Efficient Multi-Hop Communications for Software-Defined Wireless Networks

by

Yuhui Yao

A thesis submitted to the University of London for the degree of
Doctor of Philosophy

School of Electronic Engineering and Computer Science
Queen Mary University of London
United Kingdom

September 2018

Abstract

Software-Defined Networking (SDN) recently emerged to overcome the difficulty of network control by decoupling the control plane from the data plane. In terms of the wireless medium and mobile devices, although new challenges are introduced into SDN research, SDN promises to address many inherited problems in wireless communication networks. However, centralised SDN control brings concerns of scalability, reliability, and robustness especially for wireless networks. Considering these concerns, the use of physically distributed SDN controllers has been recognized as an effective solution. Nevertheless, it remains a challenge in regard to how the physically distributed controllers effectively communicate to form a logically centralised network control plane.

Dissemination is a type of one-to-many communication service which plays an important role in control information exchange. This research focuses on the strategic packet forwarding for more efficient multi-hop communications in software-defined wireless networks. The research aim is to improve the delivery efficiency by exploiting the delay budget and node mobility. To achieve this objective, existing multi-hop forwarding methods and dissemination schemes in wireless networks are investigated and analysed.

In the literature, information from the navigation system of mobile nodes has been utilised to identify candidate relay nodes. However, further studies are required to utilise partially predictable mobility based on more generalised navigational information such as the movement direction. In this research, the feasible exploitation of directional movement in path-unconstrained mobility is investigated for efficient multi-hop communications. Simulation results show that the proposed scheme outperforms the state-of-the-art because directional correlation of node movement is considered to dynamically exploit the delay budget for better selection of the relay node(s).

TO MY FAMILY

Declaration

I, Yuhui Yao, confirm that the work presented in this thesis is my own work and it is original, except where due reference has been mentioned.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis. I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature ____________________

Date ____________________

Acknowledgments

First of all, I would like to express my sincere gratitude to my supervisors, Dr. Yan Sun and Dr. Chris Phillips, for their continuous support and encouragement throughout my Ph.D. study. I am also grateful to my independent assessor, Dr. Akram Alomainy, for his valuable suggestions and comments on my research.

I would also like to thank, Dr. Yue Cao (Northumbria University) and Dr. Xiaodong Xu (Beijing University of Posts and Telecommunications), who have spent time with me discussing research ideas and writing skills.

Moreover, I offer my gratitude to the colleagues and friends at Queen Mary, who made my time in London memorable.

I would also like to thank the Chinese Scholarship Council (CSC), for their financial support over the past four years.

Finally, I would like to take this opportunity to express my gratitude to my parents and girlfriend, who gave me enormous support and couragement throughout my Ph.D.

As a summary of my Ph.D. study, there are some general experiences I wish to share as well. Firstly, about paper writing, it should be noticed that the reader and the writer think differently. The writer naturally describes a journey of problem finding and solving. However, readers are commonly more interested in why the problem is important and what can be brought to them by the proposed solution. Keep this in mind and there is a chance to be an effective researcher. Secondly, the idea should always be as simple as possible. Learn to interface your work with others so the core novelty can be focused. Meanwhile, every little progress matters and every feasible target is valuable. Last but not least, try to enjoy sorrows and pains, and then the life becomes more joyful.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CH	Cluster Head
CMOS	Complementary Metal-Oxide-Semiconductor
CRW	Correlated Random Walk
DBRS	Delay Boundary based Relay Selection
DTN	Delay/Disruption Tolerant Network
IoT	Internet of Things
LTE	Long-Term Evolution
MANET	Mobile Ad hoc NETwork
MEC	Mobile Edge Computing
MOGA	Multi-Objective Genetic Algorithm
NFV	Network Functions Virtualisation
NP	Non-deterministic Polynomial-time
OLSR	Optimized Link State Routing
PBR	Position-Based Routing
QoS	Quality of Service
QoE	Quality of Experience

RTA	Random Turning Angle
SCF	Store-Carry-Forward
SDRS	Satisfaction Degree based Relay Selection
SDTA	Standard Deviation of the Turning Angles
SDN	Software-Defined Networking
SDMSN	Software-Defined Mobile Sensor Network
SINR	Signal-to-Interference-plus-Noise Ratio
SOF	Sensor OpenFlow
UAV	Unmanned Aerial Vehicle
VANET	Vehicular Ad hoc NETwork
WCN	Wireless Cellular Network
WMN	Wireless Mesh Network
WMSN	Wireless Multimedia Sensor Network
WLAN	Wireless Local Area Network
WSN	Wireless Sensor Network
WSAN	Wireless Sensor and Actor Network

Publication List

1. **Y. Yao**, Y. Sun, C. Phillips, and Y. Cao, "Movement-Aware Relay Selection for Delay-Tolerant Information Dissemination in Wildlife Tracking and Monitoring Applications", *IEEE Internet of Things Journal*, 2018.

2. **Y. Yao**, Y. Sun, and C. Phillips, "Movement Path Modelling for Node Mobility Handling", *In Proceedings of the 4th ACM MobiHoc Workshop on Experiences with the Design and Implementation of Smart Objects (SMARTOBJECTS '18)*, 2018.

3. Y. Sun, J. Guo, and **Y. Yao**, "Speed Up-Greedy Perimeter Stateless Routing Protocol for Wireless Sensor Networks (SU-GPSR)", *IEEE 18th International Conference on High Performance Switching and Routing (HPSR)*, 2017.

Chapter 1

Introduction

Software-Defined Networking (SDN) recently emerged to overcome the difficulty of network control by decoupling the control plane from the data plane [HAG16]. Meanwhile, future networks will become increasingly heterogeneous, interconnecting users and applications over networks ranging from wired, infrastructure-based wireless networks (e.g., Wireless Cellular Networks (WCNs), Wireless Mesh Networks (WMNs)), to infrastructure-less wireless networks (e.g., Mobile Ad-hoc NETworks (MANETs), Wireless Sensor Networks (WSNs)) [NMN+14]. Despite the rapid development of wired SDN, how to apply its concept to wireless networks (especially for infrastructure-less networks) remains an open-ended question and therefore more answers are always required.

This thesis focuses on the efficient multi-hop communications to support software-defined infrastructure-less wireless networks. In this chapter, a brief background is reviewed to introduce the research motivation with summarised contributions. The use-case example is demonstrated then, followed by the structure of this thesis and a list of publications.

1.1 Research Background

Along with the fast evolution of information technologies, Internet services have now widely penetrated into our daily lives. Further accelerated by the concept of the Internet of Things (IoT), various devices are being continuously introduced into diverse application scenarios. Ranging from home automation to environmental monitoring, IoT promises a ubiquitous means to sense and control every object connected to the network. Taking habitat monitoring [CEE+01] as an example, with the deployment of numerous smart sensors/motes, close observation can be achieved remotely on the unattended environment while actions can be taken in a timely manner by automatic actuators. It can be predicted that the advances in monitoring systems will greatly benefit ecology-related research both commercially and scientifically. However, conventional communication networks are meanwhile increasingly challenged by the services emerging from ubiquitous computing technologies [Pos11]. In IoT contexts, many items can be embedded with network interfaces and any of them may request to exchange data with any other; the unprecedented interactions between devices (or so called machine-type communication) is a critical issue to be addressed. Moreover, smooth integration is required to efficiently transform the 'big' raw data into meaningful information, where dedicated algorithms and schemes are indispensable. To finally support the interconnection of 'anything anytime anywhere', future networks have been envisioned to be much more heterogeneous, versatile, and pervasive.

Within these next-generation networks, the core functionality is usually formed of a flexible packet forwarding mechanism. Data is encapsulated into packets and then forwarded in a hop-by-hop manner via intermediate devices called nodes. The intermediate nodes establish a route from the source node to the destination node, where routing is a crucial function of control to determine the route for the data traffic. Traditionally, the forwarding behaviour is controlled by a program written into the device hardware. However, studies have revealed that the hardware-defined approach is too rigid to meet the increasing needs of network management and innovation [NMN+14]. As a widely-accepted novel paradigm, SDN effectively enhances the network flexibility by centralising the decision intelligence into single logical entity called the controller. With SDN the network becomes easily programmable, because details of network control are decoupled from the device implementation and are specified by the scripts running on controller(s). The SDN structure provides great opportunity for integration with big data related technologies, such as cloud computing, data mining and machine learning. In contrast to the success of SDN in wired networks especially at the enterprise and data centre level, the use of SDN in wireless networks is still controversial. So far some pioneering works have shown that the SDN concept is equally feasible in a wireless context [HAG16], despite the open-ended standardisation issues. In terms of the wireless medium and mobile nodes, further research is being conducted to address the distinctive challenges of programmable wireless networks, especially considering reliability, efficiency, and scalability.

With regard to packet forwarding in wireless networks, greater demands are placed on adaptive forwarding strategies due to the dynamic nature of the wireless devices and channels. Radio access networks are traditionally based on single-hop communication where the wireless devices directly communicate with network infrastructures (such as a cellular base station or wireless router) within the immediate radio range. For infrastructure-based radio access networks, recent SDN-related research works, such as Network Functions Virtualisation (NFV) [NBGT17] and Mobile Edge Computing (MEC) [MB17], mainly focus on the architecture-level resource optimisation. Meanwhile, multi-hop wireless communication [Toh01] has recently gained increasing interest taking into account its ability for power-saving, traffic offloading, disruption handling, and extending of coverage. Equipped with self-organized communication (or called ad hoc networking), the wireless nodes can be deployed rapidly and operated without relying on network infrastructure. Besides typical examples such as WSNs, MANETs, and VANETs, more advanced proposals have discussed topics including wireless operation in harsh environments (underwater [APM05] and underground [ASV09]), robotic drones [AK04], and multimedia traffic [AMC07].

Nevertheless, compared with infrastructure-based networks, more attention needs to be paid to the dynamically changing infrastructure-less topology, because intermediate forwarding is performed by the wireless devices which have limited resources and may move freely. At the same time, a number of researchers have proved that node movement can effectively facilitate packet forwarding [BB15]. By allowing intermittent connectivity, the emergence of Delay/Disruption Tolerant Networks (DTN) facilitates Store-Carry-Forward (SCF) [CS13] communication. Instead of blindly discarding data, a node may keep a bundle of packets until it meets the destination or more suitable carrier. Taking advantage of the encounter opportunities caused by node mobility, SCF offers a rich research area and provides more possibilities in strategic packet forwarding.

1.2 Research Problems

As a promising IoT use-case, habitat monitoring applications [CEE+01] greatly benefit ecology-related research both commercially and scientifically. Compared with manual data investigation and collection, automatic monitoring based on sensor networks is more suitable for long-term and continuous habitat studies. In Figure 1.1, an example scenario is provided where a number of sensor nodes are deployed to cover a target field for habitat monitoring purposes. Some of the sensor nodes (e.g. attached to animals) are mobile while others (e.g. fixed on plants or environmental features) are stationary. For the surveillance of animals, plants and the environment, sensor nodes collect various information from scalar sensing data (temperature, humidity, salinity, etc.) to multimedia data (video, sound, images, etc.). The collected data are transmitted collaboratively to the sink node, which connects to outside networks, and finally arrives at end monitors.

As these sensors maybe required to work for several months or years once deployed, reconfigurations (varying from simple rules to bulk code) are usually inevitable to adapt to the changing environment and requirements. This research considers a Software-Defined Mobile Sensor Network (SDMSN) system where there is one sink node with a

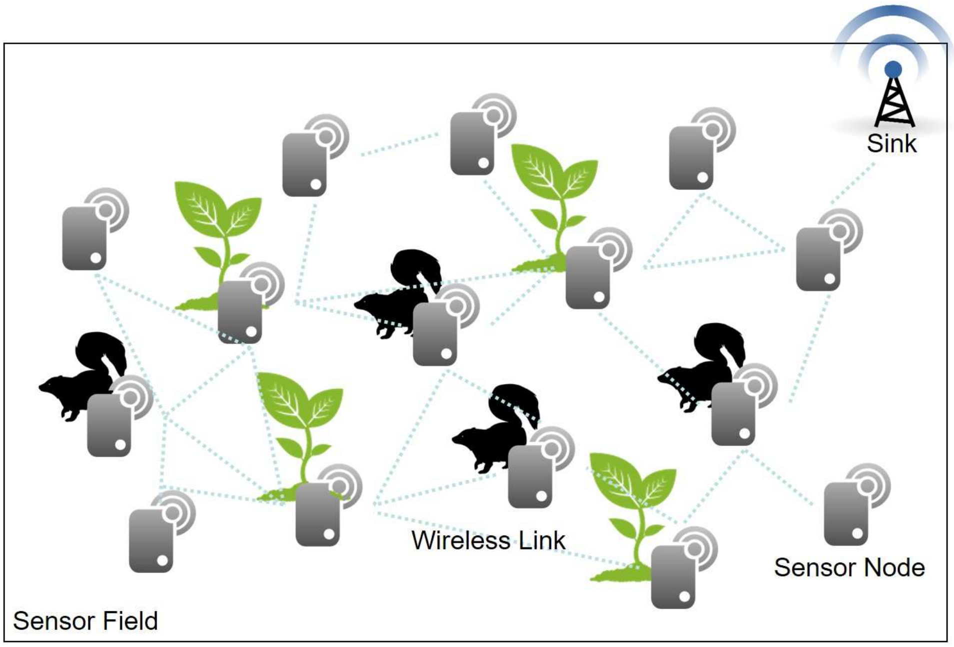


Figure 1.1: Habitat Monitoring Scenario

field of sensor nodes. Since the network functionality is defined by software, reconfigurations can be achieved more flexibly than with traditional hardware-defined systems, which is especially beneficial for long-term operations. During the system operation, the sink node (a central controller) may need to send updated control messages to all related sensor nodes (local controllers) which are called control message targets in this thesis. Specifically, we focus on the scenario where control messages are less delay-sensitive (e.g. a general instruction given like 'sending data to which kinds of node' instead of specifying 'sending data to which node'). Such partially distributed SDN control frameworks (i.e. with local controllers) have been discussed in the literature of wireless SDN research [HAG16], which can be envisioned as a development of future wireless networks.

The overall research problem is formulated as how a central controller (i.e. the sink node, which is stationary) can efficiently deliver control messages to some of the local controllers (sensor nodes) over a network consisting of stationary nodes and mobile nodes. In this thesis, the research focus is on the strategic forwarding to facilitate efficient dissemination of delay-tolerant information. Instead of immediate delivery to the mobile destination, control packets can be forwarded to a stationary relay in advance, waiting for later delivery when the destination node is directly contactable. For the habitat monitoring applications, the control information distribution can possibly be pre-planned, minutes or even hours ahead of scheduled updates (e.g. ecologists need more/less detailed data after a certain time), which provides a delivery delay budget. Considering that the radio power consumption is usually much higher than other energy expenditure [ASSC02], this strategic forwarding has the potential to bring dissemination efficiency gains by exploiting the possible delay budget.

