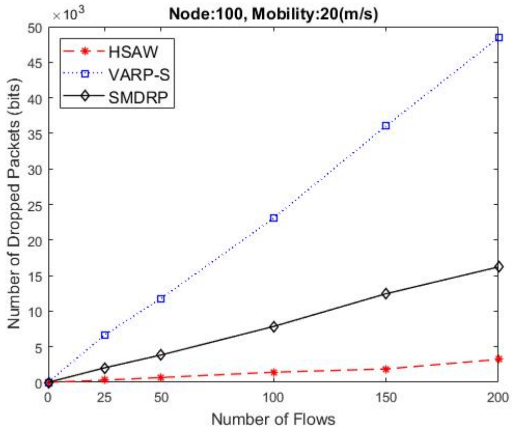


(b) Semi-dense

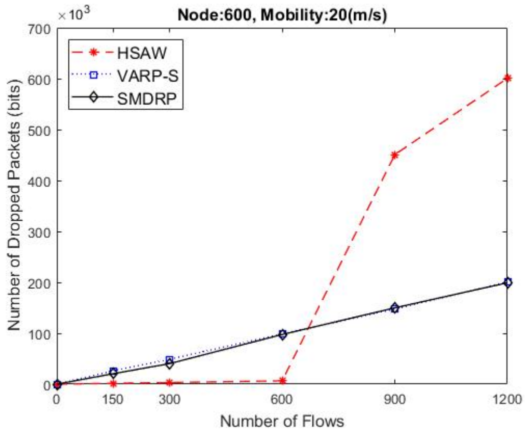


(c) Dense

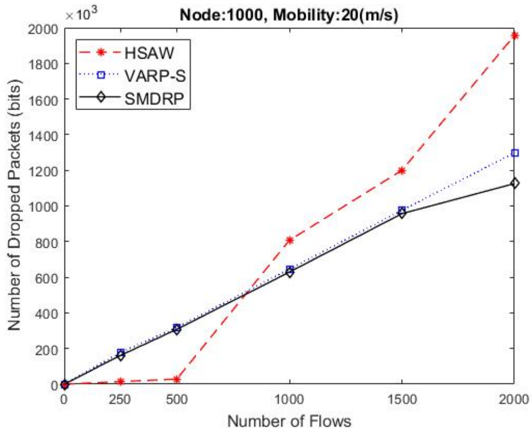
Figure 3.12: Total routing overhead in WiFi-band (bits/s) with node velocity 3m/s.



(a) Sparse

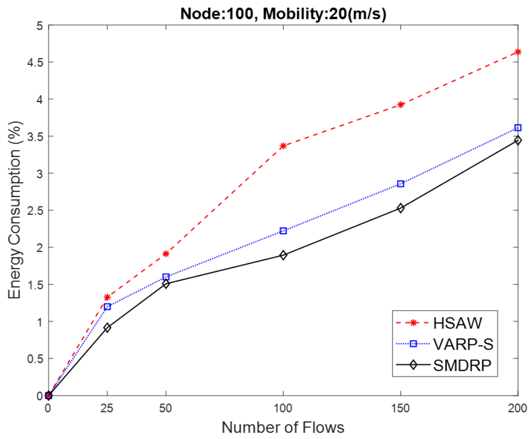


(b) Semi-dense

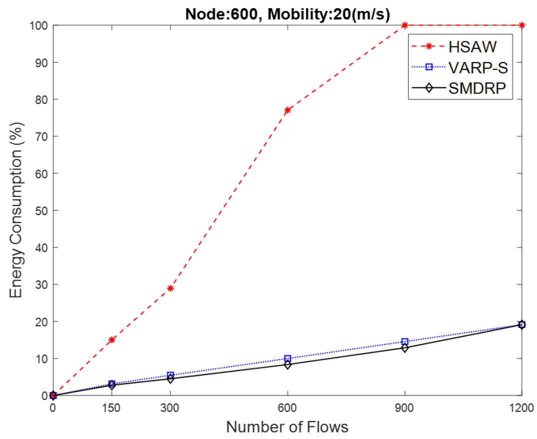


(c) Dense

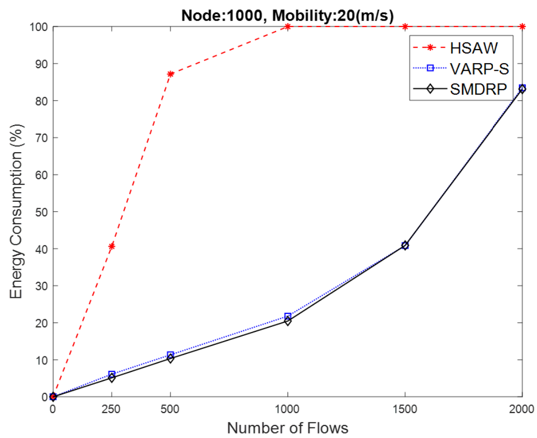
Figure 3.13: Number of dropped packets (bits) with node velocity 20m/s.



(a) Sparse

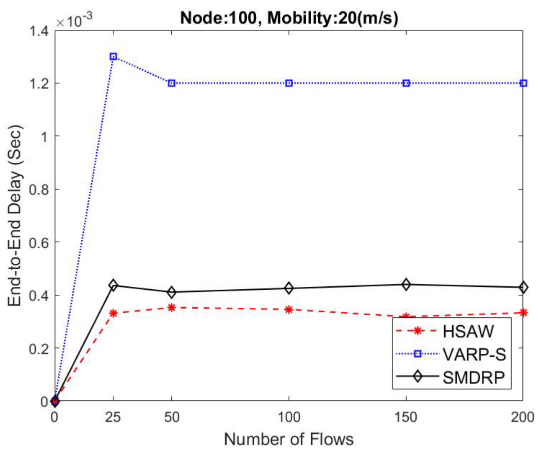


(b) Semi-dense

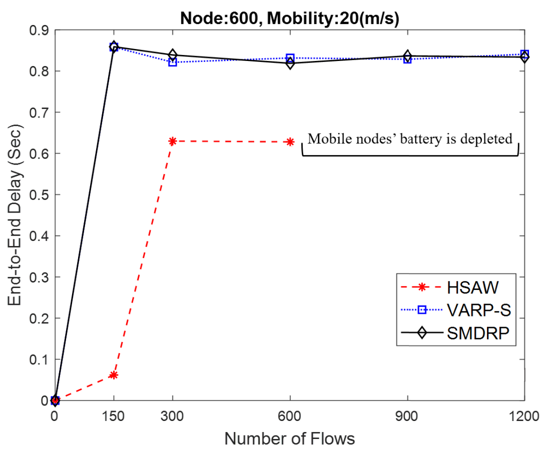


(c) Dense

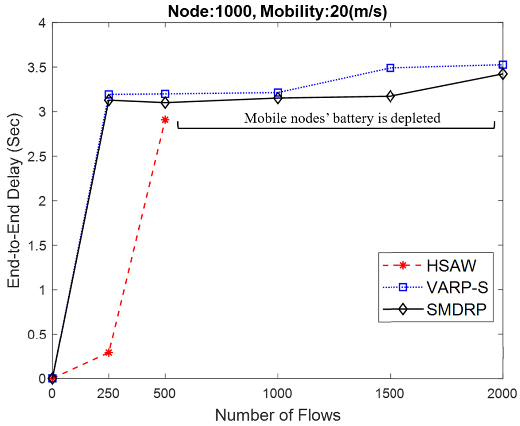
Figure 3.14: Energy consumption of nodes (%) with node velocity 20m/s.



(a) Sparse

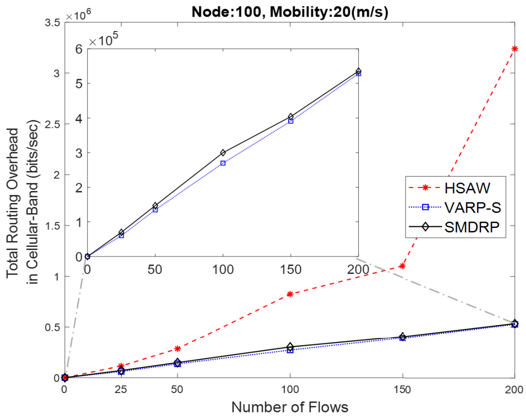


(b) Semi-dense

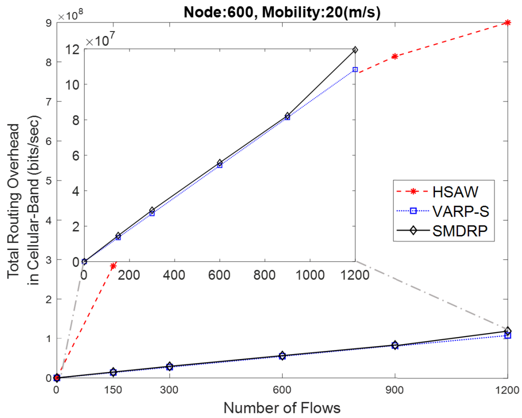


(c) Dense

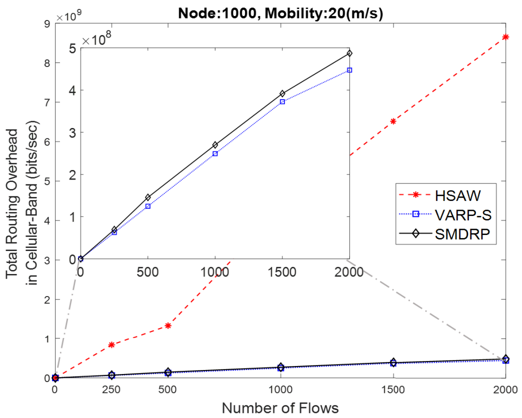
Figure 3.15: End-to-end delay (s) with node velocity 20m/s.



(a) Sparse

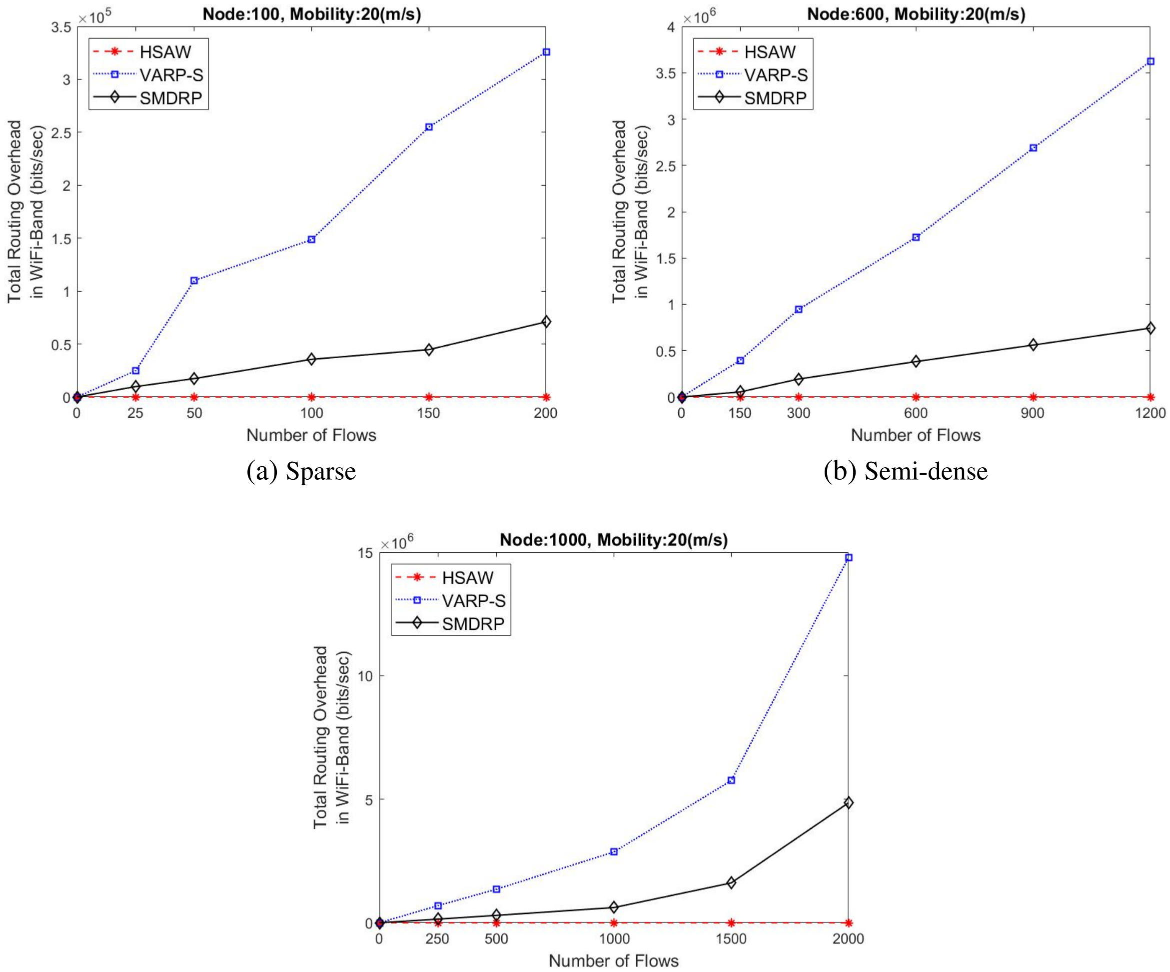


(b) Semi-dense



(c) Dense

Figure 3.16: Total routing overhead in cellular-band (bits/s) with node velocity 20m/s.



(a) Sparse

(b) Semi-dense

(c) Dense

Figure 3.17: Total routing overhead in WiFi-band (bits/s) with node velocity 20m/s.

4

Joint Mobile Node Participation and Multi-Hop Routing for Emerging Open Radio-Based Intelligent Transportation System

4.1 Overview

The upcoming research is published in IEEE Access [1]. This chapter's main objective is to propose an intelligent joint topology control and multi-hop routing protocol called FPRM. We introduce a sub-layer at the network layer that can determine nodes with the highest participation probability in routing using a fuzzy logic system, thus building a framework to create more stable routes. The proposed protocol shows superior performance compared to other routing protocols.

[1]Abolhasan, M., Lipman, J., Shariati, N., Ni, W., Jamalipour, A. (2022). Joint Mobile Node Participation and Multihop Routing for Emerging Open Radio-Based Intelligent Transportation System. IEEE Access, 10, 85228-85242.

4.2 Introduction

The widespread use of mobile devices, intelligent transportation systems, Internet-of-things (IoT) networks, and machine-type communication has led to substantial data traffic growth. This trend will most likely continue as the need for ubiquitous connectivity of people, devices, and machines follow the same trajectory. In this trend, mobile devices have significantly contributed due to advancements in new applications and services for smartphones. These applications require high data rates and perfect quality of experience (QoE) for users. The need for a faster data rate was the primary reason for wireless network evolution and the manifestation of fifth-generation (5G) cellular networks. However, with the emergence of smart applications [145], the Internet-of-everything (IoE) [146], user demands [147], and the connection of millions of people, machines, and vehicles [148], the current network paradigm requires shifting from rate-centric to ultra-reliable low latency communication. Therefore, creating an unprecedented dispute for existing 5G wireless networks. The sixth-generation (6G) cellular network is expected to overcome many associated issues in 5G by utilizing intelligence and a new radio access network (RAN) [149]. Although one can argue that available 5G systems in the market can handle basic IoE and low latency services, it is disputable whether they can deliver the scalability and reliability required for future heterogeneous networks.

Significant effort has been dedicated to enhancing the RAN architecture [150]. Most of the focus of current research is on building an operator-defined RAN (commonly referred to as Open RAN), that enables intelligent radio control and creates self-driving networks. In particular, the open-radio access network (O-RAN) allows openness in the RAN by merging the xRAN forum and centralized-RAN alliance. O-RAN uses the concept of virtualization, cloud, and intelligence to initiate agile service deliveries and enhance capabilities to end users [27]. Incorporating intelligence is one of the main reasons underlying the advantages of 6G over 5G. 6G establishes an automation and self-organization network, where most network applications will be managed by a machine rather than human intervention. One of the main building blocks in 6G is machine learning (ML) technology [151]. ML has significantly contributed to wireless network applications, such as quality-of-service (QoS), resource management, spectrum allocation, and routing [152–155]. Moreover, software-defined networking (SDN) enables the separation of control and data planes, allowing ML-based algorithms to perform optimization and automation in a centralized controller and provide intelligent decisions to devices. SDN and network functions virtualization (NFV) provide advanced features for the RAN intelligent controller (RIC) to increase the performance of future wireless networks. Furthermore, O-RAN has defined three control loops, effectively enabling NFV applications to be deployed at different locations of the cellular network architecture. For example, at the core, where the non-real-time control loop exists, non-real NFV applications could be deployed, whereas, in the near-real-time control loop, NFVs aim to perform operations at near-real-time [156, 157].

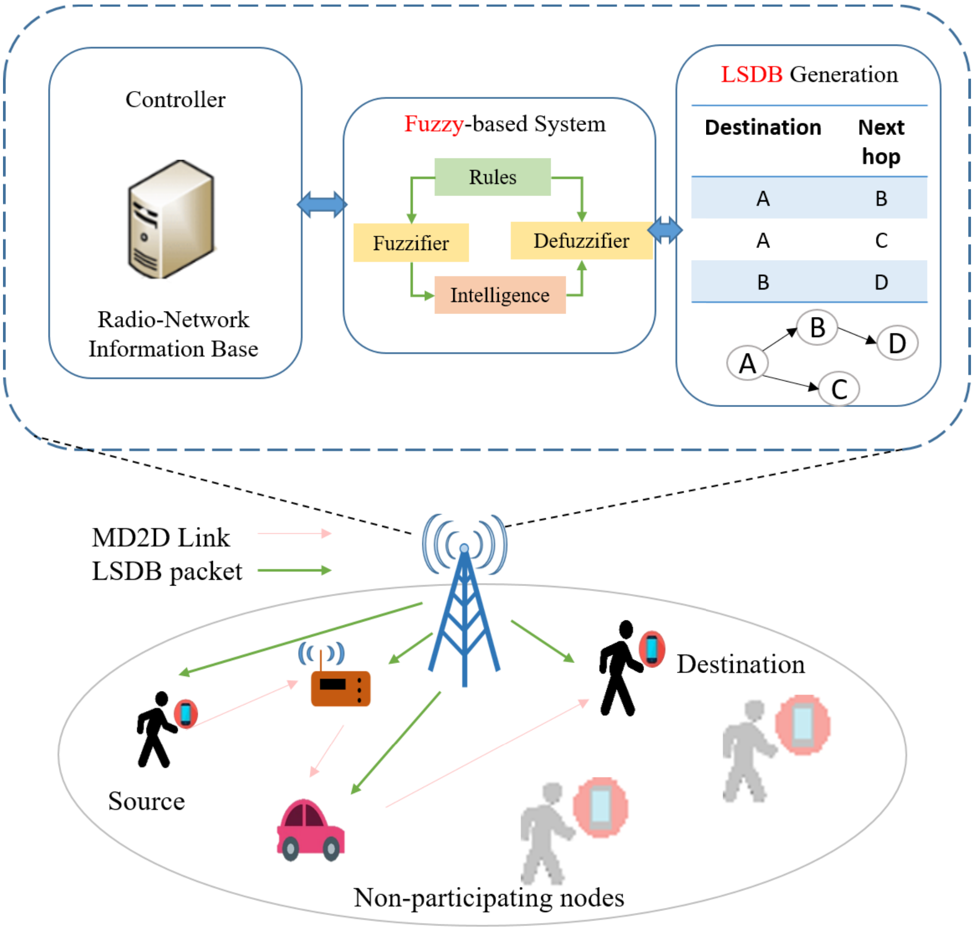


Figure 4.1: Proposed framework for the designed routing protocol, where fuzzy-logic system is adapted at edge controller.

4.2.1 Motivations

Multi-hop device-to-device (MD2D) communication is a promising technology to offload base station (BS) traffic. In MD2D frameworks, routing algorithms play an essential role, and they must be designed to provide the highest performance based on the criteria of the current and future networks. Therefore, an MD2D routing protocol must be capable of adapting to any dynamic topology changes in the network. Recent MD2D routing protocols mainly focus on optimization algorithms to increase network performance. To the best of our knowledge, no MD2D study has yet adapted a fuzzy-based topology control routing mechanism to identify the participant nodes and dynamically create efficient routes for that specific network topology. One of the advantages of our proposed framework is the intelligence of controller. The controller can collect network telemetry and network application requirements and use them to generate link-state databases (LSDBs). Moreover, the topology control mechanism enables a fuzzy system to dynamically adapt to the network changes and identify the participating nodes. Our routing protocol can create network knowledge and prescribe optimal routes using the information obtained from the topology control and acquired data from the network.

4.2.2 Contributions

This chapter proposes an intelligent joint topology control and multi-hop routing called fuzzy-based participation and routing protocol for MD2D (FPRM) to increase the network lifetime and packet delivery ratio (PDR). In our approach, a fuzzy-based participation mechanism controls and manipulates the network's topology. The fuzzy system is located in the RIC controller. The controller collects all the nodes' information and decides whether a node should participate in MD2D routing based on the fuzzy rules. Therefore, different network graphs can be obtained by manipulating which node can participate in the network. Based on the topology graphs and application requirements, network LSDBs are created. Figure 4.1 represents the fuzzy-based participation and routing framework under the management and a control. In this framework, every small BS (SBS) (e.g., picocell networks) is connected to a controller with a unique ID, and every BS can communicate to other SBSs through backhaul channels. Each SBS is responsible for providing service to every user equipment (UE) in their coverage area. The edge network is connected to the core network for Internet access and other advance processing functionalities. The controller creates a centralized fuzzy-based unit to process and store information to instruct the SBSs. The embedded fuzzy system obtains the participating nodes and creates different network LSDBs based on the network topology and application requirements. Later the LSDB is shared with participating nodes to perform MD2D routing.

The contributions of this chapter are summarized as follows:

- Proposing joint topology control and MD2D routing using an adaptive fuzzy-based learning system.
- Presenting a new MD2D knowledge-based routing framework that adapts based on the user requirements and fuzzy system to identify participating nodes and disregard nodes that can cause potential damage to network performance.
- Utilizing an intelligent O-RAN controller to collect network information and build various LSDBs for different network topologies.
- Introducing a three-step routing constraint to ensure the proposed MD2D routing can provide a reliable communication link to the maximum number of users over a long time.
- A comprehensive analysis of the proposed routing protocol with a semi-centralized routing protocol, namely hybrid SDN architecture for wireless distributed networks (HSAW) [83], and traditional distributed routing protocols, including ad hoc on-demand distance vector (AODV) and optimized link state routing (OLSR), is presented. The results show more than a 30% increase in average throughput, more than 30% reduction in end-to-end (E2E) delay, almost 8% increase in PDR, and almost 2% decrease in energy consumption compared to one of the leading MD2D routing protocols, HSAW.

The rest of this chapter is organized as follows. Section 4.3 provides a literature review of the current MD2D routing protocols. Section 4.4 briefly explains the routing framework and system model. Section 4.5 describes the route discovery and maintenance strategy over an intelligent fuzzy-based routing protocol. Section 4.6 provides a complete explanation of the fuzzy logic algorithm used for node participation. Section 4.7 introduces the constraint used after the fuzzy logic process to improve the scalability of the routing protocol. Section 4.8 evaluates the performance of the proposed routing framework using the NS-3 simulator. Finally, our conclusion and future research directions are explained in Section 4.9.

4.3 Related Works

The O-RAN provides the opportunity to create a new generation of intelligent semi-centralized routing protocols [158, 159]. This opportunity has opened the door toward developing a new set of MD2D routing protocols for future intelligent transportation systems. In MD2D routing, a controller can instruct the devices in the network and install or remove routing entries within the device's routing tables. For instance, the authors in [160] proposed a routing protocol using the SDN framework in the wireless multi-hop paradigm to increase the life span of the network. In their framework, nodes transmit their local information to the controller to generate a global network view. The SDN controller can then provide a route to a destination upon request from a source node. This is generally achieved by computing the shortest route via the shortest path first (SPF) algorithm. The SDN controller applies energy constraints and hops count limits to find the best path for each source node. Simulation results show that network lifetime is extended compared to OLSR and AODV routing protocols. Authors of [86] proposed an MD2D routing protocol for SDN-based cellular networks. The proposed method builds the LSDB of the network at the controller, and once a node requests a route, the controller uses the Dijkstra algorithm to find the shortest path to the destination. The proposed protocol provides scalability and reliability in mobile ad hoc networks (MANETs). Furthermore, in vehicular ad hoc networks (VANETs), the authors of [161] proposed a low delay and low routing overhead framework to propagate messages. Their protocol finds multiple paths using multiple network attributes, such as link stability and shortest travel time. Simulation results show that the proposed protocol significantly outperforms the conventional routing protocols regarding E2E delay and routing overhead. SDN-based routing frameworks provide scalable routing procedures among various wireless networks. However, the above strategies have not considered the dynamic changes of the network causing deterioration of network performance. One of the promising solutions to address this problem is using O-RAN near-real-time controller to make routing decisions. Fuzzy-based routing protocols have shown promising solutions for optimization of routing parameters in a self-organizing manner based on the network dynamics [162].

Recently, knowledge-based algorithms have gained widespread attention. The optimized data acquired from SDN-enabled ML-based controller can create knowledge that enables networks to adjust their parameters when necessary [163–165]. For instance, the authors of [166] used reinforcement learning (RL) in routing problems to maximize the throughput and minimize the communication delay for each source node. Their algorithm continuously predicts the network's future behavior and evaluates the most efficient path to the destination. An SDN controller is utilized to collect network information and train an RL agent to manage the data traffic among devices in the network. Simulation results illustrate that the proposed routing protocol delivers large files faster than open shortest path first (OSPF) and least loaded (LL) routing algorithms over different network scenarios. Moreover, the authors of [167] proposed an intelligent-based fuzzy routing protocol to decrease power consumption. Specifically, they solved the unbalanced distribution of cluster heads by the fuzzy c-means clustering algorithm, which categorized nodes into balanced clusters. Then the cluster heads are assigned using the Mamdani fuzzy interference system to route packets between the controller and other nodes. Obtained simulation results show an increase in network lifetime and superiority over existing clustering-based protocols. Other studies, such as [168], used fuzzy logic to improve the stability of the AODV routing protocol in MANETs. The most trusted relay nodes are selected in their framework for route generation between the source and destination nodes. The fuzzy logic method takes the node energy, mobility, and hop counts to determine the node trust level. The simulation results illustrate the proposed framework's advantages against AODV in terms of control overhead, network throughput, packet delivery ratio, and E2E delay. In vehicular networks, fuzzy learning has provided unprecedented benefits. The authors of [169] introduced an intelligent fuzzy-based routing scheme for software-defined vehicular networks (SDVNs) in urban areas. In this technique, a centralized controller maintains the routing table. These routing tables are initially constructed based on the priorities of packets using fuzzy logic and later updated based on the network changes. Then, a greedy strategy is used to obtain routing paths with the highest link stability. Simulation results demonstrate significant performance improvement in dense urban areas compared to the existing routing frameworks. As can be seen from the literature review, the ability to learn and adjust the network parameters to increase the network performance is a promising solution for future MD2D heterogeneous networks [104, 108, 170].

4.4 System Model

FPRM is a proactive routing protocol where a controller collects the network information and broadcasts the required routing information for participant nodes to perform MD2D communication. In general, the controller is responsible for separating the control and data planes such that the distributed devices only transmit the data messages, not the control messages. When the controller is adapted in the network, devices are relieved from flooding algorithms to find a route to the destination. The controller is responsible for creating paths separately and only providing devices with routing entries. Proactive routing protocols use the capability of a centralized controller to build LSDBs. However, the generated LSDB is shared with the entire network without knowing whether devices are participating in the MD2D routing or not. This can cause cellular channel overhead and reduction in the MD2D network lifetime. To mitigate problems associated with proactive routing approaches and increase the network's lifetime, we introduce a sub-layer, which can determine nodes with the highest participation probability in routing and thus build a framework to create more stable routes.

As shown in Figure 4.2 the BS collects every node's data, such as energy, number of neighboring nodes, and mobility rate, to obtain the most reliable nodes. The fuzzy system processes the information and defines the participating nodes. Furthermore, the controller creates an LSDB and broadcasts it to the participating nodes, where nodes will distributively calculate a path using the LSDB. Once a node requires to transmit a packet, it will evaluate the path using Dijkstra's algorithm and adds the route to its routing table. If a link failure occurs while sending the packet, the node can calculate a new route and delete the previous route from the routing table. Additionally, the link failure is shared with the controller, where the controller updates the fuzzy system and broadcasts the new information to nodes to update their LSDBs.

In our framework, every UE is equipped with at least two communication interfaces: cellular/licensed and WiFi/unlicensed frequencies for in-band and out-band communication. The in-band communication consists of data messages exchanged between UEs (MD2D communication) and the acknowledgment messages. Out-band communication includes Internet connectivity, control messages, and connection to other networks. At the initial stage of the network, each UE transmits a Hello message to its neighboring nodes (where neighboring nodes are nodes within each UEs communication range) via the WiFi channel. Then, UEs transmit their link-state information to the sub-controller via the cellular channel. Based on the received data, the sub-controller processes the information and generates an LSDB. Then, fuzzy logic is applied to find the participation policy in the network. The sub-controller uses defined fuzzy rules to calculate each node's eligibility index (EI). The lowest values of the EI correspond to nodes that might fail to transmit a packet or have low energy levels. These nodes are excluded from participation in routing. Moreover, once the eligible node is determined, the next step is applying the following constraints: coverage and mobility. The introduced constraints in our proposed routing framework ensure that the eligible/active nodes can support the entire network with the least interruption. It is noteworthy that The LSDB is only transmitted to active nodes in the network.

4.5 Routing Procedure in Fuzzy-Based MD2D Communication Framework

At the initial stage of our framework, nodes transmit their information, such as remaining energy, list of neighboring nodes, mobility rate, and throughput. Once the BS receives all this information, it will run them through the fuzzy logic system to obtain nodes with the least EI. These nodes are in critical conditions, and using them in the routing process may cause packet

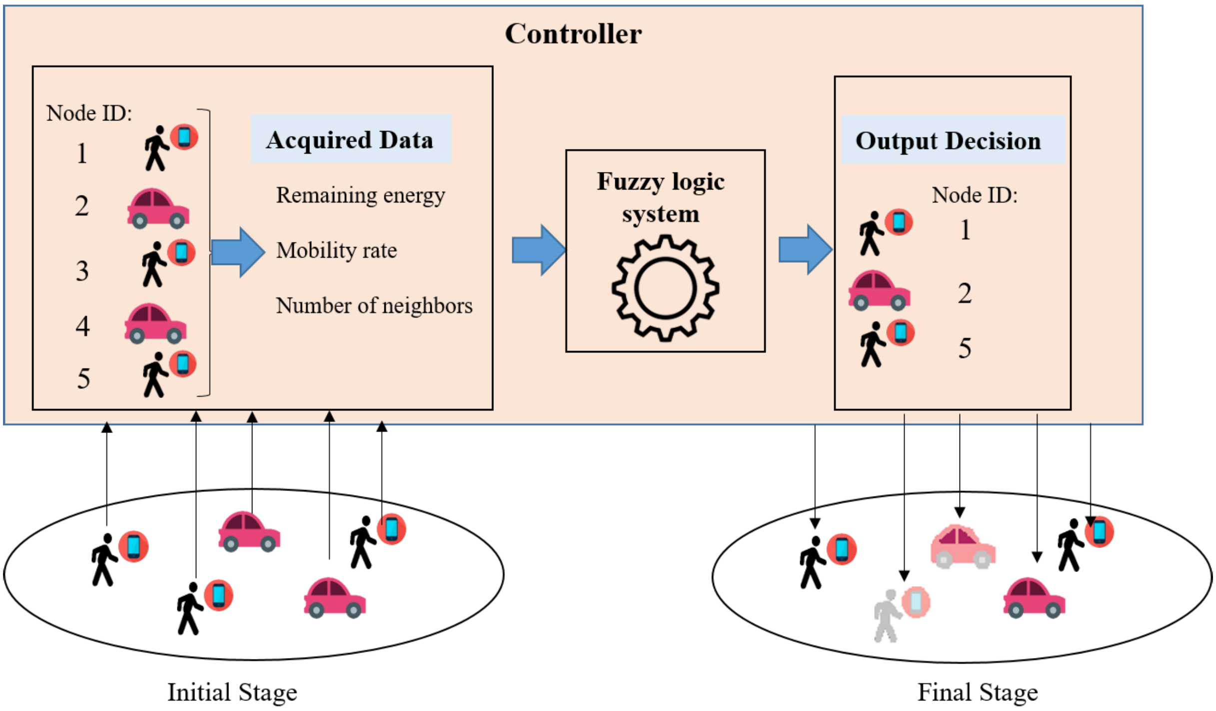


Figure 4.2: Proposed fuzzy framework, where nodes information including remaining energy, mobility rate, and the number of neighboring nodes are collected by the controller to apply fuzzy system and identify the participating nodes.

loss leading to an unstable network. After selecting eligible nodes, the next step is to compute each node's activation time duration. Then, we apply mobility and coverage constraints to enforce two rules: first, least interrupted connectivity, and second, providing service to the entire network. Then, the active nodes will receive the computed LSDB from the BS.

4.5.1 LSDB Calculation

The controller provides the entire network with link-state information, calculated in a centralized manner using the weight given to each node. The link-state information contains the reported data of nodes, including a list of neighboring nodes, the traffic arrival rate, the queue length of each node, and the channel condition between the node and the one-hop neighboring nodes. To compute the LSDB, once a node joins a BS after the authentication process by the sub-controller, it will be allowed for advance routing services. Then, each node will broadcast its existence through Hello messages to its neighboring nodes via a WiFi channel. After the adjacent nodes are identified and the link between them is established, each node will send its link-state information via the cellular channel to the sub-controller. This information consists of one-hop neighboring nodes associated with link-state, location, battery level, and throughput and sent to the BS using a topology control (TC) message. The LSDB generated at the sub-controller is also shared with the nearby sub-controllers and the main controller. Therefore, each BS has a global network view, helping them with handover decisions, traffic management, user allocation, and content caching.

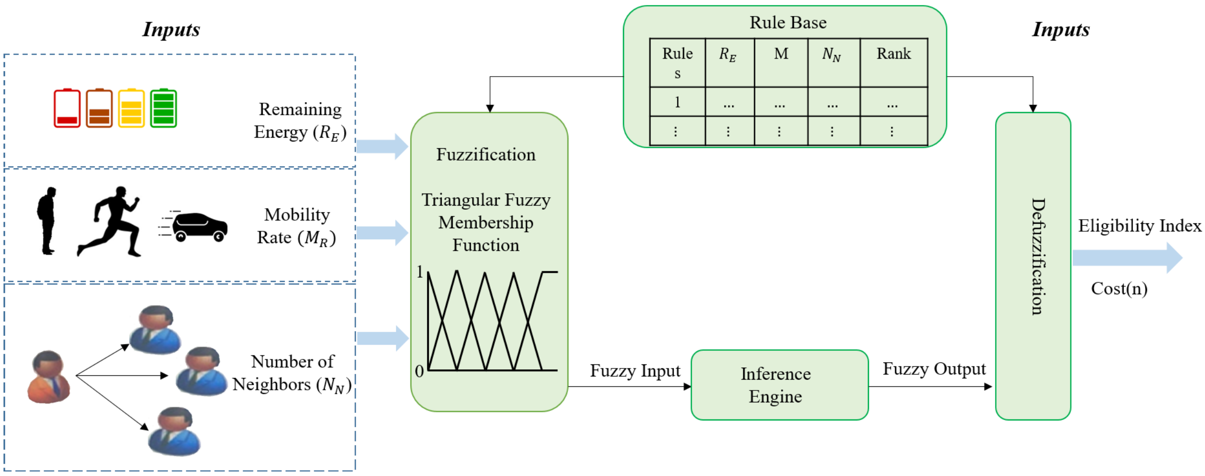


Figure 4.3: The proposed fuzzy system.

4.5.2 LSDB Update

Each sub-controller must update the generated LSDB after any changes to the network. If any changes occur to the link-state information, a TC message will be sent to the sub-controller to inform the status changes of the node. Based on the TC message, the sub-controller decides whether to change the LSDB. The new information received at the BS is processed to update the main LSDB and broadcast any link state changes. Moreover, if any node observes sudden changes or link failures, the BS is notified, and new LSDB entries are generated based on the information. The proposed algorithm only updates specific entries of the node's LSDB. However, the entire LSDB is usually updated at a constant rate.

4.5.3 Route Discovery

Our fuzzy-based routing framework is a hop-by-hop routing protocol, meaning that the data packets only carry the node destination ID. Each intermediate node or active node between the source and destination is the relay node. The relay nodes are the forwarding devices that check the destination field ID and then apply Dijkstra's algorithm to discover the next forwarding node. Each forwarding device has the entire network LSDB enabling them to find the least-cost path to the destination. After receiving LSDB, an active node performs distributed route discovery using Dijkstra's algorithm [171]. In particular, each node generates a routing table that specifies a path to any other node in the network with minimum total cost. In this stage, no further assistance is required from the controller. Nodes are now able to perform route discovery purely in a distributed manner. Multiple routes can be established between any two nodes as primary and backup routes based on the lowest cost in case of any link breakage. If both paths are independent, the data traffic can be split between the two routes to deliver the packet in parallel. In the case of heavily congested networks with a high number of nodes, route discovery is more challenging, and split multipath routing helps distribute the load. On the other hand, if multiple paths are stored in each node for any pairs of nodes, then node storage and

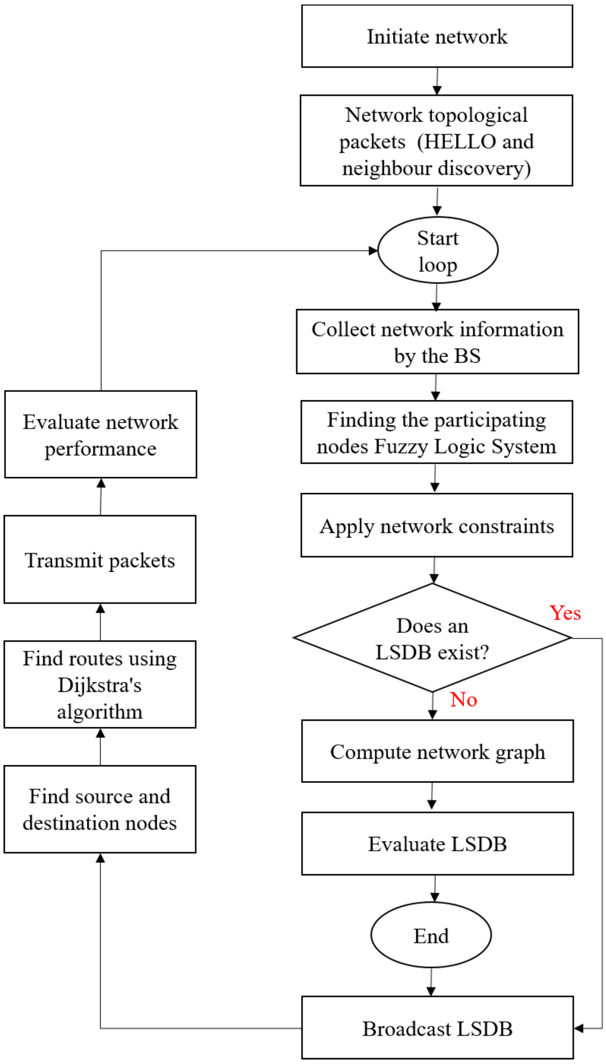


Figure 4.4: Flowchart of the proposed fuzzy-based routing protocol.

network scalability will be another challenge. To further accomplish scalability, we can deploy sub-controllers and divide the network into small cells similar to the idea in [172, 173]. Hence, each node only needs to maintain the routing table of its own cluster. However, this is the point of our future study to monitor the effect of split routing in congested networks.

4.5.4 Route Maintenance

Each relay node is responsible for forwarding packets to the next least-cost hop. After any reception, relay nodes must acknowledge the packet delivery. If a node detects a failure or error during transmission, it will initiate a flow error (FERR) message to the sub-controller containing the error type. If the error type has the link broken message, the sub-controller broadcast a FERR to notify other nodes about the broken link. Once nodes receive a FERR from the BS, any transmission to that particular node will be canceled, and the node ID is removed from the LSDB after the flow error waiting time (FERR-WT) period. FERR-WT is the associated time for deleting any entry and updating the LSDB. This time duration is required if the node is only deactivated for a short period due to bad reception. After this time is passed and nodes have not yet received an active route entry, a new route will be obtained, and the routing table will be updated. At the same time, the previous node of the forwarding path uses the LSDB and applies Dijkstra's algorithm to find the alternative path and transmit the rest of the data packet using the new route. During the route failure, nodes can find a new path due to the availability of the LSDB. However, if a new path can not be found, then there are two possibilities:

- Destination is unreachable: If the node can not find a destination, the sub-controller will receive a destination ID with the data packet.
- Low channel quality: In this case, the sender transmits a quality check message to the receiver to check the possibility of transmission. If not, a new route will be established.

In parallel to LSDB computation, the BS is also responsible for obtaining the eligible nodes through a fuzzy system. Figure 4.3 shows the overall process of the proposed protocol and all the required steps to compute the participating nodes and the routes. A fuzzy system is equipped with rules and can evaluate the eligible nodes. After the fuzzy system is applied, the three-step constraint is performed to ensure the eligible nodes computed by the fuzzy system are satisfactory if any specific condition occurs. The three constraints are further explained in Section 4.7. Then, LSDB is shared among the eligible nodes, and the MD2D communication is established between the active flows. The following section explains how nodes are segregated into two groups: active and deactivated nodes—followed by applying constraints for the final eligibility check.

4.6 Node Participation Mechanism Using Fuzzy-Based Algorithm

The main idea of fuzzy logic is to manage a system with pre-defined rules. The results of fuzzy logic consist of values ranging from 0 to 1. There are four main procedures in fuzzy logic: fuzzification, fuzzy rules, fuzzy inference system, and defuzzification. The fuzzy logic in our study aims to construct the initial stage for the participation of nodes. The fuzzy system captures network information in our framework and computes each node's EI. This index will identify the probability of a node being able to participate in a route.

4.6.1 Fuzzification

The fuzzification process measures the node cost (NC) to estimate the node's EI. The cost is calculated using three decision parameters: remaining energy, mobility rate, and the number of neighbors. These parameters are transformed into triangular fuzzy membership functions, where each input value to the system is mapped to the associated membership function to reveal

Table 4.1: The proposed fuzzy rules.