This thesis mainly focuses on a specific scenario where there are a number of fixed grid nodes and a number of mobile nodes. Such a research scenario is motivated by the fact that habitat monitoring applications commonly deploy numerous sensors over a large area, as suggested in [CEE+01]. Although animals do not always keep moving, many of them spend a lot of time moving between points of interest (such as water holes), which provides the same mobility to their attached sensors. The maintenance of these sensors should be minimised after deployment, as animals are supposed to live freely without disturbance. In contrast, the sensors deployed on stationary objects (e.g. fixed to trees) can be less restricted (e.g. size/weight) and are more accessible for maintenance. It is therefore necessary to leverage the contact opportunities between such mobile and stationary sensors for the benefit of overall monitoring system operations.

Given the movement trajectory reported by destination nodes (e.g. the vehicular navigation system), it has been shown to be feasible (see [NF16] and [CYY16]) to find a better forwarding route (usually employing fewer hops) by devising a suitable forwarding strategy by exploiting the delivery delay budget. However, habitat monitoring applications place high demand on deploying devices (attached or implanted) to animals, so the movement trajectory of these mobile nodes becomes less possible to be precisely predicted for relay selection. Meanwhile, animals usually move with a certain purpose (such as foraging or migrating) and therefore their movement exhibits more-or-less predictability. To this end, further studies are required to utilise this partially predictable mobility based on more generalised navigational information such as the movement direction. Given the habitual behaviour that usually occurs in wildlife monitoring, this research proposes to utilise directional movement in path-unconstrained mobility for the strategic forwarding of delay-tolerant information.

1.3 Research Contributions

In the literature, multi-hop wireless communication schemes have been proposed for efficient message delivery. However, these existing works either cannot make full use of node mobility or can suffer from the uncertainty of node movement. Further studies are required to utilise partially predictable mobility based on more generalised navigational information such as the movement direction.

In this research, the feasible exploitation of directional movement in path-unconstrained[1] mobility is investigated for efficient multi-hop communications, which contributes to the following two attributes specifically:

- Contact prediction models are proposed and applied to fully utilise directional movement in path-unconstrained mobility for efficient delivery. Constrained paths (e.g. the roads in city) and known travel plans (e.g. reported routes from navigation systems) can help the schemes make the most of node mobility for delivery. However, these assumptions may not stand for sensor nodes deployed in the wild, because the movement of animals is uncontrolled and less constrained. On the other hand, most animals do move in certain natural patterns, and therefore it is unnecessary to adopt costly delivery schemes which treat their movement as completely random. Nevertheless, partially predictable mobility has not been sufficiently investigated. As existing schemes rely on accurate trajectories to make delivery decisions, the unexpected movement of mobile nodes can cause perfor-

[1] Mobile individuals independently move in relatively unobstructed environment (such as a forest or ocean).

mance degradation. In the proposed scheme, the node movement is predicted as an probabilistic range to accommodate the uncertainty of mobility prediction for optimised relay selection.

- Delivery strategy models are proposed and discussed to dynamically exploit the delay budget and node mobility. The exact delay requirement is widely used as a constraint for providing delay-guaranteed delivery services. However, sometimes it may be too costly or even impossible to deliver the message to certain roaming nodes within the given deadline. Moreover, messages usually have different delay-sensitivity, therefore the impact of overdue delivery varies and the decision should be made on a case-by-case basis. Nevertheless, the effect of overdue delivery has not been well considered in optimising the delivery strategy. As the solutions are simply searched under a delay bound, some better options can be overlooked. In the proposed scheme, a soft budget boundary is considered in group relay selection to reduce the overall forwarding hops while retains suitable degree of delivery satisfactory.

1.4 Thesis Organisation

Chapter 2 reviews the literature of SDN, WSNs, and Multi-hop Wireless Communication. With the network flexibility provided by the SDN concept, WSNs can be more adaptive to changing environmental conditions, operational circumstances or user requirements, which is especially important for the monitoring applications in the wild. A further review of movement modelling and exploiting suggests that the state-of-the-art is insufficient to reasonably predict contact caused by directional movement for efficient dissemination. Besides, a brief review of related methods about computer simulations is provided as well.

Chapter 3 investigates the directional movement in path-unconstrained mobility and the quantification of contact opportunities is formulated and evaluated. The directional movement is firstly modelled and investigated on the basis of correlated random walks. A statistical experiment is conducted and the results reveal that predictability of directional movement exists. Given the described contact model, an approach is proposed and shown to be feasible to quantify contact opportunities created by the directional movement.

Chapter 4 investigates the strategic forwarding based on directional movement, and the optimisation of delivery utility is solved and discussed. Specifically, two kinds of dissemination cases are researched. The first case assumes that every delivery is independent and one relay node is selected for each target node. The second case assumes that a relay can be shared by a (sub)group of target nodes. Research issues are addressed in both cases for the scenario where mobile nodes move with directional correlation

Chapter 5 summarises the thesis with potential future work.

Chapter 2

Literature Review

In this chapter, literature concerning SDN, WSNs, multi-hop wireless communication, and computer simulations are reviewed. Finally a summary of the literature is presented.

2.1 Software-Defined Networking (SDN)

SDN is a promising candidate technology for next generation networks [NMN+14]. SDN provides the flexibility by defining network functionalities via software rather than hardware. Although concerns exist about its feasibility and effectiveness [VN11], real-world demonstrations have proven that SDN is a promising solution for future networks. In this section, the SDN technology is introduced with its developments.

2.1.1 SDN Architecture

The essential idea of SDN is decoupling the network control plane from the data forwarding plane. As shown in Figure 2.1(a), each traditional network device is a closed system and takes charge of both routing and data forwarding. For the SDN-enabled network in Figure 2.1(b), the devices are simplified by clear separation between the control and data planes. The SDN controller achieves centralised control and deploys customised routing based on an updated view of network. For the data plane, SDN-enabled routers just perform data forwarding according to the informed routes, which are stored as entries in the Flow Table. In this way, networking can be modified easily and dynamically.

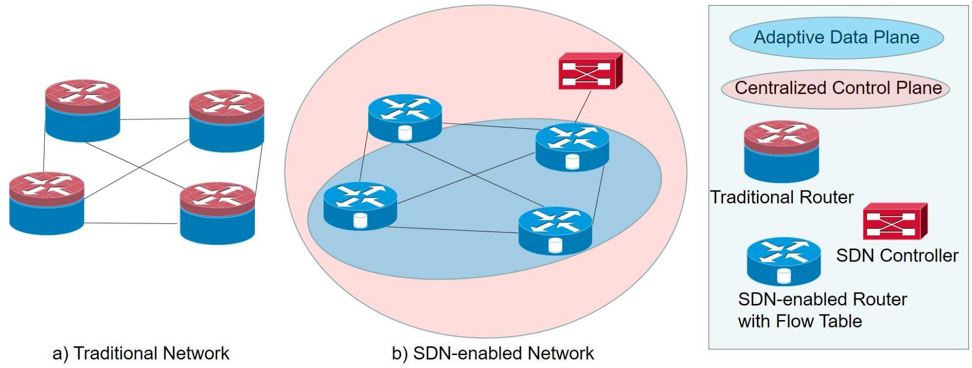


Figure 2.1: Software-Defined Networking Concept

OpenFlow [MAB+08] is the most widely-accepted paradigm of SDN for now. Although its authors initially promoted the deployment of OpenFlow for running experimental protocols or architecture, OpenFlow has become a commercial standard especially in data centres and enterprise networks [NMN+14]. Before the success of OpenFlow, many attempts had been made to create a programmable network [FRZ14]. OpenFlow is mainly based on Ethane [CFP+07] which can be seen as a VLAN technology from a certain perspective [NMN+14]. OpenFlow assumes a wired networking architecture and relies on Ethernet switches. It exploits "the fact that most modern Ethernet switches and routers contain flow-tables" [MAB+08], to achieve a compromise between generality and flexibility.

Under the SDN architecture, the network consists of the controller(s) and OpenFlow-enabled switches. An OpenFlow-enabled switch (Figure 2.2) should include at least three parts: a flow table, a secure channel, and the OpenFlow protocol. Based on the OpenFlow protocol and a secure channel, the remote controller communicates with an OpenFlow-enabled switch to form the control plane. The flow tables operate forwarding physically, whilst entries are controlled by the software application. To achieve software-defined control, the flow table of an OpenFlow-enabled switch is manipulated by the controller(s) via OpenFlow protocol through a secure channel. Unless otherwise stated, the SDN research of this thesis is based on the original OpenFlow standard.

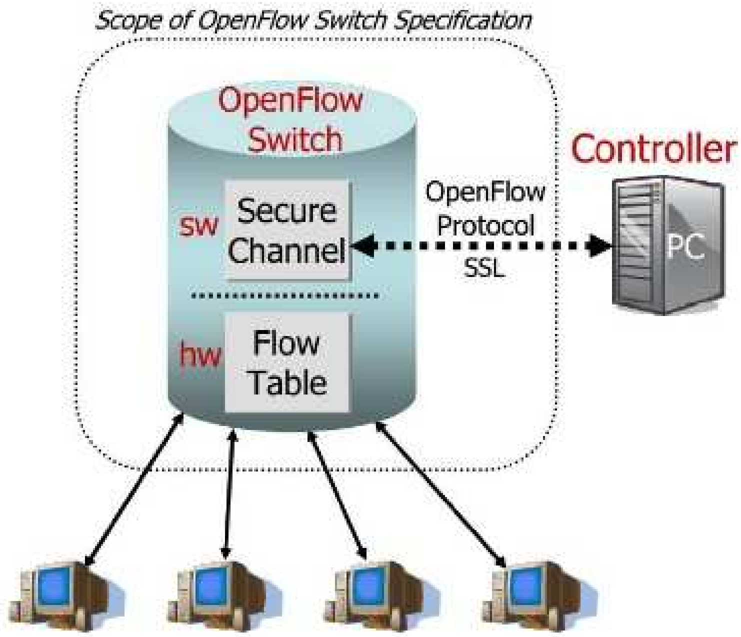


Figure 2.2: OpenFlow-Enabled Switch [MAB+08]

There are already many implementations of the SDN architecture in enterprise networks, such as Google data centres [JKM+13]. Meanwhile, since the proposal of OpenFlow, much research has been conducted to complete the SDN architecture. Many comprehensive surveys, such as [KRV+15] [XWF+15], continue to summarise and analyse the recent works. The research directions include network management [KF13], security [SHNS16], NFV [JP13] [MGT+15] [LC15], service quality [JKK12] [KSKD+12], cloud computing [YS14], network state [PJT+14], controller placement [RRNS14] etc. The difference of physical medium also brings more research issues, which makes applying the SDN concept into optical/wireless networks different [ASBSP14]. For example, an overview of software-defined optical networks is provided by [EA12]. More details about the development of wireless SDN will be provided in Section 2.1.3, as it is research focus of this thesis.

2.1.2 SDN Control Plane

The SDN control plane can be divided into four parts: the controller, the north interface, the south interface, and an optional east/west interface. The controller is a network operation system. The north inferface is for the network applications. The south inferface is for communicaiton between the controller and forwarding devices. The control channel is called "out-of-band" if the south interface is supported by a separate dedicated network; otherwise it is an in-band control as discussed in [LTQ12] and [HAG16]. The east/west interface is for the controller-to-controller communication. It is worth noting that most existing SDN architectures do not have the east/west interface as multiple controllers are usually not required.

There are two control modes for the operation of an SDN controller such as NOX [GKP+08]. A controller passively replies to device requests if it operates in a reactive manner. However, this reactive control mode may bring large first-packet delay due to the passive insertion of rules. By pushing rules in advance, proactive control reduces new flow delay while increases new flow throughput. Although a controller in proactive control mode can avoid the first-packet delay, the unpredictable link failure can also impair the performance [MV15]. Responsiveness, robustness and scalability are still an critical concern for SDN controller [NS15].

Meanwhile, the decentralisation of SDN control plane is also an important research issue. It even has been recognised that the logically-centralised SDN controller must be physically distributed [NMN+14], considering the scalability, reliability and robustness. For logical decentralisation, Flowvisor [SCC+10] is a kind of proxy controller that slices network resource for multiple controllers. Other distributed control planes are meanwhile proposed such as Hyperflow [TG10] and Onix [KCG+10]. Although it is proved that physical decentralisation can reduce the lookup overhead and increase the response speed, how much decentralisation is the most efficient is still an issue [LWH+12]. To this end, Kandoo [HYG12] proposes a hybrid mode to achieve controllable decentralisation.

Meanwhile, physical decentralisation also brings the problem of inconsistent control-plane data. The inconsistency happens when a controller wrongly thinks it has an accurate view but actually it does not, which can affect the correctness of actions to different degrees depending upon the situation. The inconsistency problem is not just for SDN networks, but becomes increasing important with the development of SDN. Many mechanisms and alogrithms have been proposed for correct and efficient updating while there is still issue in this area [FSV16] [FMW16].

In this thesis, the decentralised control framework is further researched for the multi-hop wireless networks. Because of the distributed controllers, the exchange of control information can be infrequent and delay-tolerant, which provides the opportunity to achieve more efficient control-plane communications.

2.1.3 SDN in Wireless Networks

Compared with increasing SDN deployment in data centres and enterprise networks, SDN development in a wireless network settings is significantly less [HAG16]. In terms of the wireless medium and mobile devices, although new challenges are introduced into SDN research, SDN promises to address many inherited problems in wireless communication networks [HAG16].

2.1.3.1 SDN in Infrastructure-Based Wireless Networks

Network infrastructures require specialized network devices to be deployed in advance to support networking such as the base stations in WCNs or wireless routers in Wireless Local Area Networks (WLANs). Wireless networks such as WCNs and WLANs are called infrastructure-based wireless networks because they rely on network infrastructures. In contrast, infrastructure-less wireless networks are networks without dedicated network infrastructure [RMTO11].

Adopting SDN into WCNs can reduce hardware costs while improving network flexibility and manageability. For example, Bell Labs proposes software-defined cellular networks [LMR12] to simplify the design and management of cellular data networks. Another example is the SoftAir [AWL15] architecture for the next generation (5G) wireless systems. The adoption of SDN in WCNs not only enhances core network functionality like traffic engineering, but also empowers novel solutions for particular problems in the radio access network such as seamless roaming [XZD+14].

For WLANs, the centralised control plane of SDN brings more effective coordination among devices. An important trial is the Stanford OpenRoads deployment [YKU+09]. [YYK+11] proposes an architecture that can flexibly define resource requirements as slices. For multiple access points, the concept of a lightweight virtual access point is introduced by [SSZM+12] so that the network can be managed from a global view. Considering gateway congestion in a mesh of routes, [YGH+14] proposes an SDN-based load balancing solution. [DPSBM13] introduces a local controller as a backup control mechanism in case of controller failure.

Besides, SDN naturally integrates different wireless communication technologies, which potentially improves wireless network capacity, coverage and resource utilisation. A systematic approach is proposed in [BMKL12] for supporting and optimising different protocols such as 3G, 4G and WiFi. [YKS+10] [YSK+10] proposes a SDN architecture supporting seamless roaming between Long-Term Evolution (LTE), WiFi, and WiMAX. With later attempts that adds further features such as Quality of Service (QoS) [YLJ+13], the wireless SDN architecture is more complete.

2.1.3.2 Infrastructure-Less Wireless Network

However, whichever base stations are used in WCNs or wireless routers in WLANs, all the aforementioned wireless networks rely on specialised network devices deployed in advance; this kind of wireless network is called an infrastructure-based wireless network.

While infrastructure-based wireless networks provide more stable performance and higher capacity, these networks have their own intrinsic limitations. For example, sometimes more infrastructures like base station have to be deployed just to provide enough signal strength over a fixed area (e.g. some small areas covered by trees/water [LFT+12] which can have extra interferences). In contrast, a more cost-effective way is using self-organized communication like a MANET [Toh01]. With an infrastructure–less network, any device can act as router while has its own tasks. This feature extends network range and improves flexibility but also makes the situation even more complex.

Typical infrastructure-less wireless networks include WSNs, MANETs, and VANETs. Infrastructure-less wireless networks “may form to extend the range of infrastructure-based networks or handle episodic connectivity disruptions”, and thus “enables a variety of new applications such as cloud-based services, vehicular communication, community services, healthcare delivery, emergency response, and environmental monitoring, to name a few” [NMN+14].

SDN brings network flexibility to wired networks and infrastructure-based wireless networks, which improves performance and reduces cost. In terms of infrastructure-less wireless networks, SDN is capable of providing particular merits as well:

- The development cycle and cost of devices can be decreased as a normalised hardware platform which can be adapted to different scenarios by simple software modification.

- The data transmission can be tailored to fulfil application-specific requirements, which will maximise network lifetime while guaranteeing QoS.

- Systems can be extended easily to meet increasing needs. This extensibility is important as the application demands may vary with interests and change over time.

Recently, some initial attempts [KLG+14b] [GAM14] [QDG+14] have shown that the concept of SDN is equally feasible in infrastructure-less wireless networks. However, there is no standard of SDN in the self-organized area and further research is being conducted [NMN+14]. Especially for the generality that most devices cannot guarantee energy and performance, more issues are still to be researched. Therefore, this thesis focuses on the infrastructure-less wireless networks for further SDN research.

2.1.3.3 State-of-the-Art SDN Architectures in Infrastructure-Less Wireless Networks

[LTQ12] suggests a traditional application-specific WSN lacks abstraction and contains too much complexity, resulting in resource underutilisation, counter-productivity, rigidity to policy changes and difficulty to manage. In contrast, a software-defined WSN can be programmable "by manipulating a user-customizable flow table on each sensor via Sensor OpenFlow (SOF)." Thus, the network is versatile, flexible and easy to manage. Although the authors propose SOF as the protocol for software-defined WSN, this research does not provide a specific network architecture and a wireless node design.

[GAM14] proposes a centralised network management for WSN. A base station with control-logic is the key of their architecture. They suggest a centralised controller will reduce energy consumption in a sensor node because the controller has finished the most tasks like topology maintenance and network optimisation. On-the-fly routing decisions will decrease convergence time and thus benefits mobility handling. For network management, they argue that the on-the-fly programming is much more flexible and that "highly-accurate location information could be achieved by employing a centralised localisation algorithm". Without simulation, they describe the benefits when deploying a new monitoring task or scaling up the system with new devices. However, they do not consider the control overhead which increases with node density and speed. Besides, they also do not consider the case when sensor node(s) fail to connect with the base station.

[DPSBM13] proposes using a distributed controller (based on the Optimized Link State Routing (OLSR) protocol) which has two tasks: 1) it routes the control traffic over a multi-hop path between the controller and routers; 2) it works as a backup controller in case of central controller failure. However, as this work targets the Wireless Mesh Network (WMN), heterogeneity and mobility of wireless nodes are not considered.

[KLG$^+$14b] and [KLG14a] propose three operation modes "based on the degree of control of SDN controller". For the distributed and hybrid control modes, an agent will fully or partly delegate a controller as a fallback. Their results demonstrate the effectiveness of a fallback for controller failure under centralised operation mode. However, they do not explain the details of the hybrid control nodes. Although the data channel is multi-hop ad hoc network, the control channel is still based on an LTE base station, which is a one-hop mobile network and relies on infrastructure.

[OCF15] claims they are first that "proposed how to place the controller in cluster architecture of WSN." However, their proposal do not explain how the SDN Cluster Head (CH) is selected, which is an important issue for the cluster architecture. Besides, discussions are not provided about the means of interaction between SDN CHs and how is a domain of sensor nodes managed by the SDN CH. Although their proposal is aimed at a large-scale network that requires clustering mechanisms, the network size is not specified to clarify the necessity of multiple SDN CHs.

In summary, for software-defined infrastructure-less wireless networks, how to efficiently form the control framework remains a key challenge to be addressed. Aiming at long-term network operation based on battery-supplied devices such as wireless sensors, the efficiency of control-plane communications is a crucial issue to be considered. As WSNs usually have to rely on the multi-hop wireless communication, this efficiency issue is crucial and challenging. In this thesis, the efficient multi-hop communications are researched to support software-defined infrastructure-less wireless networks (focused on WSNs). Before providing more details about multi-hop wireless communications, a brief review of WSNs is provided in Section 2.2 to examine the architecture and features. Later sections will discuss existing research works for achieving efficient multi-hop wireless communications.

2.2 Wireless Sensor Networks (WSNs)

This section introduces the concept of WSNs which is a important application type of infrastructure-less wireless networks. As mentioned in Section 2.1, how to apply the concept of SDN in an infrastructure-less context remains a challenge to be addressed. Beginning with the WSN architecture and features, the development and specific design issues regarding the habitat monitoring applications are discussed.

2.2.1 WSN Architecture

A WSN is typically composed of a number of sensor nodes which are deployed to cover a target field. The emergence of WSNs can be dated back to the Cold-War Era [SMZ07], which originally proposed them to serve military missions. Around the late 1990s, lower-cost sensors support a wider range of applications "from environmental sensing to vehicle tracking, from perimeter security to inventory management, and from habitat monitoring to battlefield management" [SMZ07].

The architecture of a WSN can be analysed from two aspects: sensor nodes and the interconnecting network. A sensor node is usually a multifunctional device. Note that the nodes can be heterogeneous which means they may have different sensing, storage, endurance and computation capabilities. For the network, Figure 2.3 gives a general architecture of a WSN. As the figure shows, information are sensed by sensor nodes and then converge to the sink point, where data can be acquired by end user. Note that this WSN architecture naturally supports the SDN concept: the sink can serve as a centralised controller to instruct the packet forwarding performed by sensor nodes.

Meanwhile, the cross-layer management of power/mobility/tasks is necessary for

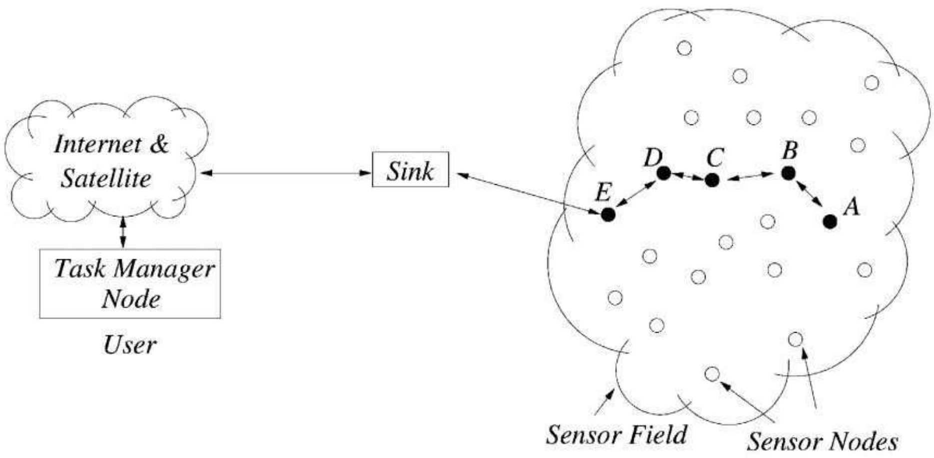


Figure 2.3: WSN Architecture [ASSC02]

WSNs as discussed in [ASSC02]. Figure 2.4 provides such a brief view of the protocol stack used by sensor nodes. It is worth noting that Figure 2.4 is just a conceptual diagram to show the general relationships between five protocol layers and three management planes. Although the five-layered model is similar to a traditional computer network, this is still a rich research area because the management planes should be specified to adapt to different scenarios. With the SDN concept, it can be easier to meet application-specific requirements, as the network control is more flexible.

2.2.2 WSN Features and Applications

Compared with traditional sensors, a network of sensors can work much more effectively and efficiently. The collaborative effort allows sharing of resources between nodes so that duplicated messages can be avoided and lifetime of the whole system can be prolonged [SMZ07]. Ad hoc networking is one important feature of WSNs. Equipped with a self-organised communication capability, the sensors can be deployed rapidly without too much pre-determination. Considering a common fact that neighbour nodes are usually deployed closely, multi-hop communication is another key technique which decreases power consumption and overcomes propagation effects [ASSC02]. In summary, a WSN usually has following features:

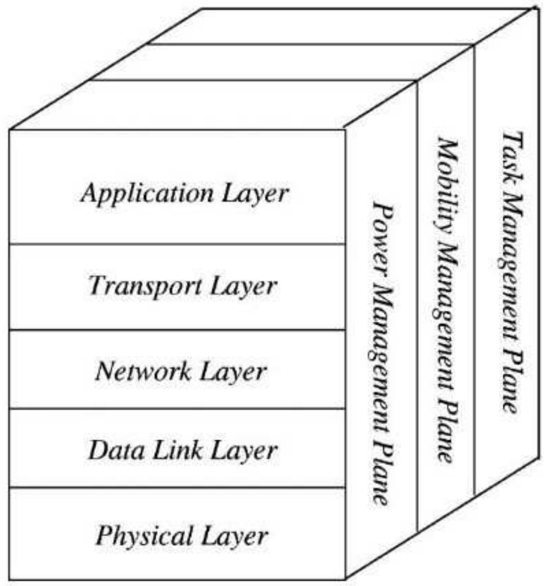


Figure 2.4: WSN Protocol Stack [ASSC02]

- Rapid deployment: Equipped with self-organized communication capacity, the sensor nodes can be deployed rapidly and work without communication infrastructure [ASSC02]. So WSNs can be directly put on field without deploying network infrastructures in advance.

- Lightweight and low-cost: The network infrastructures are usually expensive. Wireless sensors can be much smaller and cheaper compared with a base station. Because of the small size and low cost, the sensors can be scattered randomly over a large area. In this way, the network can be scalable and provide wide coverage with little cost.

- Adaptive: As sensors are usually deployed in a complex environment, they have to be flexible to adapt to the environment including the channel and mobility. A WSN is suitable for communication in inaccessible areas.

In this thesis, the aforementioned WSN features are considered in the system design to reflect more realistic scenarios.

As people pay more attention to our living planet, research like ecology and zoology become more important. Habitat monitoring, one of the well-accepted techniques that support environment-related research, is therefore gaining more interest. In recent years, modern technology have provided more convenience and possibilities for habitat monitoring. Compared with manual data investigation and collection, automatic monitoring is more suitable for long-term and continuous habitat studies. For the study of habitats at very large scales, tools like remote sensing from satellite has proved to be tremendous. But when it gets down to where the complexity is, a ground-based monitoring system is indispensable [CEE+01]. A ground-based monitoring system for habitat monitoring should be easy to deploy, flexible, and durable. Considering these demands, WSNs can be a promising candidate technology.

Figure 2.5: Sensor Node in Habitat [MCP+02]

In the study of [PSM+04], sensors were deployed on a small island off the coast of Maine to collect live data for biologists. [MCP+02] proposes a habitat-monitoring kit that "enables researchers worldwide to engage in non-intrusive and non-disruptive monitoring of sensitive wildlife and habitats". The WSN system in [MS13] collects movement information of grazing animals. Assisted by human observations, "it can be used to deduce animal behaviour (e.g. resting, grazing, foraging, etc.) remotely in real time." [MS13]. Figure 2.5 is a photo of a sensor node deployed in a habitat and Figure 2.6 is a sensor attached to a sheep.

Figure 2.6: Attached Sensor Nodes [MS13]

As suggested in [ASSC02], multiple factors are required to be carefully considered as a whole for WSN design. Firstly, the hardware of sensor has more constraints than ordinary devices. Because of the restricted size, its battery, memory and transmission power all are more limited. Meanwhile, its operating environments are harsher. Especially for networks which are usually deployed outdoors with vegetation coverage or sometimes use unusual transmission media (e.g. water), the condition of communication channel is more complex. What is more, robustness and scalability are also critical issues for a long-term system. Cost of production and maintenance overhead should be taken into account as well. Considering these issues, the challenges of WSN design for habitat monitoring are summarized as follows:

- Limited resources: Firstly, the types of equipment are restricted, for example they have to be attached or implanted on most occasions. Meanwhile, the infrastructures such as cables and base stations are unlikely to cover everywhere. Consequently, communication resources (frequency band, transmission range, power, storage and etc.) are distributed unevenly and can be scarce.

- Complex situations: Although many objects in the environment stay still, animals move with different speeds and patterns, individually or gregariously, periodically or unpredictably. Together with diversified habitats (forest, lake or even underground) and changing climate, it makes network structure and topology even more complex and dynamical.

- Various demands: The demands of observers may vary with interested areas and change over time. The associativity[1] and quality of data should be guaranteed according to real-time requirements. Moreover, the system should be extended easily to meet increasing needs.

By introducing the SDN concept, WSNs can better support habitat monitoring applications as the enhanced network flexibility contributes to long-term system operation. However, the SDN control also brings additional overhead which should be carefully considered due to the limited resources in WSNs. To this end, the efficient control-plane communications in WSNs is further researched in this thesis.

2.2.3 Future WSNs

With WSNs, it is possible to make multiple sensors work collaboratively. However, WSNs still have certain problems. For a specific scenario, the development of a WSN requires multiple steps. Although the sensors can form a network automatically by self-organizing, the initial topology may be non-ideal. In fact, the actually connectivity should be considered to optimise energy efficiency while improving system robustness, such as avoiding network failure. Therefore, it is usually necessary to change the initial deployment.