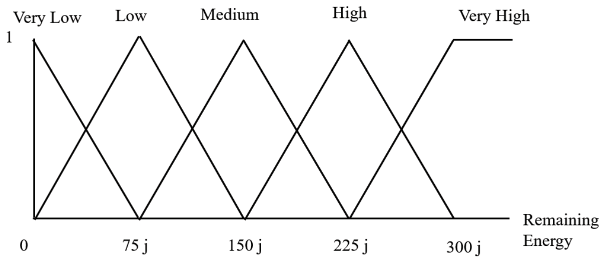
Rules	Remaining Energy	Mobility Rate	Number of Neighbors	Rule Cost	Numerical Representation of Cost
1	Very Low	Very High	Very Low	Very Low	0
2	Very Low	Very High	Low	Very Low	0
3	Very Low	Very High	Medium	Very Low	0
4	Very Low	High	Very Low	Low	0.25
5	Very Low	High	Low	Low	0.25
6	Very Low	High	Medium	Low	0.25
7	Very Low	Medium	Very Low	Low	0.25
8	Very Low	Medium	Low	Low	0.25
9	Very Low	Medium	Medium	Low	0.25
10	Low	Very High	Very Low	Very Low	0
11	Low	Very High	Low	Very Low	0
12	Low	Very High	Medium	Very Low	0
13	Low	Medium	Very Low	Very Low	0
14	Low	Medium	Low	Low	0.25
15	Low	Medium	Medium	Low	0.25
16	Low	High	Very Low	Very Low	0
17	Low	High	Low	Very Low	0
18	Low	High	Medium	Very Low	0
19	Medium	Very High	Very Low	Low	0.25
20	Medium	Very High	Low	Low	0.25
21	Medium	Very High	Medium	Low	0.25
22	Medium	High	Very Low	Low	0.25
23	Medium	High	Low	Low	0.25
24	Medium	High	Medium	Low	0.25
25	Medium	Medium	Very Low	Low	0.25
26	Medium	Medium	Low	Medium	0.5
27	Medium	Medium	Medium	Medium	0.5

the fuzzy degree. Figure 4.4 presents the fuzzy logic flowchart.

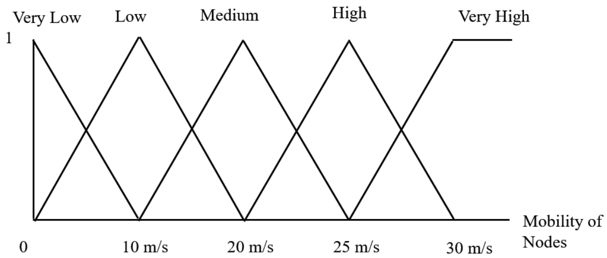
Figure 4.5 illustrates the membership diagrams for obtaining the fuzzy value of remaining energy, mobility rate, and the number of neighboring nodes, respectively. The first input parameter to the fuzzy system is the remaining energy of a node. As shown in Figure 4.5a the membership function consists of five levels, including very low, low, medium, high, and very high. We assume that the energy of nodes is randomly distributed with a maximum value of 300 Joules. The high remaining energy value represents a high-value node, which means that the node with higher remaining energy has a low risk of packet transmission failure. The amount of the remaining energy is converted to linguistic values in one or two possible levels.

As shown in Figure 4.5b the second input parameter is the mobility rate of the nodes. The higher the nodes' mobility, the larger the probability of packet failure is due to link failure or signal fluctuations. Therefore, when the mobility rate of nodes is low, the node's reliability is high, which means there is a better chance of relaying packets to the destination with minimum interruption and packet loss.

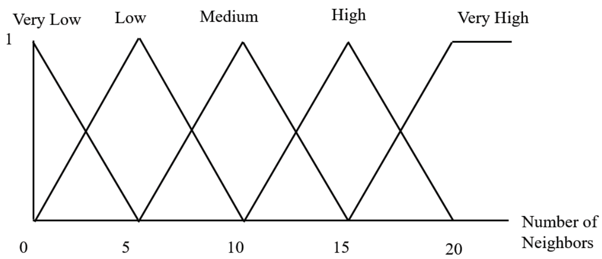
Finally, the third input parameter to the fuzzy system is the number of neighbors, as shown in Figure 4.5c. A node with a higher number of neighbors has a higher connectivity degree and a higher chance of participating in the routing.



(a) The fuzzy membership function for remaining energy.



(b) The fuzzy membership function for mobility rate.



(c) The fuzzy membership function for number of neighboring nodes.

Figure 4.5: Proposed fuzzy diagrams for evaluating the potential candidate for participation.

4.6.2 Fuzzy Rules

Fuzzy rules are obtained using the degree of importance and representation of a value into a meaningful explanation. As previously mentioned, the highest node energy is assumed to be 300 Joules, and if, for instance, a node reports an energy level of 270 Joules, the node is still at a very high energy level. As a result, we can assign each input parameter to the fuzzy system to the corresponding linguistic values in a reasonable belief. As illustrated in Figure 4.5 each parameter, remaining energy, mobility rate, and the number of neighbors are divided into five levels. Therefore, we have three fuzzy parameters leading to 5^3 states. However, there are only 27 rules counted in our system to distinguish the least qualified nodes and eligible nodes for participating in the routing. These 27 rules are demonstrated in Table 4.1, where each rule is associated with a cost. The associated cost values will be explained later in this section.

4.6.3 Fuzzy Inference Engine

After the fuzzification process and converting input values to linguistic values, the values are sent to the inference engine. Together with the rules and the input data, the inference engine forms inferences and draws conclusions. The fuzzy inference system considers all the possible states of the applied rules to evaluate the fuzzy inference output. The Mamdani fuzzy inference system [174] is used in this step to find the fuzzy matching rules and calculate the fuzzy inference output. The system's output is sent to the defuzzification module for final processing.

4.6.4 Defuzzification

In the defuzzification process, the output values of the fuzzy inference engine are used to calculate the crisp value of the EI as a final crisp value of the fuzzy learning system. EI is calculated using (4.1), where $Rule_i$ is computed via the corresponding values of the x-axis into the y-axis. Rule cost, also known as the cost function, is represented by a triangular function. The numerical representation of rule cost C is obtained using Figure 4.6.

$$EI(n) = \frac{\sum_{i=1}^{n} Rule_i C_i}{\sum_{i=1}^{n} Rule_i}, \tag{4.1}$$

Note that to calculate the y-values from the fuzzy functions, some values of the x-axis might cross two points in the y-axis. For instance, for the mobility value of 15m/s, the corresponding y-value is 0.5. On the other hand, for the mobility rate of 17m/s, the y-values intersect at two points; the low-line at 0.4 and the medium-line at 0.6. This process increases the number of states while calculating the EI.

Let us consider an example for node A reporting the remaining energy of 125 Joules, mobility rate as 27m/s, and 4 neighboring nodes. Using fuzzy functions these values correspond to: $R_{E(low)} = 0.3$, $R_{E(medium)} = 0.7$, $M_{R(medium)} = 0.4$, $M_{R(high)} = 0.6$, $N_{N(verylow)} = 0.1$, and $N_{N(low)} = 0.9$. Each parameter corresponds to two values, meaning there are $2^3 - 1$ states to calculate rules. Table 4.2 illustrates the numerical values of rules for each of the states. All three parameters of input (Remaining energy, mobility rate, and the number of neighbors) are now multiplied together. Finally, to compute whether node A can participate, EI is calculated as follows:

$$\begin{aligned} EI_A &= \frac{\sum_{i=1}^{n} Rule_i C_i}{\sum_{i=1}^{n} Rule_i} \\ &= \frac{Rule_{13}C_{13} + Rule_{14}C_{14} + Rule_{16}C_{16}}{Rule_{13} + Rule_{14} + Rule_{16}} \\ &+ \frac{Rule_{17}C_{17} + Rule_{25}C_{25} + Rule_{26}C_{26}}{Rule_{17} + Rule_{25} + Rule_{26}} \\ &+ \frac{Rule_{22}C_{22} + Rule_{23}C_{23}}{Rule_{22} + Rule_{23}}, \end{aligned} \tag{4.2}$$

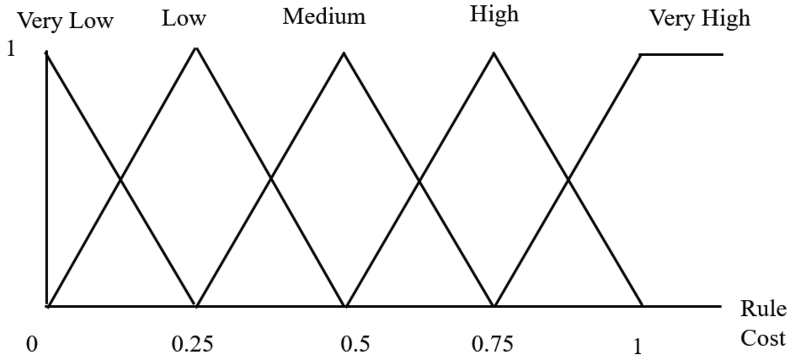


Figure 4.6: Fuzzy diagram for calculating the rule cost of each combined rules from Table 4.1.

Table 4.2: Numerical representation of rules for provided example to calculate EI.

Rules	**Values**	**Results**
$\mathrm{R}_{E(low)} \times M_{N(medium)} \times N_{N(verylow)}$	$0.3 \times 0.4 \times 0.1$	0.012
$\mathrm{R}_{E(low)} \times M_{N(medium)} \times N_{N(low)}$	$0.3 \times 0.4 \times 0.9$	0.108
$\mathrm{R}_{E(low)} \times M_{N(high)} \times N_{N(verylow)}$	$0.3 \times 0.6 \times 0.1$	0.018
$\mathrm{R}_{E(low)} \times M_{N(high)} \times N_{N(low)}$	$0.3 \times 0.6 \times 0.9$	0.162
$\mathrm{R}_{E(medium)} \times M_{N(medium)} \times N_{N(verylow)}$	$0.7 \times 0.4 \times 0.1$	0.028
$\mathrm{R}_{E(medium)} \times M_{N(medium)} \times N_{N(low)}$	$0.7 \times 0.4 \times 0.9$	0.252
$\mathrm{R}_{E(medium)} \times M_{N(high)} \times N_{N(verylow)}$	$0.7 \times 0.6 \times 0.1$	0.042
$\mathrm{R}_{E(medium)} \times M_{N(high)} \times N_{N(low)}$	$0.7 \times 0.6 \times 0.9$	0.378

where logistic values shown in Table 4.2 represent the rules in Table 4.1. For instance, the first row in Table 4.2 corresponds to rule 13 in Table 4.1. After substituting the values, the EI is 0.26. Similarly, the network's EI is evaluated, and the lowest values will represent the nodes with the lowest probability of participating in a route. In contrast, the highest values of EI represent nodes with a higher probability of participation. Network constraints will further process the eligible or active nodes to ensure the network's coverage area is not disturbed and find how long an eligible node can stay active.

4.7 Applied Constraints for the Proposed Fuzzy-based Routing Framework

In this section, the energy model is first presented, followed by the evaluated active time for the eligible nodes. To ensure all active nodes are capable of providing service to the entire network, a coverage constraint is adopted in our system. Finally, the mobility constraint maximizes the active time and guarantees no transmission failure during activation. After applying the regulations, the final decision on the eligible nodes for participation is made.

4.7.1 Proposed Time Frame for Node Participation

In our system, we assume a simple energy dissipation model of radio hardware [143], where the transmitter dissipates energy to send k-bits of packets through a power amplifier and radio elec-

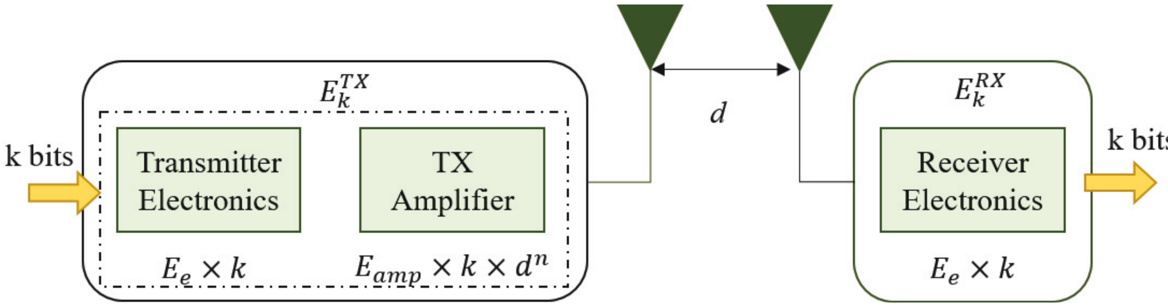


Figure 4.7: Energy consumption model.

tronics, and the receiver dissipates energy to run the received k-bits of a data packet. Then, the energy to transmit the k-bits message is E_{TX}, and E_{RX} is the consumed energy to capture the message at the receiver. As shown in Figure 4.7 the energy expenditure is the energy consumed by the network equipment (antennas, electronic components, etc.) to transmit k-bits packet over a distance d is obtained as follows:

$$E_k^{TX} = \begin{cases} E_e k + E_{amp} k d^4, & if\ d > d_0; \\ E_e k + E_{fs} k d^2, & if\ d < d_0; \end{cases} \tag{4.3}$$

$$E_K^{RX} = kE_e, \tag{4.4}$$

where d_0 is the threshold distance computed by the $\sqrt{E_{fs}/E_{amp}}$, E_e is the power consumed by the electronic devices, E_{amp} and E_{fs} are the amounts of energy per bit dissipated in the RF amplifier.

Based on (4.3) and (4.4), we can approximate that the energy consumed to transmit a packet is almost three times higher than receiving a packet. Considering this fact, let us only consider the highest energy consumption a node can have every second. Assuming that the largest packet size is equivalent to the data rate (R) and distance is less than the threshold, then the maximum consumed energy in every second is equal to:

$$E_{max} = E_e R + E_{fs} R d^2, \tag{4.5}$$

Moreover, the total energy consumption E_T of the network is evaluated using the number of the packets K sent in every flow $F \in \{1, ..., f_n\}$ as follows:

$$E_T = \sum_{k=1}^{K} \sum_{i=1}^{F} E_{k,i}^{TX} + E_{k,i}^{RX}, \tag{4.6}$$

Having the maximum amount of energy a node might dissipate, we can calculate the maximum time t_{max} a node can stay active. Once the participating nodes are discovered, the next step is to evaluate the time they can remain active. We assume that the activation time duration must not exceed more than the $\alpha\%$ threshold of the initial energy of a node. Based on this assumption, the maximum time a node can stay active can be calculated as follows:

$$t_{max} = \frac{E_{RM}\alpha}{E_{max}}, \tag{4.7}$$

where E_{RM} is the remaining energy of the node. For instance, if we assume the remaining energy of a node is 200j, the data rate of a link is 10Mbps, α is 25%, and assuming the worst case scenario maximum range $d = 100$ (Constants: $E_e = 50nj/bit$, $E_{fs} = 10pj/bit/m^2$, and $E_{amp} = 0.0013pj/bit/m^4$), the maximum time the node can stay active is $t_{max} = 33.33s$.

4.7.2 Mobility Constraint

Once the node's maximum activation time t_{max} is calculated, it is a question of whether the node can remain active with minimum transmission failure considering the neighboring nodes and its own mobility rate. Knowing how fast a node is moving and whether movement might impact the initial coverage area and the number of neighbors is crucial for selecting the best performing node to maximize coverage, capacity, and stability. Therefore, it is essential to declare constraints to determine whether a node can still support its neighbors after a particular movement.

This chapter considers a 2-D system with N number of heterogeneous nodes located randomly at positions $[x_i, y_i]^T$, and $i, j \in N$. The distance between each node is defined by their Euclidian distance: $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$. Each node's maximum transmission area is considered a fixed disk with a given range. All the nodes should be divided into two groups; the first group contains the active nodes, and the second group keeps the temporarily deactivated nodes. Let $K = \{(K_1, w_{k_1}), \ldots, (K_m, w_{k_m})\}$ be the set of deactivated nodes, where K_m represents the node ID and w_{km} is the corresponding weight to the closest active node. Now let $A = \{(A_1, t_{a_1}), \ldots, (A_n, t_{a_n})\}$ be a set of active nodes with the corresponding active time frame. Consequently, based on the maximum activation time, we can derive the final activation time as follows:

$$\sum_{l=1}^{L} d_{A_l, A_{l+L}} < max(W_{K_{A_l}}^{A_l}), \tag{4.8}$$

where L is the number of steps a node will take during the activation time, and $W_{K_{A_l}}^{A_l}$ is the weight of a set of deactivated K_{A_l} nodes allocated to active node A_l. Constraint (4.8) ensures that the active node is capable of supporting its neighboring nodes or the deactivated nodes in its neighbor (K_m) during the activation time while it is moving.

Along with checking whether the active node movement does not interrupt the communication with neighbors, the movement of K_m must also be monitored for any possible link failures. Z_k is responsible for the movement of K_m, whether the deactivated neighboring node is moving towards or away from the active node A_l.

$$Z_k^i = \frac{d_{A_i,N_j}}{Range}, \quad i \in n, j \in m \tag{4.9}$$

The movement of nodes directly depends on the mobility rate (v). As shown in Figure 4.8 once nodes start to depart from the initial position, they will move to different indicated areas (shown as rings). Each marked circular ring represents a 20% gap to the initial place, and different transmission ranges represent a different radius. In our study, the transmission range of nodes is limited to 100m. Hence every 20% gap represents a 20m deviation from the initial position. To avoid link failure between neighboring nodes due to movement, the following constraints are established:

$$0 < max(Z_k^i) < 1, A_i : t_{active} = t_{max} - \frac{t_k}{h}, \quad i \in n \tag{4.10}$$

$$max(Z_k^i) \geq 1, A_i : t_{active} = 0, \quad i \in n \tag{4.11}$$

$$t_k = max(Z_k^i) 100 v_{avg}^K \quad i \in n \tag{4.12}$$

$$h = \begin{cases} v_{max}(\frac{v_{max}}{100})^v, & for\ 15 < v \leq 30 \\ -\frac{v_{max}}{2}v^2 + v^2 + \frac{v_{max}}{2}, & for\ 0 \leq v \leq 15; \end{cases} \tag{4.13}$$

To obtain the activation time (4.10) is used, where h is evaluated from (4.13) and v_{max} equals to the maximum mobility rate, assumed to be 30m/s. As shown in Figure 4.9 if the average node mobility rate is between 0-15m/s, the probability of a node staying active is higher. Therefore, as the average velocity of neighboring nodes (K_m) decreases h-index increases faster to maximize the activation time. However, if the average mobility rate of K_m is greater than 15m/s, the activation time of an active node will decrease dramatically. That means the K_m has a higher chance of leaving the active node's coverage area. Hence, the activation time will be minimized to ensure packet loss is minimized during any transmission. t_k ensures the active time duration provides the lowest communication link breakages by considering the weight of the furthest K_m and the average node velocity. (4.11) limits the activation time to ensure deactivated nodes from set K are not out of the transmission range of A. This procedure is conducted at the RIC before finalizing the list of active nodes and allocating the corresponding activation time.

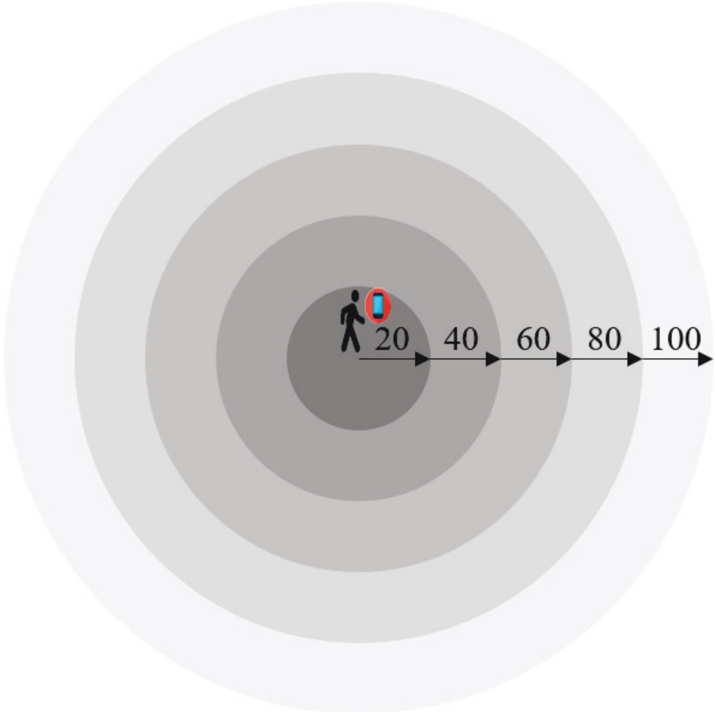


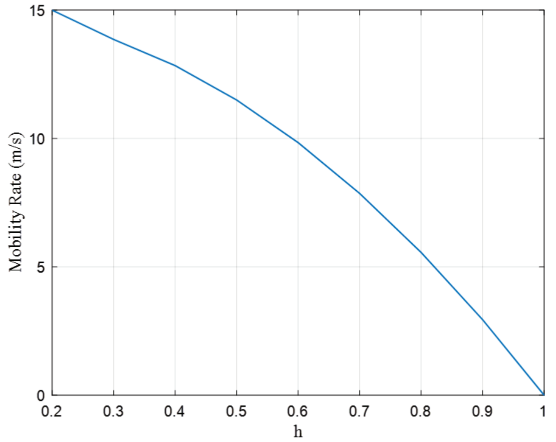
Figure 4.8: Five movement events based on the distance a node can travel over the activation period.

4.7.3 Coverage Constraint

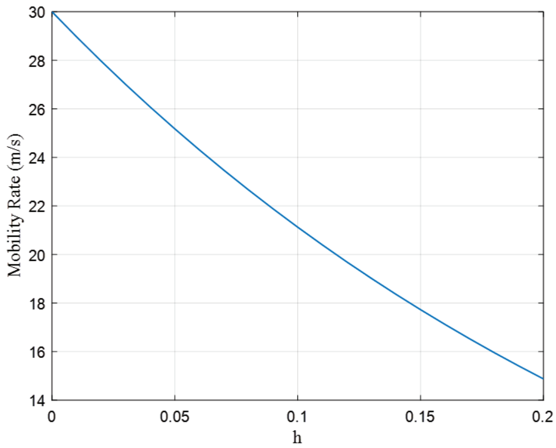
Another essential constraint applied in our routing protocol is the coverage constraint to guarantee the initial network coverage by nodes remains almost the same after identifying the active nodes. After active nodes are separated from deactivated nodes, deactivated nodes will not receive the whole LSDB. We can imply that the deactivated nodes are shut down. Hence, the network's coverage area by all the nodes might be affected when the deactivated nodes are shut down. We introduce a coverage area constraint to solve this problem to keep the final coverage area almost identical to the initial coverage area. In order to formulate the coverage constraint problem a weight (q) is assigned to each node in the network $N = \{(n_i, q_i), ..., (n_i, q_i)\}$. This weight represents the strength of the link between two nodes, which means how much overlapping coverage area two nodes might have. The weight of each node is taken as the ratio of the coverage area, a number between 0 and 1.

$$q_{ij} = \frac{A(N_i \cap N_j)}{A(N_i)}, \tag{4.14}$$

If we assume the coverage area of each node follows a disk model with a fixed radius, $A(N_i \cap N_j)$ is the intersection area between the coverage areas of node i and j, where the intersection is computed as follows:



(a) h-index for mobility rate between 0-15m/s.



(b) h-index for mobility rate between 15-30m/s.

Figure 4.9: h-index for computing the activation time of a node.

$$A(N_i \cap N_j) = \frac{1}{2}r_i^2\left[2\cos^{-1}(\theta_i) - \sin\left(2\cos^{-1}(\theta_i)\right)\right] + \frac{1}{2}r_j^2\left[2\cos^{-1}(\theta_j) - \sin\left(2\cos^{-1}(\theta_j)\right)\right], \quad (4.15)$$

where θ_i and θ_j are defined as follows:

$$\theta_i = \frac{d_{i,j}^2 + r_j^2 - r_i^2}{2d_{i,j}r_i}, \quad (4.16)$$

$$\theta_j = \frac{d_{i,j}^2 + r_j^2 - r_i^2}{2d_{i,j}r_j}, \quad (4.17)$$

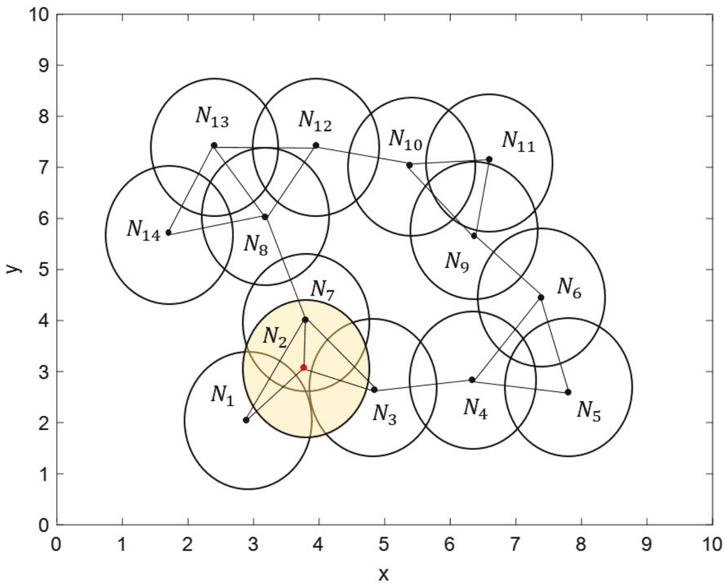


Figure 4.10: A simple illustration of the coverage area of fourteen nodes with their intersections.

r_i and r_j are the transmission range of node i and j, respectively. In our simulation, $d_{i,j}$ is the distance between N_i and N_j defined by Euclidean distance. Finally, $A(BS_i) = \pi r_i^2$ is the coverage are of N_i.

Figure 4.10 provides an example to illustrate the described algorithm. This example presents an area of $10m \times 10m$ with 14 randomly scattered nodes. To calculate the weight of the graph, the weights for all nodes q_i are added.

$$Q_i = \sum_{j \in N} q_{i,j}, \tag{4.18}$$

Based on the weights, nodes are separated into two groups, A' and K', where A' is the set of nodes that must stay active, and K' is the set of nodes with high Q. That means the list of active nodes found in the previous section must now match with set K'. Then, the list of deactivated nodes must match with online nodes A' to check whether any deactivated node must remain online due to the coverage area restriction. In both cases, the node's action will be reversed if there is any mismatch. This technique enforces the selected nodes during the fuzzy learning process to be able to provide service to the entire network. In Figure 4.10 after the weights are all computed, N_2 will have the maximum weight and will be selected as the potential node in set K. If this node is also listed in the fuzzy system deactivated nodes, this node will be selected in the final list of deactivated nodes. Moreover, as observed from Figure 4.10 the overlapping coverage area of N_1, N_3 and N_7 supports the fact that these nodes are the substitutes in that region, and by coverage constraint, N_2 is part of K'. Similarly, this technique is applied to the entire network to ensure full connectivity.

Table 4.3: Simulation parameters.

Parameters	Value
Simulation environment	500m$\times 500m$
Number of UEs	50,100,150,200,250
Packet transmission size	15 kbits
Protocols	FPRM, HSAW, AODV, OLSR
Propagation model	Rayleigh fading
Mobile node transmission range	100m
Mobile node movement model	Random waypoint mobility
Total simulation time	3000s

4.8 Simulation Results

This section presents our proposed routing protocol's simulation setup and performance evaluation. First, the simulation environment with the parameters and assumptions are explained. Second, the performance of the proposed routing protocol is thoroughly examined and compared with other routing protocols.

4.8.1 Simulation Setup

The proposed FPRM protocol is implemented using the Network Simulator-3 (NS-3). NS-3 simulator supports both IP-based and non-IP-based networks, but we are adapting the IP-based network in our simulation. IP-based simulation in NS-3 involves models for long-term evolution (LTE)/5G, WiFi, worldwide interoperability for microwave access (WiMAX), etc., for layers 1 and 2. This simulator provides different testbeds and protocols for users to run and test their proposed frameworks. For instance, routing problems in MANETs can use protocols such as AODV, OLSR, and dynamic source routing (DSR). NS-3 supports several random mobility generators and also SDN-based networks.

Figure 4.11 illustrates our simulation environment expanding over a $500m \times 500m$ area. The BS is located in the center of the network. For simplicity, only one network cell is considered with randomly generated nodes between 50 and 250, with 50 nodes increment in every simulation run. Heterogeneous nodes are considered in this simulation, including mobile nodes, IoT devices, and vehicles. However, all the nodes are specified as UE with different mobility rates. Two separate IP-based networks are set for cellular communication and WiFi communication. The cellular communication band is set to apply the LTE/5G specifications using the NS-3 modules. At the same time, the WiFi band is IP-based with IEEE standards. Simulation parameters are shown in Table 4.3. The simulation results were run for 3000s, and each simulation was averaged over multiple simulations running. We use the Monte Carlo simulation technique under 50 runs to validate our results, and the final results are averaged and data plotted with 95% confidence intervals.

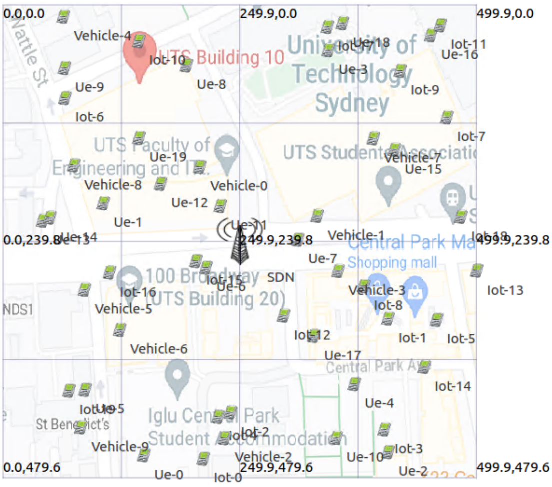


Figure 4.11: Simulation terrain setup.

4.8.2 Simulation Analysis

This section provides the simulation results of the proposed fuzzy-based node participation routing using the NS-3 simulator. Three types of nodes with different mobility rates and power are considered to evaluate our proposed routing protocol. First, we assumed that 40% of the network is filled with mobile nodes, 40% with IoT devices, and 20% vehicles. Nodes' power is randomly distributed between 0-300 joules (the value of power can be changed based on the standards; however, we chose this value to see the impact of depleted nodes). Nodes' velocities are randomly selected between 1-30 m/s. Two channels are considered for data communication, cellular and WiFi. For WiFi or MD2D communication, IEEE802.11n-5GHz is used, and LTE/5G broadcast control channel (BCH) is utilized for cellular communication. We assume that the link quality between nodes is updated every 1 second. In every scenario, the number of active nodes is increased to analyze the performance of the proposed routing protocol. MD2D routing has an average length of 4 hops between any source and destination. We compared our proposed routing framework with our previously proposed routing protocol, HSAW, and two conventional MANET routing protocols: AODV and OLSR. The signal propagation model is Friis free space model [144]. The simulation analysis generates random source and destination nodes. Finally, the assumptions taken during the simulation scenarios are introduced as follows:

- Assuming a random velocity rate from 1-30 m/s second is allocated to each node.
- The network consists of 40% pedestrian, 40% inner-city mobile nodes, and 20% outer-city mobile nodes.
- We consider a simple energy model for our system, randomly distributed among nodes between 0-300 joules per UE. This value is assigned to calculate the amount of consumed energy during the packet transmission process and realize the effect of energy consumption.

This study's main objective is to introduce a new fuzzy-based routing participation protocol to increase network lifetime, PDR, and throughput. The aim is to use fuzzy logic to identify nodes with the least capabilities to stay deactivated for a specific time. Network constraints are used to check the participant nodes regarding activation time, network coverage, and mobility patterns.

Figures 4.12 and 4.13 illustrate the number of depleted nodes and the energy consumption, respectively. FPRM shows small improvement over HSAW in network lifetime. There are two reasons why FPRM has a longer lifetime than HSAW and two other conventional routing protocols. First, in HSAW, the entire nodes in the network receive the LSDB, which causes significant energy consumption compared to FPRM. Second, FPRM keeps track of nodes and their energy consumption, and if the energy of nodes passes a certain threshold, the controller stops the node from participating in the routing. FPRM prevents unintentional and unnecessary energy consumption. Therefore, in long-term scenarios, more nodes will be active in the network than in the HSAW. Furthermore, compared to the conventional ad hoc routing protocols, such as AODV and OLSR, FPRM performs much better.

Figure 4.14 represents the average total throughput of the entire network, including the cellular and MD2D communication channels. Our proposed routing protocol performs significantly better than the three other routing protocols. This is because not all the nodes are active in the network, making the controller use the maximum available throughput. Therefore, the average network throughput increase compared to HSAW, AODV, and OLSR.

Figure 4.15 shows the E2E delay of FPRM, HSAW, AODV, and OLSR. FPRM and HSAW perform similarly at low network density, but FPRM accomplishes a better E2E delay once node density increases. In FPRM, nodes have more reliable links because of the active nodes in the network. As described in previous sections, active nodes have maximum energy, lowest mobility rate, and maximum throughput. Therefore, packets are transmitted over more reliable links faster than HSAW. Compared to traditional ad hoc protocols, namely AODV and OLSR, both HSAW and FPRM archive better results. However, FPRM has superior performance overall.

Figure 4.16 shows the PDR of the network, where FPRM achieves significantly higher PDR compared to HSAW, OLSR, and AODV. FPRM has slightly better PDR performance than OLSR and AODV in low-density networks but considerably higher in more congested networks. FPRM has the lowest packet failure due to the fuzzy participation algorithm. In the participation technique, active nodes are selected based on their performance. Active nodes are responsible for routing the packets, and as long as their performance stays high, the packet loss will be minimum.

4.9 Conclusion

This chapter proposed a new joint participation and routing protocol using a fuzzy-based routing framework called FPRM with mobility, energy, and coverage constraints. A new topology control mechanism was presented using the fuzzy-logic-based approach to identify participating

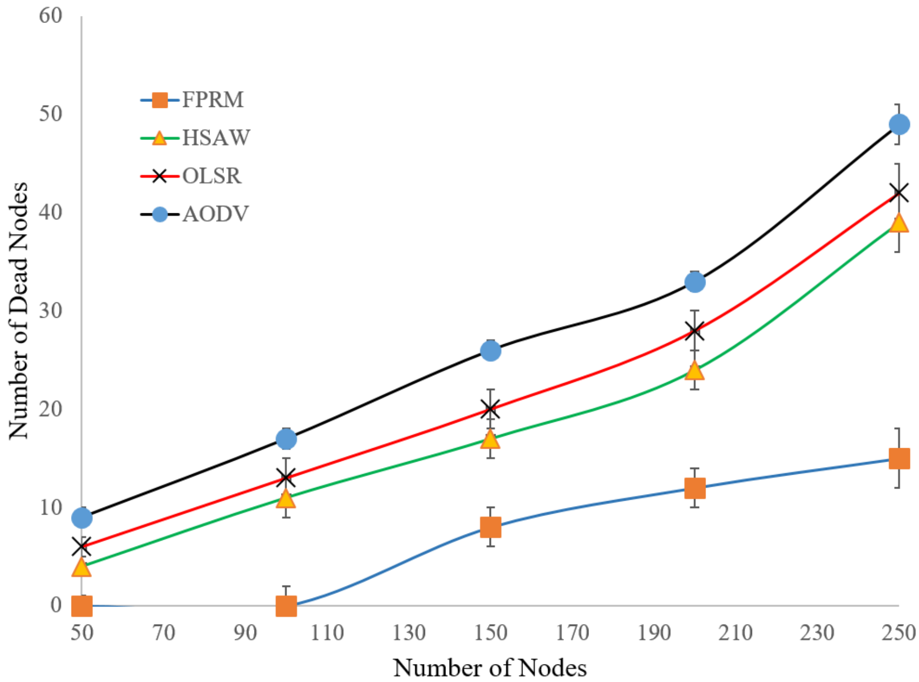


Figure 4.12: Number of depleted nodes.

nodes and create different network graphs. Our routing protocol uses the information collected from the network and application layers as knowledge. This knowledge is used to create an LSDB. In our approach, an O-RAN intelligent controller can be used to separate the control plane decision from the data plane. The controller is responsible for creating various LSDBs based on the network information and application requirements. The data plane is only responsible for relaying the data from one end to another. The controller only shares the LSDB with the participating nodes for MD2D routing to reduce cellular channel overhead and energy consumption. Any node with information to transmit is capable of processing the LSDB and obtaining the most efficient route to the destination. The simulation results show that the FPRM protocol is superior in network lifetime, E2E delay, PDR, and throughput than HSAW. Moreover, our protocol significantly improved performance compared to the purely distributed benchmark routing protocols, AODV and OLSR.

This study's future direction is to explore the timing mechanism of how often the topology control update is required to trigger and optimize this time. Moreover, LSDBs can be optimized by machine learning algorithms to contain only a specific form of knowledge based on the topology graphs.

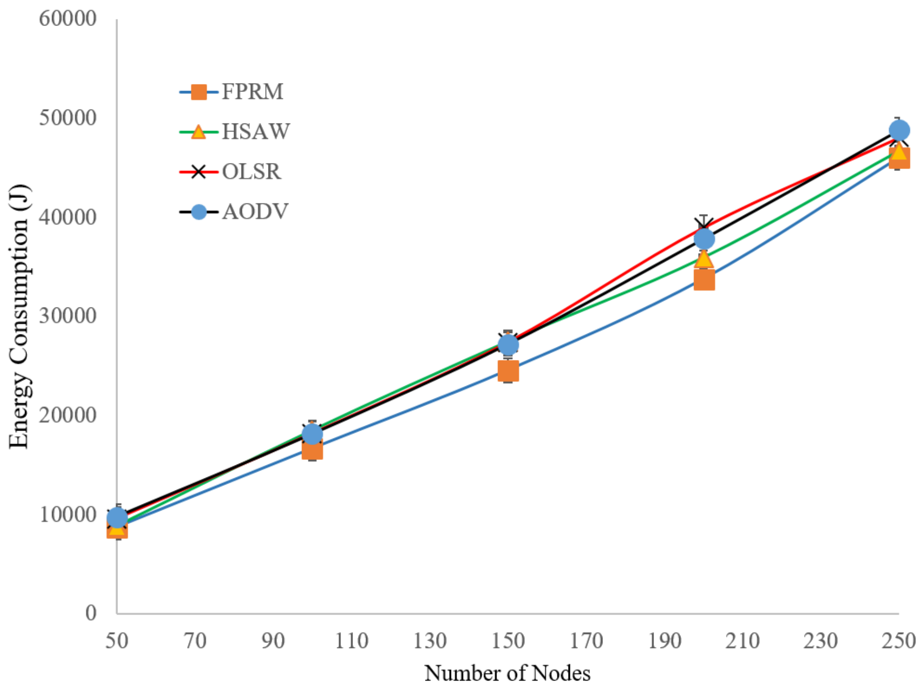


Figure 4.13: Nodes energy consumption.

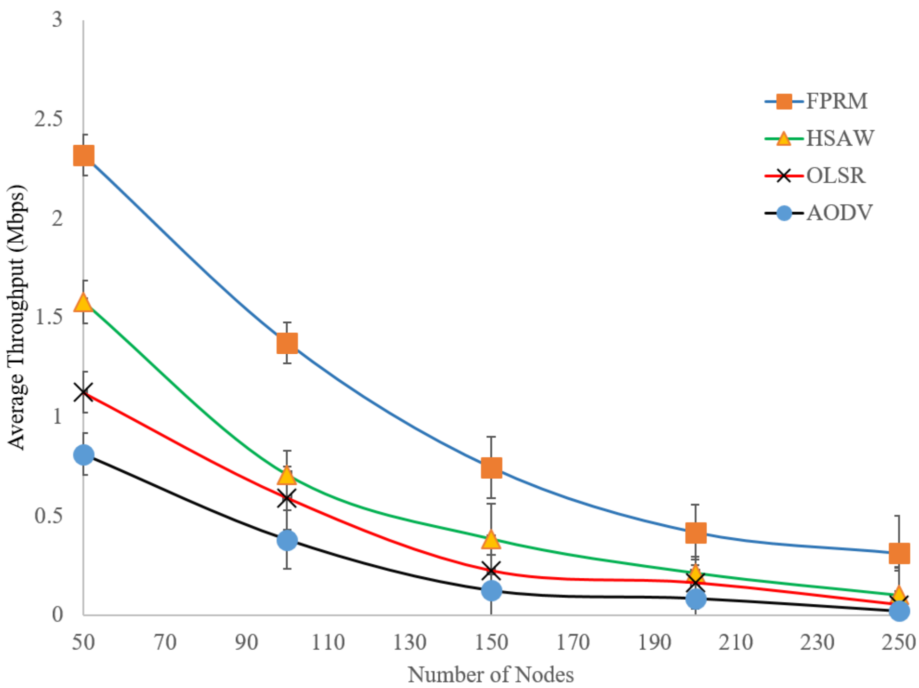


Figure 4.14: Average network throughput.

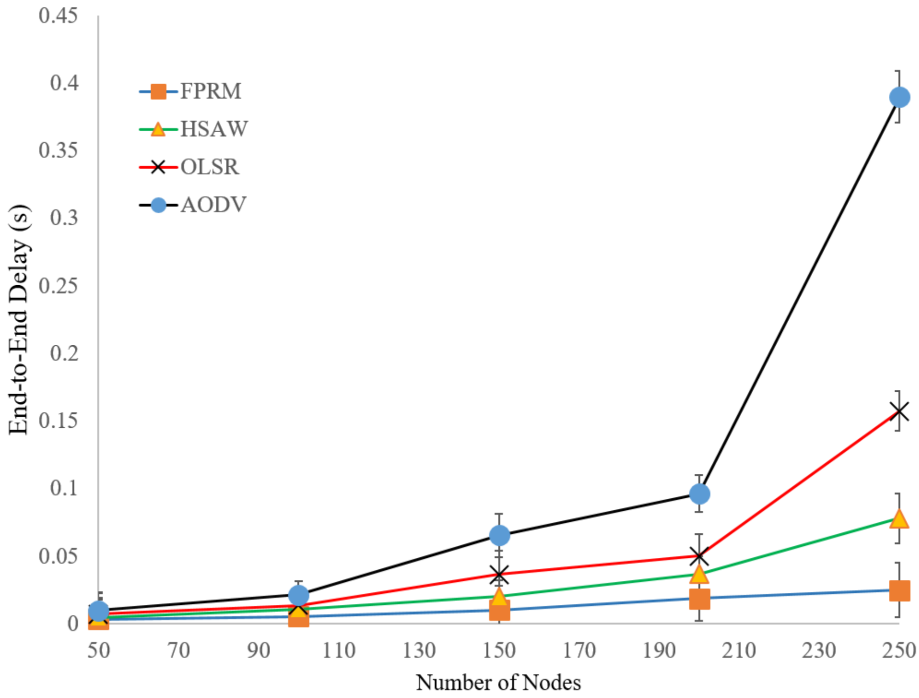


Figure 4.15: End-to-end delay.