Meanwhile, the network conditions are prone to vary with dynamic traffic and the environment [ASSC02]. For example, it is better to assign multiple nodes to share peak burden or work as a backup during rain. What is more, additional nodes can be deployed

[1] This associativity means the relevance between multiple data sources (e.g. sound recorded from different sensors for the same event).

anytime but at the cost of network re-organization. For a network with a large number of nodes, continuous extension brings considerable complexity.

Besides, future ecological research requires a more comprehensive perspective and deeper insight into the ecology system [CEE+01]. The development of Complementary Metal-Oxide-Semiconductor (CMOS) cameras and microphones inspires the concept of Wireless Multimedia Sensor Networks (WMSN), where multimedia contents and scalar data can be collected and transited together by heterogeneous sensors [AMC07]. In Wireless Sensor and Actor Networks (WSAN) [AK04], sensors gather information and actors perform actions. WSAN enables more interactive monitoring and introduces more technologies such as robots and Unmanned Aerial Vehicle (UAV) into monitoring scenario.

In sum, the lower-cost, more ubiquitous, and more versatile monitoring network is the trend of the next generation networking. For the future development of WSN, SDN is an indispensable topic to be discussed. However, as summarised in the end of Section 2.1, the efficiency of control-plane communication is an important issue to be considered in the design of software-defined WSNs. Because of the underlying multi-hop wireless communications, this issue is still challenging for WSNs. Meanwhile, communications in the control-plane (especially for a distributed control framework) is relatively infrequent with certain delay tolerance, which brings the opportunity of efficiency improvement. Section 2.3 will discuss the basics of multi-hop wireless communications and the feasibility of improving communication efficiency.

2.3 Multi-Hop Wireless Communications

As discussed in Section 2.1 and 2.2, WSNs commonly rely on multi-hop wireless communications while it is still an open-ended issue to achieve efficient control frameworks under multi-hop wireless communications. To investigate efficient SDN control for WSNs, related works about the multi-hop wireless communications are reviewed in this section.

2.3.1 Multi-Hop Communication Systems

Ad hoc networking is a technology that provides wireless devices a communication capability with devices who are outside the immediate radio range. Such multi-hop communication feature relies on forwarding packets by intermediate wireless devices from the source toward the destination(s). Because these intermediate nodes are resource-constrained wireless devices, compared with traditional computer networks such as Ethernet, more dynamics are introduced by the wireless medium, energy efficiency and node mobility. Over recent decades, abundant research has been conducted to cope with the issues caused by a dynamically changing topology.

By allowing intermittent connectivity in ad hoc networking, a more recent concept of DTN is proposed to support the SCF approach. In DTN, a node can carry the message until it meets destination(s) or more suitable intermediate node(s). Taking advantage of the encounter opportunity caused by node mobility, DTN enables opportunistic forwarding which improves the flexibility in message delivery. However, due to the complexity of node mobility and buffer management, providing the guarantees to meet specific delivery requirements by the opportunistic forwarding is still an open area [CRMS16].

Meanwhile, routing is an important issue in multi-hop communication networks. The multicast routing problem requires finding suitable intermediate node(s) to reach a group of destination nodes. Compared with the unicast routing, which only has a unique destination, the comprehensive consideration of multiple destinations provides the possibility to make full use of network resources. Nevertheless, with multiple destinations, not only is the complexity of routing problem increased, but it introduces additional routing overheads. For a wireless ad hoc network, the efficiency and reliability of multicast routing are even more critical.

2.3.2 Basics of Multi-Hop Forwarding

Flooding and epidemic transfer are the basic forwarding methods in multi-hop communications. The idea of flooding is that every node rebroadcasts the received information to its neighbours and finally the information will spread to the whole network [SMZ07]. To improve the reliability and avoid the broadcast storm problem, many revised flooding schemes have been proposed [NTCS99], including probabilistic, counter-based, distance-based, location-based, and cluster-based. Inspired by epidemic theory, forwarding methods based on node encounters are proposed to make use of the intermittent connectivity, where an encounter occurs when one node moves into another node's transmission range. Although the forwarding based on epidemic transfer does not require additional route information as with flooding, it is necessary that the node is capable to keep and carry the received message and therefore buffering and the tolerance of delay are the main constraints of encounter-based strategies [CS13]. For better performance of epidemic-based methods, studies have been conducted to control message duplication and utilise contact information between nodes [CWSC15].

To improve the forwarding performance, both unicast and multicast routing play important roles. During the routing procedure, the next hop towards the destination(s) is specified within a forwarding table to instruct packet forwarding. In this way, the route(s) can be established between nodes which is more efficient as redundant transmissions are reduced. Based on the means of establishing routes, routing approaches are categorised into proactive table-driven, reactive source-initiated on-demand, and hybrid [Toh01]. Considering efficiency, unicast forwarding will perform worse with the increasing of destination numbers. Thus the multicast schemes are proposed to improve the utilisation in the communication with multiple destinations. Based on the delivery structure, multicast schemes can be classified as flooding-based, tree-based, mesh-based, and group-based. For multicast routing, the group membership management is a critical issue as maintainance overhead may be significant. Instead of requiring multicast forwarding states to be stored in all participating nodes, stateless multicast protocols encode the routing information into the packet header which is suitable for serving a small-scale group of receivers in dynamic networks [JC01]. Based on unicast forwarding, an overlay-based management framework is another option which employs standard unicast routing and forwarding to fulfil the multicast functionality [GM04] [GM07].

By using geographic position information in routing decisions, geographic routing is an alternative principle for forwarding. Instead of using the network address, the geographic routing decision can be made solely based on the location of the current node, the neighbour node(s), and the destination node. Geographic addressing can improve routing scalability and efficiency because neither network topology maintenance nor prior route discovery are required. Besides, in geographic routing, it is possible to give the destination address as a geographic area which is called a geocast. As a specialized form of multicast, the geocast can deliver the message to some or all nodes within a geographic region [Mai04]. However, the performance of geographic routing depends on the position accuracy and their geographic distribution.

2.3.3 State-of-the-Art in Wireless Dissemination

Dissemination is a process of spreading information over a communication network [KS15]. In message dissemination, a single source node delivers messages to one or multiple interested destination nodes [Mor12]. This type of one-to-many communication service enables various applications especially network reprogramming and reconfiguration, which are necessary for many scenarios such as the management of deployed sensor nodes [GZB+16] [FS16].

2.3.3.1 Mobility Awareness

With the development of DTN and VANET, many schemes have been proposed to leverage mobility for message dissemination. These schemes can be divided into relationship-aware and movement-aware.

The relationship-aware schemes rely on the history of encounter information to analyse the sociality of nodes. If certain relationships exist between nodes, their social characteristics can be utilized for dissemination. In [LJHC16], the exponential distribution property is discovered for the contact interval between vehicles and roadside units. Without the detection of communities, the authors of [NdML16] proposed a scheme that is aware of group meetings and their evolution over time. The virtual social connections between vehicles are studied in [YZW16] and the intrinsic sociality of users and their interests are considered in [GLG16] to control and reduce the packet copies. In addition, the geographical information and physical connections are integrated into social analysis in [LNJ+16] and [ZSJ+17] respectively.

Although relationship-aware schemes take advantage of node mobility in statistics, the contribution of specific geographical and movement information to the dissemination strategy is indispensable. Based on the movement directions and velocities between vehicles, a scheme is proposed in [LW16] to make full use of intermittent connections. The authors of [NF16] predict a meeting point based on the known trajectory for efficient and reliable forwarding. In [CYY16] and [KLK16], the planned routes of vehicles are utilised in the delivery decision strategy assisted by roadside units. However, the uncertainty of node movement is overlooked in exisitng movement-based schemes. Although the scheme proposed in [HCCP16] considers the uncertainty of taxis' arriving interval, the investigation of node movement is still far from sufficient. In terms of future dissemination schemes, it is necessary to integrate both relationship and movement information for comprehensive decision making.

2.3.3.2 Design Objective

For the design of a dissemination scheme, it is important to clarify the objective to be achieved. The design objectives can be categorized as requirement-based and utility-based.

An unconstrained requirement usually focuses on certain aspect(s) of performance. For example, the scheme proposed in [HCCP16] aims at delay minimisation while the authors of [PSM17] try to maximise the delivery ratio. In [BSS+16] and [BZ16] the one-hop delay is decreased and the reception rate is increased for the dissemination of emergency messages. When there are specific bounds of requirements, it is possible to balance the performance from multiple aspects. The schemes proposed in [BS16] [CLL+17] [LWL16] all aim at meeting the delay requirements of minimising energy consumption. In [ZLW16] the minimum Signal-to-Interference-plus-Noise Ratio (SINR) threshold is given as the QoS requirement to minimise link cost consumption. The delay budget is considered in [CYY16] to construct a tree with minimised forwarding costs. In [ZWD+16] [ZWD+13], the completion time in a single hop is minimised based on a reliability requirement.

As schemes simply based on requirements are not enough to achieve the optimum use of network resources, various utility-based metrics are proposed to improve system utilisation. In [SKLR16] the utility metric is formulated as the weighted sum of dissemination cost and delay cost. The authors of [XHC17] proposed a utility model which is the weighted sum of the delay utility, video quality utiliy, and storage cost. A Quality of Experience (QoE) model is proposed in [ZCZY16] which is characterised by mean opinion score based on the channel state estimation. In [LQZ16] the utility is modelled as the overall forwarding capability, and the product of utility and distance is used to find pivot point for making the decision of message delivery. In the payoff function of [DMR16], a message-specific value is considered but does not directly reflect the delivery requirement. For various kinds of message, it is necessary, but has not been achieved by existing utility models, to comprehensively consider the heterogeneity of importance and urgency nature in the optimisation of dissemination. Besides, the cooperation of bounded requirements, such as acceptable delivery latency, have not been well researched in existing utility modelling.

2.3.3.3 Forwarding Strategy

Another critical issue in the improvement of dissemination performance is devising an appropriate forwarding strategy. From the focus of decision-making, these forwarding strategies can be divided into relay-oriented and target-oriented.

Relay-oriented strategies focus on control and scheduling of relay nodes for efficient and reliable transmission. In [LK16], a relay node will be probabilistically deactivated if meeting another node having already received the message. The scheme proposed in [BZ16] selects a one-hop helper node to deal with poor channel conditions. Based on residual energy and delay requirements, adaptive adjustments of transmission power [XLH16] [CLL+17] and node duty-cycle [LWL16] [BS16] are proposed to make a tradeoff between dissemination speed and energy consumption. In [ZWD+16] [ZWD+13] the proposed scheme adaptively adopts negotiation and leverages flooding opportunistically to guarantee delivery reliability whilst do not incurring a long dissemination time. The authors of [YNP17] analyse the performance of two relaying modes (one-hop and two-hop) and two relaying methods (first-create-first-relay and last-create-first-relay). Adaptive retransmission and rebroadcasting are studied in [ZLW16] and [wCsL16] respectively and periodic scheduling is considered in [PiAAB16] to address the problem when a device storing data leaves. In [LQZ16] the relay strategy is investigated from the aspect of message placement.

For the target-oriented strategies, the characteristics of message targets are important in making a forwarding strategy. [BSS+16] proposed a mechanism which incorporates a directional broadcast based on the position of message source to seamlessly cover the target area. In [ZYL16] the persistent dissemination of information is provided by carrier sets to an area within a period of time. The data spread time is analysed in [WWH+16] for when the scenario destination region is far from the data origin. A delayed activation mechanism is proposed in [GLG16] for an infection recovery procedure, depending on the number of nodes, destinations, and the time. In [PSM17] and [DMR16], coalitions are formed for service-specific groups to achieve cooperative message sharing. Although optimised group delivery has been investigated in [CYY16] and [KLK16], it is still an open issue as to how to efficiently assign targets to selected relays based on delivery requirements. Besides, supporting mechanisms are needed to identify suitable forwarding methods from various options for a given scenario.

2.3.3.4 Summary

With regard to multi-hop wireless communications, the limitations of existing dissemination schemes can be summarised from the perspective of mobility utilisation, design objective, and forwarding strategy:

- For mobility utilisation, relationship-aware schemes analyse statistical sociality between nodes while movement-aware schemes directly consider the movement vector and/or trajectory in decision making. However, the uncertainty of node movement is overlooked in existing movement-aware schemes and the possibility of integrating relationship and movement information for dissemination have not been investigated.
- For the design objective, requirement-based schemes are aimed at meeting the delivery requirement in certain respect(s) whilst utility-based schemes improve the overall resource utilisation. Although the heterogeneity of messages are considered in some existing solutions, they still lack a comprehensive model which reflects the message-specific nature such as the importance and urgency together with a specific requirement bound for guidance.
- For the forwarding strategy, relay-oriented schemes focus on the control and scheduling at relay nodes while target-oriented schemes take the characteristics of target nodes into consideration. Nevertheless, efficient group delivery for given situation is still an open-issue and dynamical mechanisms are needed to intelligently identify suitable forwarding methods.

In this thesis, the efficient multi-hop wireless communications is researched to apply the SDN concept into WSNs. Considering the aforementioned limitations, this research proposes a dissemination scheme which can improve the (group) delivery efficiency based on tolerable delivery delay and predictable node movement. The proposed scheme can be applied for the SDN control-plane communications (especially for a distributed control framework) which are relatively infrequent and have delay tolerance.

Specifically, considering the use-case example described in Section 1.2, mobile nodes deployed on animals tend to move with more-or-less directional correlation while their movement trajectory cannot be exactly predicted. In this case, existing works either cannot make full use of node mobility or suffer from the uncertainty of node movement. Further studies are therefore required to utilise such partially predictable mobility for more efficient multi-hop communications. To investigate the feasibility of exploiting directional movement in path-unconstrained mobility, Section 2.4 reviews related works about movement and contact modelling.

2.4 Movement Modelling and Exploiting

In Section 2.3, it is shown that the existing studies are insufficient to make full use of node movement for efficient dissemination. To this end, a systematic investigation of movement modelling is conducted in this section. A brief review of mobility research and the random walk model is provided as well.

2.4.1 Overview

Efforts have been made over decades to explore more realistic representations of node mobility [BB15]. As a general basis for mobility modelling, the whole observation period of a given mobile node is usually divided into many epochs, where the movement speed and direction are treated as constants during each epoch. With epoch-based motion models the movement process can be reproduced by time-location pairs recorded in mobility traces. Apart from using collected real-life traces, mathematical models are developed to generate synthetic traces. Typically, for simple random mobility models (such as the random waypoint, random walk, and random direction), a node's movement direction is governed by isotropic randomness, i.e. in each epoch every possible direction may be chosen with equal probability. However, the independence of epochal direction does not accord with the motion correlation commonly exhibited by mobile nodes [BB15]. Unlike completely unpredictable movement imitated by memoryless random mobility, real nodes move smoothly in most cases and very sharp turns rarely occur. This reveals that the node motion is correlated over time. To reflect the temporal dependency in movement, a gradual change in node velocity is considered in some mobility models. For classic mobility models with temporal dependency, the degree of correlation is controlled by a memory level parameter [LH03] or a maximum allowed change per epoch [Haa97], which simulates the physical/logical laws of motion. Although further improvements are accomplished by introducing acceleration/deceleration phases [Bet01] and pause/smooth phases [ZW06], the pattern of directional correlation is as yet undeveloped, particularly in regard to the random turning angle following a non-uniform distribution.

To understand the influence of mobility on information dissemination, theoretical modelling of contact opportunities is essential. In a stochastic analysis of mobility traces, contact-related properties can be measured from a temporal dimension or a spatial dimension. Temporal metrics mainly capture the regularity of movement, for example, the habitual routine of mobile nodes. Given that opportunistic communication relies highly on the sporadic contacts between mobile nodes, the contact interval of a node pair (called the inter-contact time) has been recognised as a dominant factor influencing the performance of opportunistic networking [BB15]. Owing to the popularity of research on social networks and device-to-device communications, distribution laws of the inter-contact time have been identified for many scenarios. For instance, the invariant properties of the contact interval are discovered between vehicles and roadside units in [LJHC16]. Different from the logical connectivity relationship expressed by temporal metrics, spatial metrics focus on depicting the movement trajectory features within the physical space. A basic spatial metric is the displacement distance over a period of time, which characterises the rate of node relocating (called diffusive behaviour). Super-diffusive behaviour is investigated in [KLE10], and a conclusion is drawn that the degree of diffusiveness is proportional to the epochal displacement distance, where the scale parameter varies with the variance of the step length. In another research work concerning spatial metrics, the representative movement area is studied as the radius of gyration [CCLS11]. In addition to the lack of investigations into non-isotropic motion patterns, it is worth noting that there have been only limited studies regarding spatial metrics despite their significance in mobility management and location awareness.

Evidently, it is necessary to devise appropriate strategies that leverage the node mobility in packet forwarding. Given that mobile devices are usually supplied by battery power, energy efficiency is an important issue for long-term operation. In [WBS14], the authors propose a pre-forwarding scheme to prolong the network lifetime by reducing the times of packet transmission. Based on a knowledge of the movement trajectory, packets are pre-forwarded to the expected meeting point for later last-hop delivery, which can be interpreted as an 'interception'. In this way, the shortest-possible forwarding route can be adopted and therefore the energy consumption can be lower[2] than with an immediate forwarding scheme [CCC+10]. However, the interception strategy may lead to quite a long waiting period (that could be minutes to hours). In addition, complex optimisation problems arise when considering the delivery requirements and resource utilisation. In terms of the group delivery with delay constraints, an algorithm is proposed in [CYY16] for solving the efficient multicast tree construction problem. It is noticed that individual forwarding routes can be merged if multiple targets can all arrive a rendezvous point before the delivery deadline. By exploiting the delay budget, fewer tree branches are required to reach the whole group of destinations. However, as with [WBS14] and other

[2]The interception is adopted only if less times of packet transmission can be expected, so the energy consumption is considered at least no higher.

similar works such as [KLK16], the uncertainty of node movement is overlooked [CYY16], which potentially limits the practical achievable performance and feasible application scenarios.

In this thesis, one of the research tasks is the exploitation of node mobility for packet delivery. Among the state-of-the-art forwarding schemes, six types of mobility description for contact prediction are classified from the literature (as summarised in Table 2-A). For these existing contact prediction schemes, the underlying movement estimation models are first discussed. Next, related studies on mobility modelling are reviewed, particularly in regard to the description and estimation of directional movement. Finally a short review about mobility research and random walks is provided as well.

2.4.2 Modelling of Predictable Contact

Table 2-A classifies six types of mobility description from the literature for contact prediction, which are to be discussed respectively.

Table 2-A: Classification of Mobility Description for Contact Prediction with Underlying Movement Estimation

Type	Mobility Description	Movement Estimation	Contact Prediction	Example Work(s)
I	Complete travel plan	Defined waypoint model	Contactable locations (and the waiting time before contact)	[CYY16][KLK16]
II	Explicit return period	Cyclic visit model	Contact interval (with/without uncertainty)	[LW09][SGY17]
III	Historical traces collection	Stochastic arrival model	Distribution of inter-contact time (and contact duration)	[HCCP16][LJHC16]
IV	Value of estimated speed	Circular range model	Contactable area given delay	[CSW$^+$15][PRR16]
V	Estimated speed with direction	Linear trajectory model	Expected contact location given delay	[WBS14][NF16]
VI	The diffusion exponent	Diffusive behaviour model	Contact likelihood given relative location	[SK11]

For the ease of contact prediction, sufficient knowledge of node movement is usually assumed according to the first three types of mobility description listed in Table 2-A. For the Type-I class, a complete travel plan is assumed to be known in advance and therefore future contactable locations can be directly predicted. Knowledge of predefined waypoints usually relies on access to the navigation system, for example to obtain a report of route information from vehicles, as assumed in [CYY16] and [KLK16]. For the Type-II class, although the knowledge of whole movement path is not required, strict periodicity is assumed in the contact interval (such as in [LW09] and [SGY17]). For the Type-III class, the node mobility assumptions are further relaxed as a stochastic process. For instance, the invariant properties of the inter-contact time are discovered between vehicles and roadside units in [HCCP16] and [LJHC16]. However, stochastic properties can vary with cases and their analysis requires traces recorded over a long term and/or of a large quantity.

When the mobility knowledge is limited, the estimation of node movement plays an important role in contact prediction. For the last three types listed in Table 2-A, the latest location of the mobile node is taken as the reference point and the contact is predicted by estimating the node movement. For the Type-IV class, a maximum (or average) movement speed is estimated for the mobile node, so a circular movement range can be drawn to predict contact (as in [CSW+15] and [PRR16]), where the radius depends on the delay since the location report. For the Type-V class, with additional knowledge on movement direction, the location of a future contact can be expected given the delay (such as in [WBS14] and [NF16]). However, the speed and direction are not enough to differentiate heterogeneous mobility, which limits the contact prediction based on the Type-IV/V mobility description. For the Type-VI class, a diffusion constant is introduced by [SK11] based on the exponent parameter of the step-length distribution in a Lévy walk model. Compared with the movement speed, the diffusion constant is a better characterisation of diffusive behaviour, as the randomness of step length is taken into consideration. Nevertheless, to the best of our knowledge there remains no prior work that considers differentiation of directional movement for contact prediction.

2.4.3 Modelling of Directional Movement

Unlike the completely unpredictable movement imitated by memoryless random mobility, mobile nodes tend to move smoothly in many cases and sharp turns rarely occur [BB15]. To reflect this directional movement behaviour, correlated motion patterns are researched in the literature [BCC+17]. The existing pattens for directional correlation can be roughly divided into two branches: velocity-based or step-based. For the first branch, correlation patterns focus on the gradual change of velocity which leads to persistence in the movement direction. In [LH03], the degree of temporal correlation is governed by a memory level parameter, but difficulty arises from controllable correlation of movement direction because the velocity component in each spatial dimension is independently modelled as a Gauss-Markov process. In [ZW06], despite the modelling of acceleration, the probability in choosing movement direction is basically isotropic because the authors purely focus on dividing changes into several small increments. In contrast, the correlation patterns in the second branch are based the characterisation of a step direction. In [BCF07] and [FZZ07], the directional correlation between random steps is investigated but their research is restricted to an n-dimensional lattice or grid. In [ZLZY16], the directional preference of a random step is described by a probability distribution model, but an additional assumption is required for fixing a bias of movement direction. In [Haa97], the angle of maximum turn is introduced to bound the Random Turning Angle (RTA), which results in a directional correlation between steps. However, the influence of RTA in movement is not analysed by [Haa97] or any later reference, as their focus is on the network simulation rather than movement analysis.

From biophysics-related fields, further studies on the RTA are found concerning the Correlated Random Walk (CRW) model [CPB08]. In [WLSN00], circular distribution functions are employed as the model of RTA, which is theoretically rigorous but introduces additionally complexity. In [Bye01], both the angle of maximum turn and the Standard Deviation of the Turning Angles (SDTA) are considered to characterise the probability distribution of RTA. In [NBW09], the probability distribution model of RTA is linked to the diffusion degree to analyse the movement of animals. For the aforementioned research on CRW, distance-based metrics, such as the mean square displacement, have been commonly analysed and formulated. However, the calculative method to estimate the displacement angle based on the distribution characteristics of RTA cannot be found in literature.

2.4.4 Mobility Research and Random Walks

Mobility is a significant topic for analysing natural phenomenon. The nature of mobility is mainly about the motion of particles, organisms or other mobile objects. However, the achievements in mobility research are widely applicable to various research fields. For example, a famous study of mobility is known as Brownian motion [RY13] which is based on the observation of pollen particles suspended in a fluid. Interestingly, this kind of random movement was later found to be commonly feasible for quantitative analysis of the fluctuating prices in the financial market [KSKS98]. With regard to electronic engineering and computer science, further understanding of mobility inspires better development of artificial systems and intelligent algorithms [SW14].

Due to the popularity of wireless devices, node mobility in communication networks has gained increasing research interest. Decades ago, the age of mobile communications began with the emergence of cellular networks. As a major functionality in cellular networks, mobility management guarantees communication services by tracking mobile terminals. Research on terminals' mobility has been conducted since then to realise efficient and reliable mobility management [MZSS16]. With the rapid development of ad hoc networking in recent years, considerable attention has turned to opportunistic communications [BB15]. Considering the greater challenges brought by diverse application scenarios, more advanced intelligence has been developed to utilise smart node mobility [LMO16].

Apart from collecting real-life traces, synthetic models have been developed to facilitate mobility-related studies [BB15]. Synthetic mobility models are not only beneficial for simulation purposes but also necessary for methodical prediction. Originating from the study of Brownian motion, the random walk model [CPB08] provides a generalised methodology to describe mobility. Under the description of a random walk model, walkers take random steps within defined space. Based on different space definitions and step randomness, different classes of random walks have been proposed.

For simple random walks, the walking space is usually defined as n-dimensional regular lattices so the choices of step direction are finite. With the same probability of choosing each possible direction, the lattice-based random walks are uncorrelated and unbiased. In this simplest context, it is straightforward to derive the diffusion equation even in two (or more) dimensions [OL13]. A lattice-based random walk becomes correlated if unequal probability exists among the step directions. This directional correlation in a lattice-based random walk has been investigated by [BCF07].

More realistic random walks are allowed for infinite choices of step direction so the whole two- or three-dimensional[3] Euclidean space are literally reachable. When the isotropic randomness is assumed for the step direction, the movement behaviour of random walks depends on the distribution of step length. In [KLE10], the effect of variable step lengths on diffusive behaviour is investigated, where Lévy walks is a famous model with heavy tailed step-length distribution. To further investigate the correlation between successive step orientations, random turning angles are introduced for correlated random walks. In [Bye01], the probability distribution of random turning angles is discussed to reflect the directional correlation.

The diffusion process of random walks has been mathematically analysed in the literature. However, as suggested by [CPB08], it is a non-trivial problem to derive the analytical solution for correlated random walks in two-dimensional Euclidean space. The modelling of movement paths for correlated random walks remains an unsolved issue, despite the existing analysis concerning mean squared distance [Ben06], mean dispersal

[3] Three-dimensional random walks are beyond the scope of this thesis but can be similarly researched.

distance [Bye01], and tortuosity [PBD13, LXJ15].

Linking to the use-case example described in Section 1.2, a correlated random walk is usually suitable to represent animals' mobility [CPB08], and therefore proper modelling of its movement path can be used to predict the contact opportunity caused by the movement of mobile nodes (which are deployed on animals). For this unsolved problem of movement path modelling for correlated random walks, Chapter 3 provides a feasible solution which is applied and evaluated in Chapter 4 for strategic forwarding.

2.5 Computer Simulations

This section provides a brief review related to the computer simulations used in Chapter 3 and 4. In this thesis, simulations are implemented in MATLAB (version R2017b) which is a numerical computing environment for engineers and scientists. Matrix computations [GVL12] are used to simulate scenarios and solve mathematical problems. Specifically, there are two topics to be discussed in this section: numerical analysis and computational optimisation.

The field of numerical analysis provides convenient methods that obtain solutions for mathematical problems [Hil87] [SB13]. In Chapter 3, two methods of numerical analysis are used for the modelling of directional movement. The first one is the Monte Carlo method [KBTB14] which was proposed as an experimental approach to obtain numerical results by repeated random sampling. The Monte Carlo method is used in Section 3.2.3.1 to overcome the difficulty in mathematical analysis of correlated random walks. The second one is the least-squares method [Wol05] which is a common approach to find the functional relationships between variables, by minimising the sum of squared residuals (difference with data samples). In Section 3.2.3.2, the results obtained from the Monte Carlo experiments are further analysed and formulated for movement modelling, based on the standard implementation of least-squares method provided by the MATLAB Curve Fitting Toolbox.

Computational optimisation techniques are developed to solve an optimisation problem which requires finding the best solution from a number of feasible solutions [KY11]. In Chapter 4, optimisation algorithms are designed to solve the group relay selection problem formulated in Section 4.3.2. As proved in [LJHC16] and [CYY16], the group relay selection problem is at least NP-hard (i.e. non-deterministic polynomial-time hardness) and it is infeasible to search for an optimal solution [PS82]. Heuristic algorithms are therefore required to find an approximation to the optimum in a reasonably short time [CLRS09].

The evolutionary algorithm and the greedy algorithm are two representative types of heuristic algorithm. The principle of a greedy algorithm is choosing the best solution for each sub-problem until the whole problem can be solved [CLRS09]. Therefore, it is a straightforward approach which is fast and simple especially when solving complex problems. Despite its simplicity, the performance of a greedy algorithm should be carefully evaluated because greedy searching may not lead to global optimisation but can instead converge to a local optimum. In previous works, such as in [LJHC16] [KLK16] and [CYY16], although the greedy algorithm is commonly considered to solve the group relay selection problem, none of these existing studies have evaluated the performance of their greedy algorithms in regard to multi-objective optimisation.

Unlike greedy algorithms, an evolutionary algorithm randomly searches for the global optimum based on genetic operators (e.g. crossover and mutation) [Bac96]. Although evolutionary algorithms usually require more iterative computation, which can be inefficient, they can better solve multi-objective optimisation problems as evolutionary algorithms can produce a number of potential solutions for making further decisions [CLVV+07]. When these solutions are Pareto optimal (i.e. an improvement in any objective requires a degradation in other objective(s)), a Pareto front (i.e. the set of all Pareto optimal solutions) can be approximated to analyse the trade-off between objectives [Cen77].