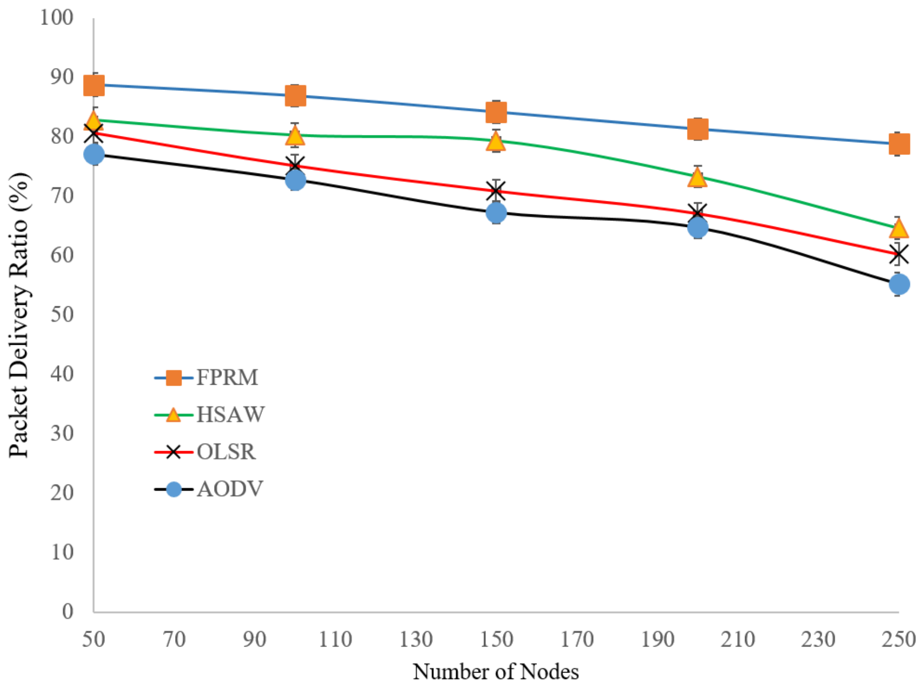


Figure 4.16: Packet delivery ratio.

5

A Cross-Layer Multi-Hop Device-to-Device Routing Protocol for Future Heterogeneous Wireless Cellular Networks

5.1 Overview

The upcoming research is submitted to IEEE Internet of Things Journal [1]. This chapter's main objective is to propose a new joint utility-based routing protocol called application-driven cross-layer MD2D routing protocol (ACMRP). We use the O-RAN paradigm to create network intelligence by incorporating application and network layer knowledge. Our protocol finds the optimal route depending on the user application and requirements. The proposed protocol shows superior performance compared to other routing protocols.

5.2 Introduction

Traffic congestion in cellular networks is a growing problem as the demand for data services continues to increase. With more and more people using smart devices, the amount of data being transmitted over cellular networks is also increasing [1]. The rise of the Internet of Things

[1]Ashtari, S., Abolhasan, M., Lipman, J., Ni,W. (2023). A Cross-Layer Multi-Hop Deviceto-Device Routing Protocol for Future Heterogeneous Wireless Cellular Networks. IEEE Internet of Things Journal.

(IoT) has led to an increase in traffic congestion in cellular networks [175]. IoT devices, such as smart home appliances and wearable devices, are constantly sending and receiving data, leading to an increase in network traffic. This increase in traffic puts pressure on the existing cellular network infrastructure and can lead to network congestion, resulting in slower data transfer speeds and reduced network reliability. To help alleviate traffic congestion, cellular providers may implement various strategies such as deploying more cell towers [176], upgrading existing infrastructure, and managing network resources more effectively [177]. Additionally, the development of new technologies such as multi-hop device-to-device (MD2D) communication, which is designed to increase network capacity and reduce latency, may also help to alleviate traffic congestion in the future [178].

MD2D communication is a method of wireless communication in which devices communicate directly with each other without the need for a central intermediary such as a router or a base station [179]. This type of communication can be used to improve network efficiency, reduce latency, and increase overall network capacity. Examples of MD2D communication include peer-to-peer file sharing and direct mobile-to-mobile communication [3]. MD2D communication can also be used in a variety of other applications such as gaming, social networking, and location-based services. Routing protocols are an important aspect of MD2D communication, as they help to ensure that data is transmitted efficiently and reliably between devices. Routing protocols for MD2D communication are responsible for determining the best path for data to travel from one device to another.

5.2.1 Motivations

The motivation behind MD2D communication is to improve the overall efficiency and capacity of wireless networks [180]. It is important for cellular network providers to address the challenges associated with increasing demands from IoT devices. These devices are constantly transmitting and receiving data, leading to a significant increase in network utilization. By allowing devices to communicate directly with each other, MD2D communication enables devices to communicate directly with each other, reducing the need for data to be routed through a central BS. This can lead to several benefits, including increased network capacity, reduced latency, improved energy efficiency and improved coverage [3]. In addition to these technical benefits, MD2D communication can also enable new and innovative applications such as peer-to-peer file sharing, location-based services, and social networking.

There are several routing protocols that have been proposed for MD2D communication, each with their own strengths and weaknesses [79]. Among them utility-based routing protocols are used to determine the best route based on the current network conditions [181]. Utility-based routing protocols can be used to optimize different network objectives, such as energy Utility function are used to integrate the collected information. The use of utility functions in MD2D routing protocol can enable several key novelties compared to traditional routing protocols.

First, utility functions can be used to adapt the routing decisions to the current network conditions, such as such as available bandwidth, E2E delay and energy level. This can lead to more efficient use of network resources and improved network performance. Second, it can be used to enable self-organization of the network, allowing devices to dynamically adapt to changes in the network conditions and to form new routes as needed. This can lead to more robust and resilient networks. Moreover, utility functions can be used to take into account the energy constraints of devices when making routing decisions. This can lead to more energy-efficient networks and extend the lifetime of devices. This paper uses a utility-based routing strategy to a formulate multi-objective cost function by ensuring efficient, reliable, secure, scalable, and energy-efficient communication between devices.

To the best of our knowledge, no MD2D routing protocols have been proposed to jointly consider the application and network layer information. Existing MD2D routing protocols use network layer information only to create stable routes. This paper proposes a routing protocol that leverages cross-layer information to create multi-objective cost function using utility metrics. The release of the open-radio access network (O-RAN) has created new opportunities to build and generate new routing protocols with global knowledge and intelligence across the medium. MD2D routing has not been given consideration in the literature of the emerging centralized-based O-RAN paradigm, despite the commercially increasing popularity of O-RAN. The key challenges are the heterogeneity of applications and the massive increase in heterogeneous devices. Different applications, such as vehicular networks, IoT, and smart devices, require different resources and management techniques. This paper is motivated to address these challenges and allow MD2D routing to be driven by applications and cross-layer knowledge in an O-RAN environment.

5.2.2 Contributions

This paper introduces a utility-based routing protocol called the Application-Driven Cross-Layer MD2D Routing Protocol (ACMRP), which adjusts routes based on both application requirements and packet-level information. ACMRP leverages application information and historical network statistics to optimize energy consumption, latency, throughput, and packet delivery ratio (PDR). As depicted in Figure 5.1, a Mobile Edge Computing-Enabled Base Station (MECBS) collects information from various nodes, including vehicles, mobile phones, and Internet of Things (IoT) devices. This information consists of remaining energy, mobility, locations, and data rates of nodes. The collected infromation is processed to create multiple link-state databases (LSDBs). We use the concept of software-defined networking (SDN) and network function virtualization (NFV) to create a centralized network function where it produces multiple LSDBs that can be used to evaluate different route strategies. The knowledge-based entity is responsible for creating the master LSDB (MLSDB) using utility functions and the collected cross-layer information. The stored knowledge in MLSDB includes LSDB-E, LSDB-M,

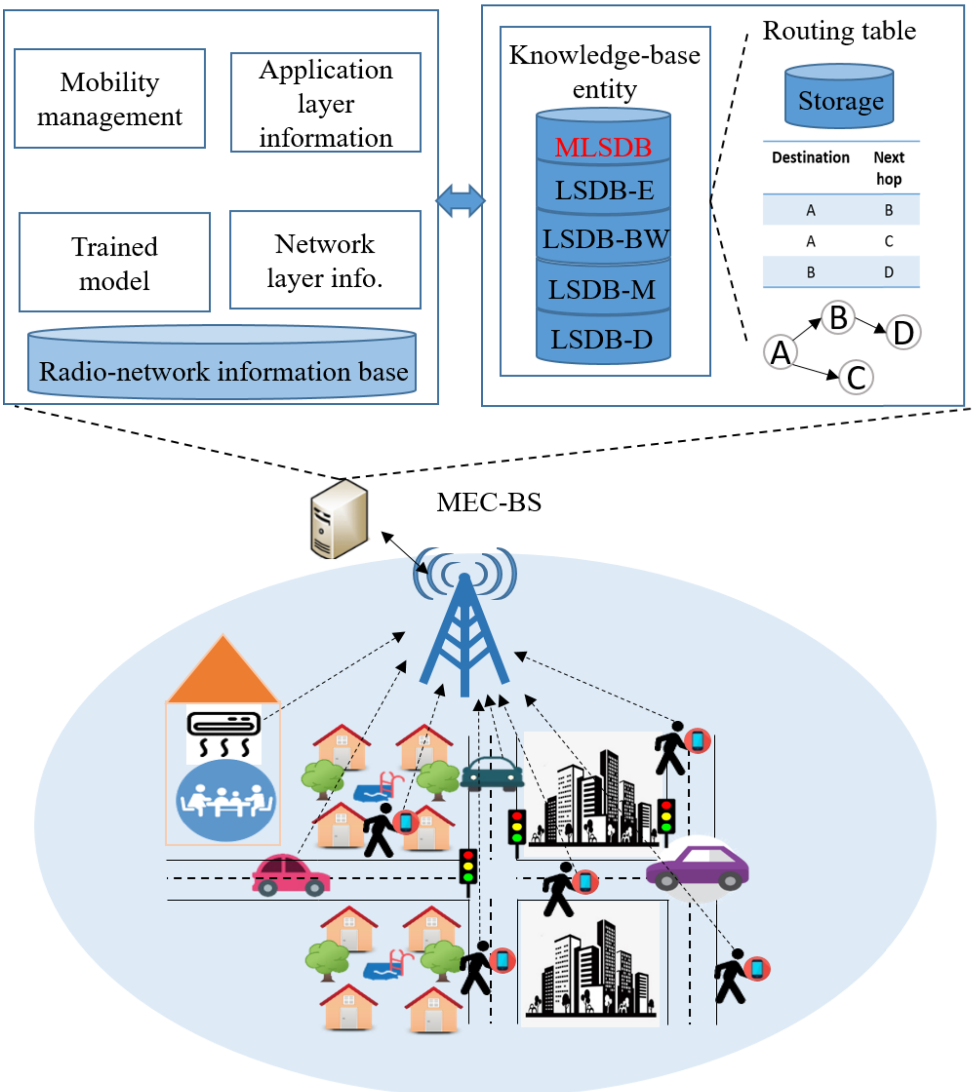


Figure 5.1: Modular design for Knowledge-based cross-layer routing framework for heterogeneous networks, where application requirements and network analytics are collected as a form of knowledge in LSDBs.

LSDB-D, and LSDB-BW, corresponding to energy, mobility, distance, and bandwidth, respectively. Using the utility function our proposed routing protocols can adapt to the current network conditions and make dynamic routing decisions based on application requirements. A multi-objective cost function selection-based algorithm is used to make efficient routing decisions.

Following are the key contributions of this chapter.

- We propose a cross-layer MD2D routing framework that leverages upper-layer information and packet-level data to create intelligent LSDBs and determine the most efficient route for various applications.

- We propose an adaptive routing algorithm, which utilizes utility-based functions with multi-objective cost functions to provide dynamic route selection based on the application and user requirements.

- We present a three step cost function to maximize the network performance. For instance, energy constraint is used to ensure the nodes' energy levels are acceptable for MD2D communication, channel constraint is deployed to guarantee the channel is ideal during transmission and packet delivery constraint is provided to keep the PDR level consistent in different network environments.
- We implement our algorithm in NS3 and conduct simulations to compare its performance to existing SDN-based and distributed routing protocols. Our results indicate that our algorithm outperforms the state-of-the-art HSAW protocol with a 10% increase in throughput, 13% reduction in end-to-end delay, 5% reduction in energy consumption, and 7% increase in packet delivery ratio.

The rest of this chapter is organized as follows. Section 5.3 provides a brief overview of current studies in the literature. Section 5.4 introduces the architecture and topology discovery in our proposed routing framework. In Section 5.5 detailed explanation of the proposed ACMRP is provided. Finally, in Section 5.6 a comprehensive simulation analysis of the proposed routing protocol is performed. Our conclusion and future research directions are clarified in Section 5.7.

5.3 Related Works

Several proposed MD2D routing protocols exist in both distributed and centralized frameworks [79, 182–184], and centralized protocols have shown superior performance to the distributed protocols [3]. Distributed frameworks use flooding algorithms to establish a route from source to destination. Centralized routing frameworks use a centralized controller such as an SDN controller to assist routing packets. Centralized routing protocols increase scalability and reliability by eliminating the need for multi-hop flooding in route discovery, thereby avoiding broadcast storms [79]. This section investigates related work and summarizes existing MD2D routing protocols' advantages and challenges. Then, the benefits of utility-based routing protocols are discussed.

By using the distributed MD2D routing protocol, all user equipment (UE) in the network can construct routing tables without the help from a centralized controller. Authors in [185] outlined the advantages of MD2D using a weighted sum of interference impact and signal to interference and noise ratio (SINR) called MIIS (metric for interference impact and SINR). Distributed routing algorithms are based on minimum SINR and interference to create the list of participants and generate MD2D routes. In [186] authors showed the significance of mmWave MD2D routing by proposing a multi-hop relay probing scheme. In this scheme, nodes will collect the received signal strengths (RSSs) and estimate the signal-to-noise ratio (SNR) while considering line-of-sight (LOS) and non-LOS (NLOS) signals. In [93], the authors demonstrated the on-demand centralized routing as an enabler for flexible and efficient MD2D networks. Any request from a UE is forwarded to relaying nodes by the SDN-based BS. Source nodes then transmit packets to destinations using next hop information. Similarly, authors in [86] provided insights into the significance of a centralized MD2D routing protocol. This protocol decreases the cellular channel overhead by only forwarding the routing information to the source node. In [187], network layer information, such as link quality, the number of hops, throughput, and route failure probability is used to demonstrate the advantages of MD2D routing in virtual mesh clustering. Using a mesh routing algorithm, relaying nodes are assigned based on user mobility and link failures. In conclusion, centralized MD2D routing protocols enhance the network's performance. However, most of the current proposed MD2D routing protocols [79,86,188–190] focus on the network telemetry information, but application requirements and historical data are also equally critical for overall network performance maintenance.

In other heterogeneous networks, including IoT and vehicular, there has been similar research interest in introducing new MD2D routing protocols with the coordination of a centralized controller. In [191] authors investigate optimal routing based on the trusted connectivity probability of IoT devices. A decoding messaging technique is designed to secure packages in a centralized BS to manage the routing and communication between devices. The algorithm's main challenge is using channel state information (CSI) to identify the route and relay nodes. Channel information is highly dynamic and can not be the only factor in routing decisions. There must be an intelligent network application to capture channel information and prescribe efficient routing policies. Authors in [192] used an unmanned aerial vehicle (UAVs) as a flying BS to provide coverage for IoT devices. The algorithm uses shortest-path-routing to determine MD2D routes. By considering only the shortest path to generate a route, their algorithm could adversely affect network performance. The authors in [193] presented an MD2D routing protocol for a vehicular network to maximize the data rate. Multiple clusters are established using a BS and communication between vehicles to optimize resource utilization. While creating routing in MD2D networks, most proposed protocols fail to take into account an essential factor. As the environment is dynamic and network devices are versatile, it is essential to assign an intelligent network application to monitor and adjust routes. O-RAN architecture offers enhanced features for managing and building intelligence and storing refined data. Together with real-time data, this non-real-time information can produce highly efficient outputs that can enhance overall network performance.

One successful technique for generating routing protocol is utilizing utility functions. Utility functions have long been used across different domains because they can provide knowledge about the past and draw conclusions about the current state [194–199]. Authors of [200] used the SDN paradigm to address energy efficiency issues. The proposed algorithm selects a ratio for energy saving in SDN (RESDN) to distinguish the amount of energy each path consumes using a link utility. The utility of each link is used to create a link utility-based energy efficiency metric that continuously quantifies efficiency and performance in terms of energy consumption.

Authors in [201] provided insights into the importance of utility-based routing protocol to calculate the next hop utility score to decide the next forwarding hop. As a result, a new routing protocol for delay tolerant network (DTN) is developed for nodes to exchange information over multiple hops. Authors in [202] investigated the importance of geographical routing protocol using the utility of traffic load in vehicular ad hoc networks (VANETs) to find one-hop neighboring information. The routing algorithm facilitates the utility functions to acquire knowledge of two-hop neighbor links using packet loss rate and residual bandwidth.

Hence, utility-based routing protocols are capable of improving network efficiency and flexibility. In utility-based routing protocols, historical and essential knowledge is provided to expand the local view of the topology of the network. However, to the best of the authors' knowledge, most MD2D routing protocols only use network metrics for introducing a new routing strategy without considering the application requirements and historical data (knowledge). A commonality among these routing strategies is using a predetermined routing technique to optimize energy, throughput, E2E delay, or PDR. If only one or two network metrics are taken into consideration when discovering routes, the performance of the network may deteriorate over time. Paths must be able to be adjusted in accordance with traffic type, application requirements, and network statics to enhance network performance.

5.4 Proposed Architecture

This section presents a detailed explanation of our proposed routing architecture for heterogeneous wireless networks. The ACMRP framework utilizes application requirements and network telemetry to create MD2D routes. A set of cost functions translates the knowledge into different LSDBs to provide various paths based on application requirements. As an example, Figure 5.2 shows possible paths taken to forward a packet between the source node (A) and destination node (I) using different utility functions. These paths are dynamically selected based on the application requirements and network dynamics to provide maximum efficiency. In the remainder of this section, we describe our proposed protocol and its architecture.

5.4.1 ACMRP Architecture

The most important factors that distinguish ACMRP from other MD2D or legacy routing protocols are as follows. First, our proposed protocol uses a centralized controller to manage the network traffic and prescribe routing tables. Second, integration of the application requirements and network telemetry together to create an MLSDB, and third mitigating the greedy-based property of routing strategies by using four sub-LSDBs (as explored in the literature review, many of the existing routing protocols rely on one or two optimization parameters, e.g., energy, E2E delay, etc.). Finally, we identify the most suitable route based on the traffic type using the MLSDB.

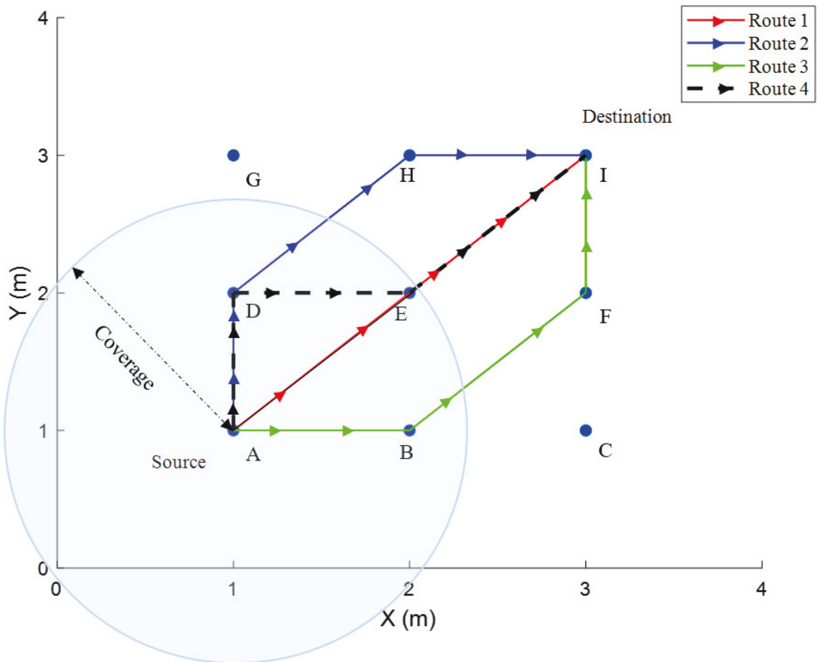


Figure 5.2: An illustration of the route discovery operation based on different cost metrics from source node A to destination node I.

Figure 5.3 illustrates the ACMRP architecture, where many nodes are scattered across a large geographic area. A controller is at the edge of the network with the ability to collect network information. This can have several advantages over distributed MD2D routing protocols, where each device manages its own routing. Centralized MD2D routing protocols can support a larger number of devices with better scalability. This is because a central entity can manage the routing of data for all devices, rather than each device having to manage its own routing. Moreover, the centralized controller can make more informed routing decisions based on the global network state, rather than each device having to make its own routing decisions based on local information. In our protocol, the controller is responsible for processing network and application layers information and evaluating the most efficient MLSDB. The generated MLSDB is broadcast into the network, enabling nodes to discover their own route in a distributed manner. Therefore, nodes can evaluate a new path without the controller's intervention in case of link failure.

In this chapter, traffic flows are classified into four different categories: mission-critical (MC) data, non-critical (NC) data, high definition (HD) multimedia, and high definition real-time (HDR) multimedia. MC traffic requires a fast and reliable link (hence, the shortest distance and lowest mobility). NC data does not need a fast route or mobility constraints. Therefore, for this type of data, we can increase the network's lifetime. HD requires downloading a large amount of data, which requires bandwidth and link reliability. HDR packets, such as online gaming and online streaming, need a high QoS, which incorporates bandwidth, energy consumption, and link reliability. In our framework, each node computes a route based on one of the above traffic

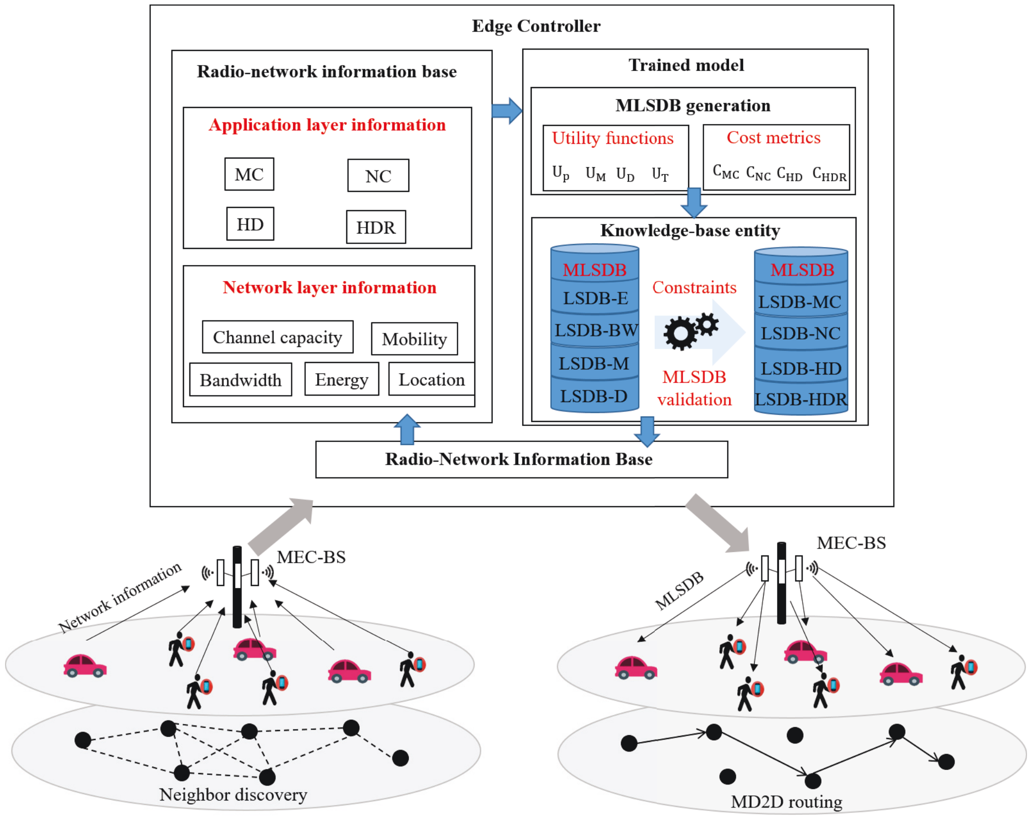


Figure 5.3: Detailed illustration of the framework for the proposed ACMRP routing protocol.

types. Once the source node identifies the traffic type, the appropriate sub-LSDB is chosen, and the route is obtained using that particular sub-LSDB.

Each node is equipped with licensed and unlicensed frequency bands for cellular and WiFi communication, respectively. A licensed frequency band is employed in cellular communication to exchange data traffic and control information between users and sub-controllers. The unlicensed WiFi frequency band is utilized for MD2D communication for forwarding data and control information between distributed nodes. Each distributed node plays two main roles: a forwarding element from the sub-controller point of view and an end-user from a cellular perspective.

5.4.2 Route Discovery in ACMRP

Figure 5.3 also illustrates the routing discovery in the ACMRP protocol, where there are three main stages; node condition reporting stage, the centralized route computation stage, and finally, the distributed route configuration. The detail of each step is presented below.

- Node condition reporting (NCR): The node condition reporting is the stage where nodes report their condition changes and network information to MBS. Network information consists of the remaining energy, bandwidth, channel capacity, mobility, and location of nodes. Other technical information is shared with the controller, including the number of neighboring nodes, queue length, and traffic arrival rate. Two separate actions are performed after receiving node conditions. First, the controller calculates the MLSDB. Second, the condition data is shared with the main controller for storage and advanced processing.

- Centralized route computation (CRC): Once the controller receives the control information from all the nodes, it will start computing the MLSDB (details provided in Section IV). It is assumed that the controller has enough storage for storing and maintaining the LSDB data. The final MLSDB in our protocol consists of four LSDBs: LSDB-MC, LSDB-NC, LSDB-HD, and LSDB-HDR. Once the MLSDB is computed, it will be sent to all the nodes in the cell.
- Distributed route configuration (DRC): Each node will perform a distributed active route configuration after receiving the MLSDB. Specifically, each node generates a routing table based on the minimum total weight of the traffic type from a node to any other node in the network, with no further assistance from the controller. Once the table is completed, nodes can transmit and receive data. For MD2D transmission IEEE 802.11 [203] WiFi protocol will be used.

5.4.3 Exchanged Packets

Generally, there are two packet types: data and control packets. The data packets are the information or the data needed to be transmitted or received. The control packets are used to maintain the structure of the network. There are three main control packets exchanged between nodes and controllers:

- The neighbor discovery packet between nodes: This packet is broadcast by each node and is only processed by the one-hop neighbors in the transmission range of a node. The interface information is generated as a Hello packet in our protocol by each node independently. As a result, each node will identify its neighbors, create a table, and share it with the sub-controller.
- The neighbor discovery packet between controllers: This packet is only generated by sub-controllers in the network. Each adjacent sub-controller generates a Hello-C packet to exchange their ID (Controller-ID) and routing information to be used when handover occurs.
- The topology control (TC) packet: Initially, when nodes come online, they must register their IP with the MBS, which allows the MBS to maintain a list of authenticated and authorized nodes in its network. This registration process is critical due to security concerns and the potential presence of malicious nodes, which must be authenticated before they can participate in MD2D communication. The authentication is performed through the TC Packet, where every new node must receive authorization from the MBS. In the ACMRP protocol, once nodes have been authorized to participate in MD2D, the sub-

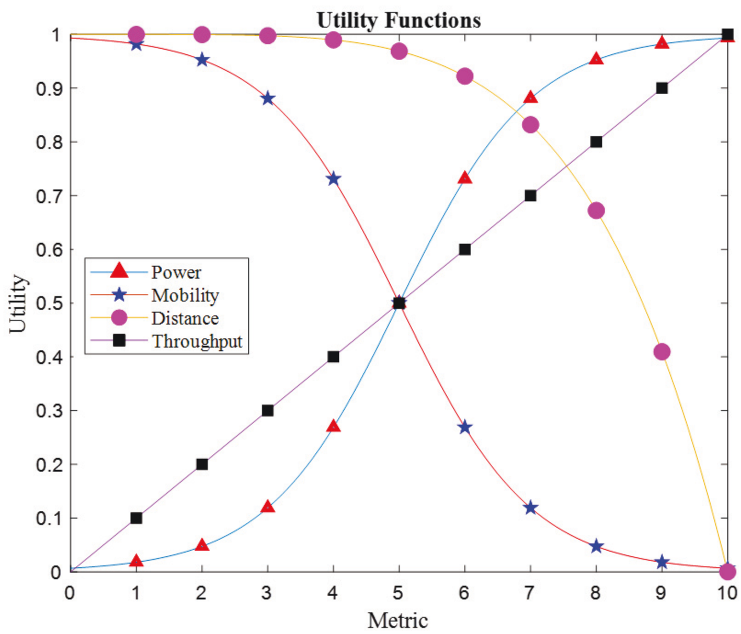


Figure 5.4: Illustration of four utility functions for remaining energy, distance, mobility rate, and throughput.

controller sends a Topology Control Request (TCREQ) to the nodes and requests their Node Information Base (IB). IBs are exchanges as unicast packets that also includes node interface information, link quality, remaining power, link capacity, and mobility rate.

5.5 Proposed Routing Protocol

There are two main steps for route discovery in ACMRP: The first is to generate the MLSDB at the controller using the metrics and utility functions. Second, is to broadcast MLSDB to nodes for final route discover and route validation using network constraints. In the second step, nodes apply the idea of generating a path to the destination based on the application requirements.

5.5.1 MLSDB Generation Phase I

Each controller keeps two different LSDBs: The intra-cell-LSDB and the inter-cell-LSDB. The intra-cell-LSDB includes the local cell information and topology, while the inter-cell-LSDB contains the details of adjacent cells. After Hello packets are exchanged between nodes and the inter-cell-LSDBs are created at each controller, adjacent controllers will exchange their inter-cell-LSDBs to update their intra-cell-LSDBs. Hence, a controller keeps the information of the existing nodes of its cell and the adjacent cells. Another important parameter that controllers and their neighbors share is the coverage area of the routing protocol, which allows the protocol to limit the communication and the route between nodes in a cell or in-between cells.

To build the MLSDB, the controller calculates the utility of each node based on the IB data. There are four IB metrics: energy, throughput, mobility, and distance. A utility function is calculated using the reported IB value and the associated utility function. The utility functions for each utility metric are depicted in Figure 5.4. The proposed utility functions are retrieved and modeled from various sources [194, 196, 204]. The x-axis represents the reported normalized IB values from a node, and the y-axis shows the associated utility value. Note that all the IB values are normalized and scaled between 0-10. Normalizing the IB values can make it easier to compare, combine, and visualize values that are measured using different units or scales, and is a requirement for our protocol later when we combine the utility functions to create multi-objective cost functions. To represent network metrics, different utility functions are utilized. For instance, sigmoid functions are used for power and mobility metrics because this function gives a flat starting point and ending point [205]. In the middle, it increases very quickly to incentivize the power management in the active regions and slow it down when power is excessively high or excessively low. Moreover, we use the linear function for a simple representation of throughput, particularly for low signal-to-noise ratio and moderate to high data rates. Linearity is a common property of many wireless communication systems and channels, making linear functions a good representation of throughput. As the throughput of a system increases, the system performance increases linearly [206]. The distance function takes a low utility for long links and a sharp drop-off to encourage short-distance transmission. We represent our utility functions with exponential functions for the following reasons: First, exponential functions are monotonically increasing, which means that as the input (e.g., the power) increases, the output (e.g., the utility) also increases. This is a natural behavior for a utility function, as it represents the increase in the quality of service as the power level of a node increases. Exponential functions are concave, which means that the second derivative is negative. This property is useful for utility functions, as it means that the rate of increase of the utility function will decrease as the input value increases. This is also a realistic behavior for a utility function, as the increase in the quality of service will be more significant for lower values of the input, and will decrease as the input value increases. Whereas, linear functions are a good choice to represent throughput in wireless communication systems because they have several desirable properties, such as linearity, simplicity, additivity, and a good approximation of many wireless communication channels.

Power utility U_P is represented as a sigmoid function with low utility and slow change at the lower power, exponential change in the utility at medium and high, and slow change at high power. The sigmoid function is commonly used to model the behavior of a system or process that saturates over time. In the case of power consumption, a sigmoid function can capture the fact that the power consumption of a device may reach a maximum level and then stabilize, even if additional energy is supplied. The shape of the sigmoid function can be adjusted to fit the specific characteristics of the power consumption behavior being modeled. Mobility utility U_M is shown as one minus sigmoid function with slow changes and high utility at low mobility rate, sharp changes in utility at the medium mobility rate, and slow changes and low utility at high mobility rate. Distance utility U_D is maximum for up to 60% of the coverage area. However, it changes rapidly after 70% up until 100% of coverage due to signal attenuation. Throughput utility U_T has a constant rate of changes, and as the throughput increases, the utility increases accordingly. The value of each function can be changed, and the functions' values can vary depending on the requirements. For instance, based on our observation in U_D, 60% coverage range is a reasonable threshold value. However, part of the future study can investigate the impact of these values.

$$U_P = \frac{1}{1 + e^{(-P+\alpha)}}, \tag{5.1}$$

$$U_M = \frac{1}{e^{(M-\alpha)} + 1}, \tag{5.2}$$

$$U_D = 1 - (\frac{D}{\beta})^{\alpha}, \tag{5.3}$$

$$U_T = mT, \tag{5.4}$$

where P is the remaining battery power, M is the mobility speed of devices, D is the distance to the next-hop node within the transmission range, and T is the maximum throughput or bandwidth to the one-hop neighbor. α is specified as half the range of the metric scale to shift the functions, β is the range of the metric scale, and m shows how fast the utility value changes. In our experiment, we chose these constant values as above. However, these values can be changed, and future studies might include adjusting those parameters. We chose these values because they can show a simple but effective rate of changes in different IB values versus the utility function.

Algorithm 1 provides the initial stage of generating the MLSDB or Phase I, mapping IB values into their corresponding utility values. The mapping process of node's IB values is where collected information by the controller is translated to the corresponding utility value using the given utility equation. For instance, let's consider a node's IB value reported as 200 Joules and go through the mapping process. The first step is to normalize the IB value between 0-10. Normalizing a value between 0 and 10 is good because it can make the data more consistent, simple, and easy to work with, which can help improve the accuracy of predictions, patterns and trends identification and overall model performance when used with route prediction. To normalize any IB value, the reported value is divided by the maximum assumed value of the corresponding entity. The maximum energy value of UEs is assumed to be 300 Joules. Therefore, the normalized value is 200 divided by 300 and the utility value or U_p according to (5.1) is $U_p = 1/(1 + e^{(-200/300+5)} = 0.8410)$. Now the 200 Joules of power is mapped to the utility value of 0.8410.

To have a clear understanding of how MLSDB is created, 8 nodes are scattered across a network as shown in Figure 5.5. The red lines in the Figure represent the transmission range of each node to the neighboring nodes. The source node is labeled as A, and the destination node is labeled

Table 5.1: Master LSDB generation Phase I with randomly given node IBs.

Master LSDB		**LSDB-E**		**LSDB-M**		**LSDB-D**		**LSDB-T**	
Root Node	Neighbor Node	Value (j)	Utility	Value (m/s)	Utility	Value (m)	Utility	Value (Mbps)	Utility
A	B	200	0.8411	25	0.2227	60	0.9222	15	0.5
A	C	210	0.8807	3	0.9859	15	0.9999	20	0.6667
A	D	90	0.1192	15	0.7772	30	0.9975	5	0.1667
B	E	180	0.7310	3	0.9859	50	0.9657	10	0.3333
B	F	280	0.9870	3	0.9859	70	0.8319	25	0.8334
C	E	280	0.9870	3	0.9859	70	0.8319	25	0.8334
C	F	150	0.5	25	0.2227	30	0.9975	3	0.1
C	G	170	0.6607	15	0.7772	10	0.9999	22	0.7333
D	F	150	0.5	25	0.2227	30	0.9975	3	0.1
D	G	170	0.6607	15	0.7772	10	0.9999	22	0.7333
E	H	50	0.0344	15	0.7772	40	0.9897	17	0.5667
F	H	50	0.0344	15	0.7772	40	0.9897	17	0.5667
G	H	200	0.8411	15	0.7772	50	0.9687	29	0.9667

by H. The rest of the nodes are the relay nodes. Table 5.1 is an example of how MLSDB is created for the network in Figure 5.5. As shown in Table 5.1, each sub-LSDB consists of two columns. The first column represents the reported value from the neighboring node and the second column shows the utility of the corresponding value. The LSDB-x represents the link state of the given IB values using the utility functions. For instance, LSDB-E shows the IB values and the corresponding utility value using (5.1). Similarly, LSDB-M, LSDB-D, and LSDB-T are computed using (5.2), (5.3), and (5.4), respectively. Using all the sub-LSDBs together represents Phase I of generating the MLSDB. Overall, the proposed MLSDB keeps the utility knowledge of the network at the sub-controller as shown in both Figures 5.1 and 5.3.

Algorithm 1: Master LSDB generation Phase I

Input : N and N_{IB}
Output : Master LSDB
for *all* $n_i = (E, T, M, D, ID_{Neighbors}) \in N_{IB}, i, j \in N$ **do**
    $U_P = \frac{1}{1+e^{-E+5}}$,
    $U_T = 0.1T$,
    $U_M = \frac{1}{e^{M-5}+1}$,
    $U_D = 1 - \left(\frac{D}{10}\right)^5$,
    $LSDB = [i, ID_{neighbor j}, U_p^j, U_T^j, U_M^j, U_D^j]$,
end

5.5.2 MLSDB Generation Phase II

After the completion of Phase I, Phase II is to associate each utility with traffic types. To complete Phase II, forwarding cost metrics are defined for each traffic type. Forwarding cost metrics are formulated based on the packet type requirements introduced in Section 5.4. As a result, four main forwarding cost metrics are introduced to select relay nodes.

$$C_{MC} = U_D U_M, \tag{5.5}$$

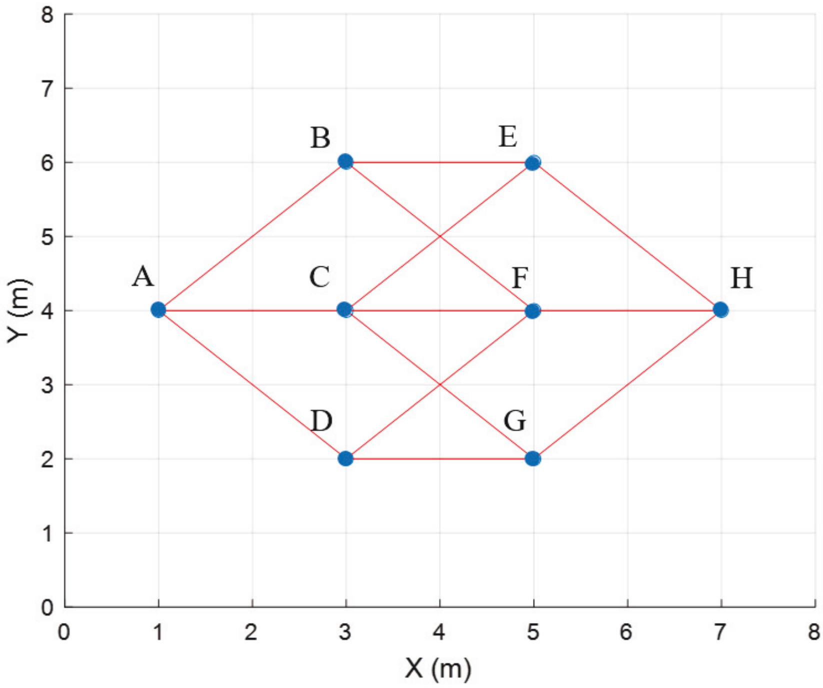


Figure 5.5: An example of the two-dimensional view of a network with 8 nodes, where each circle and line corresponds to the wireless node and the communication link, respectively.

$$C_{NC} = U_P U_D, \tag{5.6}$$

$$C_{HD} = U_T U_M, \tag{5.7}$$

$$C_{HDR} = U_T U_P U_M, \tag{5.8}$$

where C_{MC} represents the mission-critical packets with the requirements of selecting a relay node using the closest node with the least mobility rate. C_{NC} known as the non-critical data packets, have the least important requirements, choosing the next relay node based on the highest battery level and closest distance. C_{HD} are the packets that need the highest throughput, but the least-mobile node, and finally, C_{HDR} requirements of selecting the next relay node are based on the maximum throughput, least mobility rate, and highest battery level.

In this study, we introduce four traffic types, and each one's requirements depend on distance, mobility rate, throughput, or power level. However, based on the user's demand, each forwarding cost above can be changed accordingly, and this change can be based on the combination of the four utility metrics introduced earlier. The general representation of the cost metric of any traffic type C_{TT} is shown as follows:

$$C_{TT}^{k} = \begin{cases} \prod_{i \in N} U_{P_i} U_{T_i} U_{M_i} U_{D_i}, & \text{if} \quad k = 1 \\ \prod_{i \in N} U_{P_i} U_{T_i/M_i} U_{M_i/T_i}, & \text{if} \quad k = 2 \\ \prod_{i \in N} U_{P_i/M_i/D_i} U_{T_i/M_i/T_i}, & \text{if} \quad k = 3 \\ \prod_{i \in N} U_{P_i/M_i/D_i/T_i}, & \text{if} \quad k = 4 \end{cases} \tag{5.9}$$

There are four possible actions in computing the C_{TT}, if $k = 1$ the automatic reply to all the unknown traffic types. In situations when $k = 1, k = 2$ or $k = 3$ the controller orders to adjust the cost function based on the network situation, usually happening in disasters and emergency services. Finally, based on Algorithm 2, the controller generates the MLSDB. Upon completion of Phase II, the MLSDB is shared with the entire network. At the reception, nodes will start the MD2D communication by building a routing table. The complete MLSDB of the example shown in Figure 5.5 is illustrated in Table 5.2.

Table 5.2 shows the link state of neighboring nodes in terms of the proposed cost metrics based on the application requirements, which are introduced in (5.5), (5.6), (5.7), and (5.8). The first column of the table represents the link between the two nodes. In the second column, we compute the corresponding packet cost metric, specifically the cost of transmitting mission-critical packets, calculated using the values computed in Table 5.1 and (5.5). In the third column, the cost of the link to transmit non-critical messages is computed using the values in Table 5.1 and (5.6). In the fourth column, the link state for the high-definition packages is evaluated using the values computed in Table 5.1 and (5.7). Finally, the fifth column represents the cost of a link for real-time applications, computed using (5.8). Table 5.2 shows the completed MLSDB obtained by the sub-controller. This table is sent across the network, and nodes with a packet to transmit will check the application and compute the path to the destination using one of the columns in the MLSDB table. For instance, if a node sends a high-definition packet, it will use the fourth column and evaluate the path that maximizes the network performance (which means choosing the next hop links that provide the maximum cumulative values).

Once the MLSDB is transmitted, a node can decide which forwarding cost metric to select based on the traffic type. Then at any instance of time, nodes will distributively select a path. For example, from Figure 5.5 if node A has an MC packet to transmit, the MLSDB table (Table 5.2) is used to check and identify the path that maximizes the sum of utility values to the destination. Therefore, the routing problem is generally formulated for any given flow and traffic type as:

$$\underset{x_{1,2,3,4} \in D,T,M,P}{\text{Maximize}} \quad U_{x_1} U_{x_2} U_{x_3} U_{x_4}, \tag{5.10}$$

subjected to the following constraints:

- $0 < U_{X_i} \leq 1$ for $\forall_i \in i = \{D, T, M, P\}$
- $U_{x_i} \neq U_{x_{i+1}}$

The first condition ensures that the values of the utility functions fall within the normalization range, while the second condition reflects the importance of using distinct utility functions. The use of identical utility functions may reduce the optimization performance, and hence, it is imperative to avoid such redundancy in the optimization problem.

In summary, our proposed framework and the route discovery procedure can be explained by the following steps:

- ACMRP is a hybrid routing protocol (it uses a centralized controller and distributed route discovery) in which each relay node distributively selects its route to the destination. To this end, at the initial stage of the network, the sub-controller broadcasts a TCREQ to the entire network. Nodes will reply by sending IB to the sub-controller.

- Using IBs, the sub-controller creates sub-LSDBs based on the utility metrics defined earlier as; U_P, U_M, U_D, and U_T representing power utility, mobility rate utility, distance utility, and throughput utility, respectively. Combining all four sub-LSDBs constructs Phase I of the MLSDB. Next, using the defined cost metrics based on the different packet types, the final MLSDB table (Phase II) is generated and broadcasted to all the nodes under the coverage area.

- When a node needs to send data, it first checks if it has a valid route to the destination. If so, it will transmit the data. If not, it will assess the type of traffic and determine the most efficient route based on the corresponding forwarding cost metric. If the traffic type does not match any of the metrics, it will calculate Phase I of C_{TT} to evaluate a path. If the node is unable to find a relay node, it will send a "Destination-Unreachable" route error (DRERR) message to the sub-controller, and redirect its traffic through the cellular channel.

Bear in mind that data packets are exchanged through either the cellular or the WiFi channel (MD2D). To determine the most appropriate channel for data exchange, a node continuously assesses parameters such as the spent transmission time τ[2], maximum allowed number of hops M_H and maximum physical distance M_D. In ACMRP, route maintenance is a critical component and consists of two processes: initiation of a DRERR packet and processing of the received DRERR. Relay nodes between the source and destination node are responsible for relaying the source node's packet to the next relay node based on the traffic type requirements. If a node encounters an error during transmission or reception, it will send a DRERR packet indicating the error type (Unknown-Error, Link-Breakage, or Destination-Unreachable). The source node

[2]It is the time that makes a relay node realize how much time is spent to relay the data if this time is more than a threshold λ then it will switch to cellular transmission.

Algorithm 2: Master LSDB generation Phase II

Input : N, Phase I, Range, Traffic Type

Output : Master LSDB Phase II

```
for each i,j∈N do
    if C_TT = C_MC, C_NC, C_HD, C_HDR then
        Find Euclidean distance (D) between i and j
        if D_{i,j} ≤ Range then
            Compute
            C_MC^{i,j} = U_D^j U_M^j,
            C_NC^{i,j} = U_P^j U_D^j,
            C_HD^{i,j} = U_T^j U_M^j,
            C_HDR^{i,j} = U_T^j U_P^j U_M^j,
        end
    else
        Compute
        C_TT,
    end
end
```

or relay node must then respond appropriately. If the error occurs at the source node, it will notify the sub-controller, while if it is the relay node, the previous node will be notified. In the case of ”Link-Breakage”, a new route must be found, and the data must be rerouted to the destination.

5.6 Simulation Results and Performance Analysis

5.6.1 Simulation Setup

The ACMRP algorithm is implemented using the Network Simulator-3 (NS-3). The NS-3 simulator core supports simulation scenarios using both IP and non-IP-based platforms. The IP-based models involve WiFi, worldwide interoperability for microwave access (WiMAX), or LTE/5G for layers 1 and 2. There are inbuilt simulation modules for a variety of statics and routing protocols, including ad hoc on-demand distance vector (AODV), OLSR, and destination sequenced distance vector (DSDV) for MANETs applications.

Our proposed protocol is compared with two different routing frameworks, distributed and semi-centralized. To compare ACMRP with a similar routing protocol, HSAW was implemented in NS3. In comparing our protocol with a centralized strategy, we benchmark our protocol with distributed methods to investigate performance tradeoffs. AODV and OLSR are used to represent the fully distributed routing protocols. NS3 has built-in modules for AODV and OLSR, where one accounts for reactive and the other for proactive. The similarity between HSAW and ACMRP lies within the centralized management of O-RAN. The MBS collects net-

Table 5.2: Master LSDB generation Phase II.

Links \ Cost Metrics	C_{MC}	C_{NC}	C_{HD}	C_{HDR}
A-B	0.2053	0.7757	0.1113	0.0936
A-C	0.9858	0.8807	0.6572	0.58
A-D	0.7754	0.1189	0.1295	0.0154
B-E	0.9520	0.706	0.3286	0.24
B-F	0.8202	0.8211	0.8216	0.811
C-E	0.8202	0.8211	0.8216	0.811
C-F	0.2221	0.4987	0.0222	0.011
C-G	0.7773	0.6607	0.57	0.3765
D-F	0.2221	0.4987	0.0222	0.011
D-G	0.7773	0.6607	0.57	0.3765
E-H	0.7693	0.0340	0.4404	0.0151
F-H	0.7693	0.0340	0.4404	0.0151
G-H	0.7693	0.8148	0.7513	0.6319

work layer data and processes the information to create routing tables. In the ACMRP protocol, the MBS is responsible for authenticating nodes and continuously collects IB information to update the MLSDB.

The simulation environment consists of uniformly distributed heterogeneous nodes, such as mobile devices, IoT sensors, and vehicles. All the nodes are specified as user equipment (UE), assuming that 40% pedestrian, 40% inner-city mobile nodes, and 20% outer-city mobile nodes. The initial energy of UEs is 300 Joules. The MBS is located in the center of the network. Each node has two interfaces for communication purposes: one interface for WiFi-band communications and one LTE/5G interface for cellular communication. Node's transmission range is limited to $100m$ and follows IEEE 802.11n-5GHz[3]. It is assumed that the mobility of nodes is categorized into two main types; relatively stationary and mobile nodes with velocities of 1m/s, 15m/s, and 25 m/s.The total simulation time is 300s.

Two different networks are simulated to observe the scalability of the proposed protocol. The first network expands over $500m \times 500m$ area with nodes varying from 50 to 250. The second network expands over $1000m \times 1000m$ area with nodes varying from 100 to 500. The motivation behind changing the network density is as follows. By changing the network density and size, the simulation tests the performance of the MD2D routing protocol in different scenarios, including sparse and dense network environments. This assesses the scalability of the network and its ability to handle varying levels of traffic. The network's density can also affect the energy consumption of the devices. For example, a high density of devices can lead to increased energy consumption due to increased interference and collisions, while a low density of devices can lead to decreased energy consumption due to reduced interference and colli-

[3]The proposed architecture is not limited to using IEEE802.11n, as some other ISM/802.11-based radio signals can also be used. However, for this study, we chose 802.11n as a potential ISM-based radio technology.

Table 5.3: Simulation parameters.

Parameters	Value
Simulation environment	500m$\times 500m/1000m \times 1000m$
Initial power of UEs	300j
Number of UEs	50,100,150,200,250 100,200,300,400,500
Packet transmission size	15 kbits
Protocols	ACMRP, HSAW, AODV, OLSR
Propagation model	Rayleigh fading
Mobile node transmission range	100m
Mobile rate	1 m/s, 15 m/s, 25 m/s
Mobile node movement model	Random waypoint mobility
Total simulation time	300s

sions. Moreover, varying the network density can provide a more realistic representation of an MD2D network, which can be used to make more accurate predictions about the network's performance. For example, a network with a high density of devices can be more representative of a crowded urban environment. In comparison, a network with a low density of devices can be more representative of a rural environment.

As can be seen from the simulation results shown in Figures 5.6-5.10, two sets of line graphs depict the simulation results, with one set for a 500m^2 environment that starts from 50 nodes and ends at 250, and another set for a 1000m^2 environment that starts from 100 and ends at 500. Both environments demonstrate the superior scalability of the proposed protocol. A specific number of active flows are defined with randomly selected source and destination nodes. Simulation parameters are shown in Table 5.3. The simulation was run for 300s, and each simulation was averaged over multiple simulations running using different seed values. To prove the effectiveness of the proposed protocol and validate our results, the results are validated with the Monte Carlo simulation technique under 50 runs. The final results are averaged and plotted with 95% confidence intervals.

5.6.2 Simulation Results

Figures 5.6 and 5.7 illustrate the network lifetime based on the number of depleted nodes and energy consumption, respectively. To analyze the performance of our proposed protocol, we have simulated HSAW, the closest routing protocol to our framework. In addition to investigating our protocols against an existing MD2D protocol (Namely HSAW), we also explored its performance with purely ad hoc protocols (AODV and OLSR) to trade-off the efficiency of MD2D protocols with Ad hoc protocols. AODV and OLSR are still in use for comparison of ad hoc based routing protocols [207, 208]. The three comparison protocols always choose the next hop without knowing the energy levels of the next hop, while in our protocol, nodes know the next hop's energy levels and can choose the node with the highest power. Figure 5.6 shows the number of depleted nodes versus the number of nodes. We can observe that ACMRP has the lowest number of depleted nodes in both networks. Figure 5.7 represents the network's energy consumption versus the number of nodes. When the number of nodes in the network is low, all the protocols have similar energy consumption. However, once the number of nodes increases, ACMRP performs better because in AODV and OLSR flooding process increases the node's power consumption due to acknowledgment, route maintenance, and route update. Moreover, in HSAW, the size of the LSDB is larger than ACMRP, and every time nodes get the LSDB update, they will consume more power than in ACMRP. ACMRP has an optimization mechanism in both LSDB table creation and the route identification process, which reduces the overall energy consumption of the network.

Figure 5.8 shows the E2E delay versus the number of nodes. In AODV and OLSR routing protocols, nodes compute a route in a distributed manner.To compute a path in these two distributed routing protocols, nodes use Hello packets and flooding algorithms to populate their routing tables, and only then can they transmit a packet. AODV uses a blind flood for route discovery and OLSR uses MPF (multipoint flooding) to share neighbor tables with all nodes in the network so that shortest path routing (Djikstra) can be performed. However, in centralized protocols such as HSAW and ACMRP, the controller collects and stores refined information about the entire network and can compute the route based on the proposed framework. HSAW and ACMRP are both centralized routing protocols where nodes have the LSDB of the network and can immediately find a route. Although HSAW and ACMRP performance is similar in a low-density network, once node density increases, ACMRP accomplishes a better result because nodes facilitate network intelligence to identify the most reliable and efficient path. In HSAW, nodes select the next hop based on the shortest distance without knowing whether the next relay node is reliable. Therefore, in some instances selecting the closest node for relaying traffic may cause link failure due to mobility rate and QoS, which causes extra computation time, resulting in extra time consumption.

Figure 5.9 represents the network throughput, including the cellular and WiFi communications. ACMRP has a significant advantage over other routing protocols because it has the throughput utility of every node, where it will select the node with the highest throughput. Our proposed routing protocol can identify the node's behavior and select the node with the best QoS. In HSAW, AODV and OLSR nodes will determine a path based on the shortest path first (SPF) using hop counts. Therefore, they have no knowledge about the nodes and cannot decide which next-hop has the maximum transmission capacity. Moreover, our proposed protocol is application-aware, improving the overall network throughput because nodes know the route that maximizes the throughput. The other three comparison protocols cannot predict the quality of the link. Hence, they have lower throughput.

Figure 5.10 depicts the PDR versus the number of nodes. The ratio of the delivered packets and the total number of packets sent by the source node is obtained to simulate PDR. PDR is

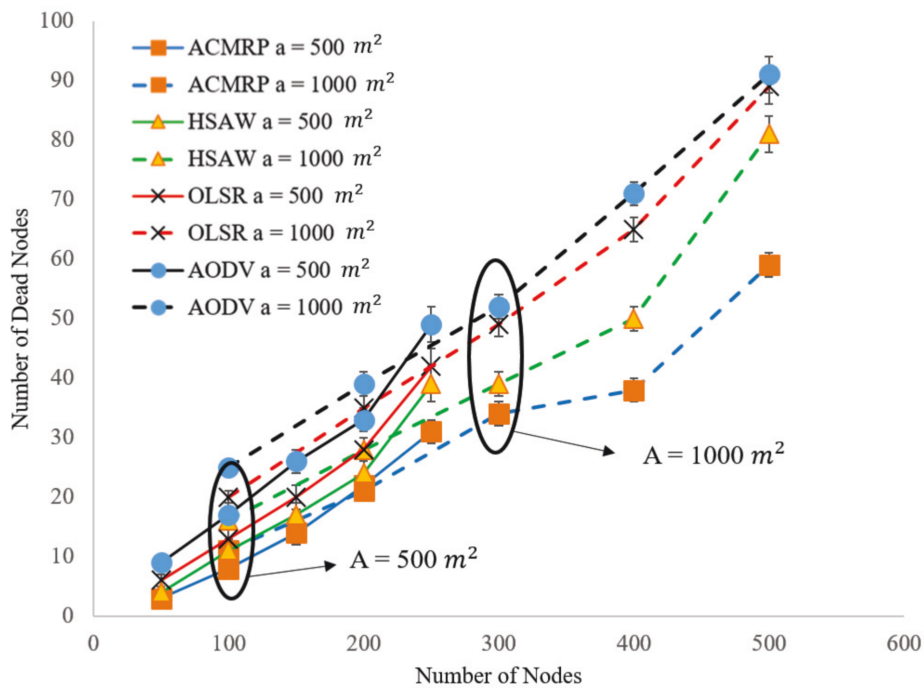


Figure 5.6: Number of depleted nodes for network sizes of $500m^2$ and $1000m^2$.

affected by several factors, including mobility, quality of the link, queue size, and distance to the neighbor. These factors can cause link or packet failure and cause retransmission at any relay node. Therefore, relay nodes may receive multiple copies of the same packet. ACMRP can predict the next hop with the most reliable link, with the lowest mobility rate, high power levels, and closer distance. Therefore, nodes have a higher probability of sending a packet without retransmitting due to link breakages caused by these factors. On the other hand, other protocols have no information about the next relay's mobility, distance, or power. Hence, the probability of choosing the wrong relay node is higher, making packet retransmission inevitable. As shown in Figure 5.10, ACMRP has the highest PDR, which remains almost 80% as the network density increases. The difference between all the routing protocols is not significant in low-traffic networks. However, once the node density and network size increase, the difference becomes significant, and ACMRP has the highest PDR compared to the rest. Next-hop selection in ACMRP depends on the lowest mobility rate, shortest distance, and highest bandwidth. Therefore, a data packet will be exchanged faster with a more reliable link.

5.7 Conclusion

The proliferation of smart devices and the Internet of Things (IoT) has resulted in a substantial increase in cellular network traffic. To address this challenge, we propose an application-driven cross-layer MD2D routing protocol that integrates application layer requirements with network analytics to obtain the most efficient path based on the application requirement. Our solution improves network lifetime by a minimum of 25% compared to HSAW protocol and by over 40%

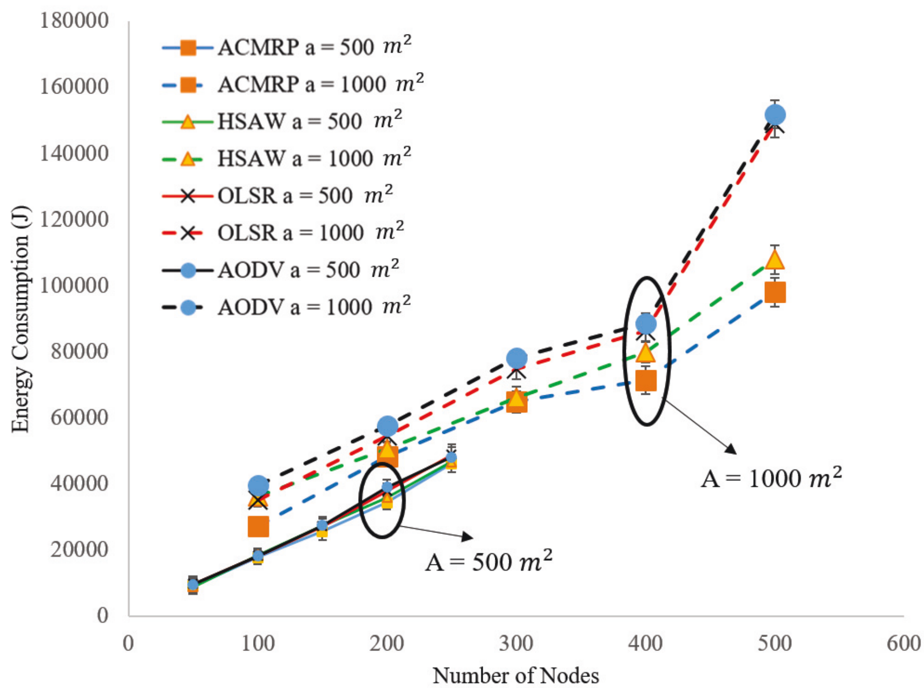


Figure 5.7: Nodes energy consumption for network sizes of 500m^2 and 1000m^2.

compared to the AODV and OLSR protocols. Furthermore, our proposed solution enhances network throughput by approximately 20% compared to HSAW and over 50% compared to AODV. Finally, the PDR of our solution demonstrates a more stable and reliable delivery compared to HSAW, OLSR, and AODV.

In the future, artificial intelligence (AI) or machine learning (ML) based algorithms can be used to further optimize or introduce new protocols and frameworks to offload the BS traffic. For instance, packet sizes and headers can be modified based on the network conditions and further adjusted by AI/ML-based NFV modules.

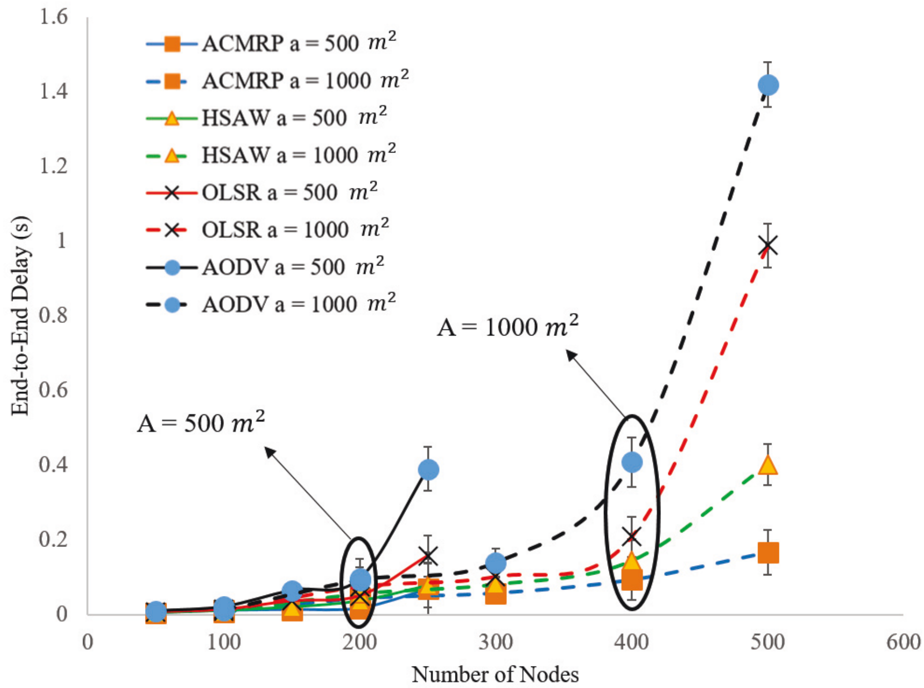


Figure 5.8: End-to-end delay for network sizes of 500m^2 and 1000m^2.

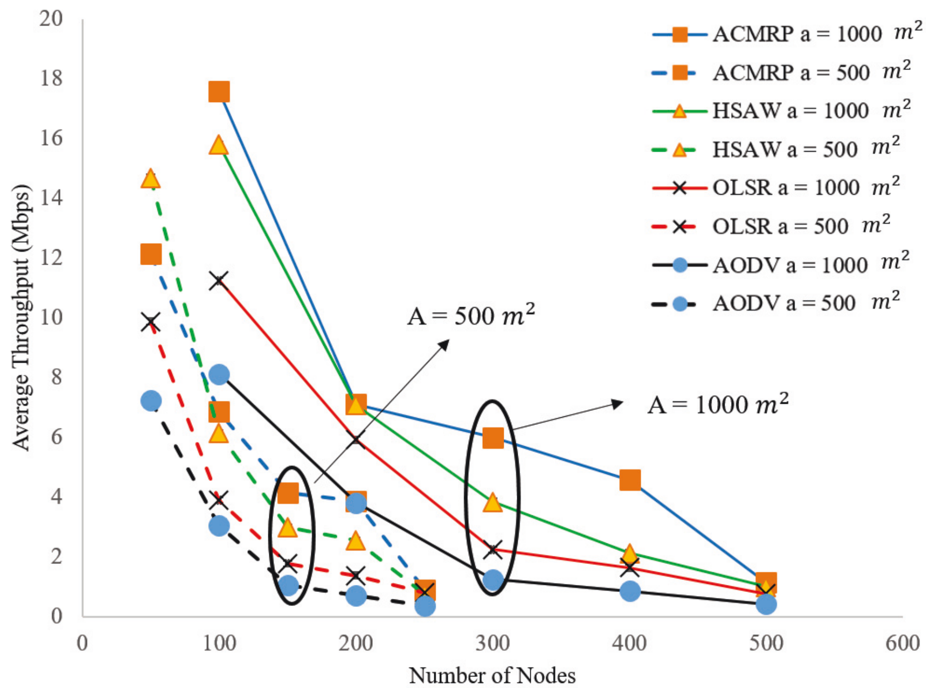


Figure 5.9: Average network throughput for network sizes of 500m^2 and 1000m^2.

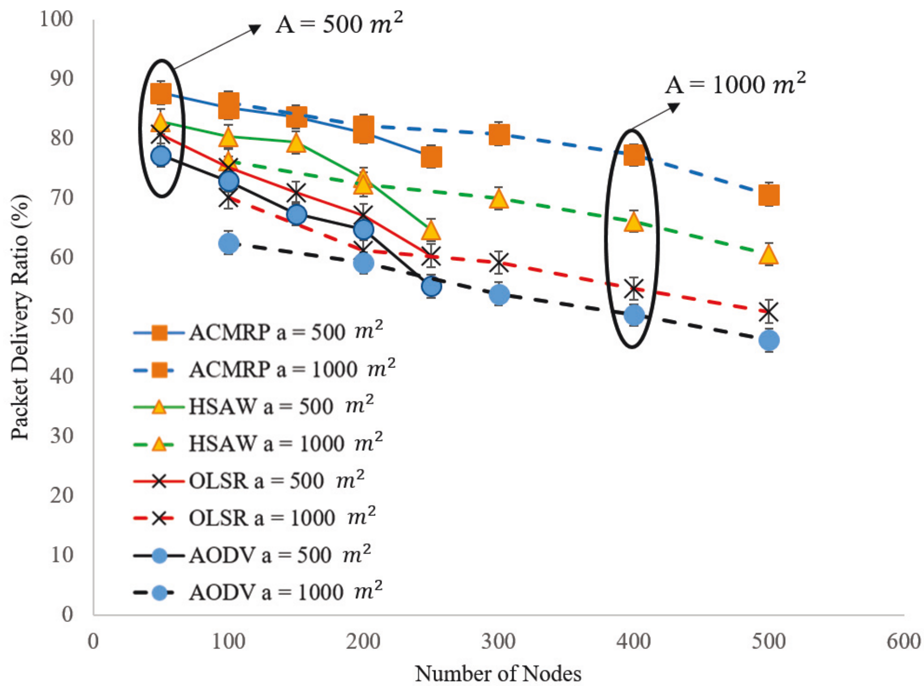


Figure 5.10: Packet delivery ratio for network sizes of $500m^2$ and $1000m^2$.

6

An Adaptive Multi-hop Device-to-Device Routing Framework for Future Cellular Networks

6.1 Overview

The upcoming research is submitted to Elsevier Journal of Network and Computer Applications [1]. This chapter's main objective is to introduce the concept of virtual network slicing for WiFi channels and proposes a mechanism to enable multiple MD2D routing protocols to be deployed over each virtual slice. The improved self-adaptive particle swarm optimization (IDPSO) algorithm is utilized to cluster data and form virtual slices. Our proposed framework, called ASDR, enables network slicing in WiFi channel, which performs better than single-route frameworks.

6.2 Introduction

In a cellular network, communication between a user and a base station (BS) is established through a direct link. This direct communication provides a low end-to-end (E2E) packet trans-

[1]Ashtari, S., Abolhasan, M., Lipman, J., Ni, W., Jamalipour, A. (2023). An Adaptive Multi-hop Device-to-Device Routing Framework for Future Cellular Networks. Elsevier Journal of Network and Computer Applications.

mission time, enhancing the user's quality of experience (QoE). However, relying solely on this direct communication results in excessive cellular traffic overhead, due to the limited availability of spectrum resources [74, 209]. To address the issue of massive cellular traffic overhead, various solutions have been proposed, including deploying additional base stations and dividing the network into smaller cells, which is cost-prohibitive [210]. While other techniques such as space division multiple access are available, they may impact network coverage due to low transmission power. As a promising alternative, direct device-to-device (D2D) communication has been introduced by the third-generation partnership project (3GPP) as a potential solution to enhance network capacity and coverage [211, 212].

In the 3GPP long-term evolution (LTE) release 12, two devices in close proximity can communicate with each other directly via a WiFi connection. This direct communication first enabled devices out of the coverage area to communicate with a BS through another device [213]. D2D can be between two devices in the coverage of the BS or out of the BS. This technique can alleviate the network capacity and coverage problems and further assist in disaster scenarios where the BS is not available [214]. Recently, 3GPP release 13-15 two-hop communication has been approved, which can further increase the network capacity and coverage [215, 216]. As a result of the proliferation of D2D communication and the continuous evolution of D2D in 3GPP releases, it can be predicted that multi-hop device-to-device (MD2D) will be a part of the standard in future cellular networks. The advantages of D2D communication in the cellular network can be fully realized when extended to multi-hop communication scenarios since the single-hop is usually limited to a specific geographic area.

6.2.1 Motivation

The increasing number of applications and user demands create massive data traffic on the cellular channel. Allocating such data to a secondary network (such as the MD2D network) can alleviate traffic congestion and stress on the cellular network. For example, the work in [217] proposed an MD2D-based routing protocol to support services in large geographic areas. Their algorithm targets the areas that are out of the cell boundaries and provide services using MD2D. The protocol decreases the cellular channel overhead while providing fast and reliable services with no extra resource consumption. At the same time, other MD2D-proposed protocols and frameworks focus on preserving the device energy, spectrum and coverage. To improve spectrum and energy efficiency, authors in [218] proposed an optimal adaptive forwarding strategy (OAFS) for multi-hop D2D communications. This strategy adaptively chooses between the best relay forwarding mode and the cooperative relay beamforming mode based on energy efficiency metrics. Authors in [219] improved the cellular coverage quality by proposing an enhanced MD2D scheme by jointly considering multiple quality-of-service (QoS) metrics. A new power adjustment scheme based on a game decision algorithm performs spectrum sharing and underlay D2D relaying links.

The integration of cellular and MD2D networks provides various advantages. However, if the routing frameworks and protocols are not designed carefully, they can perform even worse than traditional cellular networks [3, 220]. Routing in MD2D is a critical issue because it must take care of the node's mobility, dynamic network topology, energy consumption and network fragmentation. For instance, the work in [86] presented a routing framework to offload traffic using a source-based routing protocol. Similarly, a routing framework is proposed in [221] to identify the nodes with the highest resources to increase the network performance and improve energy consumption. Current MD2D networks operate under one framework, which means only one routing protocol is usually investigated and proposed. However, due to the heterogeneity of future network devices and the versatility of applications, it is essential to allocate specific routing protocols to different applications and create a multi-framework network architecture. This can be done by network slicing and allocating dynamic MD2D routing protocols for each slice. Such framework may be integrated into future cellular networks due to the emergence of open-radio access networks, which facilitates software-defined networking (SDN) and network function virtualization (NFV) to create an intelligent, open and programmable infrastructure. Therefore, new network functions could be developed for a specific purpose, which allows service providers to adjust their optimization techniques for network applications [7].

Another general issue in future wireless cellular networks will be data traffic management. Nowadays, the data among user equipment (UE) usually shows high spatial and temporal correlation, in which a significant amount of the data transferred between UE's and BS are similar. Such data is redundant and could be assigned to the MD2D network to get shared and downloaded by nodes. Therefore, instead of all the devices asking for and receiving the content from the BS, they can request it from nearby devices. One of the promising solutions to manage the data and resources is the dynamic partitioning of physical network infrastructure into logical networks. Network slicing is a solution to distinguish nodes with similar content requests. For instance, the BS can use clustering algorithms to identify and aggregate software updates or viral content, allowing UEs with similar requests to share and exchange data.

6.2.2 Contributions

This chapter introduces the concept of virtual network slicing for WiFi channels and proposes a mechanism to enable multiple MD2D routing protocols to be deployed over each virtual slice (VS). As shown in Figure 6.1 network data is collected by the small-cell BSs (SBSs) and then transmitted to the mobile edge computing (MEC) controller for further processing. The controller uses improved self-adaptive particle swarm optimization (IDPSO) for data clustering and genetic algorithm (GA) in conjunction with learning automata (LA) for dynamic route selection. After the process, each VS_i represents an independent sub-layer with UEs requesting the same content. As can be seen from Figure 6.1 the entire network is sliced into n number of sub-layers based on the node's content request. Each VS receives different routing protocols

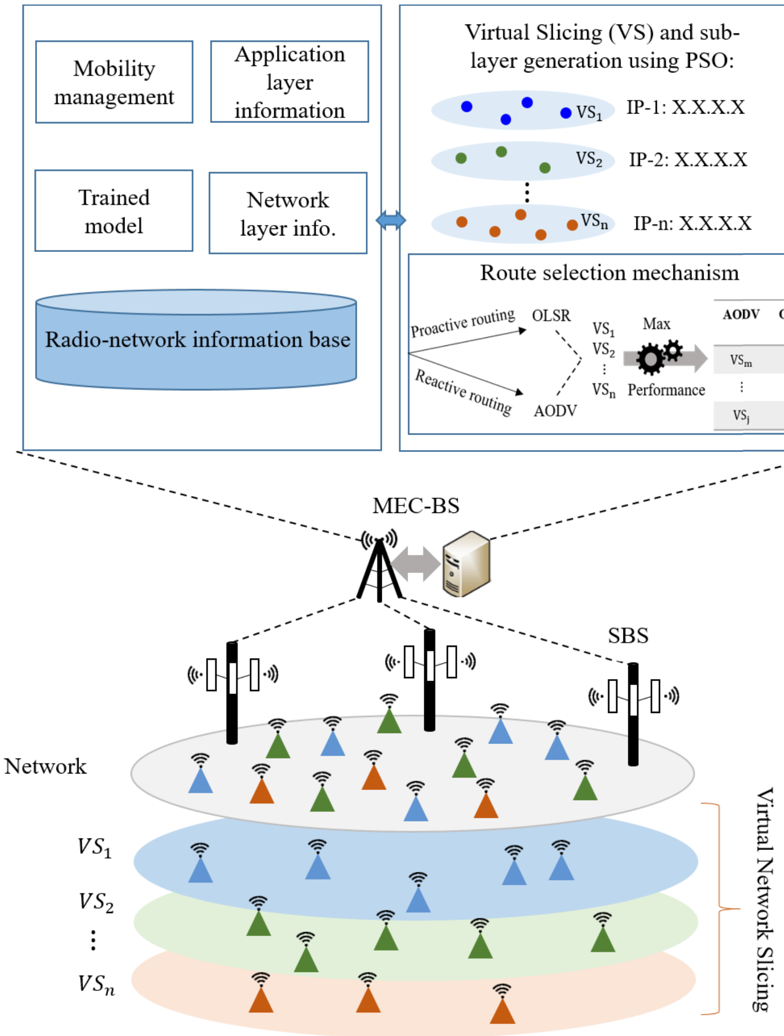


Figure 6.1: Modular design for virtual slicing and dynamic route selection.

depending on the overall network performance in terms of end-to-end (E2E) delay, throughput, remaining energy, and packet delivery ratio (PDR).

The contributions of this chapter are summarized as follows:

- We propose a new adaptive MD2D routing framework using virtual network slicing and dynamic routing protocol selection for the WiFi channel.
- We introduce a data clustering algorithm to allocate nodes into different network slices using the IDPSO algorithm. The technique is based on fitness function computations using the data similarity requested by MD2D nodes.
- We present a routing selection mechanism for sub-MD2D network slices using the genetic algorithm (GA) in conjunction with learning automata (LA). The combination of GA and LA leads to dynamic routing protocol selection through a learning process.

An extensive experimental analysis is performed to evaluate our proposed approach. The results show that dynamic virtual network slicing with a multi-routing protocol performs significantly better than a non-sliced single-routing framework. Simulation results illustrate an average increase of 53% in throughput, a 50% decrease in the E2E delay, an almost 5% reduction in energy consumption, and an 8% increase in PDR.

The rest of this chapter is organized as follows. Section 6.3 provides an extensive overview of the current research studies. Section 6.4 presents a detailed description of the proposed routing framework and problem formulation. Section 6.5 demonstrates the simulation results to validate the significance and efficiency of our proposed framework. Finally, Section 6.6 concludes the discussion and points to possible directions and future works.

6.3 Related Work

This section covers the related works on D2D and MD2D routing protocols and frameworks, primarily on clustering nodes or geographical areas. To the best of our knowledge, no study has performed data clustering and virtual network slicing with dynamically prescribed proactive or reactive routing protocol for each slice to maximize the network throughput. Most of the proposed protocols in the literature for MD2D and D2D networks are based on one framework and one routing protocol, except the work in [179]. Authors geographically partition the network into fixed areas and prescribe routing to each region based on the highest performance of the routing protocol in that specific cluster. They proposed a joint clustering and dynamic selection strategy for MD2D routing.

The majority of the related works for MD2D clustering are based on segregating the physical devices. For instance, the work in [222] clustered the devices and associated the MD2D network into a collision-free cooperative graph to identify efficient routes. The clustering technique groups the devices into non-interfering devices. A decoding delay-sensitive problem is formulated as a joint optimization over a set of transmitting nodes using graph theory to reduce the broadcast decoding delay in MD2D communication. In [223], the authors aim to identify the performance gain of cooperative MD2D communication. A macro BS (MBS) is responsible for constructing suitable clusters using a large number of single-antenna relay nodes. The MBS collects all the necessary information from the node and iteratively forms groups of nodes. The MBS creates layers of clusters, and when a source node wants to transmit the data, it broadcasts the information to the next nearest cluster toward the destination. Once the next cluster receives the information, it will broadcast to the next cluster until it reaches the destination. For privacy reasons and secure routing in MD2D communication, authors in [224] designed a cluster-head centered fast secure routing (CCFSR) for MD2D networks to force authenticated nodes to forward their packet to a secure cluster head. Cluster heads are formed based on the position and energy and deploy or use one routing protocol throughout the network. The cluster head identifies the fastest route to the destination using the breadth-first search algorithm. A non-cooperative game theory called secure routing choose game (SRCG) helps the cluster head to identify the malicious nodes.

The Internet of things (IoT) has gained considerable attention over the past few years, and it is expected that the number of connected IoT devices to the Internet will rise by billions every year. With this massive growth, having individual nodes connected to the Internet will cause serious scalability issues in cellular and wireless networks. One way to reduce the impact of this problem is to cluster nodes and have a representative node to collect, distribute and fetch data from the Internet on behalf of other nodes. Furthermore, the integration of MD2D with IoT networks can increase network coverage and capacity. Authors in [225] proposed a multi-hop clustering mechanism to minimize the number of nodes connected to the Internet. Their objective is to reduce the number of cluster heads while minimizing the total distance between nodes and the cluster heads. Nodes in the same cluster are allowed to go through each other to reach the cluster head until the maximum number of hop constraints. By allowing MD2D communication between nodes and between cluster heads, the proposed scheme reduces the number of connected devices by more than 80%.

In disaster scenarios where part of the communication infrastructure is unavailable, D2D networks are highly desirable. For instance, authors in [226] proposed an energy-efficient D2D scheme for disasters with clustering algorithms to ensure that devices can communicate for a long time until they are rescued. Their proposed algorithm establishes clusters until every node in the area of failed infrastructure can communicate with an available nearest BS. Candidate cluster heads are determined based on the location and residual energy. The modified ant colony algorithm (MACA) is adapted to increase the routing efficiency. Similarly, the work in [227] proposed an efficient multicast routing protocol for D2D communication to help improve the transmission power efficiency and QoS. The cluster heads are dynamically changed based on the packet transmission sequence and data demand factor (frequency of requests as a transmitter). The cluster heads are chosen based on the power levels.

Apart from the proposed protocols in [179, 227], many existing MD2D or D2D clustering algorithms are based on fixed frameworks. That means the clustering algorithms do not have an adaptability mechanism based on the network condition. Moreover, most of the clustering methods in D2D and MD2D networks operate to cluster devices or geographical areas. However, in our proposed framework, the data requests are clustered to construct virtual network slices. The controller groups the devices based on similar content requests. Then, based on the mobility rate and the number of nodes in every VS, the most efficient route is prescribed, which makes our proposed technique intelligent and adaptable based on the network and user requirements.

6.4 Proposed Routing Framework

In Figure 6.2, a multi-access edge server or controller processes and collects all the data from the nodes $N = \{n_1, n_2, ..., n_i\}$ at the edge network. We only consider one edge server and the

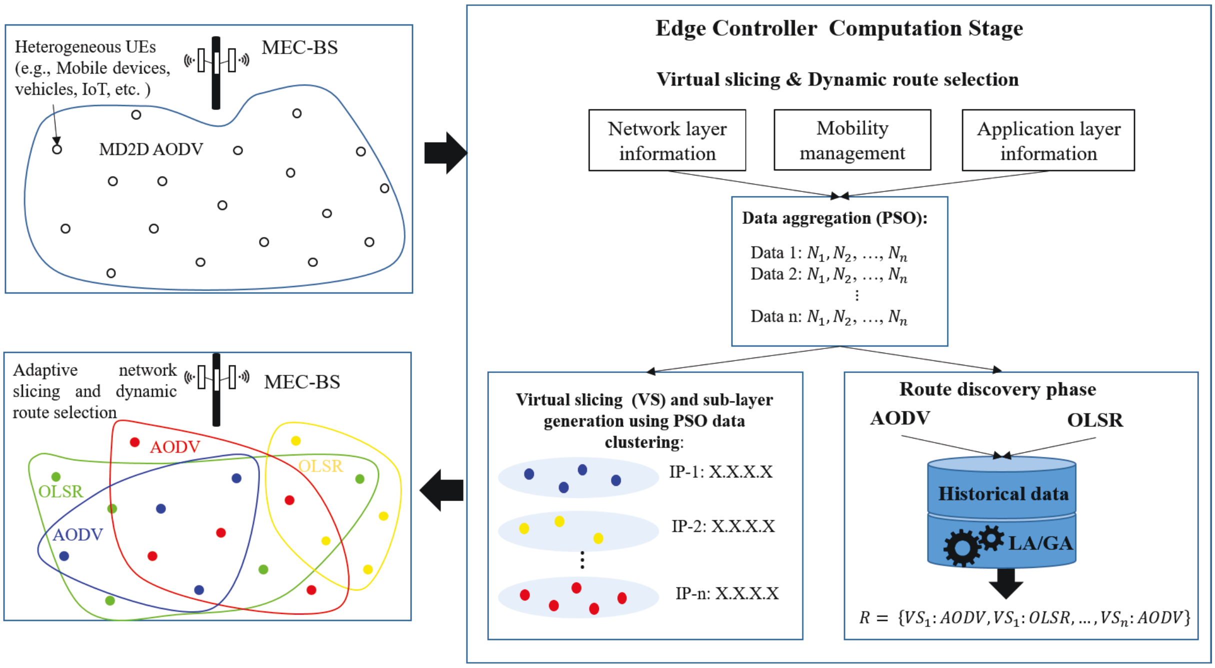


Figure 6.2: Proposed PSO network slicing and dynamic route selection.

nodes under its coverage for simplicity. However, our proposed framework can be extended to multiple edge servers that are all synchronized with the main MBS to share and allocate a pool of resources. The MEC-BS is responsible for dynamically changing the network slices based on node data requests. At the initial stage, the MEC-BS receives all the requests from the nodes and based on the requested contents, the PSO algorithm clusters the data. Then the VSs are formed with nodes requesting the same content. In each VS, there are representative nodes to share the data with the nodes in that slice using MD2D communication. The network consists of n slices $C = \{c_1, c_2, ...c_n\}$, and each one shares the same resources. Each network slice is assigned a routing policy and strategy based on the number of nodes, mobility rate and average remaining energy. In our framework, two routing protocols are prescribed for each VS, AODV or OLSR, depending on which routing performs efficiently. These two routing protocols represent proactive and reactive benchmark MD2D routings. Each slice receives routing information, IP addresses, and the operated frequency band. This approach will provide necessary slice isolation even in dynamic scenarios due to associated frequency and IP address. The detail of our framework, including the problem formulation and the proposed constraints, are presented in the following sub-sections.

6.4.1 Improved Particle Swarm Optimization

The PSO is a swarm intelligence-based stochastic algorithm modeled after the social behavior of a swarm of birds. This algorithm mimics complex global swarm patterns and exploits them to solve complex optimization problems. In PSO, each particle $P = \{p_1, p_2, ..., p_n\}$ of the population follows the best position found by any individual (known as global best or P_{gd}) and its local best position (known as local best or P_{id}). Each particle behaves like an intelligent agent in a highly decentralized environment. The PSO aims to find the particle position that leads to efficient optimization of the fitness function. The advantage of the PSO is that its accuracy is higher than some of the known clustering-based algorithms, such as the K-means algorithm [228]. Each particle represents a position in N-dimensional (N_d) space, the velocity of the particle is calculated in a multi-dimensional space and adjusted towards its individual local best position, $y_i(t)$, and the global best position, $\hat{y}(t)$, found so far in the neighborhood of that particle. The particle's position in the populated swarm G is adjusted based on the following equations, where $1 \leq i \leq P$, $1 \leq j \leq G$:

$$v_{i,k}(t+1) = wv_{i,k}(t) + c_1\phi_1(y_{i,k}(t) - x_{i,k}(t)) + c_2\phi_2(\hat{y}_k(t) - x_{i,k}(t)), \tag{6.1}$$

$$x_i(t+1) = x_i(t) + v_i(t+1), \tag{6.2}$$

where $v_i(t)$ and $x_i(t)$ are the particle velocity and position, respectively. c_1 and c_2 represent the acceleration constant, w is the inertia weight and ϕ_1, and ϕ_2 are random numbers $\phi_1, \phi_2 \sim U(0,1)$. In (6.1), the first part of the equation is the previous speed of particles. The second part of the equation is cognition, which represents the particle's own thinking. The last part of the equation is known as the social component, which corresponds to the information exchanged and cooperation between particles. The velocity of a particle is calculated by combining a fraction of its previous velocity with the distances of the particle from its personal best position and the global best position (i.e., the best among all personal bests). The new global best position $\hat{Y}(t+1)$ and the new local best position $Y_i(t+1)$ are adopted from [228] and defined as follows:

$$Y_i(t+1) = \begin{cases} y_i(t), & \text{if } f(x_i(t+1)) \geq f(y_i(t)) \\ x_i(t+1), & \text{if } f(x_i(t+1)) < f(y_i(t)); \end{cases} \tag{6.3}$$

$$\hat{Y}(t+1) = argbest\{Y_i(t+1)\}, \tag{6.4}$$

There are two approaches to solving a PSO problem. The first approach formulates the entire network as the particle search area. In the second approach, the swarm is partitioned into smaller areas to limit the particle's search area. In the first approach, the social component remains the same as in (6.1). However, in our framework, we use the second approach to change the social part in (6.1) from searching the entire swarm to a neighborhood, as follows:

$$c_2\phi_{2,k}(t)(\hat{y}_{j,k} - x_{i,k}(t)), \tag{6.5}$$

where y_j is the best particle in the neighborhood of the j^{th} particle.