In this thesis, the Multi-Objective Genetic Algorithm (MOGA) [DK01] is adopted because it is a standard evolutionary algorithm which is suitable for solving combinatorial optimisation problems (as formulated in Section 4.3.2). Based on the standard MOGA implementation provided by the MATLAB Global Optimization Toolbox, an evolutionary approach is proposed in Section 4.3.2.2 to solve the formulated problem. Additionally, referring the algorithm given by [CYY16], two variants of the greedy algorithm are proposed in Section 4.3.2.3 as the greedy approach for problem solving. Finally, the performance of the greedy approach and evolutionary approach are compared.

2.6 Summary

With the network flexibility provided by the SDN concept, WSNs can be more adaptive to changing environmental conditions, operational circumstances or user requirements, which is especially important for the monitoring applications in the wild. To alleviate the negative influence of centralised SDN control in infrastructure-less wireless networks, physically distributed control frameworks have been recognised as a promising solution. However, efficient control-plane communication in WSN remains a key challenge to be addressed, concerning limited radio range, energy consumption, and node mobility.

A further review of movement modelling and exploiting suggests that the state-of-the-art is insufficient to reasonably predict contact caused by directional movement for efficient dissemination. Referring to the use-case example described in Section 1.2, animals' directional movement can be utilised to reduce forwarding hops given a delay budget, which contributes to higher energy efficiency for long-term network operation. By introducing directional correlation as an index of node mobility, an approach is proposed in this thesis for the modelling of predictable contact to facilitate strategic forwarding.

Chapter 3

Quantification of Contact Opportunity for Strategic Forwarding

In this chapter, the contact opportunities created by directional movement is analysed and modelled to support strategic forwarding. This directional movement is supposed to be exhibited by mobile nodes (e.g. resulting from the animals' behaviour in habitat monitoring applications as described in Section 1.2). The utilisation of contact opportunities will be investigated to facilitate energy-efficient dissemination of control packets.

3.1 Problem Overview

Multi-hop wireless communication [Toh01] has gained increasing interest taking into account its capability for power-saving, traffic offloading, disruption handling, and extending coverage. Besides classical examples such as MANETs, VANETs, and WSNs, more advanced proposals have discussed topics including operation in diverse environments (underwater [APM05] and underground [ASV09]), robotic actuators [AK04], and multimedia traffic [AMC07]. However, packet delivery to mobile destination(s) remains a challenging problem within these multi-hop wireless communication networks, especially considering efficiency, reliability, and scalability [WCZL16].

Instead of relying on end-to-end connectivity, the emergence of Delay/Disruption Tolerant Networks (DTN) facilitates Store-Carry-Forward (SCF) [CS13] communication. By taking advantage of the encounter opportunities created by node mobility, the SCF strategy offers a rich research area and provides more possibilities in multi-hop wireless packet forwarding. As a feasible delivery solution inspired by SCF, packets can be forwarded to the selected relay node(s) in advance and wait for later delivery when the destination node is directly contactable. For the control packet dissemination in SDMSN, the destination node is called a control message target node.

By intelligently selecting the relay node, optimisation of the forwarding route can be realised, which is beneficial for prolonging network lifetime [NF16] and/or improving resource utilisation [XHC17]. Because an end-to-end route may not be established directly, this relay-based delivery solution is different from conventional routing. In contrast with the DTN research, the existing network topology can be utilised rather than being fully dependent on encounter-based connectivity.

The focus of this research is a typical scenario where delay-tolerant information is disseminated from a special node which serves as the central controller or distribution centre. Under such kinds of scenarios, the information source is considered able to acquire global knowledge of the underlying network for assisting with relay selection. In the SDMSN scenario, this dissemination means downstream communication that delivers the control information (commands and updates) from the central controller (sink node) to control message target nodes (sensor nodes).

To avoid reliance on maintaining globally topological dynamics at the information source, relay candidates are limited to stationary nodes for the last-hop direct delivery. These stationary nodes form a stable topology framework, for example the sensor nodes fixed to plants or installed in the environment[LFT+12]. Besides, we further narrow down our research to the case when the only mobile nodes are control message target nodes, for example sensors attached to or implanted within animals [LFT+12], to make the problem more tractable.

In the literature, information from the navigation system of mobile destinations has been utilised to identify candidate relay(s) from a set of stationary nodes. Given a predefined movement trajectory of the destination node, one or several stationary nodes can be selected for the last-hop direct delivery when the mobile destination node passes by. Unfortunately, it is unrealistic to assume accurate information of travel plans is always available in every scenario, for instance when animals are involved. The lack of mobility knowledge can lead to performance degradation in either traditional routing or DTN routing. Meanwhile, nodes usually move with a certain purpose (such as searching or migrating [YLK07]) rather than completely randomly, and therefore further studies are required to utilise partially predictable mobility based on more generalised navigational information such as the movement direction.

In this research, the feasible exploitation of directional movement in path-unconstrained[1] mobility is investigated for strategic forwarding. The research premise is straightforward: a degree of persistence in the future direction of travel is expected on the basis of the current movement direction of a mobile node. As a consequence, there is a fair chance that a stationary node lying along this path will encounter the mobile node. Given an adequate budget of delivery delay and enough confidence of movement persistence (i.e. the tendency to keep moving in a certain direction), the risk associated with opportunistic delivery can probably be balanced by the desire to reduce the number of forwarding hops. To the best of our knowledge, no existing research has provided such a quantitative analysis for strategic packet forwarding based on the comprehensive consideration of node mobility, network topology, and delivery requirements.

[1] Mobile individuals independently move in a relatively unobstructed environment (such as a forest or ocean).

In this chapter, analysis is conducted on the modelling of directional movement and corresponding contact opportunities, to fit in the habitat monitoring scenario that mobile nodes (animals) usually move with a certain purpose (such as searching or migrating). As the result, an approach is proposed (and validated) to approximately describe the contact certainty, based on the directional correlation of node movement. This proposed approach is further applied in Chapter 4 for strategic forwarding.

3.2 Analysis of Directional Movement

Before analysing the contact opportunity, it is necessary to formulate and investigate the directional movement. The mobility model is specified first in this section, followed by a discussion of correlating the movement direction. Finally experimental simulations are conducted to discover statistical characteristics of the directional movement.

3.2.1 Mobility Model

Devising a suitable mobility model is the first step in creating a successful dynamic delivery mechanism. We consider a mobile target moves within a two-dimensional space $\mathbb{R}^2$.

The fluctuations of movement speed are assumed to be negligible compared with the mean value, so that the average speed (denoted by V) is considered to be a constant.

The movement direction (denoted by $\theta \in [-\pi, \pi]$) is assumed to be fixed during a step which is a constant period defined as

$$T = t_{i+1} - t_i \tag{3.1}$$

where $i \in \mathbb{N}$ denotes the discrete step number corresponding to the time point t_i.

Between two successive steps, the movement direction is updated by an angle of random turning

$$\Theta_i = \theta_i - \theta_{i-1} \tag{3.2}$$

where θ_i denotes the movement direction during time step i, and $\Theta_i \in [-\pi, \pi]$ denotes the random turning angle between step $i-1$ and step i.

For the movement made during time step i, Λ_i denotes the step movement distance and θ_i denotes the step movement direction. Based on equations (3.1) and (3.2), the step movement distance and direction can be derived as

$$\Lambda_i = V \cdot T \tag{3.3}$$

$$\theta_i = \theta_{i-1} + \Theta_i = \theta_0 + \sum_{\zeta=1}^{i} \Theta_\zeta \tag{3.4}$$

where θ_0 denotes a given initial movement direction.

Consequently, the real-time movement of a mobile target can be specified given the initial direction and the probability distribution of the random turning angles. Under such a mobility model, the persistence of the initial movement direction progressively diminishes over successive time steps, which will be further discussed in the Section 3.2.2.

3.2.2 Directional Correlation

Based on the mobility model, the predictable contact with a mobile target can be utilised to select relay(s) from a set of stationary nodes. However, because the movement direction changes randomly in a correlated pattern, it is not easy to obtain a simple expression for the contact potential. To this end, further analysis is required on the stochastic process of movement.

As assumed in equations (3.1) and (3.2), the uncertainty of movement results from the Random Turning Angle (RTA) between two successive steps. In Figure 3.1, a comparison is shown between three types of distribution for RTA. Traditionally, it is assumed that the current movement direction may change to any other direction with equal probability. This uncorrelated change can be represented by an isotropic uniform distribution. By restricting the maximum allowed direction change, the bounded uniform distribution [Haa97] leads to correlation in the step direction. However, this bounded uniform distribution fails to reflect the natural phenomenon that smaller deviations tend to have higher frequency of occurrence. Therefore, in this research, the truncated Normal distribution is proposed as the probability model for RTA.

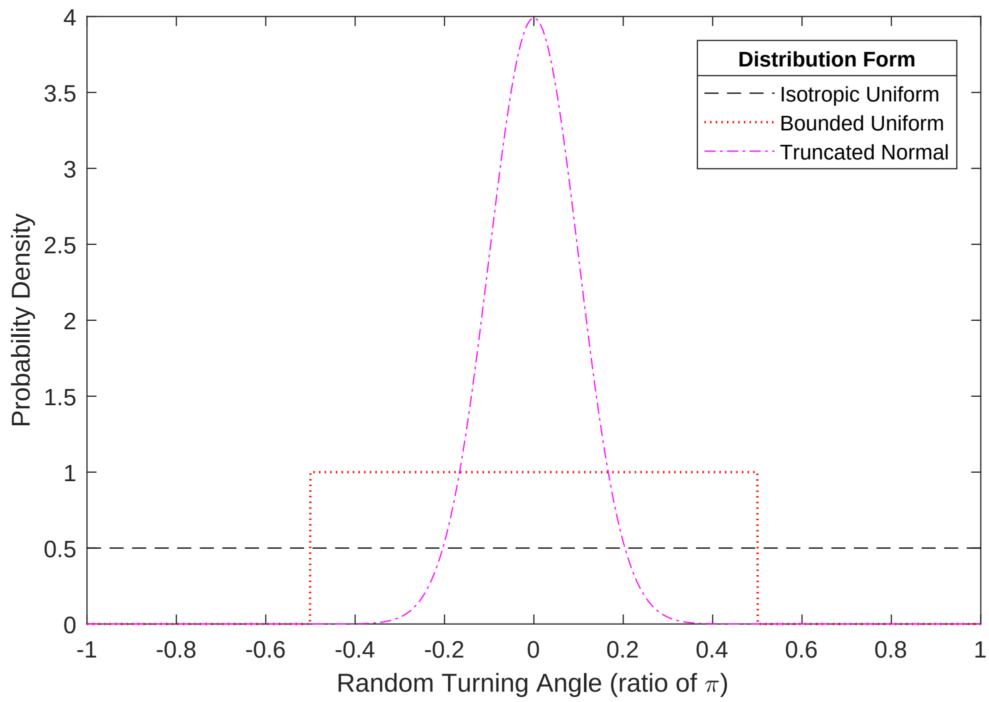


Figure 3.1: A Comparison between Three Types of Probability Distribution

The probability density function of general normal distribution is defined as

$$\phi(\mu, \sigma^2; x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3.5}$$

where $\sigma > 0$ denotes the standard deviation and μ denotes the expectation of the distribution.

Accordingly, the cumulative distribution function of general normal distribution is defined as

$$\Phi(\mu, \sigma^2; x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\delta-\mu)^2}{2\sigma^2}} d\delta. \tag{3.6}$$

Based on the zero-centred Normal distribution with symmetric truncation, the probability density of the RTA (denoted by f_Θ) is given as

$$f_\Theta = \frac{\phi(0, \sigma^2; \Theta)}{2\Phi(0, \sigma^2; \epsilon) - 1} \tag{3.7}$$

where $\epsilon \in [0, \pi]$ denotes the truncation bound, and Θ denotes the random turning angle as defined in equation (3.2).

Then the probability of a turning angle within a given interval can be expressed as

$$\begin{aligned} Pr\left[\varrho_1 \leqslant \Theta \leqslant \varrho_2\right] &= \int_{\varrho_1}^{\varrho_2} \frac{\phi(0, \sigma^2; \delta)}{2\Phi(0, \sigma^2; \epsilon) - 1} d\delta \\ &= \frac{\int_{\varrho_1}^{\varrho_2} \phi(0, \sigma^2; \delta) d\delta}{2\Phi(0, \sigma^2; \epsilon) - 1} \\ &= \frac{\Phi(0, \sigma^2; \varrho_2) - \Phi(0, \sigma^2; \varrho_1)}{2\Phi(0, \sigma^2; \epsilon) - 1} \end{aligned} \tag{3.8}$$

where $-\pi \leqslant \varrho_1 < \varrho_2 \leqslant \pi$ denote the two ends of the interval.

Due to the truncation, the probability is zero outside the bounded range, which can be proved as follows

$$\begin{aligned} Pr\left[\epsilon < |\Theta| \leqslant \pi\right] &= 1 - Pr\left[-\epsilon \leqslant \Theta \leqslant \epsilon\right] \\ &= 1 - \frac{\Phi(0, \sigma^2; \epsilon) - \Phi(0, \sigma^2; -\epsilon)}{2\Phi(0, \sigma^2; \epsilon) - 1} \\ &= 1 - \frac{\Phi(0, \sigma^2; \epsilon) - [1 - \Phi(0, \sigma^2; \epsilon)]}{2\Phi(0, \sigma^2; \epsilon) - 1} \\ &= 0 \end{aligned} \tag{3.9}$$

Meanwhile, the effect of truncation depends on the relative relationship between σ and ϵ. As shown in Figure 3.2, the possible turning angle is represented as a random variable having a continuous distribution between truncation bounds. Given a certain truncation bound ϵ, the shape of the distribution curve varies with the value of σ. With an increasing σ value, the density difference within the bounded range becomes smaller. For the special case that σ is extremely large, a uniform distribution is approximated.