To improve the PSO and create a better balance between the local best and global best search, we calculate the inertia weight using the proposed work in [229]. They proposed a generational process with a detection function to calculate the inertia weight value. This process leads to an improved self-adaptive particle swarm optimization (IDPSO) technique, resulting in an intelligent and more accurate PSO. As a result, the original PSO values of c_1, c_2 and w are now changed to c_{1i}, c_{2i} and w_i and updated as follows:

$$w_i(t) = \frac{w_{ini} - w_f}{1 + e^{\varphi_i(t)(t-(1+ln(\varphi_i(t))t_{max})/\mu)}} + w_f, \tag{6.6}$$

$$c_{1i}(t) = c_{1i}(t-1)(\varphi_i(t))^{-1}, \tag{6.7}$$

$$c_{2i}(t) = c_{2i}(t-1)\varphi_i(t), \tag{6.8}$$

where w_{ini} and w_f are the initial and final values allowed for the inertial weight, t_max is the maximum number of generations, and μ is the adjustment factor to prevent wrong convergence of w, w_ini and w_f. The φ function is defined as follows:

$$\varphi_i(t) = d(\hat{Y}(t-1), X_i(t-1))/d(Y_i(t-1), X_i(t-1)), \tag{6.9}$$

As a result, the new velocity equation for IDPSO for particle $X_i(t)$ is evaluated as follows:

$$V_i(t+1) = w_i(t)v_i(t) + (c_{1i}(t)\phi_1)(Y_i(t) - X_i(t)) + (c_{2i}(t)\phi_2)(\hat{Y}(t) - X_i(t)), \tag{6.10}$$

Some of the advantages of IDPSO over the conventional PSO are:

- Increased Convergence Rate: IDPSO has a higher convergence rate than the conventional PSO algorithm. This means that the algorithm is able to reach the optimal solution faster than the conventional PSO.
- Reduced Premature Convergence: IDPSO reduces the problem of premature convergence, which is a common issue in the conventional PSO. Premature convergence occurs when the algorithm gets trapped in a local minimum and is unable to explore the search space to find the global optimum. IDPSO uses a population-based search mechanism to reduce the effect of local minimums.
- Improved Diversity: IDPSO improves the diversity of the population by using a modified search mechanism that allows particles to explore the search space more efficiently.
- Better Performance: IDPSO outperforms the conventional PSO in terms of both solution quality and computational efficiency.

6.4.2 PSO Data Clustering

In the context of PSO for data clustering, each particle represents a cluster centroid with N_c elements, and each particle x_i is comprised of the following elements $x_i = m_{i,j}, m_{i,l}, ..., m_{i,N_c}$. Where $m_{i,j}$ corresponds to the j^{th} cluster centroid of the i^{th} particle in cluster $C_{i,j}$. Therefore, if we execute the PSO algorithm, a swarm will represent different candidate clusters for a number of data sets $S = \{s_1, s_2, ..., s_n\}$. The fitness of particles can be computed by the quantization error as follows:

$$J_e = \frac{\sum_{j=1}^{N_c} [\sum_{\forall s_p \in C_{i,j}} d(s_p, m_j)/|C_{i,j}|]}{N_c}, \tag{6.11}$$

$$d(s_p, m_j) = \sqrt{\sum_{l=1}^{k} |s_{p,l} - x_{j,l}|^2}, \tag{6.12}$$

where d is the distance to the closest centroid for each data set s_p, $|C_{i,j}|$ is the number of data vectors that belong to cluster $C_{i,j}$, and N_c represents the number of cluster centroids. As a result, Algorithm 3 shows the data clustering using the standard global best IDPSO.

Algorithm 3: PSO Data Clustering Algorithm

Input : P particles, t_{max} maximum number of iterations
Output : Data Clustered C

Initialize each particle to contain N_c randomly selected centriods
for *t=1 to* t_{max} **do**
  for *particle* $i \in P_n$ **do**
    for *data set* $s_p \in S$ **do**
      $d(s_p, m_j) = \sqrt{\sum_{m=1}^{k} |s_{p,m} - x_{j,m}|^2}$ *Assign* s_p *to cluster* $C_{i,j}$ *such that* $d(s_p, m_{i,j}) = min_{\forall C \in N_C} d(s_p, m_{i,c})$
      $J_e = \frac{\sum_{j=1}^{N_c} [\sum_{\forall s_p \in C_{i,j}} d(s_p, m_j)/|C_{i,j}|]}{N_c}$,
    end
  end
  Update global best and local best positions Update the cluster centroids
  $V_i(t+1) = w_i(t)v_i(t) + (c_{1i}(t)\phi_1)(Y_i(t) - X_i(t)) +$ $(c_{2i}(t)\phi_2)(\hat{Y}(t) - X_i(t))$ $x_i(t+1) = x_i(t) + v_i(t+1)$,
end

In the context of IDPSO for data clustering, each particle represents a cluster centroid with a set of parameters. At the start of the algorithm, each particle is initialized with a random set of parameters. During the iteration process, each particle utilizes its personal best and global best position information to optimize its parameters and maximize the likelihood of reaching the global optimum solution, thus improving its fitness. The IDPSO technique employs a population-based search mechanism, which mitigates the impact of initial constraints and conditions present in other techniques, such as K-mean clustering where the search process starts from multiple positions simultaneously. The IDPSO algorithm can identify multiple optimum solutions when they exist in a given problem space. Therefore, when the optimum individual best fitness is found in every iteration of optimization, the value is used to update the candidate solution accordingly. The summary of the advantages of PSO data clustering over K-means data clustering is provided as follows [228, 229]:

- Ability to find global optimum: PSO data clustering can find the global optimum for a given problem, while K-means clustering can only find local optima. This is because PSO searches the entire search space to find the Moreover, best possible solution, whereas K-means starts from random initial conditions and can get stuck in local optima.

- Robustness: PSO data clustering is less sensitive to the initial conditions than K-means data clustering. This is because the PSO algorithm uses a population-based search mechanism, which ensures that the algorithm is less likely to get stuck in a local optima.
- Flexibility: PSO data clustering is more flexible than K-means data clustering as it can be used to cluster data in a variety of spaces, including continuous, discrete and mixed spaces. K-means data clustering can only be used in continuous spaces.
- No need for prior knowledge of cluster number: In PSO data clustering, the number of clusters does not need to be predetermined. The algorithm can determine the optimal number of clusters automatically based on the data.

6.4.3 Virtual Network Slicing

After IDPSO data clustering, virtual network slices must be established. Representation nodes or cluster heads (CHs) should be defined to form a VS. In our proposed framework, nodes join different slices in each round, or new slices are formed at the beginning of the phase. Initially, the MEC-BS identifies nodes with a similar content request to obtain the number of slices. Then, CHs are determined based on the candidate CH identification algorithm. In this algorithm, nodes with the highest data similarity, energy and link capacity are identified as potential CHs. Determining the optimum CHs and the associated nodes to a cluster forms a VS. The optimal CHs are obtained using the following cost functions, given that N is the number of particles and S is the matrix of transmitted data that each column is allocated to a particle $i, j \in N$:

$$\underset{P_i}{\text{minimize}} \quad \frac{1}{n}\sum_{i,j}\frac{1}{\sigma_{P_i}\sigma_{P_j}}\Big[(P_i(x,y)-\mu_{P_i})*(P_j(x,y)-\mu_{P_j})\Big], \tag{6.13}$$

Given that the below function is maximum

$$\frac{\sum_{l=1}^{K} E_{P_i}^{l}}{\sum_{m=1}^{L} E(CH_{P_i}^{m})} > \frac{\sum_{l=1}^{K} E_{P_j}^{l}}{\sum_{m=1}^{L} E(CH_{P_j}^{m})}, \quad i,j \in N, \quad K, L \in S \tag{6.14}$$

where f_1 is the maximum correlation of particle P_i data content request and particle P_j data content request. σ and μ are the standard deviation and mean of the data. $P_i * P_j$ indicates the cross-correlation between two data vectors, which measures the similarity of the two data sets. The function f_2 represents the particles' total energy over the CHs' total energy. Finally, α is a constant that defines the importance of each function in every iteration. The defined fitness functions can identify nodes with the potential to be the CH. Different nodes are selected as CH in each iteration, acting as a local controller to coordinate the data transmission in the clus-

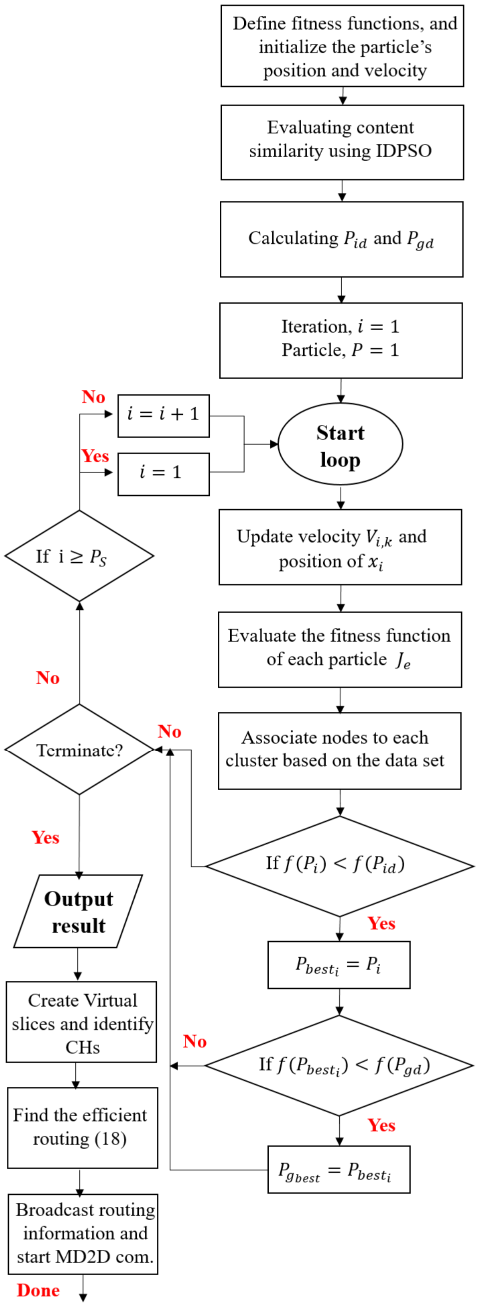


Figure 6.3: Flowchart of the proposed adaptive MD2D routing framework.

ter. Nodes must obey the TDMA scheduling protocol enforced by the BS, reducing collisions among the data messages. After VSs are ready and the CHs are identified, the network is ready for MD2D communication. However, the next step is broadcasting the routing information to each slice. The routing information includes IP addresses, WiFi channel frequency, and the optimal MD2D routing protocol. Based on the performance of the routing protocols in each slice, the most efficient one is prescribed for the VS. The process of finding the optimal route is explained in the following subsection.

Table 6.1: PSO and IDPSO algorithm parameters setup.

Algorithm Parameter	Value
Number of Particles	150
w	0.75
w_1	0.9
w_2	0.4
r_1	1.45
r_2	1.45

6.4.4 PSO and IDPSO Clustering Comparison

PSO is a population-based optimization algorithm that seeks to find the optimal solution by iteratively searching the problem space using a group of candidate solutions called particles. Each particle represents a potential solution and moves through the problem space in search of the best solution. PSO data clustering is a specific application of PSO algorithm that can cluster a data set. Data clustering involves grouping similar data points together to discover underlying patterns or structures in the data. PSO data clustering aims to automatically partition data points into clusters based on their similarity.

Improved self-adaptive particle swarm optimization (IDPSO) is an extension of PSO that introduces dynamic parameter adaptation, allowing for better exploration and exploitation of the search space. This adaptability can potentially lead to faster convergence and improved solution quality compared to the original PSO algorithm. Some of the advancements of IDPSO algorithm are as follows:

1. Velocity and Position Update:
   - PSO: In PSO, the velocity and position updates are calculated using fixed parameters for inertia weight, cognitive weight, and social weight.
   - IDPSO: In IDPSO, the velocity and position updates are calculated based on dynamic parameters. The inertia weight, cognitive weight, and social weight are adaptively adjusted based on the performance of each particle and the overall swarm.
2. Convergence Speed:
   - PSO: PSO has a relatively fast convergence speed, especially in simple optimization problems. However, in complex or multimodal problems, it may struggle to quickly converge to the global optimum.
   - IDPSO: IDPSO aims to improve the convergence speed by dynamically adjusting the parameters. By considering both local and global search information, IDPSO can effectively guide the particles towards promising regions in the search space, potentially leading to faster convergence.

At first we compared the convergence of both algorithm to find the ideal number of iteration. Table 6.1 shows the PSO and IDPSO algorithm parameters setup.

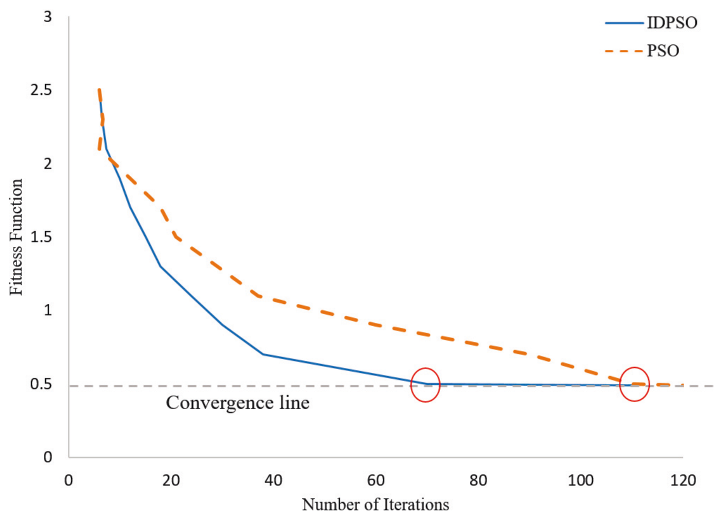


Figure 6.4: Normalized algorithms convergence for IDPSO and PSO.

Figure 6.4 illustrates the normalized convergence behavior of the algorithms. The IDPSO converges to optimum position by almost 70 iterations, whereas PSO takes almost 110 iterations to reach to the optimum convergence area. Both PSO and IDPSO balance exploration (searching for new solutions) and exploitation (exploiting known good solutions). However, IDPSO's adaptive mechanisms provide a more refined balance between exploration and exploitation, enabling particles to effectively navigate the search space and converge faster. IDPSO introduces improvements to enhance the convergence speed and performance compared to traditional PSO. It incorporates adaptive mechanisms to dynamically adjust the parameters, such as the inertia weight, cognitive weight, and social weight. These adjustments are based on the performance of each particle and the overall swarm. By considering both local and global search information simultaneously, IDPSO aims to guide particles more effectively towards promising regions in the search space, potentially leading to faster convergence

To further compare the two algorithms during clustering a data set, we have introduced 4 randomly distributed data sets. Each data set is a vector that represents a package in a wireless network, where N_1 size is 1×30, N_2 size is 1×250, N_3 size is 1×100 and N_4 size is 1×500. Algorithm 4 shows the steps of the complete process on how PSO and IDPSO cluster the nodes into 4 different classes based on their packets. In this Algorithm particles in a swarm represent potential cluster centroids or cluster assignments. The particles are iteratively updated based on their personal best positions and the global best position found so far, guiding their movement towards optimal clustering solutions. In each iteration, the particles' velocities are adjusted according to a combination of their previous velocities, the cognitive component (distance to personal best), and the social component (distance to global best). The positions are updated accordingly, representing new cluster centroids. The fitness of each particle's clustering solution is evaluated using a fitness function that measures the quality of the clustering. Through repeated iterations, PSO and IDPSO dynamically explore and exploit the search space, aiming to converge to an optimal clustering configuration where data points are grouped into meaningful

Pseudocode 1: Evaluating PSO and IDPSO Clustering

1. Initialization of the particles
   i. Initialize $V_i(t), X_i(t), V_{max}, w, w_{initial}, w_{final}, c_1$ and c
   ii. Initialize the swarm size, generation
   iii. Initialize particle to input data set
2. Initialization of the particles
   i. Start iteration in swarm
      a) Find data vectors
      b) Update velocity and position for PSO
      c) Update velocity and position for IDPSO
   ii. Calculate the distance to each centers and assign each data patterm to the nearest center
      a) Calculate fitness function for PSO and IDPSO
      b) Update velocity and positions
      c) Cluster the data based on the centroid and the data set
3. Evaluate the time it takes for PSO and IDPSO to cluster data and exit

clusters based on their similarity.

Figure 6.5 demonstrates the the running time of the PSO and IDPSO to cluster nodes' data cluster between 50 to 400. While the number nodes are low both protocols preform similar in terms of convergence. However, as the number of nodes increases from 50 to 400, the computational complexity of clustering also grows. PSO, being a traditional optimization algorithm, may require more time to converge and reach an acceptable clustering solution as the number of nodes increases. On the other hand, IDPSO, with its improved convergence speed and adaptive parameter adjustments, exhibits faster convergence and significantly outperforms PSO in terms of clustering time for larger node populations.

6.4.5 Route Discovery Phase

The controller prescribes the most efficient routing protocol for network slices. At first, a fixed set of number of flows $F = \{f_1, f_1, ..., f_n\}$ is generated for a random source $SN = \{sn_1, sn_1, ..., sn_n\}$ and destination nodes $DN = \{dn_1, dn_1, ..., dn_n\}$, where $F_i = \{SN_i, DN_i\}$ consist of a source and a destination node. Then, based on the performance of network metrics, the most efficient route is discovered. The performance metrics are based on the network

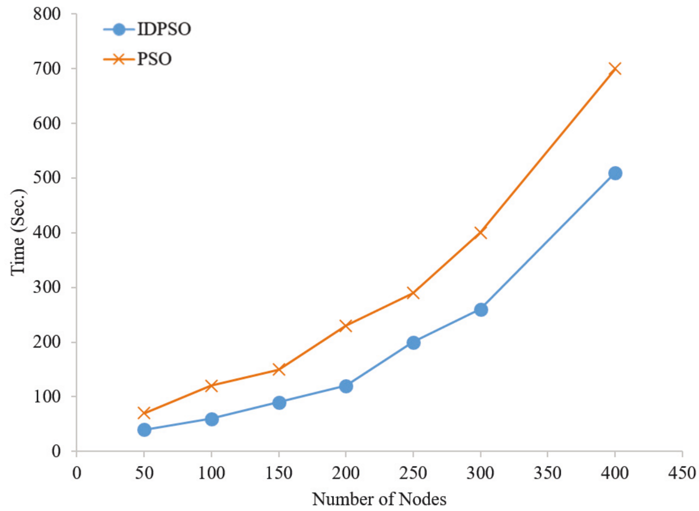


Figure 6.5: Effect of the number of nodes on the clustering time using four different dataset.

operator's preferences: energy efficiency, maximum data rate, link reliability, least E2E delay and PDR. In this chapter, the output of remaining energy (R_E), link stability ratio (L_Z), wireless channel traffic (T_W), and hop count (H) are considered to compare AODV and OLSR. The genetic algorithm (GA) in conjunction with learning automata (LA) is used to identify which routing protocol has the highest efficiency based on the network metrics. The LA is an adaptive process that runs concurrent to the GA to initialize, optimize, and adjust its coefficients using the network feedback. The LA is based on the reward/penalty algorithm that, together with GA, defines the cost function based on the performance metrics. A fitness function is introduced to formulate a mathematical model and select the best route. The fitness function is adopted from [230, 231] and depends on the number of hops, average traffic rate T, remaining energy E, and link stability ratio Z.

$$F_{AODV,OLSR} = \frac{Z+E}{H+T}, \tag{6.15}$$

where the routing protocol R is selected based on the value in (6.15).

$$R = \begin{cases} AODV, & if\ F_{AODV} > F_{OLSR} \\ OLSR, & if\ F_{AODV} < F_{OLSR}; \end{cases} \tag{6.16}$$

the optimal route must have high remaining energy and stability but with the lowest traffic generation and fewer hops to decrease the E2E delay and wifi channel traffic. We now expand the fitness function into the following equations:

$$E_i = \frac{\sum_{i=1}^{N_{hops}} R_{E_i}}{H_i}, i \in F = \{f_1, f_1, ..., f_n\} \tag{6.17}$$

$$T_i = \frac{\sum_{i=1}^{N_{hops}} T_{W_i}}{H_i}, i \in F = \{f_1, f_1, ..., f_n\} \tag{6.18}$$

$$S_i = \frac{\sum_{i=1}^{N_{hops}} L_{Z_i}}{H_i}, i \in F = \{f_1, f_1, ..., f_n\} \tag{6.19}$$

where E is the average remaining energy level of the nodes in flow i, T is the average traffic rate of the same flow i, and Z is the average stability ratio. Also, T_{W_i} and L_{Z_i} are evaluated as follows:

$$T_{W_i} = \frac{P_{TX}}{1 + P_{RX}}, \tag{6.20}$$

$$L_{Z_i} = 1 - \frac{n_{c_i}}{1 + n_{n_i}}, \tag{6.21}$$

where P_{TX} and P_{RX} are the number of transmitted and received packets, respectively. The link stability directly relies on the mobility rate, which depends on the number of neighboring nodes that have changed over time. n_{c_i} and n_{n_i} are the number of changed nodes and the initial number of neighbors.

Figure 6.3 presents the flowchart of our proposed framework, which includes PSO data clustering, VS and adaptive routing protocol selection. In this flowchart, S number of particles are initialized, which contains C randomly selected data clusters based on (6.11). Then the cost function of each particle is computed using (6.13), which contains the similarity of content between the CH and node i. Each particle's personal and global best positions are updated with their velocity using (6.4), (6.3) and (6.10). After the procedure, the swarm with the updated positions and particles' nearest coordinates to the corresponding CHs are recorded and set as a new VS. Our strategy is performed iteratively during the network operation time.

6.5 Simulation results and performance analysis

6.5.1 Simulation Setup

The proposed adaptive network slicing and dynamic routing selection protocol is implemented in the Network Simulator-3 (NS-3). The NS-3 IP-based module includes WiFi and LTE/5G for layers 1 and 2. There are inbuilt routing protocols available for mobile networks. For instance, AODV, OLSR, and destination sequenced distance vector (DSDV) are employed for mobile ad hoc networks (MANETs). These protocols can be extended to work under the MD2D framework to enable multi-hop routing.

Our framework is investigated under a partial section of an entire cellular network covered by an MBS. We consider an MBS located in the center of a cell with randomly generated heterogeneous nodes such as mobile devices, IoT sensors, and vehicles. All the nodes are uniformly distributed and specified as user equipment (UE), following the Random Waypoint mobility model. The simulation environment is expanded over a $500m \times 500m$ area. The channel throughput for cellular transmission is set to have LTE/5G specifications, and WiFi channel bandwidth is determined by IEEE802.11n-5GHz. We assumed a simple energy model

Table 6.2: Simulation parameters.

Parameters	Value
Simulation environment	$500m \times 500$ m
Initial power of UEs	300 j
Number of UEs	50,100,150,200,250
Packet transmission size	10 kbits
Routing protocols	AODV, OLSR
Propagation model	Rayleigh fading
Mobile node transmission range	100 m
Mobile rate	10 m/s - 30 m/s
Mobile model	Random waypoint mobility
Total simulation time	3000 s

[79] for each UE to see the effect of battery depletion. Node's velocity varies between 10m/s and 30 m/s. We analyzed our proposed protocol in two different scenarios. First, we assume a fixed network density with a variable mobility rate. Second, variable network density with the constant mobility rate. Simulation parameters are shown in Table 6.2.

Figures 6.6a-6.7a illustrate the simulation analysis of our proposed protocol with different mobility rates. ASDR represents adaptive slicing and dynamic routing, and AODV/OLSR describes non-slicing and single routing. Figures 6.6b-6.7b investigate the significance of ASDR in different network densities. A predefined set of active flows with randomly chosen source and destination nodes are selected. The simulation results were run for 3000s, and each was averaged over multiple simulations using different seed values. The simulation results are validated with the Monte Carlo simulation technique under 50 runs, and the final results are averaged and plotted with 95% confidence intervals.

6.5.2 Simulation Results

This section discusses simulation results and performance analysis of dynamic network slicing and MD2D route selection. First, we analyze the network metrics against the mobility rate. Then, we observe the network performance versus the network density. The objective of this chapter is to introduce network slicing for WiFi channels, thereby opening up new possibilities for a multi-protocol routing framework. In particular, two routing protocols, AODV and OLSR, were employed to demonstrate the performance of our proposed framework. These protocols serve as fundamental examples of proactive and reactive routing protocols, forming the basis for the development of other routing protocols. The utilization of these protocols was driven by the fact that they are widely employed by researchers in the field. Our analysis highlights that the adoption of a multi-protocol routing framework yields superior performance compared to single-protocol routing frameworks which is the main objective of the chapter. Notably, these two classes of routing protocols continue to be extensively explored by researchers [232, 233].

Figure 6.6a presents the PDR versus different mobility rates. ASDR performs better than networks that either use OLSR or AODV. The reason is that ASDR adapts based on the user applications and chooses the ideal routing protocol for each slice to increase the overall network performance. One of the main advantages of adaptive virtual slicing is the reliability of links. Link reliability is higher in ASDR because each VS is only responsible for one specific user application. Therefore, traffic congestion for relaying nodes is less due to the independency of VSs. For instance, mission-critical and non-critical applications must use a different routing protocol. Mission-critical applications require a fast and responsive protocol. Whereas, the non-critical application can use the route that increases the throughput and reduces the traffic flow on a low battery-level node. The fastest routing protocol must be used if a node requires an emergency broadcast to inform other nodes. ASDR recognizes the application and prescribes the most efficient MD2D routing protocol for every application.

Figure 6.8a illustrates energy consumption versus the mobility rate. The graph shows the difference in energy consumption in slicing and dynamic route selection frameworks against a single routing framework. Our proposed protocol consumes less energy due to the dynamic adaptation of routing protocols. The reason is that ASDR chooses either AODV or OLSR as routing protocols for different slices based on their performance. For instance, when the node's mobility rate is low, AODV performs better, and when the network is highly dynamic, OLSR is better than AODV. Our framework can adjust the routing protocols for each of the VSs based on a multi-objective fitness function involving energy consumption. The benefit of ASDR is it can identify the most efficient routing protocol for each network slice in terms of energy consumption.

Figure 6.9a shows the E2E delay versus the mobility rate, where ASDR is superior to AODV/OLSR. ASDR accomplishes a better result because the whole network is now virtually sliced. Each slice facilitates the most efficient routing protocol based on the mobility rate to minimize the E2E delay. Our proposed framework divides the network into smaller virtualized clusters to split the network traffic from one slice to multiple. This will help to differentiate user applications and make MD2D transmission inherently faster. The application of ASDR will be in heterogeneous networks consisting of mobile, vehicular, drones, etc. For example, vehicular networks usually require similar traffic flows (e.g., incident reports, traffic congestion, network updates, etc.). Therefore, the routing protocol that reduces the E2E delay in highly dynamic environments will be deployed. In terms of E2E delay, OLSR routing performs much better than AODV because OLSR is a proactive link-state routing protocol that keeps a routing table. As a result, once failure occurs, a new route can be instantly identified. Whereas in AODV, a new path must be constructed, which costs time. Hence, the E2E delay of ASDR is much less than AODV for two reasons, adaptability and virtual slicing. Adaptability is identifying the routing protocol that performs best (always choosing the protocol with the lowest E2E delay), and virtual slicing is the clustering of the network into smaller size networks (decreasing the network size and data transmission). ASDR has a lower E2E delay than OLSR due to network virtualization, leading to smaller network densities and faster data transmission.

Figure 6.7a presents the average throughput of the whole network versus different mobility rates. ASDR has a significant advantage because each VS is allocated a frequency band with a sharing pool of resources. Therefore, nodes can send a maximum amount of data over the medium. In AODV/OLSR, the whole network uses the shared resources and frequency band, and nodes can not use the maximum allocated throughput. No matter the mobility rate, the benefit of ASDR is that for each VS, the controller uses our proposed algorithms to identify the protocol that can provide the maximum throughput. Hence, network slices allow access to maximum resources, which means nodes can use the highest throughput allocated to the slice. As the mobility rate increases, packet loss rates rise due to wireless channel errors. Hence, in highly mobile environments, the throughput decreases dramatically.

Now, we analyze the network performance in different network densities while the mobility rate is constant. Figure 6.6b presents the PDR in different network densities. As the number of nodes increases, the network PDR drops because as the number of traffic flows increases, there is more chance of packet loss. ASDR performs better than the networks that either use OLSR or AODV. The reason is that ASDR uses multi-objective route selection mechanisms that facilitate learning automata and genetic algorithms to select the most efficient routing protocol. Link reliability in ASDR is higher than AODV/OLSR in all the network densities because of two main reasons. First, the ASDR divides the network into smaller sub-networks that enable nodes to access more reliable links. Therefore, traffic congestion at relaying nodes is minimal, and nodes transmit the packet to a more reliable node with more space in their queues. Second, we maximize the link reliability based on (6.19). Whereas, if the entire network uses AODV/OLSR, all the applications work under one network slice, and there will be a higher chance of losing a packet.

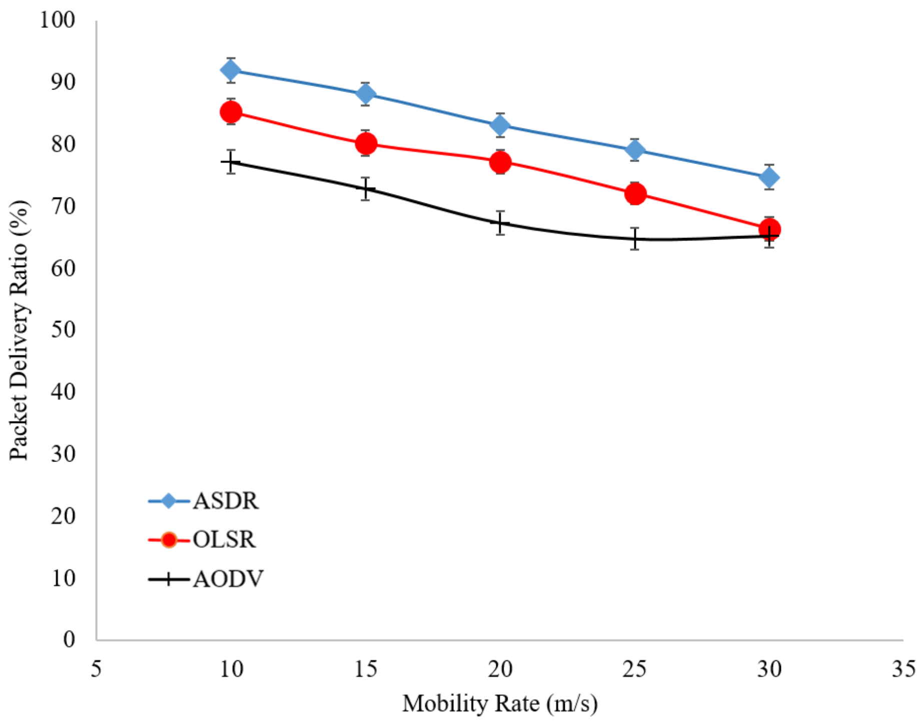
Figure 6.8b illustrates energy consumption in different network densities while running ASDR, AODV, and OLSR. Our proposed protocol shows better energy consumption than the other two non-slicing and single-routing protocols due to the adaptive routing mechanism at the controller. Based on the slice density, the most efficient routing protocol is identified at every iteration using the multi-objective fitness function, which consists of energy, link reliability, and traffic rate. Therefore, our proposed protocol's routing selection mechanism adapts to the network density and selects the protocol that consumes less energy. In contrast, running AODV or OLSR in the entire network without network slicing makes both protocols perform worst. Once the network and traffic flow increase, there will be more chance of packet failure and route reconstruction, causing additional energy consumption.

Figure 6.9b shows the E2E delay versus the node density, where ASDR performs significantly better than AODV/OLSR. ASDR accomplishes a consistent result in all the network densities. The best routing protocol is prescribed based on the VS's node population. As a result, the routing protocol that provides the lowest latency is chosen for VSs for packet delivery. OLSR performs much better than AODV in all the network densities because OLSR is a proactive link-state routing protocol that keeps a routing table and finds new routes immediately after packet failure. However, in large densities and low traffic AODV would generally perform better. According to the protocol behavior and performance, the ASDR adopts the most reliable protocol that provides the lowest E2E delay in that specific network density. Therefore, decreasing the E2E delay of the overall network. Furthermore, the significant increase in End-to-End (E2E) delay observed between 200 and 300 nodes, or at speeds ranging from 25m/s to 30m/s, can be attributed to the performance differences between proactive and reactive routing frameworks. While no notable performance degradation is observed in OLSR and ASDR, which are proactive and hybrid frameworks, respectively, AODV exhibits a substantial increase in E2E delay beyond a certain point. In reactive routing protocols, nodes need to acquire the route to the destination node. Consequently, as the mobility rate and number of nodes increase, the route to the destination undergoes continuous changes, resulting in frequent route requests and failures. On the other hand, proactive routing protocols like ASDR and OLSR maintain a routing table, enabling nodes to swiftly identify alternative paths when a neighboring node fails to transmit the packet.

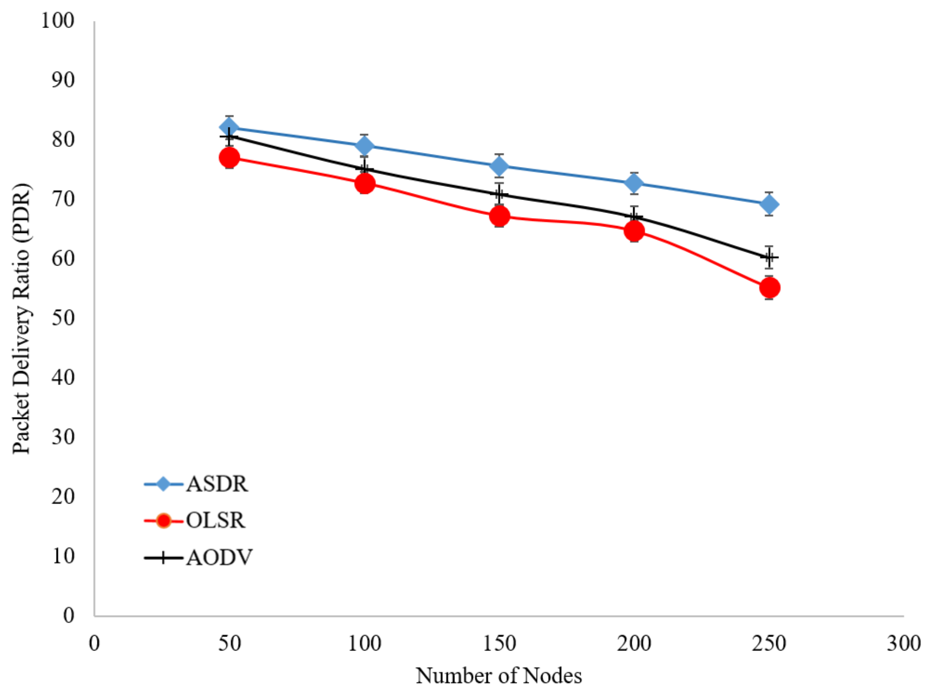
Figure 6.7b presents the average throughput of the whole network in different network densities. ASDR has a significant benefit over the AODV/OLSR for its capability to associate independent WiFi channel bands to each slice. The advantage of the ASDR in terms of network throughput is in the network slicing that provides independent virtual sub-layers enabling access to more WiFi channel capacity. Therefore, individual slices can use the maximum available resources to increase the throughput. We allocate a pool of resources to the entire network where the mission-critical services are given priority, then the rest of the resources are associated with different slices. Once the transmission is finished, the slice resources can be allocated to other slices. AODV/OLSR cannot perform network slicing and use the maximum network WiFi throughput, which is why ASDR has superior performance.

6.6 Conclusions

This chapter presented an adaptive slicing mechanism for WiFi channels with a dynamic routing protocol selection technique for future MD2D and wireless cellular network integration. We utilized the IDPSO algorithm for slicing the network into efficient VSs. Then for every VS, the most efficient routing protocol was identified using LA/GA. Our proposed framework illustrates the significance of the multi-framework MD2D routing protocol in terms of energy efficiency, capacity and E2E delay. We compared our framework with non-slicing and single-routing protocol-based frameworks. The simulation results showed significant advantages in congested and highly mobile networks. According to the results, on average, E2E delay decreased by half, the energy consumption decreased by 8%, and PDR and throughput increased by 5% and 53%, respectively.

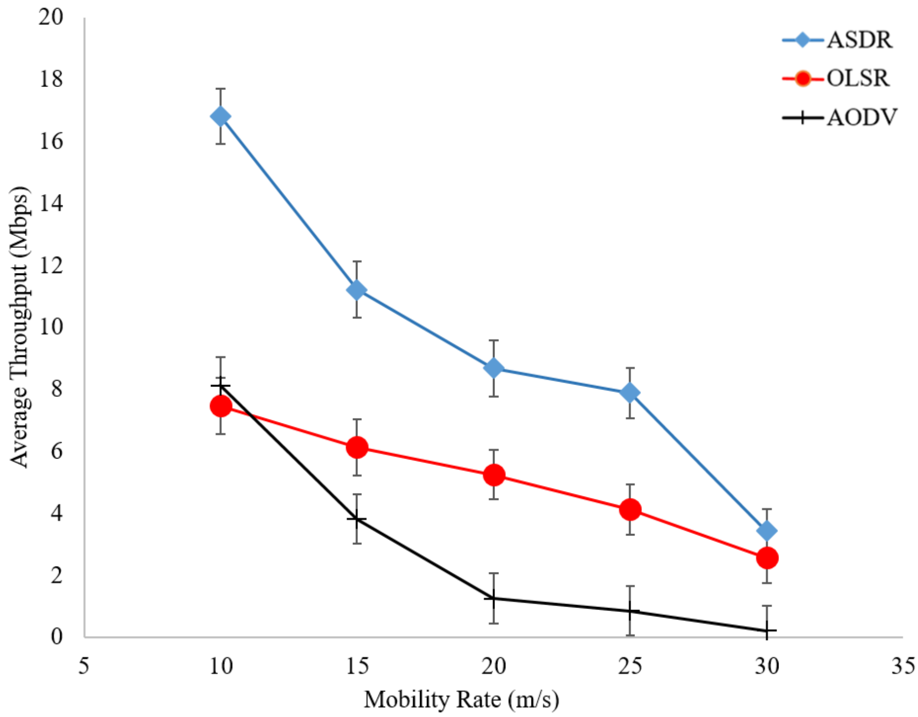


(a) Packet delivery ratio versus mobility rate.

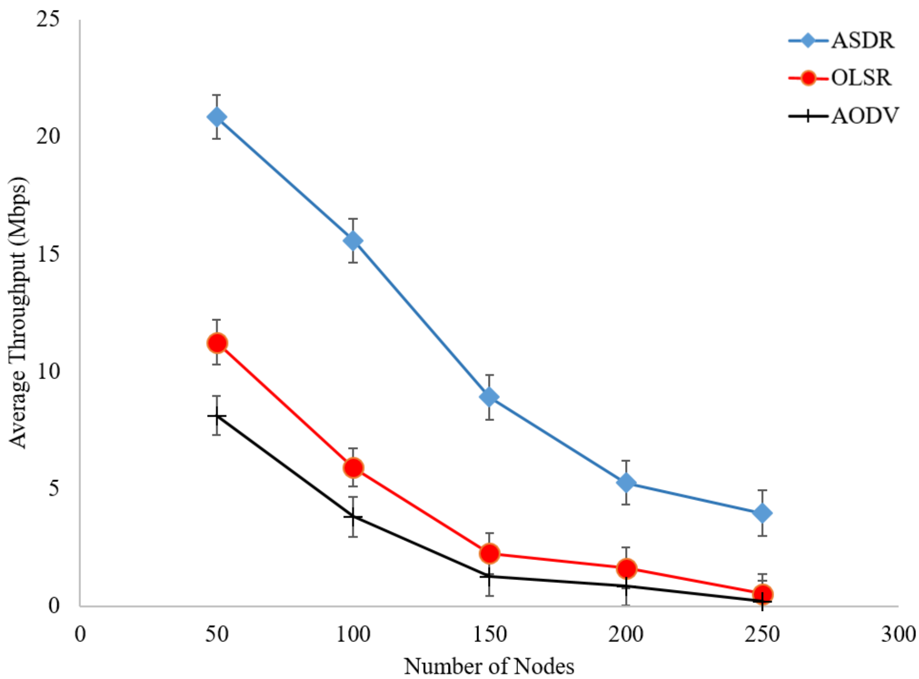


(b) Packet delivery ratio versus number of nodes.

Figure 6.6: ASDR simulation results.

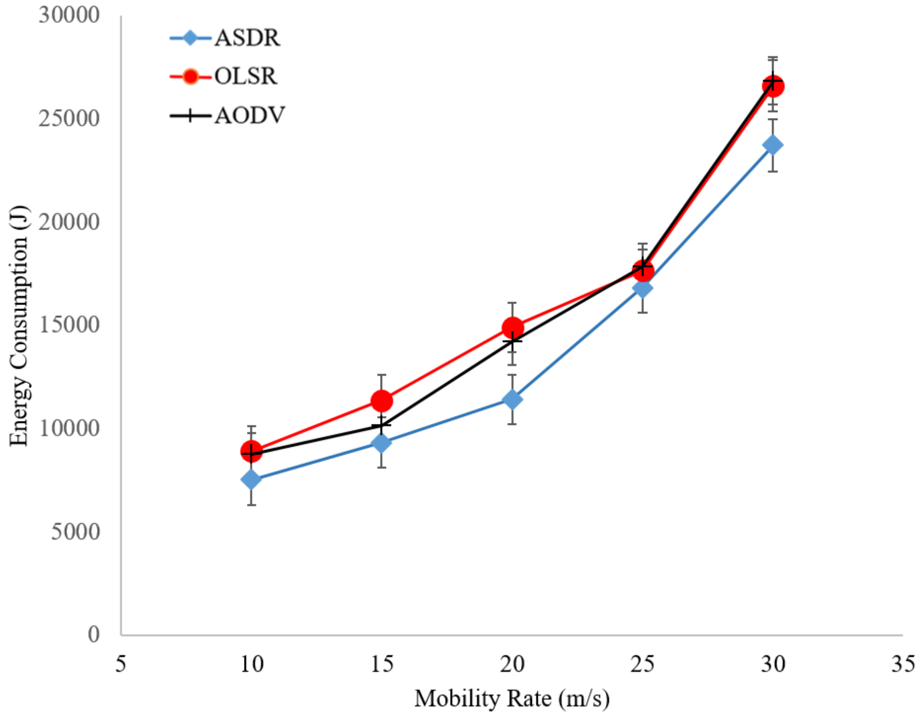


(a) Average network throughput versus mobility rate.

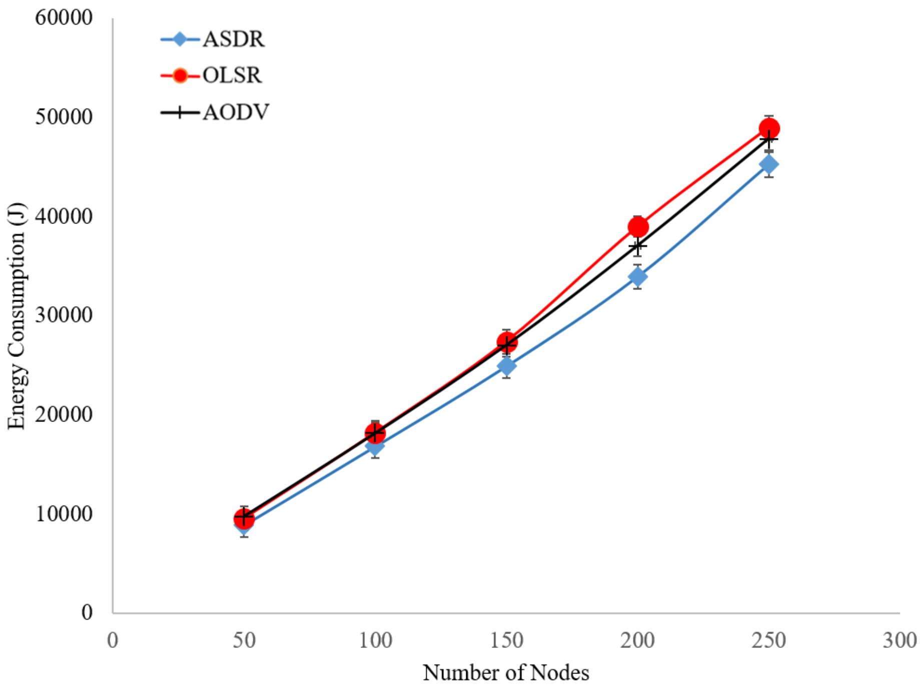


(b) Average network throughput versus number of nodes.

Figure 6.7: ASDR simulation results.

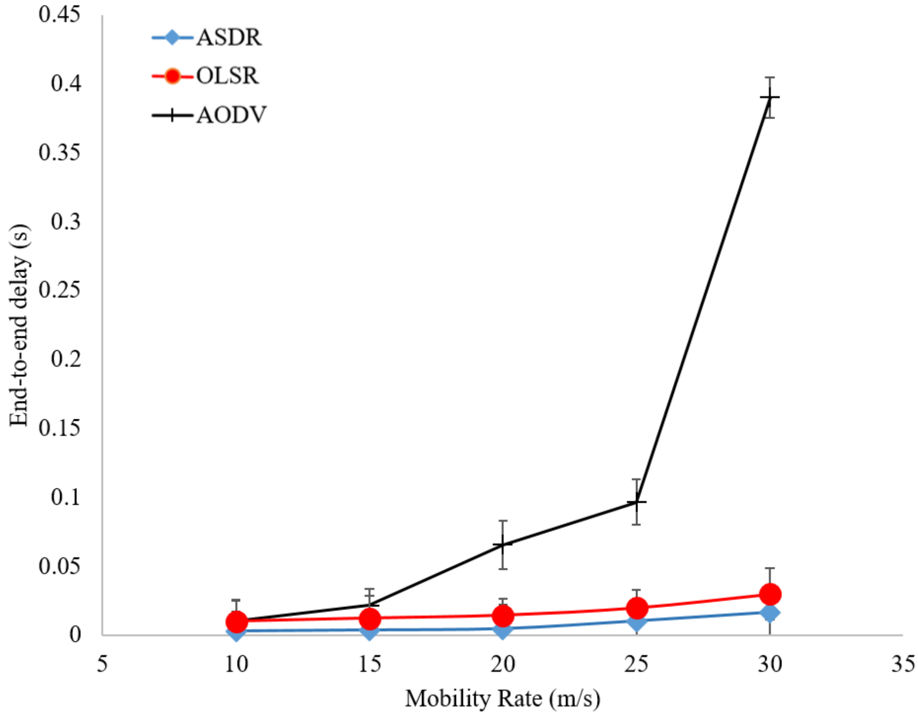


(a) Nodes energy consumption versus mobility rate.

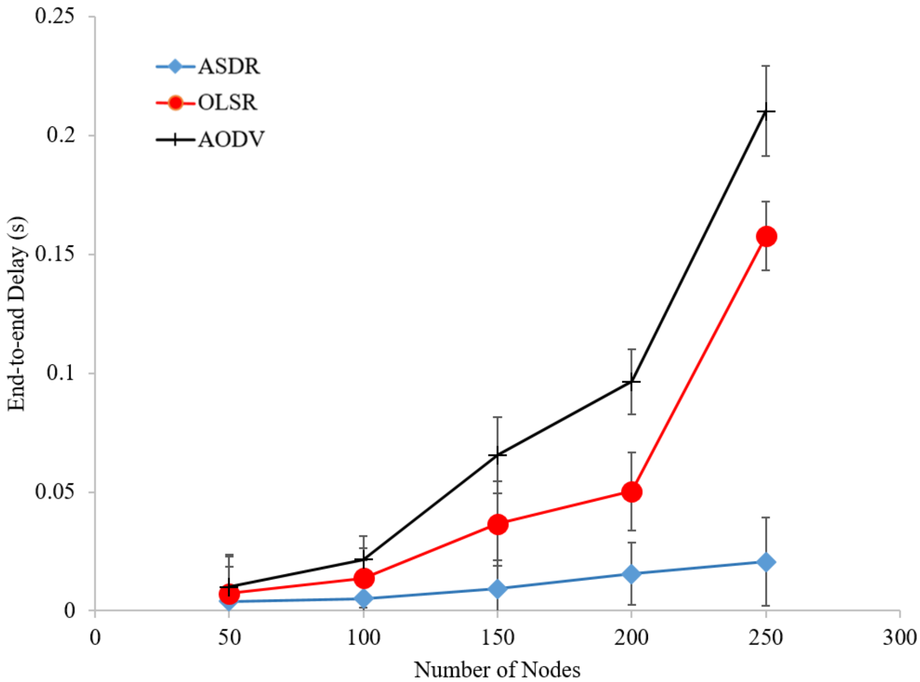


(b) Nodes energy consumption versus number of nodes.

Figure 6.8: ASDR simulation results.



(a) End-to-end delay versus mobility rate.



(b) End-to-end delay versus number of nodes.

Figure 6.9: ASDR simulation results.

7

A Location-based Multi-hop Routing Protocol for Future Wireless Cellular Networks

7.1 Overview

The upcoming research is submitted to Elsevier Computer Networks [1]. This chapter's main objective is to present a new MD2D position-based routing (MPRR) protocol using the coordinates of nodes with zone-prescribed neighbor discovery to make fast and reliable routing decisions. The presented work uses an encrypted cluster-based algorithm for data privacy in multi-hop routing. The proposed protocol was compared with a purely link-state routing protocol.

7.2 Introduction

Device-to-device (D2D) communication was introduced as one of the enabling technologies in the fifth-generation (5G) cellular networks [73]. D2D consists of heterogeneous devices that can communicate and exchange data without any fixed infrastructure or the base station (BS). D2D networks allow cellular networks to allocate some of the data traffic to a secondary infrastructure

[1] Ashtari, S., Abolhasan, M., Lipman, J., Shariati, N., Ni, W. (2023). A Location-based Multi-hop Routing Protocol for Future Wireless Cellular Networks. Elsevier Computer Networks.

to reduce traffic congestion and increase network capacity and coverage. The 3rd generation partnership project (3GPP) has been working on D2D standardization, and recently in release 15, two-hop D2D communication has been approved [211, 234]. From the proliferation of D2D and advancements in radio network architecture, multi-hop D2D (MD2D) communication will soon be part of future wireless cellular networks.

As a result of the exponential growth of heterogeneous devices and applications, the radio access network (RAN) is evolving to incorporate intelligence and openness into the network. The open networking foundation (ONF) and open-radio access network (O-RAN) are among the alliances that are working towards intelligent and self-adaptable radio platforms [7, 150]. O-RAN is considered a promising solutions for future wireless cellular networks to allow vendors to create intelligent virtualized applications [27]. O-RAN facilitates the concept of software-defined networking (SDN) and network function virtualization (NFV) to make the network flexible, self-manageable and programmable. The openness of network functions in future radio networks allows the integration of MD2D into the network architecture.

MD2D is an essential technology for existing and future wireless cellular services such as self-driving vehicles, unmanned aerial vehicles (UAVs), live data streaming, and video sharing. Routing protocols play a crucial role in MD2D communication [3]. Without efficient routing, the MD2D network can impact the overall network performance. In recent years, various routing protocols have been introduced to address the challenges of multi-hop routing [86, 89, 235]. In general, routing protocols are centralized, distributed or hybrid (a combination of centralized and distributed) frameworks. Centralized frameworks use a controller to coordinate and instruct nodes on how to relay packets. Whereas distributed routing protocols use flooding algorithms to obtain routing information. Centralized routing protocols have shown superior performance compared to distributed [79]. As shown in Figure 7.1 in centralized frameworks, two approaches could be used to obtain a route: topologic-based (link-state) or location-based.

7.2.1 Motivations

Topologic-based approaches use information from neighbor discovery to identify routes [3]. Nodes can locate and register neighboring nodes using link-quality or a signal-based metric. The benefits of topologic-based/link-state routing are the reliability of the route and the lower packet failure. The disadvantages of link-state routing are in highly dynamic networks, such as vehicular. As the mobility rate increases, the probability of route failure or packet loss rises. Location-based routing have been proposed to address such challenges associated with the link-state approaches. Location-based routing are especially advantageous because, first, it only requires the coordinates of the nodes, which is easy and fast to get; second, it reduces the burden of obtaining the neighboring information on the WiFi channel [236]. However, location-based routing can be inefficient in urban areas, where there are many obstacles causing signal deterioration and link failures. This will cause repeated retransmission of packets, which can

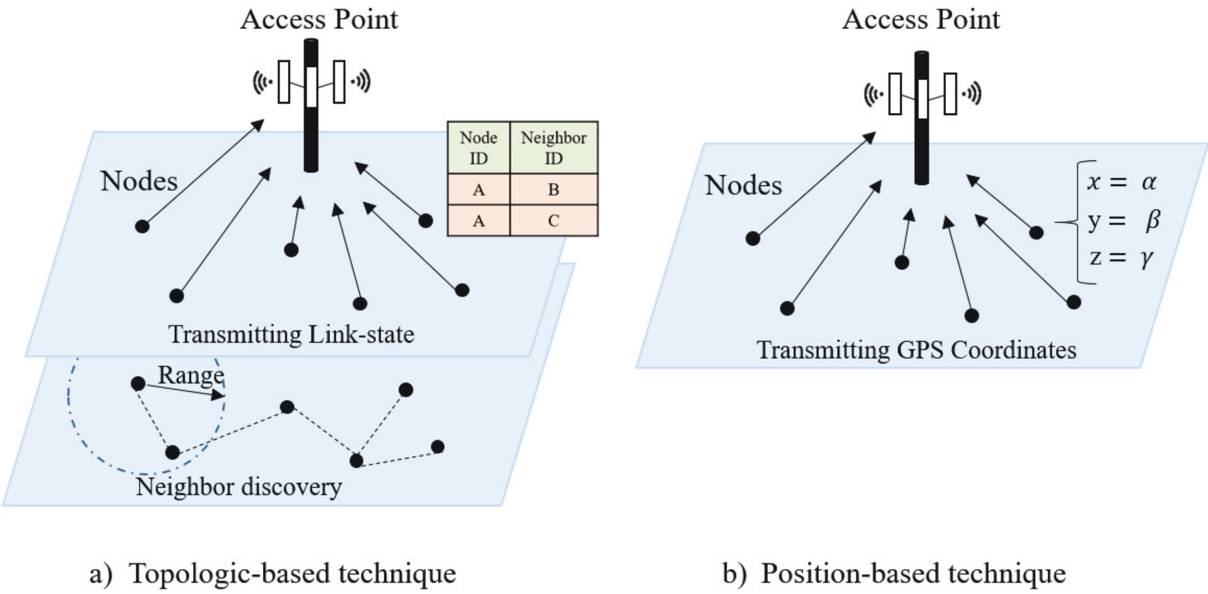


Figure 7.1: Abstract comparison of topology-based and location-based routing protocols.

lead to exposure to various security and privacy threats. Therefore, there is a requirement for a high-performance, link-reliable and secure location-based routing protocol. Different privacy-based routing protocols have been proposed for wireless networks [237–239], which require some standard cryptography and authentication methods. However, they need extra capacity with fast processing units, which leads to additional costs.

To the best of our knowledge, no previous study has introduced a privacy-preserving location-based MD2D routing protocol along with a zone-prescribed neighbor discovery mechanism. This combined approach aims to enhance the performance of location-based routing while safeguarding the confidentiality of nodes' identities and locations during multi-hop communication. Our proposed protocol offers two primary advantages. Firstly, by employing a zone-prescribed neighbor discovery mechanism, we enhance link reliability while minimizing unnecessary procedure execution to specific areas only. Secondly, our privacy-preserving multi-hop routing scheme ensures location privacy alongside agile service response during WiFi communication. While existing WiFi standards, such as WiFi-Protected Access II (WPA2) or Advanced Encryption Standard (AES), address data privacy concerns, our protocol incorporates an additional encryption method to ensure the security of both location and multi-hop routing. This ensures that intermediate nodes cannot access the source node's data or location information.

7.2.2 Contributions

This paper proposes a new MD2D position-based routing (MPBR) protocol using only the GPS coordinates of the nodes. Our protocol uses network footprint and decides on the zones that need to perform neighbor discovery. Then, a privacy-preserving scheme is used in the WiFi channel to prevent MD2D users from being affected by data and location breaches. The BS uses node locations to create geographical link-state databases (GLSDB). The GLSDB is

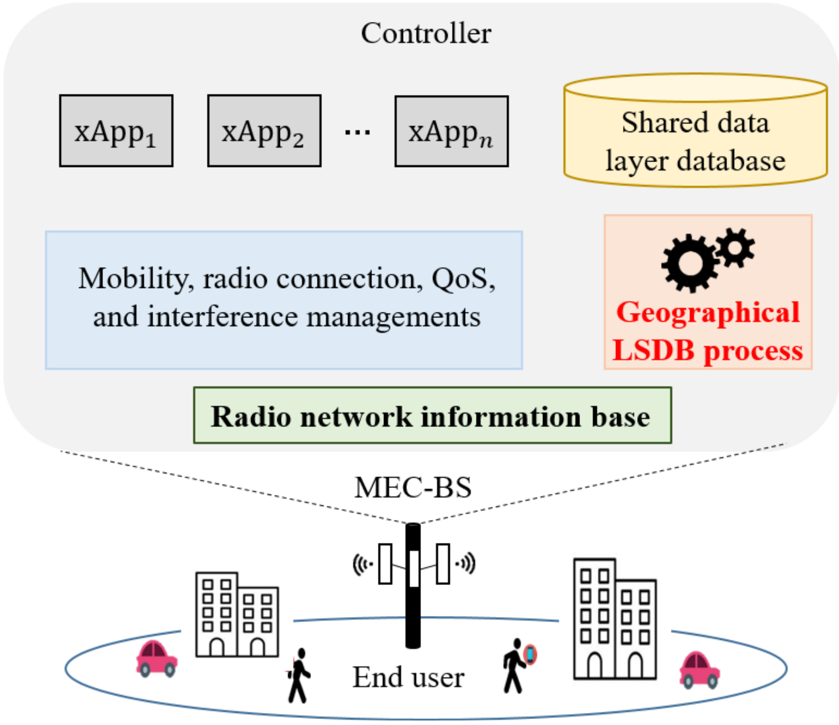


Figure 7.2: Network architecture.

constructed at the controller attached to the mobile edge computing (MEC)-BS, as shown in Figure 7.2. The controller identifies the zones required to perform neighbor discovery using the network footprint and the obstacle density in the urban environment. With this information, the controller completes its GLSDB. Then the GLSDB is transmitted to the nodes, where the nodes calculate their path to the destination using the multi-objective shortest path (MOSP) algorithm. The routing objectives include link reliability, distance, delay, and channel overhead to identify the most efficient route. The nodes must obey two constraints to obtain the optimal next hop: the distance and number of hops to the destination node.

The following are the key contributions of this chapter:

- We present a new location-based routing protocol for MD2D communication in future cellular networks. The protocol utilizes network footprint to determine the zones for performing neighbor discovery, and creates a GLSDB routing table.

- We introduce a privacy-preserving messaging framework to develop a reliable E2E connection between MD2D nodes. An XOR-based privacy preservation scheme protects the node's location and data. The advantages of XOR scheme are fast and low in complexity.

- We present a multi-objective shortest path (MOSP) algorithm to identify routes with the highest efficiency. MOSP uses the combination of different routing objectives, such as link reliability, distance, delay, and channel overhead to identify the best route.

- We implemented our protocol in NS3 and conducted a comparative simulation analysis to provide insights into the performance difference between the location-based and link-state MD2D routing protocol. The simulation results show, on average 15% reduction in energy consumption, a 20% reduction in E2E delay, a 5% increase in PDR and almost 5% increase in throughput compared to fuzzy-based participation and routing protocol for MD2D (FPRM) [240] and hybrid SDN architecture for wireless distributed networks (HSAW) [83]. Moreover, MPBR performs significantly better than other investigated ad hoc protocols in terms of PDR, E2E delay and energy consumption.

The rest of this chapter is organized as follows. Section 7.3 provides a brief overview of current studies in the literature. In Section 7.4, the system model of MPBR is presented. Section 7.5 provides detailed explanation of route discovery procedure. Next, in Section 7.6, the details of MPBR problem formulation and identification of routes are explained. This is followed by Section 7.7, where simulation results of the proposed routing protocol are illustrated. Finally, Section 7.8 summarizes the study.

7.3 Related Works

The main goal of this study is to create an MD2D location-based routing protocol that preserves the location of users with higher performance compared to MD2D link-state-based routing. Therefore, we present a brief overview of the current routing protocols proposed for the two frameworks and summarize the advantages and challenges.

7.3.1 Link-state Routing Protocols

The primary distinction between location-based and link-based routing algorithms lies in their approach to determining the next hop node. Location-based routing relies on the GPS coordinates of nodes to make this decision, while link-state routing utilizes neighbor discovery to identify the most suitable next hop candidate. However, it is important to note that two nodes may be in close proximity geographically, yet face communication obstacles that disrupt the link. In this regard, link-state routing frameworks enable nodes to establish a neighboring table, which reveals the communication link status between nodes. Consequently, link-state routing protocols offer enhanced reliability for mobile-to-mobile (MD2D) communication. Within link-state routing, two primary frameworks exist: centralized and distributed [79].

Authors of [86] a centralized on-demand Mobile-to-Mobile (MD2D) routing protocol with the aim of minimizing routing overhead and improving network scalability. They propose a source-based routing protocol that leverages a Software-Defined Networking (SDN) controller. Under this protocol, each source node initiates a path request to the controller, which then identifies the most optimal route and communicates it back to the source node. Similarly, the authors of [93] introduced a semi-source-based routing protocol in their proposal. In this protocol, the source node initiates a path request to a centralized controller, which subsequently provides instructions regarding the packet and the next relay node to both the source node and the intermediate nodes along the route. A Software-Defined Networking (SDN) controller is responsible for managing the network information and constructing a centralized routing table based on the link-state information of the nodes. In [241], the authors used an adaptive load-balancing mechanism based on the predicted link-state of nodes to calculate optimal paths between source and destination nodes. They use a neural network in the SDN controller to indicate the link-state and enable load-balancing. Then, the predicted values are used as the Dijkstra weight to calculate the optimal path and proactively add flow table entries.

While link-state routing protocols demonstrate efficient and reliable performance, their response agility is limited by the neighbor discovery process. The evolving wireless cellular network demands rapidity to accommodate diverse applications such as autonomous vehicles, smart transportation, and more. In these specific use cases, the link-state routing protocol falls short in terms of accident mitigation and response time optimization. To address this concern, location-based routing protocols present a viable solution by providing swift service responses. In the subsequent section, we delve into an exploration of several contemporary location-based routing protocols.

7.3.2 Location-based Routing Protocols

Location-based routing protocols have the potential to address the limitations of link-state Mobile-to-Mobile (MD2D) protocols, particularly in scenarios characterized by highly dynamic network topologies and nodes with high mobility. Link-state routing protocols tend to exhibit increased energy consumption due to the network overhead stemming from the neighbor discovery process, as indicated by Kumar et al. in their survey [242]. This, in turn, leads to an increase in end-to-end (E2E) delay, a critical factor in vehicular networks where timely updates on safety and road hazards are essential and must be efficiently broadcasted throughout the network. Consequently, location-based routing has emerged as a preferred choice in most vehicular networks as it helps alleviate the challenges associated with link-state protocols.

Authors of [243] proposed a predictive geographic routing protocol (PGRP) to improve link reliability and E2E delay in vehicular ad hoc networks (VANETs). In PGRP, every vehicle exchanges a weight with its neighbor nodes, including direction and the vehicle's angle. Their algorithm can predict the future location of cars according to the acceleration and location information. Once the location of nodes is obtained, a network graph is created, and nodes start transmitting. Their algorithm shows higher PDR and lower E2E delay than other VANET routing protocols. Similarly, authors of [244] proposed an intersection-based geographic routing with transmission quality (IGRTQ) guaranteed in Urban VANETs. They use historical and real-time information to assign weight to each road segment to choose the best path. This weight is based on the delay and data congestion at each road segment. The most efficient road segment is dynamically selected using the allocated weight from segments, and the ideal next-hop nodes are chosen accordingly. Their algorithm guarantees fast and efficient packet transmission in a highly dynamic environment.

Authors in [245] proposed a fuzzy-based routing protocol using geographical information. Their algorithm considers several factors to select the most suitable next-hop for packet forwarding and route generation, such as the vehicle's location, direction, link quality, and throughput. The simulation result shows better PDR, E2E delay, and throughput than other location-based routing protocols. To improve the network performance and overcome geographical routing challenges in terms of network lifetime and resource optimization, some authors used ML techniques in their protocols. For instance, authors in [246] proposed an efficient particle swarm optimization (PSO)-based resource optimization geographic routing (PS-ROGR). The PSO is applied to enhance the geographical routing by adding an optimization technique and considering the location and velocity of each particle. According to the best fitness value in the update stage, particles are selected for multi-hop routing. The PS-ROGR protocol shows the significance of location-based routing in highly dynamic environments. There are other location-based MD2D routing protocols in wireless communication networks, including IoT [247], MANET [248], and wireless sensors [249].

7.4 System Model

The MPBR protocol consists of three main steps before MD2D is operational. First, the BS identifies the zones where the neighbor discovery must happen based on the network footprint and the node's locations. Second, generating the GLSDB using the geographical location of nodes together with the link state of nodes in the designated zones. Finally, the BS clusters the nodes into groups to distribute the secure management keys for data and location privacy. The following subsections describe the operation of each step in detail.

7.4.1 Zone-Prescribed Neighbor Discover

In our protocol, neighbor discovery is essential in areas characterized by high congestion due to the presence of large buildings and other obstacles. Figure 7.3 illustrates a geographical area where we recommend nodes in two distinct zones to conduct neighbor discovery based on their respective locations. The determination of these zones relies directly on the network footprint and the spatial distribution of buildings and nodes. Depending on the movement patterns of the nodes, as discussed in Section 7.6.4 regarding location prediction, different nodes are assigned the task of performing neighbor discovery. In our protocol, we employ image processing techniques to capture the coverage area under a single base station (BS) and extract information about the buildings. This process can be achieved using edge computation or advanced machine learning (ML) algorithms. Inspired by the approach presented in [250], we utilize the Canny edge detection algorithm to identify the boundaries of the buildings. Typically, when observed from a top-down perspective, buildings exhibit a rectangular shape or cover a rectangular area.

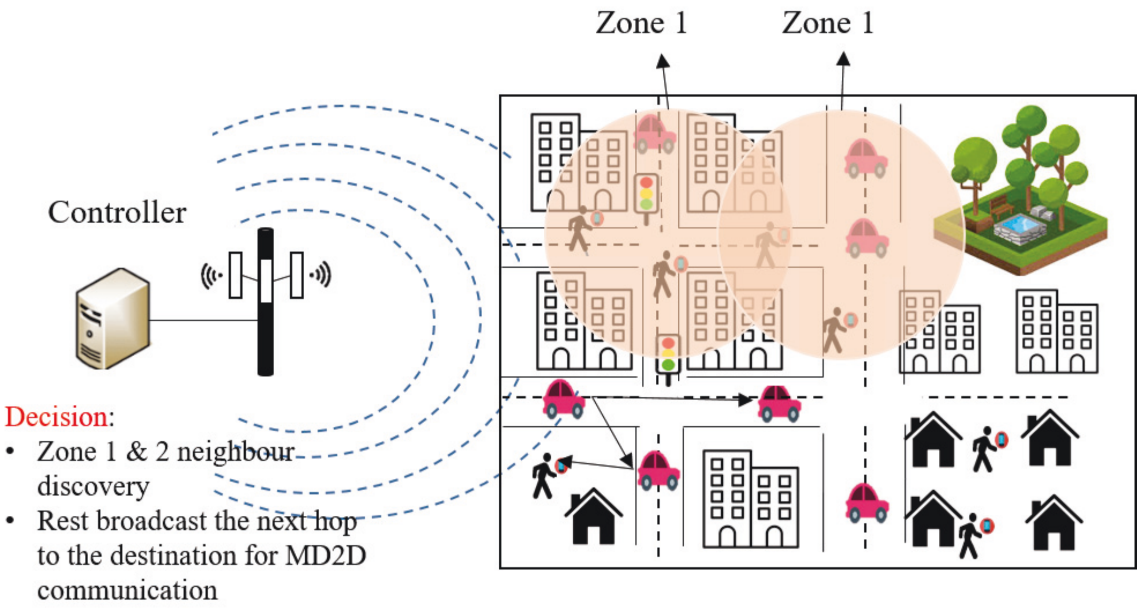


Figure 7.3: Zone selection to perform neighbor discovery.

By employing the Canny edge detection algorithm, we outline the edges of the buildings and subsequently utilize the Hough transform to vectorize the detected edges. The Hough transform aids in identifying analytically defined shapes, such as circles, lines, and ellipses, within the image. These procedures are executed prior to network setup, and once the network footprint is extracted, it is inputted into a network simulator to distribute the nodes accordingly. Figure 7.4 shows our building extraction method and network footprint detection under a BS's coverage. This simple but effective method can be successfully used to detect buildings from a noisy image to extract overlapping edges. Figure 7.4 shows the final network footprint, where the controller initiates the configurations of MD2D communication.

7.4.2 GLSDB Calculation

In a designated area where a base station (BS) is operational, two distinct Geographic Location Service Databases (GLSDBs) are maintained: the intra-cell GLSDB and the inter-cell GLSDB. The intra-cell GLSDB includes local cell information, while the inter-cell GLSDB contains details pertaining to neighboring cells. Once the nodes share their respective locations with the BS, the inter-cell GLSDBs are established. These GLSDBs are subsequently shared with adjacent cells to facilitate seamless handover procedures. To construct the GLSDB, the controller assesses the cost associated with each node, utilizing the node information stored in the node information base (IB). Four different cost metrics are employed to populate the GLSDB table, namely: distance, number of hops, remaining energy, and link expiry time. The mathematical intricacies of each metric are elaborated upon in Section 7.6. Next-hop candidates for each link are identified using the MOSP, such as the device's distance, cost, delay, load, reliability, and transmission range. After populating the GLSDB, we check the link expiry time to ensure links

a) Satellite image b) Building extraction process c) Building footprint

Figure 7.4: Network footprint extraction from satellite image.

are reliable after a particular time due to the mobility of nodes.

7.4.3 Encryption Clustering Technique

The proposed encryption clustering technique is inspired by [251], where the BS groups the authorized MD2D nodes into clusters to provide secure keys. In our framework, the clustering algorithm is translated to consecutive rings formed by different values of the received signal strength indicator (RSSI) received at the BS. As shown in Figure 7.5, each ring corresponds to a cluster and a region. Each region is marked with a different encrypted key starting from the first ring R_1. If we assume the number of given rings is i, then the ring widths is $R = \{R_1, R_2, ..., R_i\}$. The width of each ring can be represented by a range of received signal strength (RSS) which can ultimately create a list of $RSS = \{RSS_1, RSS_2, ..., RSS_n\}$, where RSS_n is between the initial and max threshold value of rings ($\alpha_{R_1} \leq RSS_n < \beta_{R_1}$). Therefore, the set of RSS decides the ring width while constructing the cluster structure. In this chapter, we assume a fixed number of rings and that the width of each cluster ring is assumed to be given.

To distinguish the border nodes and associate them with a ring, the BS broadcasts a border decision message (BDM). The BDM contains the identity of a newly-formed ring, the RSS_R for the ring and a threshold δ. Nodes in the first region will receive the signal and measure the RSS of received BMD (RSS_I). If the difference between RSS_I and RSS_R is less than δ, the nodes will consider themselves members of the first ring. The threshold is a margin that helps nodes during the ring association process and more importantly, decreases the overhead and time of the node's association process at the border of two rings. In cases where a node receives multiple BDMs, the strongest correlation to the actual RSS is considered, identifying the ring to which the node must associate. The ring creation and cluster formation process repeats until the entire network under the BS coverage is clustered. During this process, nodes in the same ring will receive a similar encrypted key to use during MD2D communication.

7.5 Route Discovery Procedure

This section explains the route discovery process in MPBR protocol. The details of network policies and the exchanged packets are discussed. In the end, we briefly overview the FPRM and HSAW routing protocols as high-performance routing protocols to serve as benchmarks for

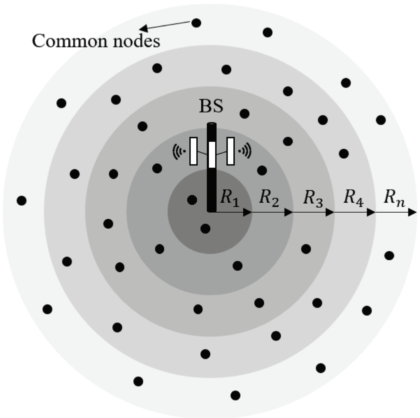


Figure 7.5: Geographical clustering for distribution of secure keys for MD2D communication.

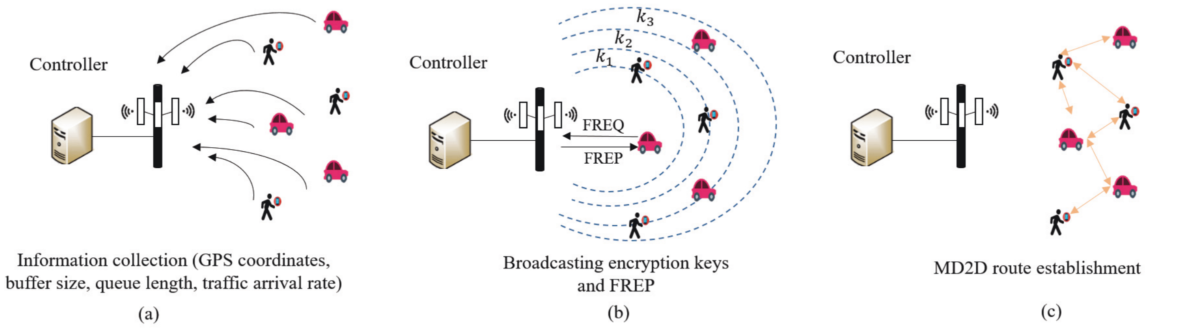


Figure 7.6: An Illustration of the route discovery operations, where a single edge control is shown for illustration convenience. However, the operations can be straightforwardly extended to multiple controllers.

MPBR.

7.5.1 Route Discovery in MPBR

MPBR facilitates a centralized framework where a controller processes geographical locations and creates a GLSDB. Nodes use the GLSDB to acquire a route to their intended destination. Two separate frequency bands are assigned for data communication. 5G standards are used for cellular channel communication, which allows forwarding control packets for routing information, route maintenance, and packet failure. The IEEE 802.11 [203] WiFi frequency bands are used for exchanging data packets in the WiFi channel.

Figure 7.6 provides insights into MPBR architecture, where various heterogeneous nodes are scattered over a large geographic area. Figure 7.6a illustrates the interconnection between the controller and the BS, where the link can be wired or wireless. Network information consists of the mobility rate, remaining energy, bandwidth, and location of nodes. Other technical information, including the number of neighboring nodes, queue length, and traffic arrival rate, is also shared with the controller. It is also assumed that the BS obtains the satellite image under its coverage to extract the building and the area's footprint. Two separate actions are performed after receiving node coordinates. First, calculate the cost of each node to be selected as a forwarding relay. A table consisting of all the calculated cost corresponding to a particular path to a neighboring node is created (GLSDB detail in Section 7.6). Second, the controller prescribes zone-based neighbor discovery using the satellite image and nodes' location.

Later, the controller generates N random virtual clusters and sends K encrypted keys to each cluster, $K = \{k_1, k_2, ..., k_N\}$. Therefore, each cluster member will receive a similar code but different from the others. As shown in Figure 7.6b, nodes in the same cluster/ring have received a unique key. Once a node has a packet to transmit, it will send a flow request (FREQ) message to the controller. The controller sends a flow reply (FREP) message containing the routing information to the source node. Then, the source node starts the transmission by attaching the encrypted code to the packet.

Figure 7.6c shows MD2D communication after the routing information has been received. Remember that if a designated zone needs to perform a neighbor discovery and a relay node is within that zone, then that node must have completed the prescribed zone-based neighbor discovery. We are using the 802.11 MAC for the MD2D WiFi communication. Finally, each node is responsible for acting as a forwarding element from the controller's point of view and an end-user from a cellular perspective.

7.5.2 FPRM Overview

This section and the following section explain the route discovery procedure of two routing protocols compared to the MPBR. FPRM is an energy-efficient MD2D routing protocol that uses neighbor discovery and the fuzzy logic system to obtain routes. In the FPRM protocol, nodes use Hello packets to acquire information about the link state of their neighbors. Then, each device reports its neighbor discovery table, remaining energy, mobility rate and the number of neighbors. Then the fuzzy logic system identifies nodes with minimum eligibility index. The eligibility index relies on the energy of nodes, the number of neighbors and the mobility rate. For instance, if a node is highly mobile and has low battery levels, then this node might cause performance deterioration to the network because it has a higher probability of losing a packet. Once the eligible nodes are defined, the LSDB is created and broadcast to the whole network. Then nodes use a cost function to find their route to the destination. Therefore, only a selected number of nodes participating in the MD2D communication. The proposed protocol increases the network lifetime and average throughput.

7.5.3 HSAW Overview

The HSAW protocol utilizes neighbor discovery, also known as Hello packets, to gather information about the link state of neighboring nodes. The devices in the network then share this information with the central controller via topology control (TC) messages. The controller compiles an LSDB of its coverage area and distributes it to all devices in the network, providing a comprehensive view of the network and enabling the creation of routing tables. The controller also broadcasts traffic policies, which include the traffic type, a maximum cost based on the data packet size, and metric allocation for each traffic type. The transmission process starts by checking the routing table for a pre-existing route. If such a route exists, the packet is forwarded to the next hop. If not, Dijkstra's algorithm is employed to find the most efficient path to the destination. If the found path meets quality requirements, it is added to the routing table. If it does not, the packet is sent to the BS for transmission over the cellular network. If a node fails to respond within a set threshold, other nodes will notify the BS, which will then send a delete entry to update the LSDB for all devices in the network.

7.6 Problem Formulation

This section provides the mathematical modeling and constraints used to solve the routing problems and increase network performance. The main objectives of this paper are to minimize energy consumption and E2E delay while providing secure and reliable MD2D communications. Our problem formulation stage is divided into five sections. At first, the energy model is explained, followed by the route identification strategy and GLSDB population based on the multi-objective shortest path algorithm. In the GLSDB creation process, once the links are identified, they must continuously be monitored for validity through the use of a link expiry time. If a connection between two nodes is no longer valid, the algorithm performs node location prediction to facilitate node adoption when they move into the neighbor discovery zone. Afterward, nodes initiate MD2D communication by attaching a distributed encrypted key to their data packet. Algorithm 4 provides a comprehensive overview of the MPBR route protocol selection process.

7.6.1 Energy Constraints

We use the remaining energy of nodes as a cost metric to identify the best candidate next hops. A simple energy model is employed for radio hardware similar to work in [143]. We assume a directed graph $G(N, L)$ to model our wireless cellular network spread across an $M_1 \times M_2$ two-dimensional area. A controller is responsible for N number of nodes distributed with $L_{i,j}$ links between neighboring nodes $(i, j) \in N$. Each node is assumed to have an initial energy value E_{in}, where the transmitter dissipates energy (E^{TX}) by sending k-bits of data packets through radio link, and the receiver dissipates energy by demodulation and encoding the message (E^{RX}). Therefore, the total consumed energy E_T between node i and j is defined as transmitting K-bits of data and receiving M-bits of data by:

$$E_{T_{i,j}} = \sum_{l=1}^{K} \sum_{s=1}^{M} E_{l,i}^{TX} + E_{s,j}^{RX} \tag{7.1}$$

Algorithm 4: GLSDB calculation

Input : Number of iterations N_I,Number of UEs N, locations P, nodes energy E, and destination d_i

Output : Geographic LSDB

for $n_i = 1 \quad i \in N$ **do**
    Each UE broadcasts E,P, and d
end
while *nodes join the network* **do**
    for $n_i = 1 \quad i \in N$ **do**
        Compute the clusters
        $K = \{k_1, k_2, ..., k_n\}$
        Allocate nodes to clusters
        Compute zone neighbor discovery
    end
    broadcasts encryption key to clusters
    for $j = 1 \quad j \in N_R$ **do**
        Calculate Fitness function
        $F_i = w_1 \sum_{i \in N} \bar{E} + w_2 \sum_{i,j \in N} \bar{L}_{i,j} + w_3 \sum_{i,j \in N} \bar{d}_{i,j}$
        Where
        $w_1 + w_2 + w_3 \in [0 \sim 1]$
    end
    Evaluate GLSDB table
end

For any specific source and destination node, the dissipated energy at the transmitter and receiver is calculated as follows:

$$E_k^{TX} = \begin{cases} E_e k + E_{amp} k d^4, & if\ d > d_0; \\ E_e k + E_{fs} k d^2, & if\ d < d_0; \end{cases} \tag{7.2}$$

$$E_K^{RX} = k E_e, \tag{7.3}$$

where E_e is the consumed power by the electronic devices, E_{amp} and E_{fs} is the energy per bit used by the RF amplifier. d_0 is the threshold distance computed by the $\sqrt{E_{fs}/E_{amp}}$. The remaining energy of a node (E_R^i) is calculated as follows:

$$E_R^i = \begin{cases} E_{in} - E_i^{TX}, & if \quad \text{transmitting}; \\ E_{in} - E_i^{RX}, & if \quad \text{receiving}; \end{cases} \tag{7.4}$$

A relay node is chosen based on the highest amount of energy to maximize the network lifetime. Therefore, we need to identify the relaying nodes that maximize the network lifetime. The problem formulation is interpreted from [252, 253] and written as the minimum of traffic q sent for a given flow f:

$$\underset{q}{\text{Minimize}} \quad \frac{E_{T_f}}{\sum_{j \in S_j} E_j^{TX} q_j + \sum_{i \in D_i} E_i^{RX} q_i} \tag{7.5}$$

where (7.5) represents the optimization of network lifetime for any given flow, where every flow consists of a source node S with a destination D. To minimize the total transmitted traffic, the following constraints must be met:

$$\sum_{j \in S_j} E_j^{TX} tq_j + \sum_{i \in D_i} e_i^{RX} tq_i \leq E_i \tag{7.6}$$

where $tq_{i,j}$ is the amount of data transmitted for flow f from source node S_j to destination node D_i until time t. Variable t must be independent to satisfy the linearity property of the linear programming algorithm. Solving (7.6) optimizes the transmission time of a node.

7.6.2 Next Hop Candidate

Several candidate nodes are selected for transmitting data to choose the next relay node. For each node, at time t, the number of neighboring nodes will be N. Assuming at any time t there are $F_i, i \in N$ fitness functions of nodes, then the following conditions should be met to choose the best next hop:

$$F_1(t) > F_2(t) > ... > F_i(t) \tag{7.7}$$

According to (7.7), the most efficient neighboring nodes are selected based on the appropriate fitness function. That means candidate node 1 in the neighboring table has more fitness than the other candidate nodes. The fitness function is based on MOSP, which depends on the remaining energy (R_E), the quality of the link (L), and the distance (D) to the relay node [254].

$$F_i = w_1 \sum_{i \in N} \bar{E} + w_2 \sum_{i,j \in N} \bar{L}_{i,j} + w_3 \sum_{i,j \in N} \bar{d}_{i,j} \tag{7.8}$$

where $\bar{E}$ is the normalized remaining energy of node (i), $\bar{L}$ is the normalized cost of link between (i, j), and $\bar{d}$ is the normalized distance of link (i, j). w_1, w_2 and w_3 are the weight factors. The fitness function is computed using three parameters distance to the target node, remaining energy of the next hop node, and the link quality. The distance between node k and the relaying node l is calculated as follows:

$$d = \sqrt{(x_k - x_l)^2 + (y_k - y_l)^2} \tag{7.9}$$

The link quality of nodes directly depends on the mobility rate of nodes, leading to the number of neighboring nodes that have changed over time:

$$L_i = 1 - \frac{n_{c_i}}{1 + n_{n_i}} \tag{7.10}$$

where n_{c_i} and n_{n_i} represent the number of changed nodes and the initial number of neighboring nodes for node (i), respectively. The mobility prediction function enables the prediction of the number of neighboring nodes n_{c_i} a node might have in the future, which is explained in the following section.

7.6.3 Link Expiry Time

A time expiry mechanism is required to remove a specific entry of GLSDB to save processing time and energy. The communication link between the two nodes will expire at some point due to the mobility of nodes. Therefore, data transmission through nodes that have lost connection will increase the packet loss and E2E delay. To solve this issue, an expiry time mechanism is introduced to measure and remove the candidate nodes from the GLSDB. The expiration time in MD2D routing is the time to keep contact between two adjacent mobile nodes (two neighboring nodes). This time depends on different parameters, such as location and mobility rate. As a result, every link between any two nodes will be eliminated before their timers expire and before the source node starts to transmit. Consequently, the PDR will increase, and link reliability will be maximized.

To predict the expiry time, we assume that the radio range between any node in the network is R and the distance between two nodes is denoted as $d_{i,j}$. Therefore, the link expiry time (L_t) between two nodes i and j is predicted as follows:

$$L_{t_{i,j}} = \frac{R + \alpha d_{i,j}}{|v_i - v_j|} \tag{7.11}$$

where α is predicted based on the direction of movement between two nodes. α is either 1 or -1. If nodes are moving toward each other, then α is 1, and if nodes are moving far away, then α is -1. For instance, for two nodes that are moving away with a distance of 50m and mobility rates of $15m/s$ and $2m/s$, with a communication range of 100, the predicted expiry time based on (7.11) is almost $11.5s$.