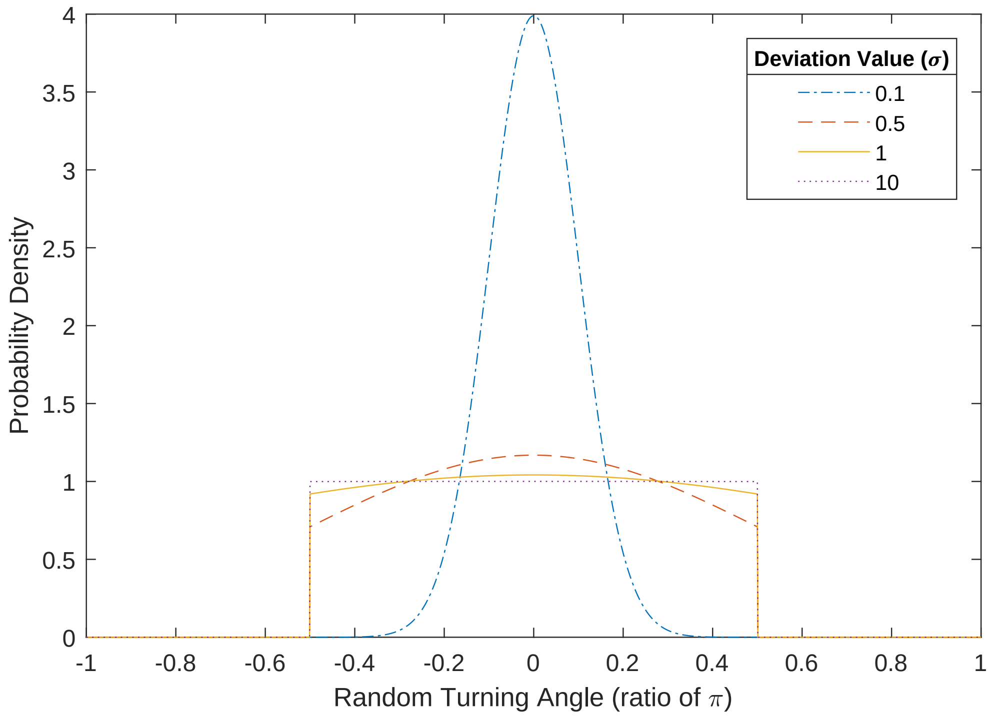


Figure 3.2: Effect of Deviation Value on the Truncated Normal Distribution

When the value of σ is sufficiently small compared with the truncation bound ϵ, the truncated Normal distribution is very close to a Normal distribution without truncation. To measure the relationship between σ and ϵ, we define a ratio as

$$\varsigma = \frac{\epsilon}{\sigma}. \tag{3.10}$$

Then equation (3.7) can be written as

$$f_\Theta = \frac{\phi(0, \sigma^2; \Theta)}{2\Phi(0, \sigma^2; \epsilon) - 1} \approx \phi(0, \sigma^2; \Theta) \tag{3.11}$$

where ς_{th} is a threshold value that satisfies $\Phi(\varsigma) \approx 1$ and typically $\varsigma_{\text{th}} \geqslant 4$.

With the condition $\varsigma \geqslant \varsigma_{\text{th}}$, the truncated Normal distribution can be simplified such that it is only affected by the single parameter σ which stands for the Standard Deviation of the Turning Angles (SDTA) in this research. In other words, the Normal distribution is considered naturally bounded within a certain range. To apply this single-parameter description, the maximum allowed value of σ is given as

$$\sigma_{\text{max}} = \frac{\epsilon_{\text{max}}}{\varsigma_{\text{min}}} = \frac{\pi}{\varsigma_{\text{th}}}. \tag{3.12}$$

3.2.3 Movement Investigation (with Validation)

The movement investigation is divided into two parts: the first part illustrates the experimental design and implementation; the results obtained from the experiment are discussed in the second part.

3.2.3.1 Predictability Investigation

Referring to the correlated random walk model given by [Bye01], the Normal distribution is assumed for random turning angles with a fixed step length (set to 30 metres in this investigation based on 30 seconds step period with 1 m/s movement speed). Given a length of simulation time (set to 3600 seconds in this investigation), a sample trace results from a succession of random steps while different traces are repeatedly sampled under same standard deviation of turning angles (set to 10 degrees as an example). As shown in Figure 3.3, sample traces are synthetically generated to investigate the movement path of correlated random walks. This figure shows that, although all the traces have same initial direction (towards the positive x-axis), the existence of directional randomness results in increasing deviations to the left or right (towards the positive or negative y-axis). Because the random walk is a stochastic process, an accurate path cannot be known before the movement occurs. However, possibility is shown by Figure 3.3 to statistically describe the movement path based on the sample traces, which will be further discussed.

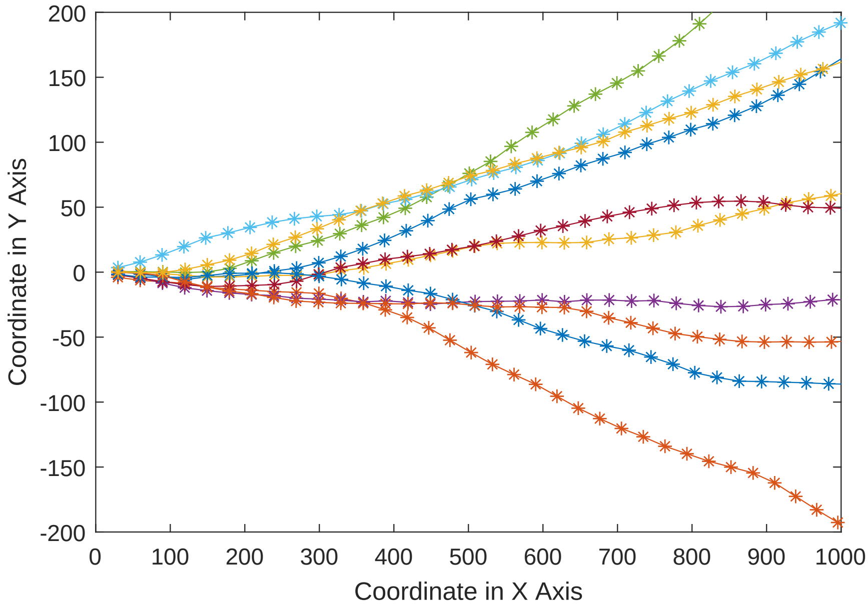
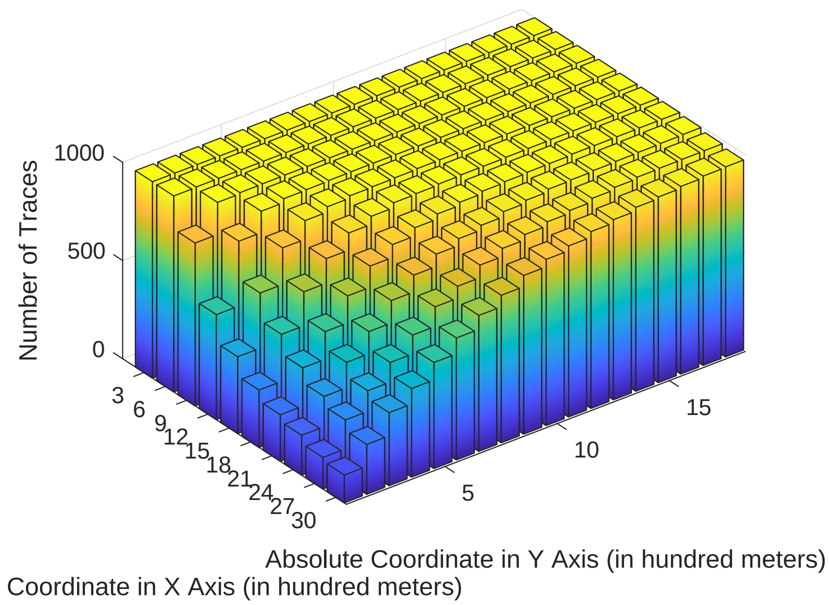
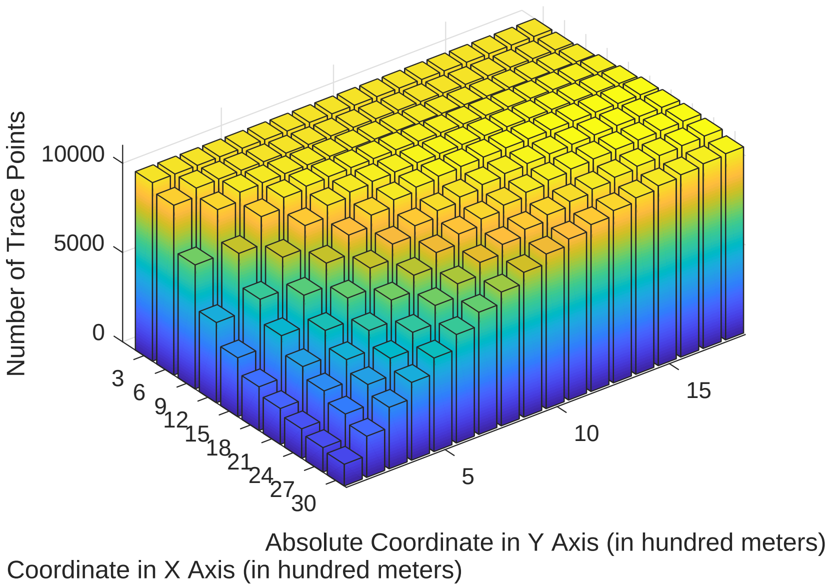


Figure 3.3: Synthetic Trace Generation (10 Example Traces)

To quantify the statistical predictability, the x-axis is sliced equally (300 metres per slice in this case) to count the number of passing traces. Given an absolute value of the y-axis coordinate, the number of traces passing the enclosed rectangular area can be calculated. As shown in Figure 3.4(a), the number of passing traces diminishes progressively, which indicates the existence of a certain spatial distribution. It is worth mentioning that the results of Figure 3.3 are essentially consistent with Figure 3.4(a); the difference is that the absolute value of the y-axis coordinate affects the size of enclosed rectangular area in Figure 3.4(a) (a larger absolute coordinate results in a larger enclosed area and therefore more traces tend to be counted). By counting the number of trace points within an enclosed area, similar trends can be obtained as shown in Figure 3.4(b). Based on above observation, it is shown that both the spatial distribution of sample traces and trace points can reflect the statistical predictability of correlated random walks. For the ease of computing, the remainder of this investigation focuses on the distribution of trace points.