7.6.4 Location Prediction

The MPBR predicts the future node's location based on a node's speed, movement direction, and current location. The advantages of location prediction are to reduce the route table update and help to establish an optimal path from source S_i to destination D_j, $i, j \in \{1, 2, ..., n\}$. We assume at time t_0 the current location of a node is (x_i, y_i) and after a time interval t_1 the location is updated to (x_{i+1}, y_{i+1}). Then the node's movement direction θ and velocity of the node V are obtained as follows:

$$\theta = tan^{-1}(\frac{x_{i+1} - x_i}{y_{i+1} - y_i}) \tag{7.12}$$

$$V = \frac{\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}}{t_1} \tag{7.13}$$

We assume that at time t_2 the location of the node is (x_2, y_2), then using velocity and the direction of the movement, we can predict the next location of the node (x_3, y_3) at time t_3 as follows:

$$x_3 = x_2 + V cos(\theta) \tag{7.14}$$

$$y_3 = y_2 + V sin(\theta) \tag{7.15}$$

In this way, the controller knows the next location of nodes and can predict the node that must perform the neighbor discovery. Moreover, by forecasting the next location of nodes, the controller can predict when to update the routing tables based on the new neighbors of a node. Therefore, the controller doesn't require to make frequent GLSDB updates and retransmissions. This will decrease cellular overhead and energy consumption.

7.6.5 XOR Privacy-Preserving Forwarding Method

The MD2D communication happens in the WiFi channel with WiFi standards, where WiFi communication has already integrated a secure encryption mechanism to secure the data package from intruders. In MD2D communication, a packet can pass through various nodes until it reaches the destination. This will open the door for malicious nodes to try and steal the source node's location or data. To mitigate this problem, we are using simple but effective XoR privacy-preserving mechanisms to protect the source node's location and data. Under the BS coverage area, N number of nodes are dispersed randomly. Each node is recognized by a unique identification number or IP address. The BS coverage is virtually clustered into fixed RSSI rings. Every node within that area's boundary is considered the cluster's member. If the total number of clusters around the BS is K, then the proposed XOR privacy-preserving mechanism for multi-hop routing is written as follows (based on [100]):

1. The controller generates K random secret keys $K = \{k_1, k_2, ..., k_n\}$, and broadcasts each key to separate clusters. For instance, if cluster 1 is given k_1, then all the nodes in that cluster will receive k_1. The length of k_1 must be the same as the data packet size l to allow the matrix multiplication.

2. The cluster encryption key is changed every minute to reduce the possibility of any intrusion.

3. If we assume a node is transmitting a random data packet D_i, then encrypted data Y with

Table 7.1: Simulation parameters.

Parameters	**Value**
Simulation environment	$1000m \times 1000m$
Initial power of UEs	300j
Number of UEs	50,100,150,200,250
Packet transmission size	10 kbits
Protocols	MPBR, FPRM, HSAW, AODV, OLSR
Propagation model	Rayleigh fading
Mobile node transmission range	100m
Mobile rate	1 m/s & 15 m/s
Mobile node movement model	Random waypoint mobility
Total simulation time	300s

the cluster secret key K_n is written as follows:

$$Y_i = K_n \otimes D_i \tag{7.16}$$

4. This encrypted data Y_i from a node in cluster K_n is then forwarded to the next node. If the next node is the destination, it knows the decryption code and can read the data. If the next node is not the destination, the data will be forwarded to the next hop, which will be in another cluster. Once the node receives the message, it will relay it to the destination.

5. The transmitted encrypted data message carries the identification (ID) of the source node. Once the destination node receives the data, it will request a decryption code from the controller using the source node ID.

6. Once the destination node receives the decryption code α, it can easily decrypt the data by an XOR operation as follows:

$$D_i = \alpha^{-1} \otimes Y_i \tag{7.17}$$

This simple XOR secret-sharing scheme can be further extended such that all the relay nodes in each cluster apply their encryption code to increase the level of privacy. However, more encryption will increase the system's complexity and packet size. This can be further studied to find the optimum number of applied encryption, which is out of the scope of this study.

7.7 Simulation Results and Performance Analysis

In this section, we first explain the simulation setup and then provide the obtained results. The simulation results are averaged and scattered within a 95% confidence interval.

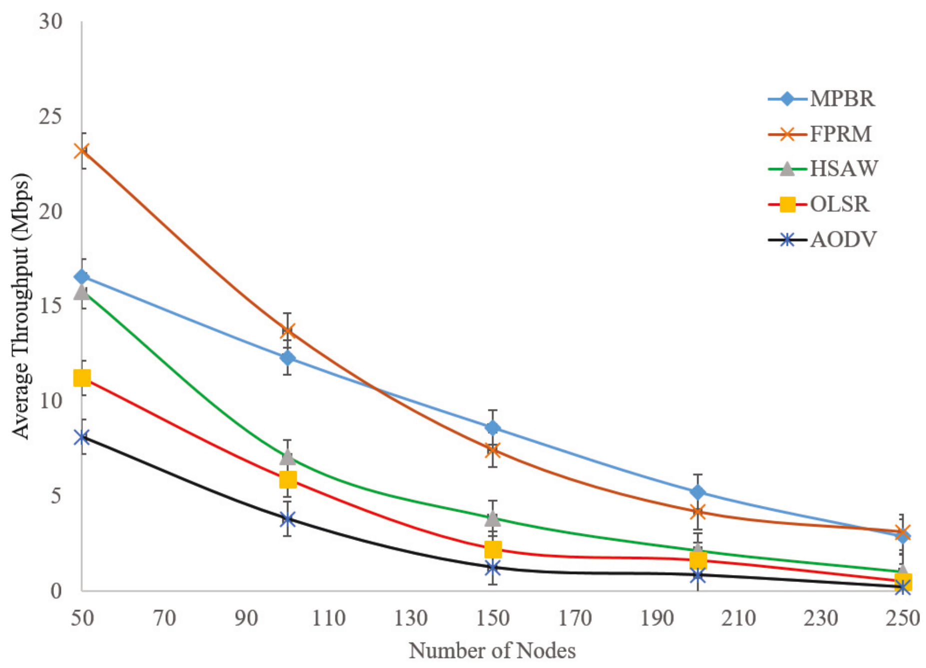


Figure 7.7: Average network throughput versus number of nodes.

7.7.1 Simulation Setup

The proposed routing protocol was implemented in Network Simulator-3 (NS-3). NS3 is the most common platform for network implementation and building routing protocols, which supports IP-based simulation. Our network environment spreads over a $1000m \times 1000m$ and consists of uniformly distributed heterogeneous nodes, including vehicles and mobile devices. The initial energy of users is assumed to be 300 Joules, and users are specified as user equipment (UE). All the nodes are inside the coverage of the BS. Each node is equipped with two wireless interfaces for cellular and WiFi communication. The LTE/5G interface is considered for cellular communication, while for WiFi or MD2D communication, the IEEE 802.11n-5GHz standard is implemented with a limited range of $100m$. For every run, there is a specified number of flows with randomly selected source and destination nodes. Each run time is 300s, and to validate our results and improve the effectiveness of our simulation results, the Monte Carlo simulation technique is under 50 consecutive runs. Simulation parameters are shown in Table 7.1.

7.7.2 Simulation Results

Figure 7.7 illustrates the network throughput, highlighting the significant advantage of FPRM over HSAW, AODV, and OLSR in terms of faster data transmission and recovery processes. FPRM achieves this by sharing the spectrum with fewer users, resulting in higher throughput. However, as the number of nodes increases, the performance of MPBR becomes comparable to or better than FPRM. Our proposed routing protocol aims to maximize link quality by selecting nodes with the best Quality of Service (QoS). In contrast, FPRM, HSAW, AODV, and

OLSR nodes determine paths based on the shortest path first (SPF) using hop counts, thus lacking knowledge of the node's link quality and the ability to select the next hop with the highest link quality. On average, MPBR demonstrates an approximately 4% higher throughput than FPRM and a 40% higher throughput than HSAW. Comparatively, our proposed routing protocols achieve an average throughput increase of 55% compared to AODV and OLSR.

Figure 7.8 presents the average network energy consumption, allowing for a comparative analysis of our proposed routing framework against two link-state-based routing protocols, namely HSAW and FPRM. To provide a comprehensive assessment, we also examine and compare our protocols with two well-established purely ad hoc protocols, AODV and OLSR, which serve as benchmark MD2D routing protocols [207, 208]. In the case of HSAW, nodes continuously receive and update the entire network topology or Link-State Database (LSDB) table, resulting in significant battery drainage and depletion. On the other hand, AODV and OLSR rely on constant flooding algorithms to acquire the link state of nodes and discover routes to the destination. AODV, for instance, utilizes flow requests to initiate forwarding flooding messages towards the destination, with subsequent replies establishing a route that is then selected for transmission. This process proves to be energy-intensive and exhibits suboptimal performance in dynamic environments. Similarly, OLSR maintains both the network topology and routing table to determine the most suitable path to the destination. In FPRM, nodes with low energy consumption, high mobility rate and a low number of neighboring nodes will be excluded from MD2D communication. Therefore, many nodes will not participate in the routing, making the network energy consumption low. However, nodes still receive the whole network topology, which decreases the network lifetime. In our proposed protocol, only the location is communicated between nodes and the BS, where HSAW and FPRM nodes must perform neighbor discovery, which is time and energy-consuming. The simulation result shows that MPBR performs better than AODV and OLSR due to the flooding process that increases the node's power consumption due to acknowledgment, route maintenance, and route update. Moreover, in HSAW and FPRM, the size of the LSDB is significant, and every time nodes get the LSDB update, they will consume more power than in MPBR. MPBR has an optimization mechanism in both GLSDB table creation and the route identification process, which reduces the overall energy consumption of the network.

Figure 7.9 shows the E2E delay versus the number of nodes. In FPRM, HSAW, AODV and OLSR routing protocols, nodes use Hello packets and flooding algorithms to populate their routing tables. FPRM and HSAW nodes use Hello packet to compute the neighbor discovery table and report it to the BS. The BS creates the LSDB for the nodes and then provides routing information and policies to them. In AODV and OLSR, nodes compute a route in a distributed manner using flooding algorithms. Our proposed protocol uses only the location of nodes and a zone-prescribed neighbor discovery process. Hence, the E2E delay will be reduced significantly. MPBR and FPRM performance are similar when the number of nodes in the network is low. Once node density increases, ACMRP achieves better results. In FPRM, several nodes

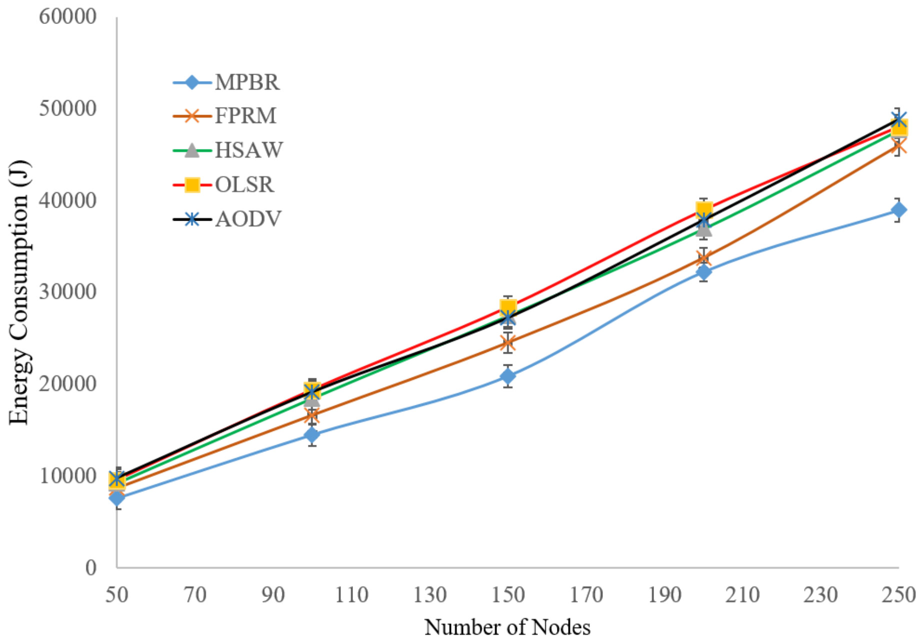


Figure 7.8: Network energy consumption versus number of nodes.

are offline and do not participate in the MD2D. Therefore, fewer nodes must perform neighbor discovery, and BS will create the LSDB quicker. That is why the performance of FPRM and MPBR is almost the same in the low-density networks, but in higher densities, MPBR shows almost 5% better than FPRM. Compared to HSAW and conventional routing protocols, MPBR has a significantly lower E2E delay.

Figure 7.10 depicts the packet delivery ratio versus the number of nodes. The percentage of the total number of delivered packets and the total number of packets sent by the source node is obtained as PDR. The succession of packet delivery is affected by several factors, including mobility, network density, link quality, queue size, and distance to the neighbor. Moreover, obstacles in an urban area can cause signal deterioration and cause link or packet failure leading to retransmission. Therefore, during an MD2D communication, relay nodes may receive multiple copies of the same packet. MPBR, FPRM and HSAW have similar PDR in low network densities. As the network gets congested, more nodes are added with varying mobility rates, increasing the probability of losing the packet. However, MPBR and FPRM keep the PDR higher than the rest of the routing protocols. MPBR has almost 2% higher PDR than FPRM, and as the number of nodes increases, the PDR will be almost 5% higher for MPBR, almost 5%. Our protocol uses the multi-objective shortest path to predict the next hop with the most reliable link, the lowest mobility rate, and high power levels. Therefore, nodes have a higher probability of sending a packet without retransmitting.

Other protocols have to use the shortest path to the destination, which means nodes with a higher probability of packet failure will still be chosen as the next relay node if they are close enough to the destination. When the network density is low, the PDR difference between all the routing protocols is insignificant; however, once the node density and network size increase, the

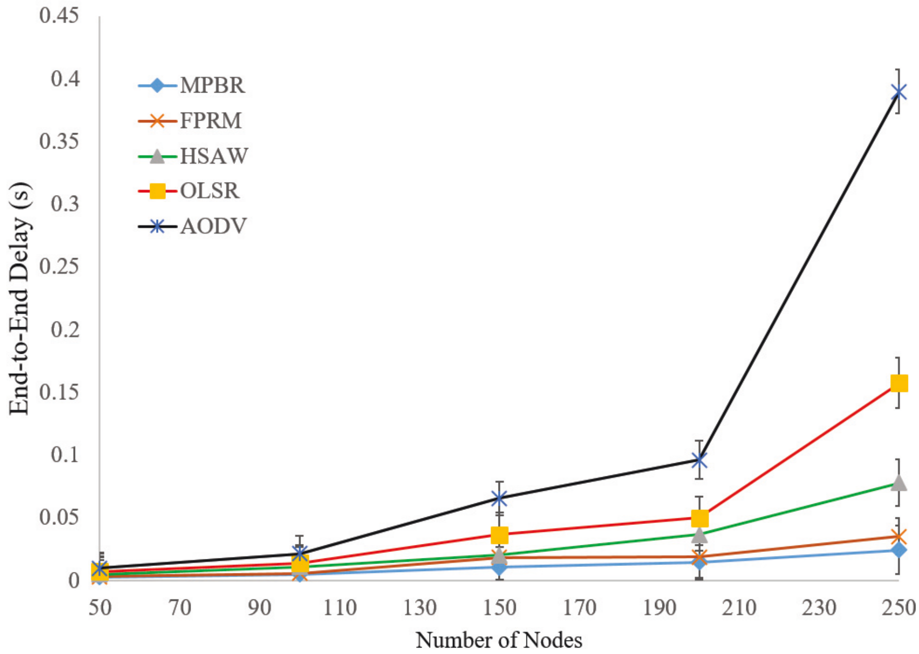


Figure 7.9: End-to-end delay versus number of nodes.

difference becomes substantial. MPBR has the highest PDR compared to the routing protocols that were tested (or used for comparison).

7.8 Conclusion

This chapter proposed a new MD2D privacy-preserving location-based routing protocol for future wireless networks. We use the GPS coordinates of nodes to generate a GLSDB table, where the introduced constraints optimize the GLSDB entries. The optimization is based on a fitness function that identifies the most efficient relay node based on the remaining energy, link quality, and mobility prediction. The proposed protocol employs a simple but effective privacy-preserving forwarding scheme to preserve the node's data privacy during MD2D communication. The introduced protocol is compared with a high-performance centralized SDN-based routing protocol, HSAW. As a result, we analyzed the advantages and disadvantages of a centralized-based geographical routing protocol versus the link-state routing protocols. The simulation results show an average of 15% reduction in energy consumption, a 5% reduction in E2E delay and a 5% increase in PDR compared to one of the leading MD2D routing protocols. In the future, we can study multi-framework routing using network application and user demands, where an adaptive routing framework can be designed and implemented to prescribe location-based or link-state routing protocols.

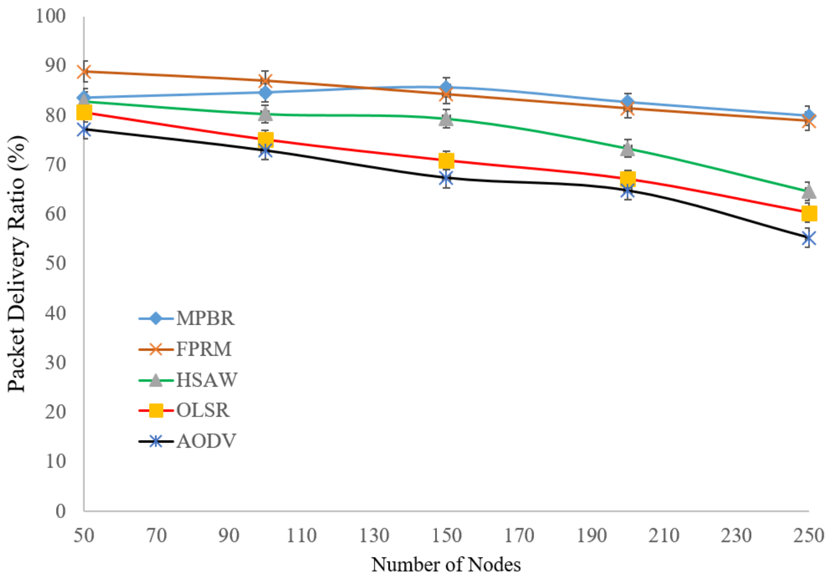


Figure 7.10: Packet delivery ratio versus number of nodes.

8
Conclusion and Future Direction

This thesis begins with a comprehensive review of the current and future challenges and limitations of wireless cellular networks. One of the most pressing challenges of these networks is excessive network traffic. Our research found that MD2D routing protocols will play a crucial role in alleviating and improving network congestion. Despite numerous proposed routing protocols, we concluded that little attention has been paid to creating intelligent and self-adaptive routing protocols in MD2D networks. In the coming years, networks must be equipped with the capability to self-manage and self-optimize routing strategies and policies. To address this issue, we proposed innovative solutions and intelligent algorithms that aim to enhance the performance of MD2D networks.

In order to determine the most suitable routing protocol for various conditions, this thesis conducted a comparative study of centralized MD2D routing protocols. Given their superior performance compared to distributed protocols, the focus was on centralized routing protocols. Three centralized MD2D routing protocols, namely HSAW, VARP-S, and SMDRP, were specifically evaluated. HSAW is a proactive routing protocol where nodes acquire routes using the full LSDB provided by the BS. On the other hand, VARP-S and SMDRP are reactive routing protocols where paths are developed on demand. Simulation results illustrated the differences when the network size and node's mobility increase. HSAW demonstrated superior PDR and E2E delay in sparse networks, whereas VARP-S and SMDRP were found to be more energy-efficient and result in lower cellular overhead in congested networks. In conclusion, proactive routing protocols are more effective in highly mobile networks, while reactive routing is a better option for networks that prioritize low energy consumption.

In Chapter four, a joint mobile node participation and routing protocol for MD2D communication in intelligent transportation systems was proposed, called fuzzy-based participation and routing protocol for MD2D (FPRM). This proposed protocol is designed to operate over a virtual application in future wireless cellular networks to alleviate network traffic and improve network performance. A sub-layer at the network layer is introduced, which uses a fuzzy logic system to determine the nodes with the highest participation probability in routing, thereby establishing a framework for the creation of stable routes. To guarantee the participating nodes' capability to handle the data traffic, two constraints are proposed: mobility and coverage constraints. The former supports the establishment of sustainable communication links, while the latter ensures complete coverage of the MD2D network by the communication service. Results from simulations show that our protocol outperforms the benchmarked MD2D protocols and other investigated ad-hoc protocols.

The creation of a new, flexible routing protocol for MD2D networks presents a challenge due to the dynamic nature of the environment. However, the openness, intelligence, and programmability of future wireless networks offer opportunities to design and develop new routing protocols with global knowledge and intelligence. As a result, Chapter Five presented a new joint utility-based routing protocol called application-driven cross-layer MD2D routing protocol (ACMRP) to find the optimal route by incorporating knowledge from the application and network layers. A utility-based master link-state database (MLSDB) is generated to determine the optimal path. Four utility metrics are used: mobility rate, bandwidth, energy, and distance. Based on the requirement of user applications, we introduced cost functions containing a different combination of utility metrics. Simulation results demonstrated that ACMRP outperforms HSAW, AODV and OLSR. There are several benefits to the proposed protocol: adaptability of route based on the user application (to provide maximum QoE), utilization of network knowledge, and removing distributed multi-hop flooding from routing decisions.

Network slicing can significantly assist cellular networks in managing traffic and providing flexibility. In Chapter Six, to incorporate the advantages of network slicing, we proposed an adaptive slicing mechanism with a dynamic MD2D routing protocol selection technique for future cellular networks. In our framework, a controller is responsible for collecting the network data and utilizing the particle swarm optimization (PSO) algorithm to virtually slice WiFi channels into various virtual sub-layers based on network traffic. In particular, our algorithm used viral content to create virtual slices where users with similar content can share and download data to or from other users. Moreover, the controller prescribed the most efficient routing protocol using the genetic algorithm (GA) in conjunction with learning automata (LA) for any virtual slices based on the content, mobility rate, throughput, number of neighbors, and network density. Each virtual sub-layer used AODV or OLSR protocol to route the packet. Hence, they illustrated the performance of reactive and proactive routing frameworks in virtual slicing problems. The simulation results indicated that our proposed framework performs significantly better than the non-sliced single-routing protocol-based frameworks. Other routing protocols can also be used, but the primary purpose is to create an adaptive slicing mechanism with a dynamic route selection method.

In Chapter Seven, we introduced a position-based routing protocol and provided insights into the performance difference between position-based and link-state MD2D routing protocols. Link-state routing protocols use the neighbor information obtained by the nodes to create routing tables. Position-based routing uses nodes' location (GPS coordinates) to generate routing tables. Both methods are used in the current MD2D communication networks. However, position-based methods proved to have fast and acceptable results in highly dynamic environments (especially VANETS) compared to link-state. Therefore, a new MD2D position-based routing (MPRR) protocol was proposed. We jointly used the coordinates of nodes with zone-prescribed neighbor discovery to make fast and reliable routing decisions. The presented work uses an encrypted cluster-based algorithm for data privacy in multi-hop routing. Our proposed routing protocol utilizes two constraints to increase the performance, link expiry time and node location prediction. We compared the proposed routing protocol with a centralized link-state MD2D routing protocol. The simulation results showed the differences and advantages of our proposed position-based routing compared to a purely link-state routing protocol.

8.1 Future Direction

The growing traffic load on BS is rapidly deteriorating network performance. It is projected that a significant portion of cellular traffic will stem from social media. To alleviate this issue, MD2D networks offer a solution by enabling direct communication and traffic harvesting among devices in the network. The routing protocol plays a crucial role in the functioning of MD2D networks. To effectively incorporate MD2D into future cellular networks, the research community must concentrate on addressing the open challenges and issues that remain to be solved.

1. **Self-organizing Networks:** Self-organizing networks (SONs) provide self-optimization, self-coordination, self-management, and self-correction for the next generation of wireless networks [255]. In particular, most researchers now consider ML techniques as an official approach to achieving self-organization in the network, providing the future MD2D network with the opportunity to self-tune the policies and routing protocols. 3GPP has already started developing protocols and technologies to automate network configurations [256]. In this context, RL is the most recognizable approach for optimizing network parameters and topology based on the environment and experience. In particular, in load balancing, handover management, routing, etc. SON is an evolution of self-driving networks, and most applications in wireless networks require intelligence to tune, correct,

and decide on behalf of human operators. Therefore, future research should focus on advancing the capabilities of SONs in Open-RAN networks, particularly for MD2D routing protocols to enable network adaptation to changing environments.

2. **Lack of Network Knowledge and Intelligence:** There are centralized algorithms and optimization techniques that can achieve high-performance criteria of the current networks. However, future wireless networks require intelligence to respond efficiently to user demands. Researchers have recognized the lack of global network knowledge and intelligence [257,258]. For instance, the baseline technique to achieve load balancing and backhaul management in [259–261] requires complete information about the traffic load and content popularity of users before execution of cache content, which is challenging to acquire precise information in advance. Therefore, we suggest prior knowledge for decision-making. Using a transfer learning (TL) algorithm, the previous experience in cache content can be utilized by BSs to guide cache management even without knowing any current traffic information. Moreover, designing a routing policy in MD2D networks faces various challenges when considering multiple factors that suggest a predetermined knowledge or intelligence is inevitable. For instance, mobility pattern, power, channel quality and signal attenuation all affect the future decision-making in optimal routing policy. Deep reinforcement learning (DRL) can be used to make routing decisions by integrating various information and formulating the long-term optimal decision-making problem as a partial observable Markov decision process (POMDP), which generates knowledge for future routing decisions. As a result, authors in [262] proposed a distributed multi-agent DRL to optimize the system routing policies to increase the overall data delivery and energy consumption. Future cellular systems must adapt intelligence and network knowledge to enable optimal decision-making.

3. **Knowledge Validation, Uncertainty and Compromises**: ML and intelligence have been envisioned by many researchers as the most important feature in 6G, as ML algorithms have been extensively used in complex scenarios. Therefore, it is evident that the O-RAN architecture can be used to address the challenges of 6G. One of the main challenges faced by all technologies is the validation of knowledge. If 6G targets an automatically configures a cellular network, there must be a mechanism to verify the confidence and certainty of knowledge. As a result, a certainty mechanism must be deployed to acknowledge the level of certainty, whether knowledge is practical or compromised. The output of the ML algorithm must be checked by the expected results to evaluate the degree of uncertainty. A threshold barrier can be used to validate the usefulness of knowledge. If the knowledge is authorized to deploy in the network, the ML output has been successful, but if the compared strategy has revealed unauthorized knowledge, then the ML's output cannot pass the threshold value. In this case, a new ML technique must compute the new knowledge and go through the same procedure, which causes a delay that affects the

system performance. In the worst-case scenario, if the knowledge is rejected again, then an extreme case must be considered. To mitigate the worst-case scenario, any proposed algorithm must undergo various experimental analyses in a real testbed to find the possible ML substitutes in any given scenario. Therefore, ML algorithms must be tested in the same environment with similar characteristics to provide insights into different ML techniques.

4. **Privacy**: One open issue in the MD2D network is user and data privacy. In position-based routing protocols, the user shares their GPS locations which might provide an opportunity for intrusion. Moreover, some users might not even be happy to share their location in the first place. Therefore, there must be a trustable privacy policy in MD2D networks for users to authorize and share their locations. Moreover, the data exchanged between users go through relaying nodes, which increase the chance of malicious node and data hijacking. In this scenario, data should be perfectly secured while transmitted to the destination.

5. **Identifying Suitable ML Algorithm**: The majority of routing problems in MD2D networks are solved within a few ML algorithms, including regression problems, classification problems, clustering problems, and Markov decision process (MDP) problems. In regression problems, the ML algorithm is required to predict a continuous value output given an input. In classification problems, the ML algorithm needs to predict a discrete value output, usually answered by a yes or no, and zero or one to essentially identify the class to which the input belongs. The clustering problems are ML techniques, where the data are grouped based on their type or value. Finally, MDP problems are ML techniques that require taking action in the current system state based on the feedback reward resulting from the previous action. Therefore, it is important to consider the advantages and disadvantages of each technique before applying them.

BIBLIOGRAPHY

[1] N. Wang, P. Wang, A. Alipour-Fanid, L. Jiao, and K. Zeng, "Physical-layer security of 5g wireless networks for iot: Challenges and opportunities," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8169–8181, 2019. 1, 93

[2] P. K. Malik, D. S. Wadhwa, and J. S. Khinda, "A survey of device to device and cooperative communication for the future cellular networks," *International Journal of Wireless Information Networks*, pp. 1–22, 2020. 1

[3] F. S. Shaikh and R. Wismüller, "Routing in multi-hop cellular device-to-device (d2d) networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2622–2657, 2018. 1, 3, 24, 94, 97, 120, 144

[4] A. Aissioui, A. Ksentini, A. M. Gueroui, and T. Taleb, "On enabling 5g automotive systems using follow me edge-cloud concept," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 6, pp. 5302–5316, 2018. 2

[5] F. Tang, B. Mao, Y. Kawamoto, and N. Kato, "Survey on machine learning for intelligent end-to-end communication toward 6g: From network access, routing to traffic control and streaming adaption," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1578–1598, 2021. 2

[6] X. You, C.-X. Wang, J. Huang, X. Gao, Z. Zhang, M. Wang, Y. Huang, C. Zhang, Y. Jiang, J. Wang *et al.*, "Towards 6g wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Science China Information Sciences*, vol. 64, no. 1, pp. 1–74, 2021. 2, 12

[7] M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "Understanding o-ran: Architecture, interfaces, algorithms, security, and research challenges," *arXiv preprint arXiv:2202.01032*, 2022. 2, 120, 144

[8] A. Nayak Manjeshwar, P. Jha, A. Karandikar, and P. Chaporkar, "Virtran: An sdn/nfv-based framework for 5g ran slicing," *Journal of the Indian Institute of Science*, vol. 100, no. 2, pp. 409–434, 2020. 2

[9] D. S. Rana, S. A. Dhondiyal, and S. K. Chamoli, "Software defined networking (sdn) challenges, issues and solution," *Int J Comput Sci Eng*, vol. 7, no. 1, pp. 884–889, 2019. 2

[10] F. Z. Yousaf, M. Bredel, S. Schaller, and F. Schneider, "Nfv and sdn—key technology enablers for 5g networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 11, pp. 2468–2478, 2017. 2

[11] A. Mestres, A. Rodriguez-Natal, J. Carner, P. Barlet-Ros, E. Alarcón, M. Solé, V. Muntés-Mulero, D. Meyer, S. Barkai, M. J. Hibbett *et al.*, "Knowledge-defined networking," *ACM SIGCOMM Computer Communication Review*, vol. 47, no. 3, pp. 2–10, 2017. 2, 17, 18

[12] D. D. Clark, C. Partridge, J. C. Ramming, and J. T. Wroclawski, "A knowledge plane for the internet," in *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*. ACM, 2003, pp. 3–10. 2, 17

[13] P. Sun, Z. Guo, S. Liu, J. Lan, J. Wang, and Y. Hu, "Smartfct: Improving power-efficiency for data center networks with deep reinforcement learning," *Computer Networks*, vol. 179, p. 107255, 2020. 2

[14] J. Navarro-Ortiz, P. Romero-Diaz, S. Sendra, P. Ameigeiras, J. J. Ramos-Munoz, and J. M. Lopez-Soler, "A survey on 5g usage scenarios and traffic models," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 905–929, 2020. 3

[15] T. Barnett, S. Jain, U. Andra, and T. Khurana, "Cisco visual networking index (vni) complete forecast update, 2017–2022," *Americas/EMEAR Cisco Knowledge Network (CKN) Presentation*, pp. 1–30, 2018. 3

[16] S. Zhang, J. Liu, H. Guo, M. Qi, and N. Kato, "Envisioning device-to-device communications in 6g," *IEEE Network*, vol. 34, no. 3, pp. 86–91, 2020. 3

[17] R. Shafin, L. Liu, V. Chandrasekhar, H. Chen, J. Reed, and J. C. Zhang, "Artificial intelligence-enabled cellular networks: A critical path to beyond-5g and 6g," *IEEE Wireless Communications*, vol. 27, no. 2, pp. 212–217, 2020. 9

[18] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6g: A comprehensive survey," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 334–366, 2021. 9

[19] T. A. Q. Pham, Y. Hadjadj-Aoul, and A. Outtagarts, "Deep reinforcement learning based qos-aware routing in knowledge-defined networking," in *International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness*. Springer, 2018, pp. 14–26. 9, 20

[20] S. Gosh, B. El Boudani, T. Dagiuklas, and M. Iqbal, "So-kdn: A self-organised knowledge defined networks architecture for reliable routing," in *2021 The 4th International Conference on Information Science and Systems*, 2021, pp. 160–166. 9

[21] J. A. del Peral-Rosado, R. Raulefs, J. A. López-Salcedo, and G. Seco-Granados, "Survey of cellular mobile radio localization methods: From 1g to 5g," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1124–1148, 2017. 11

[22] A. H. Khan, M. A. Qadeer, J. A. Ansari, and S. Waheed, "4g as a next generation wireless network," in *2009 International conference on future computer and communication*. IEEE, 2009, pp. 334–338. 11

[23] E. Ezhilarasan and M. Dinakaran, "A review on mobile technologies: 3g, 4g and 5g," in *2017 second international conference on recent trends and challenges in computational models (ICRTCCM)*. IEEE, 2017, pp. 369–373. 11

[24] H.-Y. Wei, Y.-W. P. Hong, W.-T. Shay, and T.-L. Wu, "Bridging the gap between academia and industry: Most 6g research program in taiwan," in *2021 IEEE VTS 17th Asia Pacific Wireless Communications Symposium (APWCS)*. IEEE, 2021, pp. 1–5. 12

[25] M. Series, "Imt vision–framework and overall objectives of the future development of imt for 2020 and beyond," *Recommendation ITU*, vol. 2083, p. 0, 2015. 12

[26] M. A. Habibi, M. Nasimi, B. Han, and H. D. Schotten, "A comprehensive survey of ran architectures toward 5g mobile communication system," *IEEE Access*, vol. 7, pp. 70 371–70 421, 2019. 13

[27] S. Niknam, A. Roy, H. S. Dhillon, S. Singh, R. Banerji, J. H. Reed, N. Saxena, and S. Yoon, "Intelligent o-ran for beyond 5g and 6g wireless networks," *arXiv preprint arXiv:2005.08374*, 2020. 13, 67, 144

[28] C. Pan, M. Elkashlan, J. Wang, J. Yuan, and L. Hanzo, "User-centric c-ran architecture for ultra-dense 5g networks: Challenges and methodologies," *IEEE Communications Magazine*, vol. 56, no. 6, pp. 14–20, 2018. 13

[29] N. Kazemifard and V. Shah-Mansouri, "Minimum delay function placement and resource allocation for open ran (o-ran) 5g networks," *Computer Networks*, vol. 188, p. 107809, 2021. 13

[30] B. Balasubramanian, E. S. Daniels, M. Hiltunen, R. Jana, K. Joshi, R. Sivaraj, T. X. Tran, and C. Wang, "Ric: A ran intelligent controller platform for ai-enabled cellular networks," *IEEE Internet Computing*, vol. 25, no. 2, pp. 7–17, 2021. 13

[31] C.-X. Wang, M. Di Renzo, S. Stanczak, S. Wang, and E. G. Larsson, "Artificial intelligence enabled wireless networking for 5g and beyond: Recent advances and future challenges," *IEEE Wireless Communications*, vol. 27, no. 1, pp. 16–23, 2020. 13

[32] S. Niknam, H. S. Dhillon, and J. H. Reed, "Federated learning for wireless communications: Motivation, opportunities, and challenges," *IEEE Communications Magazine*, vol. 58, no. 6, pp. 46–51, 2020. 13

[33] "Open ran - the open road to 5g," *SAMSUNG white paper*, 2019. 13

[34] J. Ordonez-Lucena, P. Ameigeiras, D. Lopez, J. J. Ramos-Munoz, J. Lorca, and J. Folgueira, "Network slicing for 5g with sdn/nfv: Concepts, architectures, and challenges," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 80–87, 2017. 13

[35] "Nr; overall description; stage-2," *3GPP*, 2020. 14

[36] S. F. Abedin, A. Mahmood, N. H. Tran, Z. Han, and M. Gidlund, "Elastic o-ran slicing for industrial monitoring and control: A distributed matching game and deep reinforcement learning approach," *IEEE Transactions on Vehicular Technology*, 2022. 14

[37] N. Sultana, N. Chilamkurti, W. Peng, and R. Alhadad, "Survey on sdn based network intrusion detection system using machine learning approaches," *Peer-to-Peer Networking and Applications*, vol. 12, no. 2, pp. 493–501, 2019. 14

[38] R. Michalski, J. Carbonell, and T. Mitchell, "Machine learning: An artificial intelligence approach, springer science & business media," 2013. 14

[39] A. Ksentini, M. Bagaa, and T. Taleb, "On using sdn in 5g: The controller placement problem," in *2016 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2016, pp. 1–6. 15

[40] D. Simeonidou, "Joint fronthaul optimization and sdn controller placement in dynamic 5g networks," in *Optical Network Design and Modeling: 23rd IFIP WG 6.10 International Conference, ONDM 2019, Athens, Greece, May 13–16, 2019, Proceedings*, vol. 11616. Springer Nature, 2020, p. 181. 15

[41] S. Khan, A. Gani, A. W. A. Wahab, M. Guizani, and M. K. Khan, "Topology discovery in software defined networks: Threats, taxonomy, and state-of-the-art," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 303–324, 2016. 15

[42] P. Chemouil, P. Hui, W. Kellerer, Y. Li, R. Stadler, D. Tao, Y. Wen, and Y. Zhang, "Special issue on artificial intelligence and machine learning for networking and communications," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 6, pp. 1185–1191, 2019. 15

[43] S. Tomovic and I. Radusinovic, "Toward a scalable, robust, and qos-aware virtual-link provisioning in sdn-based isp networks," *IEEE transactions on network and service management*, vol. 16, no. 3, pp. 1032–1045, 2019. 15

[44] W. Jiang, M. Strufe, and H. Schotten, "Autonomic network management for software-defined and virtualized 5g systems," in *European Wireless 2017; 23th European Wireless Conference*. VDE, 2017, pp. 1–6. 15

[45] B. Han, V. Gopalakrishnan, L. Ji, and S. Lee, "Network function virtualization: Challenges and opportunities for innovations," *IEEE communications magazine*, vol. 53, no. 2, pp. 90–97, 2015. 15

[46] A. Laghrissi and T. Taleb, "A survey on the placement of virtual resources and virtual network functions," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1409–1434, 2018. 15

[47] I. Tomkos, D. Klonidis, E. Pikasis, and S. Theodoridis, "Toward the 6g network era: Opportunities and challenges," *IT Professional*, vol. 22, no. 1, pp. 34–38, 2020. 15

[48] Y. Lu and X. Zheng, "6g: A survey on technologies, scenarios, challenges, and the related issues," *Journal of Industrial Information Integration*, vol. 19, p. 100158, 2020. 15

[49] S. Ashtari, I. Zhou, M. Abolhasan, N. Shariati, J. Lipman, and W. Ni, "Knowledge-defined networking: Applications, challenges and future work," *Array*, p. 100136, 2022. 15

[50] T. Li, J. Chen, and H. Fu, "Application scenarios based on sdn: an overview," in *Journal of Physics: Conference Series*, vol. 1187, no. 5. IOP Publishing, 2019, p. 052067. 16

[51] D. Kreutz, F. Ramos, P. Verissimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *arXiv preprint arXiv:1406.0440*, 2014. 16

[52] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J. A. Zhang, "The roadmap to 6g: Ai empowered wireless networks," *IEEE communications magazine*, vol. 57, no. 8, pp. 84–90, 2019. 16

[53] A. Zahmatkesh and T. Kunz, "Software defined multihop wireless networks: Promises and challenges," *Journal of Communications and Networks*, vol. 19, no. 6, pp. 546–554, 2017. 16

[54] X. Chen, W. Ni, T. Chen, I. B. Collings, X. Wang, R. P. Liu, and G. B. Giannakis, "Multi-timescale online optimization of network function virtualization for service chaining," *IEEE Transactions on Mobile Computing*, vol. 18, no. 12, pp. 2899–2912, 2018. 16

[55] P. Goransson, C. Black, and T. Culver, *Software defined networks: a comprehensive approach*. Morgan Kaufmann, 2016. 16

[56] M. Nick and R. Jen, *Clarifying the differences between P4 and OpenFlow*, 2016. [Online]. Available: https://p4.org/p4/clarifying-the-differences-between-p4-and-openflow.html 17

[57] H. Stubbe, "P4 compiler & interpreter: A survey," in *Proc. Future Internet (FI) Innov. Internet Technol. Mobile Commun.(IITM)*, vol. 47, 2017, pp. 1–72. 17

[58] P. Bosshart, D. Daly, G. Gibb, M. Izzard, N. McKeown, J. Rexford, C. Schlesinger, D. Talayco, A. Vahdat, G. Varghese *et al.*, "P4: Programming protocol-independent packet processors," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 3, pp. 87–95, 2014. 17

[59] M. B. H. WELL and B. OPTIMIZATIONS, "P4 data plane programming for server-based networking applications." 17

[60] S. Signorello, R. State, J. François, and O. Festor, "Ndn. p4: Programming information-centric data-planes," in *2016 IEEE NetSoft Conference and Workshops (NetSoft)*. IEEE, 2016, pp. 384–389. 17

[61] I. Butun, Y. K. Tuncel, and K. Oztoprak, "Application layer packet processing using pisa switches," *Sensors*, vol. 21, no. 23, p. 8010, 2021. 17

[62] C. Kim, A. Sivaraman, N. Katta, A. Bas, A. Dixit, and L. J. Wobker, "In-band network telemetry via programmable dataplanes," in *ACM SIGCOMM*, vol. 15, 2015. 18

[63] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *arXiv preprint arXiv:1809.08707*, 2018. 18

[64] H. Mao, M. Alizadeh, I. Menache, and S. Kandula, "Resource management with deep reinforcement learning," in *Proceedings of the 15th ACM Workshop on Hot Topics in Networks*. ACM, 2016, pp. 50–56. 18

[65] C. Marshall, "By 2019, 80% of the world's internet traffic will be video," *Tubular Insights. Retrieved on January*, vol. 3, p. 2017, 2015. 18

[66] S. Cass, "The age of the zettabyte cisco: the future of internet traffic is video [dataflow]," *IEEE Spectrum*, vol. 51, no. 3, pp. 68–68, 2014. 18

[67] J. Hyun and J. W.-K. Hong, "Knowledge-defined networking using in-band network telemetry," in *2017 19th Asia-Pacific Network Operations and Management Symposium (APNOMS)*. IEEE, 2017, pp. 54–57. 20

[68] D. Careglio, S. Spadaro, A. Cabellos, J. Lazaro, J. Perelló, P. Barlet, J. Gené, and J. Paillissé, "Alliance project: Architecting a knowledge-defined 5g-enabled network infrastructure," in *2018 20th International Conference on Transparent Optical Networks (ICTON)*. IEEE, 2018, pp. 1–6. 20

[69] A. Duque-Torres, F. Amezquita-Suárez, O. M. Caicedo Rendon, A. Ordóñez, and W. Y. Campo, "An approach based on knowledge-defined networking for identifying heavy-

hitter flows in data center networks," *Applied Sciences*, vol. 9, no. 22, p. 4808, 2019. 20

[70] J. Pisarov and G. Mester, "The impact of 5g technology on life in 21st century," *IPSI BgD Transactions on Advanced Research (TAR)*, vol. 16, no. 2, pp. 11–14, 2020. 20

[71] A. L. Rezaabad, H. Beyranvand, J. A. Salehi, and M. Maier, "Ultra-dense 5g small cell deployment for fiber and wireless backhaul-aware infrastructures," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 12 231–12 243, 2018. 22

[72] P. Cardieri, "Modeling interference in wireless ad hoc networks," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 4, pp. 551–572, 2010. 22

[73] R. I. Ansari, C. Chrysostomou, S. A. Hassan, M. Guizani, S. Mumtaz, J. Rodriguez, and J. J. Rodrigues, "5g d2d networks: Techniques, challenges, and future prospects," *IEEE Systems Journal*, vol. 12, no. 4, pp. 3970–3984, 2017. 22, 143

[74] P. Gandotra and R. K. Jha, "Device-to-device communication in cellular networks: A survey," *Journal of Network and Computer Applications*, vol. 71, pp. 99–117, 2016. 22, 119

[75] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE access*, vol. 3, pp. 1206–1232, 2015. 22, 37

[76] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 86–92, 2014. 22, 38

[77] X. Lin, J. G. Andrews, A. Ghosh, and R. Ratasuk, "An overview of 3GPP device-to-device proximity services," *IEEE Communications Magazine*, vol. 52, no. 4, pp. 40–48, 2014. 22

[78] L. Babun, A. I. Yürekli, and I. Güvenç, "Multi-hop and d2d communications for extending coverage in public safety scenarios," in *2015 IEEE 40th local computer networks conference workshops (LCN Workshops)*. IEEE, 2015, pp. 912–919. 22

[79] S. Ashtari, M. Abdollahi, M. Abolhasan, N. Shariati, and J. Lipman, "Performance analysis of multi-hop routing protocols in sdn-based wireless networks," *Computers & Electrical Engineering*, vol. 97, p. 107393, 2022. 22, 94, 97, 98, 135, 144, 147

[80] S. Selmi and R. Bouallègue, "Energy and spectral efficient relay selection and resource allocation in mobile multi-hop device to device communications: Energy and spectral efficient multi-hop d2d communications," *IET Communications*, vol. 15, no. 14, pp. 1791–1807, 2021. 22

[81] S. Ashtari, M. Abdollahi, M. Abolhasan, N. Shariati, and J. Lipman, "Performance analysis of multi-hop routing protocols in sdn-based wireless networks," *Computers & Electrical Engineering*, p. 107393, 2021. 24

[82] G. Kaur and P. Thakur, "Routing protocols in manet: An overview," in *2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, vol. 1. IEEE, 2019, pp. 935–941. 24, 25

[83] M. Abolhasan, J. Lipman, W. Ni, and B. Hagelstein, "Software-defined wireless networking: centralized, distributed, or hybrid?" *IEEE Network*, vol. 29, no. 4, pp. 32–38, 2015. 24, 27, 41, 69, 147

[84] G. He, "Destination-sequenced distance vector (dsdv) protocol," *Networking Laboratory, Helsinki University of Technology*, vol. 135, pp. 1–9, 2002. 25

[85] F. De Rango, M. Fotino, and S. Marano, "Ee-olsr: Energy efficient olsr routing protocol for mobile ad-hoc networks," in *MILCOM 2008-2008 IEEE Military Communications Conference*. IEEE, 2008, pp. 1–7. 25

[86] M. Abolhasan, M. Abdollahi, W. Ni, A. Jamalipour, N. Shariati, and J. Lipman, "A routing framework for offloading traffic from cellular networks to sdn-based multi-hop device-to-device networks," *IEEE Transactions on Network and Service Management*, vol. 15, no. 4, pp. 1516–1531, 2018. 25, 27, 41, 70, 98, 120, 144, 147

[87] S. Boussoufa-Lahlah, F. Semchedine, and L. Bouallouche-Medjkoune, "Geographic routing protocols for vehicular ad hoc networks (vanets): A survey," *Vehicular Communications*, vol. 11, pp. 20–31, 2018. 25

[88] A. Verma and S. Kumar, "Routing protocols in delay tolerant networks: Comparative and empirical analysis," *Wireless Personal Communications*, vol. 118, no. 1, pp. 551–574, 2021. 26

[89] R. E. Ahmed, "A low-overhead multi-hop routing protocol for d2d communications in 5g." *J. Commun.*, vol. 16, no. 5, pp. 191–197, 2021. 26, 144

[90] R. Jain and I. Kashyap, "An qos aware link defined olsr (ld-olsr) routing protocol for manets," *Wireless Personal Communications*, vol. 108, no. 3, pp. 1745–1758, 2019. 26

[91] A. K. Biswas and M. Dasgupta, "Modification of dsdv and secure routing using blockchain technology," in *2020 4th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*. IEEE, 2020, pp. 1–5. 26

[92] P. Pandey and R. Singh, "Advanced aodv routing based on restricted broadcasting of route request packets in manet," *International Journal of Information Technology*, vol. 13, no. 6, pp. 2471–2482, 2021. 26

[93] M. Abdollahi, M. Abolhasan, N. Shariati, J. Lipman, A. Jamalipour, and W. Ni, "A routing protocol for sdn-based multi-hop d2d communications," in *2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 2019, pp. 1–4. 27, 41, 97, 147

[94] S. Agrawal, N. Tyagi, A. Iqbal, and R. S. Rao, "An intelligent greedy position-based multi-hop routing algorithm for next-hop node selection in vanets," *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, vol. 90, no. 1, pp. 39–47, 2020. 27

[95] K. K. Rana, S. Tripathi, and R. S. Raw, "Inter-vehicle distance-based location aware multi-hop routing in vehicular ad-hoc network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 11, pp. 5721–5733, 2020. 27

[96] K. S. Eunice and I. Juvanna, "Secured multi-hop clustering protocol for location-based routing in vanets," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 4, 2019. 28

[97] S. R. Chavva and R. S. Sangam, "An energy-efficient multi-hop routing protocol for health monitoring in wireless body area networks," *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 8, no. 1, pp. 1–10, 2019. 28

[98] A. Rezaeipanah, P. Amiri, H. Nazari, M. Mojarad, and H. Parvin, "An energy-aware hybrid approach for wireless sensor networks using re-clustering-based multi-hop routing," *Wireless Personal Communications*, vol. 120, no. 4, pp. 3293–3314, 2021. 28

[99] E. Alnawafa and I. Marghescu, "New energy efficient multi-hop routing techniques for wireless sensor networks: Static and dynamic techniques," *Sensors*, vol. 18, no. 6, p. 1863, 2018. 28

[100] K. Haseeb, N. Islam, A. Almogren, I. U. Din, H. N. Almajed, and N. Guizani, "Secret sharing-based energy-aware and multi-hop routing protocol for iot based wsns," *IEEE Access*, vol. 7, pp. 79 980–79 988, 2019. 28, 158

[101] D. K. Sharma, S. K. Dhurandher, I. Woungang, R. K. Srivastava, A. Mohananey, and J. J. Rodrigues, "A machine learning-based protocol for efficient routing in opportunistic networks," *IEEE Systems Journal*, vol. 12, no. 3, pp. 2207–2213, 2016. 30, 36

[102] F. Tang, B. Mao, Z. M. Fadlullah, N. Kato, O. Akashi, T. Inoue, and K. Mizutani, "On removing routing protocol from future wireless networks: A real-time deep learning approach for intelligent traffic control," *IEEE Wireless Communications*, vol. 25, no. 1, pp. 154–160, 2017. 30, 36

[103] N. Kato, Z. M. Fadlullah, B. Mao, F. Tang, O. Akashi, T. Inoue, and K. Mizutani, "The deep learning vision for heterogeneous network traffic control: Proposal, challenges, and

future perspective," *IEEE wireless communications*, vol. 24, no. 3, pp. 146–153, 2016. 30, 36

[104] H. Yao, X. Yuan, P. Zhang, J. Wang, C. Jiang, and M. Guizani, "A machine learning approach of load balance routing to support next-generation wireless networks," in *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*. IEEE, 2019, pp. 1317–1322. 31, 36, 71

[105] D. K. Sharma, S. K. Dhurandher, D. Agarwal, and K. Arora, "krop: k-means clustering based routing protocol for opportunistic networks," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 4, pp. 1289–1306, 2019. 31, 36

[106] S. K. Dhurandher, D. K. Sharma, I. Woungang, and S. Bhati, "Hbpr: history based prediction for routing in infrastructure-less opportunistic networks," in *2013 IEEE 27th international Conference on advanced information networking and applications (AINA)*. IEEE, 2013, pp. 931–936. 31

[107] A. Lindgren, A. Doria, and O. Schelén, "Probabilistic routing in intermittently connected networks," *ACM SIGMOBILE mobile computing and communications review*, vol. 7, no. 3, pp. 19–20, 2003. 31

[108] Y. Tang, N. Cheng, W. Wu, M. Wang, Y. Dai, and X. Shen, "Delay-minimization routing for heterogeneous vanets with machine learning based mobility prediction," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 3967–3979, 2019. 31, 36, 71

[109] L. Zhao, J. Wang, J. Liu, and N. Kato, "Routing for crowd management in smart cities: A deep reinforcement learning perspective," *IEEE Communications Magazine*, vol. 57, no. 4, pp. 88–93, 2019. 32, 36

[110] Y. Saleem, K.-L. A. Yau, H. Mohamad, N. Ramli, M. H. Rehmani, and Q. Ni, "Clustering and reinforcement-learning-based routing for cognitive radio networks," *IEEE Wireless Communications*, vol. 24, no. 4, pp. 146–151, 2017. 32, 33, 36

[111] A. R. Syed, K.-L. A. Yau, J. Qadir, H. Mohamad, N. Ramli, and S. L. Keoh, "Route selection for multi-hop cognitive radio networks using reinforcement learning: An experimental study," *IEEE Access*, vol. 4, pp. 6304–6324, 2016. 32, 36

[112] V. Tilwari, K. Dimyati, M. Hindia, A. Fattouh, and I. S. Amiri, "Mobility, residual energy, and link quality aware multipath routing in manets with q-learning algorithm," *Applied Sciences*, vol. 9, no. 8, p. 1582, 2019. 33, 36

[113] J. Yi, A. Adnane, S. David, and B. Parrein, "Multipath optimized link state routing for mobile ad hoc networks," *Ad hoc networks*, vol. 9, no. 1, pp. 28–47, 2011. 33

[114] J. Yi and B. Parrein, "Multipath extension for the optimized link state routing protocol version 2 (olsrv2)," 2017. 33

[115] H. A. Al-Rawi, K.-L. A. Yau, H. Mohamad, N. Ramli, and W. Hashim, "Effects of network characteristics on learning mechanism for routing in cognitive radio ad hoc networks," in *2014 9th International Symposium on Communication Systems, Networks & Digital Sign (CSNDSP)*. IEEE, 2014, pp. 748–753. 33

[116] H. Yao, M. Tianle, J. Chunxiao, K. Linling, and G. Song, "Ai routers network mind: A hybrid machine learning paradigm for packet routing," in *IEEE Computational Intelligence Magazine*, vol. 14, no. 4. IEEE, 2019, pp. 21–30. 33, 36

[117] G. Stampa, M. Arias, D. Sánchez-Charles, V. Muntés-Mulero, and A. Cabellos, "A deep-reinforcement learning approach for software-defined networking routing optimization," *arXiv preprint arXiv:1709.07080*, 2017. 33, 36

[118] N. F. M. Aun, P. J. Soh, A. A. Al-Hadi, M. F. Jamlos, G. A. Vandenbosch, and D. Schreurs, "Revolutionizing wearables for 5g: 5g technologies: Recent developments and future perspectives for wearable devices and antennas," *IEEE Microwave Magazine*, vol. 18, no. 3, pp. 108–124, 2017. 38

[119] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5g: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE journal on selected areas in communications*, vol. 35, no. 6, pp. 1201–1221, 2017. 38

[120] A. A. Barakabitze, A. Ahmad, R. Mijumbi, and A. Hines, "5g network slicing using sdn and nfv: A survey of taxonomy, architectures and future challenges," *Computer Networks*, vol. 167, p. 106984, 2020. 38

[121] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, and H. Taoka, "Scenarios for 5g mobile and wireless communications: the vision of the metis project," *IEEE communications magazine*, vol. 52, no. 5, pp. 26–35, 2014. 38

[122] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5g be?" *IEEE Journal on selected areas in communications*, vol. 32, no. 6, pp. 1065–1082, 2014. 38

[123] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5g wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016. 38

[124] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 1801–1819, 2014. 38

[125] U. N. Kar and D. K. Sanyal, "An overview of device-to-device communication in cellular networks," *ICT express*, vol. 4, no. 4, pp. 203–208, 2018. 39

[126] R. Zhang, X. Cheng, L. Yang, and B. Jiao, "Interference-aware graph based resource sharing for device-to-device communications underlaying cellular networks," in *2013 IEEE wireless communications and networking conference (WCNC)*. IEEE, 2013, pp. 140–145. 39

[127] A. Osseiran, K. Doppler, C. Ribeiro, M. Xiao, M. Skoglund, and J. Manssour, "Advances in device-to-device communications and network coding for imt-advanced," *ICT Mobile Summit*, pp. 1–8, 2009. 39

[128] K. Pentikousis, Y. Wang, and W. Hu, "Mobileflow: Toward software-defined mobile networks," *IEEE Communications magazine*, vol. 51, no. 7, pp. 44–53, 2013. 39

[129] L. E. Li, Z. M. Mao, and J. Rexford, "Toward software-defined cellular networks," in *2012 European workshop on software defined networking*. IEEE, 2012, pp. 7–12. 39

[130] M. Jarschel, "An assessment of applications and performance analysis of software defined networking," 2014. 39

[131] N. Feamster, J. Rexford, and E. Zegura, "The road to sdn: an intellectual history of programmable networks," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 2, pp. 87–98, 2014. 39

[132] A. Huang, N. Nikaein, T. Stenbock, A. Ksentini, and C. Bonnet, "Low latency mec framework for sdn-based lte/lte-a networks," in *2017 IEEE International Conference on Communications (ICC)*. IEEE, 2017, pp. 1–6. 39

[133] J. Costa-Requena, "Sdn integration in lte mobile backhaul networks," in *The International Conference on Information Networking 2014 (ICOIN2014)*. IEEE, 2014, pp. 264–269. 39

[134] A. Basta, W. Kellerer, M. Hoffmann, K. Hoffmann, and E.-D. Schmidt, "A virtual sdn-enabled lte epc architecture: A case study for s-/p-gateways functions," in *2013 IEEE SDN for Future Networks and Services (SDN4FNS)*. IEEE, 2013, pp. 1–7. 39

[135] R. Jayadi and Y.-C. Lai, "Low-overhead multihop device-to-device communications in software defined wireless networks," in *2017 International Conference on Soft Computing, Intelligent System and Information Technology (ICSIIT)*. IEEE, 2017, pp. 144–149. 40

[136] M. Usman, A. A. Gebremariam, U. Raza, and F. Granelli, "A software-defined device-to-device communication architecture for public safety applications in 5g networks," *IEEE Access*, vol. 3, pp. 1649–1654, 2015. 40

[137] M. Kaplan, C. Zheng, M. Monaco, E. Keller, and D. Sicker, "WASP: a software-defined communication layer for hybrid wireless networks," in *2014 ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS)*. IEEE, 2014, pp. 5–15. 40

[138] A. Thomas and G. Raja, "Finder: A d2d based critical communications framework for disaster management in 5g," *Peer-to-Peer Networking and Applications*, vol. 12, no. 4, pp. 912–923, 2019. 40

[139] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave mimo systems," *IEEE transactions on wireless communications*, vol. 13, no. 3, pp. 1499–1513, 2014. 44

[140] P. Xia, S.-K. Yong, J. Oh, and C. Ngo, "A practical sdma protocol for 60 ghz millimeter wave communications," in *2008 42nd Asilomar Conference on Signals, Systems and Computers*. IEEE, 2008, pp. 2019–2023. 44

[141] ——, "Multi-stage iterative antenna training for millimeter wave communications," in *IEEE GLOBECOM 2008-2008 IEEE Global Telecommunications Conference*. IEEE, 2008, pp. 1–6. 44

[142] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, 2014. 44

[143] S. A. M. Ghaleb and V. Vasanthi, "Energy efficient multipath routing using multi-objective grey wolf optimizer based dynamic source routing algorithm for manet," *International Journal of Advanced Science and Technology*, vol. 29, no. 3, pp. 6096–6117, 2020. 49, 80, 154

[144] H. T. Friis, "A note on a simple transmission formula," *Proceedings of the IRE*, vol. 34, no. 5, pp. 254–256, 1946. 52, 88

[145] L. Zhu, F. R. Yu, Y. Wang, B. Ning, and T. Tang, "Big data analytics in intelligent transportation systems: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 1, pp. 383–398, 2018. 67

[146] P. K. Padhi and F. Charrua-Santos, "6g enabled industrial internet of everything: Towards a theoretical framework," *Applied System Innovation*, vol. 4, no. 1, p. 11, 2021. 67

[147] G. Liu and D. Jiang, "5g: Vision and requirements for mobile communication system towards year 2020," *Chinese Journal of Engineering*, vol. 2016, no. 2016, p. 8, 2016. 67

[148] Y. Xu, G. Gui, H. Gacanin, and F. Adachi, "A survey on resource allocation for 5g heterogeneous networks: current research, future trends and challenges," *IEEE Communications Surveys & Tutorials*, 2021. 67

[149] L. Bonati, M. Polese, S. D'Oro, S. Basagni, and T. Melodia, "Open, programmable, and virtualized 5g networks: State-of-the-art and the road ahead," *Computer Networks*, vol. 182, p. 107516, 2020. 67

[150] Y. L. Lee, D. Qin, L.-C. Wang, and G. H. Sim, "6g massive radio access networks: Key applications, requirements and challenges," *IEEE Open Journal of Vehicular Technology*, vol. 2, pp. 54–66, 2020. 67, 144

[151] W. Saad, M. Bennis, and M. Chen, "A vision of 6g wireless systems: Applications, trends, technologies, and open research problems," *IEEE network*, vol. 34, no. 3, pp. 134–142, 2019. 67

[152] F. Hussain, S. A. Hassan, R. Hussain, and E. Hossain, "Machine learning for resource management in cellular and iot networks: Potentials, current solutions, and open challenges," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1251–1275, 2020. 67

[153] Z. Chen, M. A. Masrur, and Y. L. Murphey, "Intelligent vehicle power management using machine learning and fuzzy logic," in *2008 IEEE International Conference on Fuzzy Systems (IEEE World Congress on Computational Intelligence)*. IEEE, 2008, pp. 2351–2358. 67

[154] M. Qin, Q. Yang, N. Cheng, H. Zhou, R. R. Rao, and X. Shen, "Machine learning aided context-aware self-healing management for ultra dense networks with qos provisions," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 12 339–12 351, 2018. 67

[155] H. Fatemidokht, M. K. Rafsanjani, B. B. Gupta, and C.-H. Hsu, "Efficient and secure routing protocol based on artificial intelligence algorithms with uav-assisted for vehicular ad hoc networks in intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4757–4769, 2021. 67

[156] A. Kak, "Towards 6g through sdn and nfv-based solutions for terrestrial and non-terrestrial networks," Ph.D. dissertation, Georgia Institute of Technology, 2021. 67

[157] M. M. Azari, S. Solanki, S. Chatzinotas, O. Kodheli, H. Sallouha, A. Colpaert, J. F. M. Montoya, S. Pollin, A. Haqiqatnejad, A. Mostaani *et al.*, "Evolution of non-terrestrial networks from 5g to 6g: A survey," *arXiv preprint arXiv:2107.06881*, 2021. 67

[158] K. Benzekki, A. El Fergougui, and A. Elbelrhiti Elalaoui, "Software-defined networking (sdn): a survey," *Security and communication networks*, vol. 9, no. 18, pp. 5803–5833, 2016. 70

[159] A. Filali, Z. Mlika, S. Cherkaoui, and A. Kobbane, "Preemptive sdn load balancing with machine learning for delay sensitive applications," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 15 947–15 963, 2020. 70

[160] J. Wang, Y. Miao, P. Zhou, M. S. Hossain, and S. M. M. Rahman, "A software defined network routing in wireless multihop network," *Journal of Network and Computer Applications*, vol. 85, pp. 76–83, 2017. 70

[161] K. L. K. Sudheera, M. Ma, and P. H. J. Chong, "Link stability based optimized routing framework for software defined vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2934–2945, 2019. 70

[162] C. Wu, S. Ohzahata, and T. Kato, "Flexible, portable, and practicable solution for routing in vanets: A fuzzy constraint q-learning approach," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4251–4263, 2013. 70

[163] S. Sun, Z. Cao, H. Zhu, and J. Zhao, "A survey of optimization methods from a machine learning perspective," *IEEE transactions on cybernetics*, vol. 50, no. 8, pp. 3668–3681, 2019. 70

[164] S. Ali, W. Saad, N. Rajatheva, K. Chang, D. Steinbach, B. Sliwa, C. Wietfeld, K. Mei, H. Shiri, H.-J. Zepernick *et al.*, "6g white paper on machine learning in wireless communication networks," *arXiv preprint arXiv:2004.13875*, 2020. 70

[165] D. Bega, M. Gramaglia, A. Banchs, V. Sciancalepore, and X. Costa-Pérez, "A machine learning approach to 5g infrastructure market optimization," *IEEE Transactions on Mobile Computing*, vol. 19, no. 3, pp. 498–512, 2019. 70

[166] Y.-R. Chen, A. Rezapour, W.-G. Tzeng, and S.-C. Tsai, "Rl-routing: An sdn routing algorithm based on deep reinforcement learning," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 3185–3199, 2020. 70

[167] Z. M. Zahedi, R. Akbari, M. Shokouhifar, F. Safaei, and A. Jalali, "Swarm intelligence based fuzzy routing protocol for clustered wireless sensor networks," *Expert Systems with Applications*, vol. 55, pp. 313–328, 2016. 71

[168] N. I. Abbas, M. Ilkan, and E. Ozen, "Fuzzy approach to improving route stability of the aodv routing protocol," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, pp. 1–11, 2015. 71

[169] L. Zhao, Z. Bi, M. Lin, A. Hawbani, J. Shi, and Y. Guan, "An intelligent fuzzy-based routing scheme for software-defined vehicular networks," *Computer Networks*, vol. 187, p. 107837, 2021. 71

[170] C. Ghorai, S. Shakhari, and I. Banerjee, "A spea-based multimetric routing protocol for intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 6737–6747, 2020. 71

[171] M. Abolhasan, T. Wysocki, and E. Dutkiewicz, "A review of routing protocols for mobile ad hoc networks," *Ad hoc networks*, vol. 2, no. 1, pp. 1–22, 2004. 74

[172] K. Kalkan, "Sutsec: Sdn utilized trust based secure clustering in iot," *Computer Networks*, vol. 178, p. 107328, 2020. 75

[173] T. S. Darwish, K. A. Bakar, and K. Haseeb, "Reliable intersection-based traffic aware routing protocol for urban areas vehicular ad hoc networks," *IEEE Intelligent Transportation Systems Magazine*, vol. 10, no. 1, pp. 60–73, 2018. 75

[174] C. Wang, *A study of membership functions on mamdani-type fuzzy inference system for industrial decision-making*. Lehigh University, 2015. 79

[175] W.-C. Chien, C.-F. Lai, M. S. Hossain, and G. Muhammad, "Heterogeneous space and terrestrial integrated networks for iot: Architecture and challenges," *IEEE network*, vol. 33, no. 1, pp. 15–21, 2019. 94

[176] X. Yang, X. Wang, Y. Wu, L. P. Qian, W. Lu, and H. Zhou, "Small-cell assisted secure traffic offloading for narrowband internet of thing (nb-iot) systems," *IEEE Internet of things Journal*, vol. 5, no. 3, pp. 1516–1526, 2017. 94

[177] S.-Y. Lien, K.-C. Chen, Y.-C. Liang, and Y. Lin, "Cognitive radio resource management for future cellular networks," *IEEE Wireless Communications*, vol. 21, no. 1, pp. 70–79, 2014. 94

[178] X. Liu and N. Ansari, "Green relay assisted d2d communications with dual batteries in heterogeneous cellular networks for iot," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1707–1715, 2017. 94

[179] M. Abdollahi, S. Ashtari, M. Abolhasan, N. Shariati, J. Lipman, A. Jamalipour, and W. Ni, "Dynamic routing protocol selection in multi-hop device-to-device wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 8, pp. 8796–8809, 2022. 94, 122, 123

[180] G. Bloom, B. Alsulami, E. Nwafor, and I. C. Bertolotti, "Design patterns for the industrial internet of things," in *2018 14th IEEE International Workshop on Factory Communication Systems (WFCS)*. IEEE, 2018, pp. 1–10. 94

[181] Y. Hui, Z. Su, and S. Guo, "Utility based data computing scheme to provide sensing service in internet of things," *IEEE Transactions on Emerging Topics in Computing*, vol. 7, no. 2, pp. 337–348, 2017. 94

[182] Y. Zhao, Y. Li, H. Mao, and N. Ge, "Social-community-aware long-range link establishment for multihop d2d communication networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 11, pp. 9372–9385, 2016. 97

[183] J. Huang, Y. Liao, C.-C. Xing, and Z. Chang, "Multi-hop d2d communications with network coding: From a performance perspective," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2270–2282, 2019. 97

[184] G. Chen, J. P. Coon, and S. E. Tajbakhsh, "Secure routing for multihop ad hoc networks with inhomogeneous eavesdropper clusters," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10 660–10 670, 2018. 97

[185] F. S. Shaikh and R. Wismuller, "Interference-conscious routing in multihop d2d communications," in *2017 2nd International Conference on Computer and Communication Systems (ICCCS)*. IEEE, 2017, pp. 146–151. 97

[186] E. M. Mohamed, B. M. Elhalawany, H. S. Khallaf, M. Zareei, A. Zeb, and M. A. Abdelghany, "Relay probing for millimeter wave multi-hop d2d networks," *IEEE Access*, vol. 8, pp. 30 560–30 574, 2020. 97

[187] C. Huang, B. Zhai, A. Tang, and X. Wang, "Virtual mesh networking for achieving multi-hop d2d communications in 5g networks," *Ad Hoc Networks*, vol. 94, p. 101936, 2019. 98

[188] R. Jin, X. Fan, and T. Sun, "Centralized multi-hop routing based on multi-start minimum spanning forest algorithm in the wireless sensor networks," *Sensors*, vol. 21, no. 5, p. 1775, 2021. 98

[189] A. Dusia, R. Ramanathan, and A. S. Sethi, "Corr: Centralized opportunistic reactive routing for mobile multi-hop wireless networks," in *2019 28th International Conference on Computer Communication and Networks (ICCCN)*. IEEE, 2019, pp. 1–6. 98

[190] M. Adnan, L. Yang, T. Ahmad, and Y. Tao, "An unequally clustered multi-hop routing protocol based on fuzzy logic for wireless sensor networks," *IEEE Access*, vol. 9, pp. 38 531–38 545, 2021. 98

[191] G. Chen, J. Tang, and J. P. Coon, "Optimal routing for multihop social-based d2d communications in the internet of things," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1880–1889, 2018. 98

[192] X. Liu, Z. Li, N. Zhao, W. Meng, G. Gui, Y. Chen, and F. Adachi, "Transceiver design and multihop d2d for uav iot coverage in disasters," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1803–1815, 2018. 98

[193] F. Yang, J. Han, X. Ding, Z. Wei, and X. Bi, "Spectral efficiency optimization and interference management for multi-hop d2d communications in vanets," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 6422–6436, 2020. 98

[194] J. Lipman, P. Boustead, J. F. Chicharo, and J. Judge, "Resource aware information dissemination in ad hoc networks," 2003. 98, 104

[195] A. Behzadan, A. Anpalagan, I. Woungang, B. Ma, and H.-C. Chao, "An energy-efficient utility-based distributed data routing scheme for heterogenous sensor networks," *Wireless Communications and Mobile Computing*, vol. 15, no. 16, pp. 2020–2037, 2015. 98

[196] M. Abolhasan and J. Lipman, "Self-selection route discovery strategies for reactive routing in ad hoc networks," in *Proceedings of the first international conference on integrated internet ad hoc and sensor networks*, 2006, pp. 21–es. 98, 104

[197] W. Zhao and W.-H. Kuo, "Utility-based wireless routing algorithm for massive mimo heterogeneous networks," *Applied Sciences*, vol. 10, no. 20, p. 7261, 2020. 98

[198] B. G. Assefa and Ö. Özkasap, "A novel utility based metric and routing for energy efficiency in software defined networking," in *2019 International Symposium on Networks, Computers and Communications (ISNCC)*. IEEE, 2019, pp. 1–4. 98

[199] D. Jiang, L. Huo, Z. Lv, H. Song, and W. Qin, "A joint multi-criteria utility-based network selection approach for vehicle-to-infrastructure networking," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 10, pp. 3305–3319, 2018. 98

[200] B. G. Assefa and Ö. Özkasap, "Resdn: A novel metric and method for energy efficient routing in software defined networks," *IEEE Transactions on Network and Service Management*, vol. 17, no. 2, pp. 736–749, 2020. 98

[201] S. Krug, M. Aumüller, and J. Seitz, "Hybrid scheme to enable dtn routing protocols to efficiently exploit stable manet contacts," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, pp. 1–13, 2018. 99

[202] O. Alzamzami and I. Mahgoub, "Link utility aware geographic routing for urban vanets using two-hop neighbor information," *Ad Hoc Networks*, vol. 106, p. 102213, 2020. 99

[203] S. Banerji and R. S. Chowdhury, "On ieee 802.11: wireless lan technology," *arXiv preprint arXiv:1307.2661*, 2013. 102, 152

[204] M. Abolhasan, J. Lipman, and J. Chicharo, "A routing strategy for heterogeneous mobile ad hoc networks," in *Proceedings of the IEEE 6th Circuits and Systems Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communication (IEEE Cat. No. 04EX710)*, vol. 1. IEEE, 2004, pp. 13–16. 104

[205] K. Alexandris, C.-Y. Chang, K. Katsalis, N. Nikaein, and T. Spyropoulos, "Utility-based resource allocation under multi-connectivity in evolved lte," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2017, pp. 1–6. 104

[206] C. Curescu and S. Nadjm-Tehrani, "A bidding algorithm for optimized utility-based resource allocation in ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 12, pp. 1397–1414, 2008. 104

[207] X. Wang, Y. Weng, and H. Gao, "A low-latency and energy-efficient multimetric routing protocol based on network connectivity in vanet communication," *IEEE Transactions on Green Communications and Networking*, 2021. 112, 161

[208] Q. Usman, O. Chughtai, N. Nawaz, Z. Kaleem, K. A. Khaliq, and L. D. Nguyen, "A reliable link-adaptive position-based routing protocol for flying ad hoc network," *Mobile Networks and Applications*, vol. 26, no. 4, pp. 1801–1820, 2021. 112, 161

[209] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfröjd, and T. Svensson, "The role of small cells, coordinated multipoint, and massive mimo in 5g," *IEEE communications magazine*, vol. 52, no. 5, pp. 44–51, 2014. 119

[210] J. S. Walia, H. Hämmäinen, and M. Matinmikko, "5g micro-operators for the future campus: A techno-economic study," in *2017 internet of things business models, users, and networks*. IEEE, 2017, pp. 1–8. 119

[211] U. N. Kar and D. K. Sanyal, "A critical review of 3gpp standardization of device-to-device communication in cellular networks," *SN Computer Science*, vol. 1, no. 1, pp. 1–18, 2020. 119, 144

[212] M. Noura and R. Nordin, "A survey on interference management for device-to-device (d2d) communication and its challenges in 5g networks," *Journal of Network and Computer Applications*, vol. 71, pp. 130–150, 2016. 119

[213] D. Griffith, A. B. Mosbah, and R. Rouil, "Group discovery time in device-to-device (d2d) proximity services (prose) networks," in *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*. IEEE, 2017, pp. 1–9. 119

[214] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Stochastic geometry study on device-to-device communication as a disaster relief solution," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3005–3017, 2015. 119

[215] "Study on further enhancements to lte device to device (d2d), user equipment (ue) to network relays for internet of things (iot) and wearables version 15.1.1," *Third Gener. Partnership Projec*, TR 36.746, 2018. 119

[216] P. Gandotra, R. K. Jha, and S. Jain, "A survey on device-to-device (d2d) communication: Architecture and security issues," *Journal of Network and Computer Applications*, vol. 78, pp. 9–29, 2017. 119

[217] G. Nardini, G. Stea, and A. Virdis, "A fast and reliable broadcast service for lte-advanced exploiting multihop device-to-device transmissions," 2017. 119

[218] B. Klaiqi, X. Chu, and J. Zhang, "Energy-and spectral-efficient adaptive forwarding strategy for multi-hop device-to-device communications overlaying cellular networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 5684–5699, 2018. 119

[219] J. Gui and J. Deng, "Multi-hop relay-aided underlay d2d communications for improving cellular coverage quality," *IEEE access*, vol. 6, pp. 14 318–14 338, 2018. 119

[220] R.-S. Cheng, C.-M. Huang, and S.-Y. Pan, "Wifi offloading using the device-to-device (d2d) communication paradigm based on the software defined network (sdn) architecture," *Journal of Network and Computer Applications*, vol. 112, pp. 18–28, 2018. 120

[221] S. ashtari, M. Abolhasan, J. Lipman, N. Shariati, W. Ni, and A. Jamalipour, "Joint mobile node participation and multihop routing for emerging open radio-based intelligent transportation system," *IEEE Access*, vol. 10, pp. 85 228–85 242, 2022. 120

[222] A. Douik, S. Sorour, T. Y. Al-Naffouri, H.-C. Yang, and M.-S. Alouini, "Delay reduction in multi-hop device-to-device communication using network coding," *IEEE Transactions on Wireless Communications*, vol. 17, no. 10, pp. 7040–7053, 2018. 122

[223] L. F. Del Carpio, A. A. Dowhuszko, O. Tirkkonen, and G. Wu, "Simple clustering methods for multi-hop cooperative device-to-device communication," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*. IEEE, 2015, pp. 1–6. 122

[224] Y. Lv, W. Feng, A. Xiao, and X. Bao, "Cluster-head centered fast secure routing based on game theory for device-to-device communication," *Wireless Personal Communications*, vol. 113, no. 4, pp. 2079–2106, 2020. 122

[225] Y. Sung, S. Lee, and M. Lee, "A multi-hop clustering mechanism for scalable iot networks," *Sensors*, vol. 18, no. 4, p. 961, 2018. 123

[226] H. Rong, Z. Wang, H. Jiang, Z. Xiao, and F. Zeng, "Energy-aware clustering and routing in infrastructure failure areas with d2d communication," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8645–8657, 2019. 123

[227] M. Pushpalatha, T. Ramesh, C. Giriraja, and S. K. Konda, "Power efficient multicast routing protocol for dynamic intra cluster device to device communication," in *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*. IEEE, 2016, pp. 1853–1856. 123

[228] L. D. Pacifico and T. B. Ludermir, "Hybrid k-means and improved self-adaptive particle swarm optimization for data clustering," in *2019 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2019, pp. 1–7. 125, 127

[229] Y. Zhang, X. Xiong, and Q. Zhang, "An improved self-adaptive pso algorithm with detection function for multimodal function optimization problems," *Mathematical Problems in Engineering*, vol. 2013, 2013. 125, 127

[230] M. Nabati, M. Maadani, and M. A. Pourmina, "Agen-aodv: an intelligent energy-aware routing protocol for heterogeneous mobile ad-hoc networks," *Mobile Networks and Applications*, vol. 27, no. 2, pp. 576–587, 2022. 133

[231] A. Rezvanian, A. M. Saghiri, S. M. Vahidipour, M. Esnaashari, and M. R. Meybodi, *Recent advances in learning automata*. Springer, 2018, vol. 754. 133

[232] M. K. Hasan and O. Sarker, "Routing protocol selection for intelligent transport system (its) of vanet in high mobility areas of bangladesh," in *Proceedings of International Joint Conference on Computational Intelligence: IJCCI 2018*. Springer, 2020, pp. 123–135. 135

[233] S. Malik and P. K. Sahu, "A comparative study on routing protocols for vanets," *Heliyon*, vol. 5, no. 8, p. e02340, 2019. 135

[234] M. Höyhtyä, O. Apilo, and M. Lasanen, "Review of latest advances in 3gpp standardization: D2d communication in 5g systems and its energy consumption models," *Future Internet*, vol. 10, no. 1, p. 3, 2018. 144

[235] S. Ashtari, M. Abolhasan, J. Lipman, N. Shariati, W. Ni, and A. Jamalipour, "Joint mobile node participation and multi-hop routing for emerging open radio-based intelligent transportation system," *IEEE Access*, 2022. 144

[236] S. Kumar and A. K. Verma, "Position based routing protocols in vanet: a survey," *Wireless Personal Communications*, vol. 83, no. 4, pp. 2747–2772, 2015. 144

[237] R. Dhanapal, V. Kalpana, R. Chiwariro, N. Thanagadurai, K. Sentamilselvan, and D. J. Immanuel, "An energy efficient secure routing for mobile nodes for multi hop adhoc network," *Journal of Computational and Theoretical Nano science*, vol. 17, pp. 1–4, 2020. 145

[238] U. Srilakshmi, S. A. Alghamdi, V. A. Vuyyuru, N. Veeraiah, and Y. Alotaibi, "A secure optimization routing algorithm for mobile ad hoc networks," *IEEE Access*, vol. 10, pp. 14 260–14 269, 2022. 145

[239] N. Veeraiah, O. I. Khalaf, C. Prasad, Y. Alotaibi, A. Alsufyani, S. A. Alghamdi, and N. Alsufyani, "Trust aware secure energy efficient hybrid protocol for manet," *IEEE Access*, vol. 9, pp. 120 996–121 005, 2021. 145

[240] S. ashtari, M. Abolhasan, J. Lipman, N. Shariati, W. Ni, and A. Jamalipour, "Joint mobile node participation and multihop routing for emerging open radio-based intelligent transportation system," *IEEE Access*, vol. 10, pp. 85 228–85 242, 2022. 147

[241] J. Chen, Y. Wang, X. Huang, X. Xie, H. Zhang, and X. Lu, "Alblp: adaptive load-balancing architecture based on link-state prediction in software-defined networking," *Wireless Communications and Mobile Computing*, vol. 2022, 2022. 148

[242] M. Kumar, A. K. Nigam, and T. Sivakumar, "A survey on topology and position based routing protocols in vehicular ad hoc network (vanet)," *International Journal on Future Revolution in Computer Science & Communication Engineering*, vol. 4, no. 2, pp. 432–440, 2018. 148

[243] R. Karimi and S. Shokrollahi, "Pgrp: Predictive geographic routing protocol for vanets," *Computer Networks*, vol. 141, pp. 67–81, 2018. 148

[244] L. Liu, C. Chen, Z. Ren, and F. R. Yu, "An intersection-based geographic routing with transmission quality guaranteed in urban vanets," in *2018 IEEE international conference on communications (ICC)*. IEEE, 2018, pp. 1–6. 148

[245] O. Alzamzami and I. Mahgoub, "Fuzzy logic-based geographic routing for urban vehicular networks using link quality and achievable throughput estimations," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 6, pp. 2289–2300, 2018. 149

[246] C. Nallusamy and A. Sabari, "Particle swarm based resource optimized geographic routing for improved network lifetime in manet," *Mobile Networks and Applications*, vol. 24, no. 2, pp. 375–385, 2019. 149

[247] A. R. Hameed, S. ul Islam, M. Raza, and H. A. Khattak, "Towards energy and performance-aware geographic routing for iot-enabled sensor networks," *Computers & Electrical Engineering*, vol. 85, p. 106643, 2020. 149

[248] C.-L. Hu and C. Sosorburam, "Enhanced geographic routing with two-hop neighborhood information in sparse manets," *Wireless Personal Communications*, vol. 107, no. 1, pp. 417–436, 2019. 149

[249] M. Naghibi and H. Barati, "Egrpm: Energy efficient geographic routing protocol based on mobile sink in wireless sensor networks," *Sustainable Computing: Informatics and Systems*, vol. 25, p. 100377, 2020. 149

[250] D. K. San and M. Turker, "Building extraction from high resolution satellite images using hough transform," 2010. 149

[251] S.-H. Moon, S. Park, and S.-j. Han, "Energy efficient data collection in sink-centric wireless sensor networks: A cluster-ring approach," *Computer Communications*, vol. 101, pp. 12–25, 2017. 151

[252] C.-C. Hu, Y.-L. Kuo, C.-Y. Chiu, and Y.-M. Huang, "Maximum bandwidth routing and maximum flow routing in wireless mesh networks," *Telecommunication Systems*, vol. 44, no. 1, pp. 125–134, 2010. 155

[253] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, "Impact of interference on multi-hop wireless network performance," *Wireless networks*, vol. 11, no. 4, pp. 471–487, 2005. 155

[254] D. Jinil Persis and T. Paul Robert, "Review of ad-hoc on-demand distance vector protocol and its swarm intelligent variants for mobile ad-hoc network," *IET Networks*, vol. 6, no. 5, pp. 87–93, 2017. 156

[255] Y. Tan, J. Yang, and N. Gopalakrishnan, "Self-learning, adaptive approach for intelligent analytics-assisted self-organizing-networks (sons)," Aug. 13 2019, uS Patent 10,382,979. 167

[256] J. Moysen and L. Giupponi, "From 4g to 5g: Self-organized network management meets machine learning," *Computer Communications*, vol. 129, pp. 248–268, 2018. 167

[257] N. Kato, B. Mao, F. Tang, Y. Kawamoto, and J. Liu, “Ten challenges in advancing machine learning technologies toward 6g,” *IEEE Wireless Communications*, vol. 27, no. 3, pp. 96–103, 2020. 168

[258] C. Huang, S. Hu, G. C. Alexandropoulos, A. Zappone, C. Yuen, R. Zhang, M. Di Renzo, and M. Debbah, “Holographic mimo surfaces for 6g wireless networks: Opportunities, challenges, and trends,” *IEEE Wireless Communications*, vol. 27, no. 5, pp. 118–125, 2020. 168

[259] E. Baştuğ, M. Bennis, and M. Debbah, “A transfer learning approach for cache-enabled wireless networks,” in *2015 13th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*. IEEE, 2015, pp. 161–166. 168

[260] E. Bastug, M. Bennis, and M. Debbah, “Anticipatory caching in small cell networks: A transfer learning approach,” in *1st KuVS workshop on anticipatory networks*, 2014. 168

[261] B. Bharath, K. G. Nagananda, and H. V. Poor, “A learning-based approach to caching in heterogenous small cell networks,” *IEEE Transactions on Communications*, vol. 64, no. 4, pp. 1674–1686, 2016. 168

[262] M. S. A. Muthanna, A. Muthanna, A. Rafiq, M. Hammoudeh, R. Alkanhel, S. Lynch, and A. A. Abd El-Latif, “Deep reinforcement learning based transmission policy enforcement and multi-hop routing in qos aware lora iot networks,” *Computer Communications*, vol. 183, pp. 33–50, 2022. 168