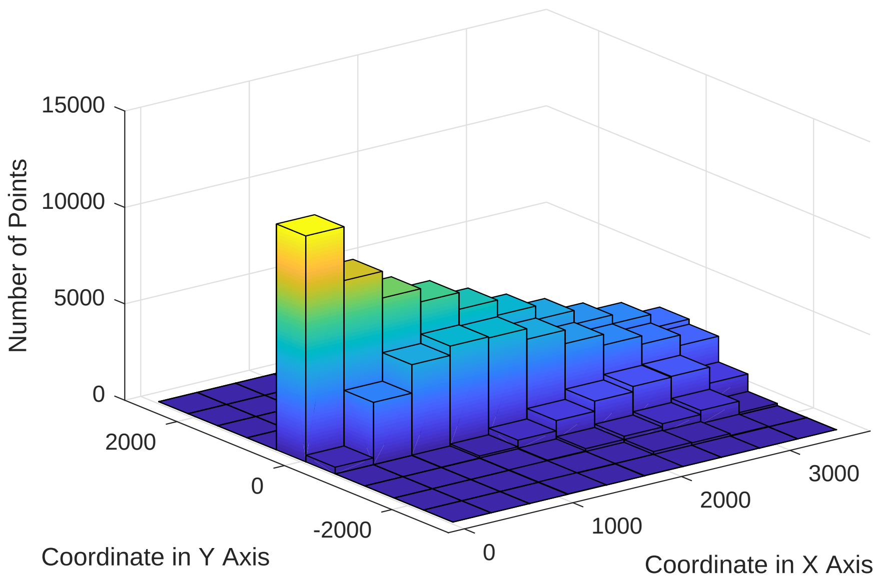


(a) Counting Traces

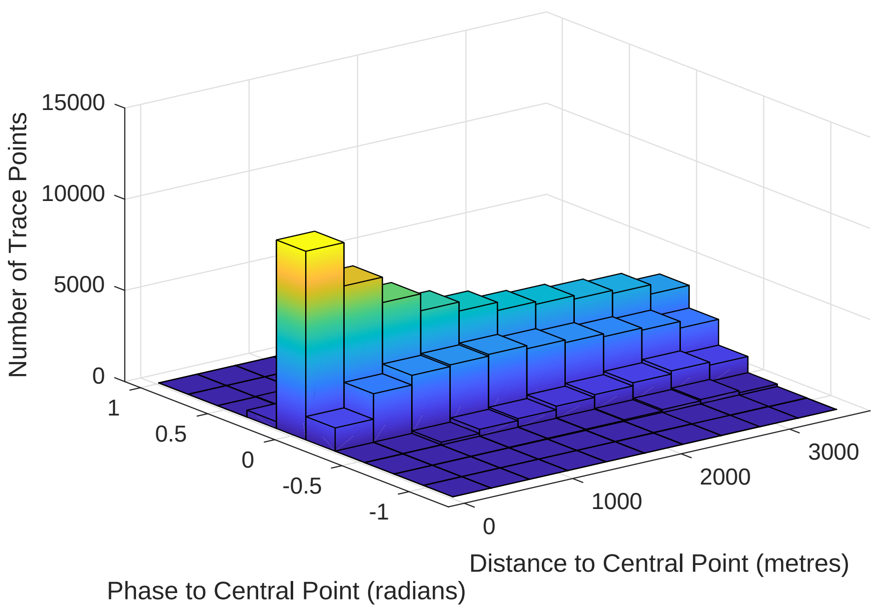


(b) Counting Trace Points

Figure 3.4: Predictability Representation of 1000 Traces

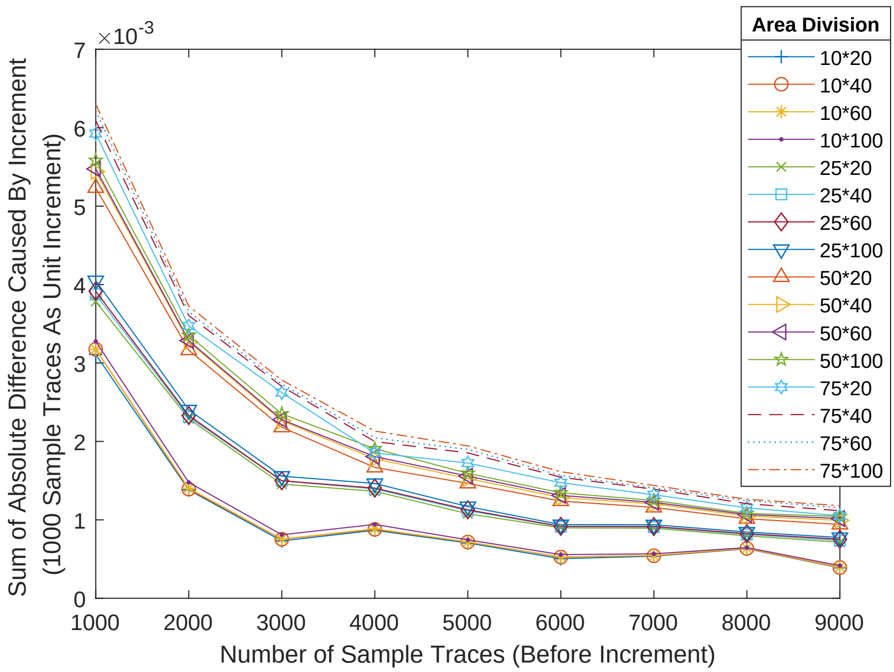


(a) Cartesian Coordinate System

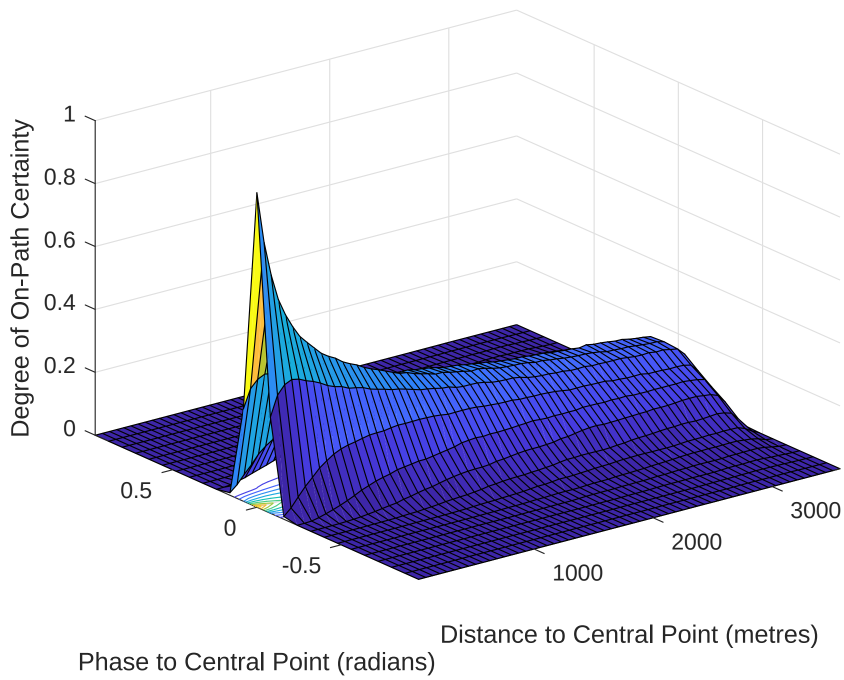


(b) Polar Coordinate System

Figure 3.5: Histogram of Trace Points from 1000 Traces



(a) Convergence with Increasing Trace Numbers



(b) Smooth Distribution

Figure 3.6: Obtaining Continuous Distribution Model

As shown in Figure 3.5, the distribution of trace points can be visualised by a two-dimensional histogram. Compared with the Cartesian coordinate system (Figure 3.5(a)), the polar coordinate system (Figure 3.5(b)) can be adopted for a better representation of the spatial displacement. To construct such a histogram, the observed area is divided into multiple sub-areas and the number of trace points is counted for each sub-area. By increasing the number of sample traces, trace points are increased at the same time. After each increment of sample traces, the absolute difference of spatial distribution (the change of trace points number at each sub-area) can be calculated and the sum of absolute difference (over all sub-areas) is shown convergent in Figure 3.6(a), regardless the specific values of bin size.

Moreover, as shown in Figure 3.6(b), a smooth distribution of trace points can be obtained by appropriately dividing the geographical area (60 bins for the distance dimension and 25 bins for the phase dimension are found suitable and adopted for the remainder of this investigation). The on-path certainty is defined here as the degree of confidence that a spatial point is located on the movement path, which is calculated by normalising the spatial distribution of trace points. The corresponding contour map for Figure 3.6(b) is given by Figure 3.7, which demonstrates the movement path can be described with different thresholds of certainty degree.

Based on above observations, the feasibility of a fuzzy representation for a movement path is preliminarily shown. However, more analytical studies are still required to obtain the equation relating the quantities involved. To this end, experimental results are fitted with mathematical models for further analysis.

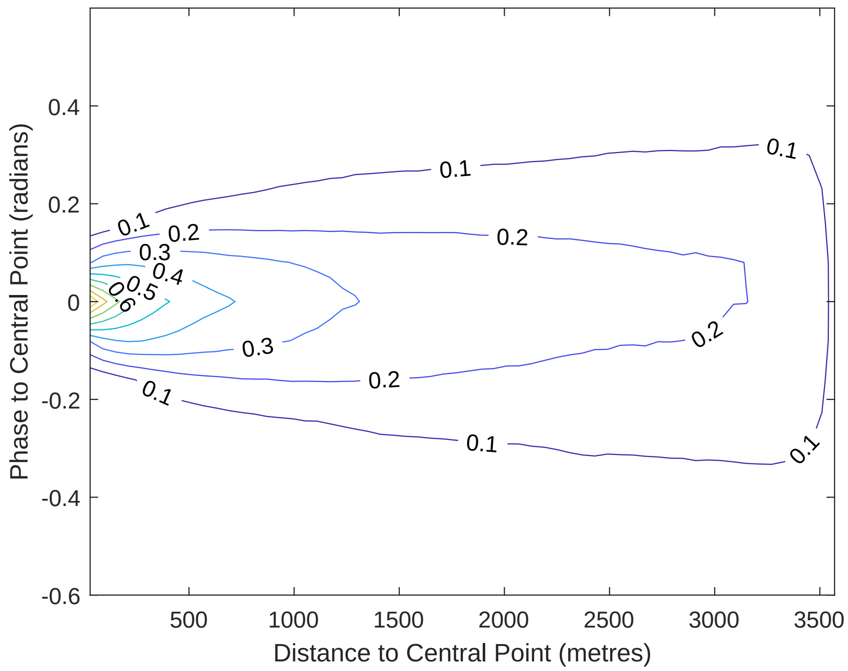


Figure 3.7: Contour of Normalised Distribution as Path Model

3.2.3.2 Predictability Validation

Although the feasibility of an ambiguous representation of a movement path is preliminarily shown in Section 3.2.3.1, observational studies cannot reflect the quantified relationships of path model. Further analysis is therefore required to fit experimental results with mathematical functions. In this section, the experimentally obtained spatial distribution of trace points (on-path certainty) is analysed for two dimensions separately, based on the standard implementation of least-squares method [Wol05] provided by the MATLAB Curve Fitting Toolbox. The experimental distribution is obtained from 10000 sample traces where the SDTA value is 10 degrees.

For the phase dimension, the normalised bell curve[2] is fitted with a Gaussian function over distance. The residuals show that the growing scale factor of the Gaussian function

[2]This bell curve means the shape observed from phase dimension in Figure 3.6(b) given a certain distance value. Each observed bell curve is normalised by dividing its maximum, so no multiplier is required in the fitting function.

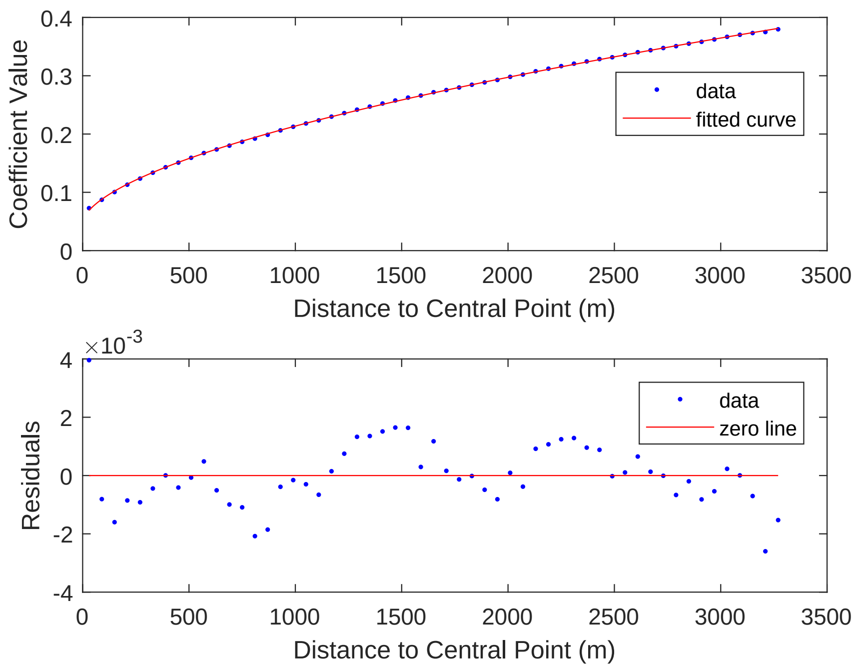


Figure 3.8: Fitting the Bell Curve to the Gaussian Function (Scale Factor as Two-term Power Series over Distances)

can be nearly modelled by a two-term power series as shown in Figure 3.8. Therefore the bell curve function is formulated as

$$f_{\text{bell}}(r, \varphi) = \exp\left[-\left(\frac{\varphi}{q_1 \cdot r^{q_2} + q_3}\right)^2\right] \tag{3.13}$$

where where r denotes the distance to the central point, φ denotes the phase to the central point, and the q_1,q_2,q_3 are functional parameters.

For the distance dimension, the normalised decay curve exhibits a dichotomy phenomena (shown by Figure 3.9): fitted with a two-term exponential series model with a certain value of distance but then fitted with a two-term power series model. A similar dichotomy phenomenon (fitting to two possible theoretical distributions) was discovered by [KBV10] when modelling the distribution of inter-contact time. Consequently, a

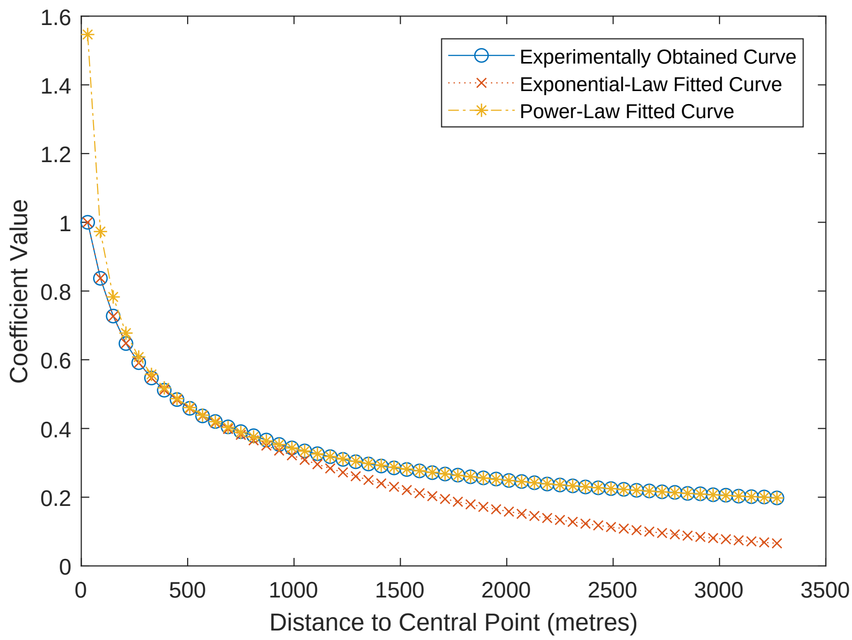


Figure 3.9: Fitting the Decay Curve in Dichotomy Phenomena

characteristic distance is introduced to formulate the decay curve as

$$f_{\text{decay}}(r) = \begin{cases} c_1 \cdot e^{c_2 r} + c_3 \cdot e^{c_4 r} & r \leqslant r_{\text{char}} \\ d_1 \cdot r^{d_2} + d_3 & r > r_{\text{char}} \end{cases} \tag{3.14}$$

where the c_1-c_4, d_1-d_3 are functional parameters and r_{char} denotes the characteristic distance.

To complete the piecewise form of the decay curve, the characteristic distance is further investigated to identify key factors influencing it. An experimental method is adopted to find the characteristic distance which minimises the sum of the squared residuals between the decay curve and the data samples. While no specific influence can be observed by changing the SDTA value, experiments show that the characteristic distance seems to be linearly related to the step length. By multiplying the default step length (30 metres), the characteristic distance scales up at the same time which roughly results in a linear relationship (as shown in Table 3-A). Based on these additional experiments, the equation of the characteristic distance is given as

$$r_{\text{char}} = g_1 \cdot \Lambda^{g_2} \approx g_1 \cdot \Lambda \tag{3.15}$$

where g_1 and g_2 denotes the functional parameters and Λ denotes a constant of step length.

Table 3-A: Investigation of Characteristic Distance

Step Length (m)	**Characteristic Distance (m)**	**Multiple**
30	450	1/1
60	875	2/1.94
90	1350	3/3
120	1825	4/4.06
240	3950	8/8.78
480	8000	16/17.78

With the functional relationships revealed by curve fitting, further analysis is provided about the surface model. The spatial distribution of trace points (on-path certainty) are fitted to the proposed surface model and the goodness of fit is reflected by the residuals (the difference between surface model and the experimentally obtained distribution).

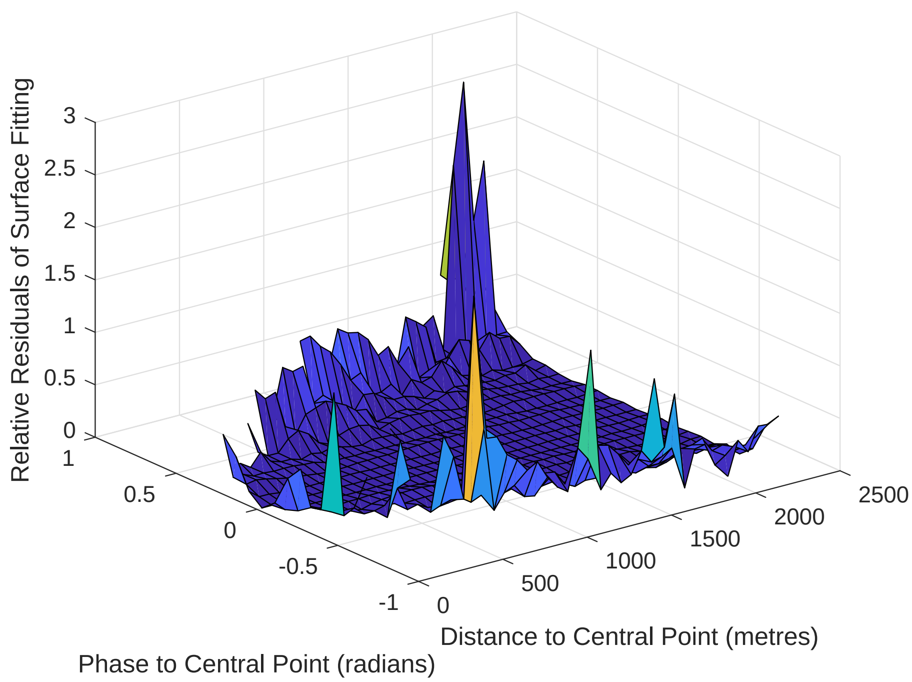


Figure 3.10: Relative Residuals of Surface Fitting

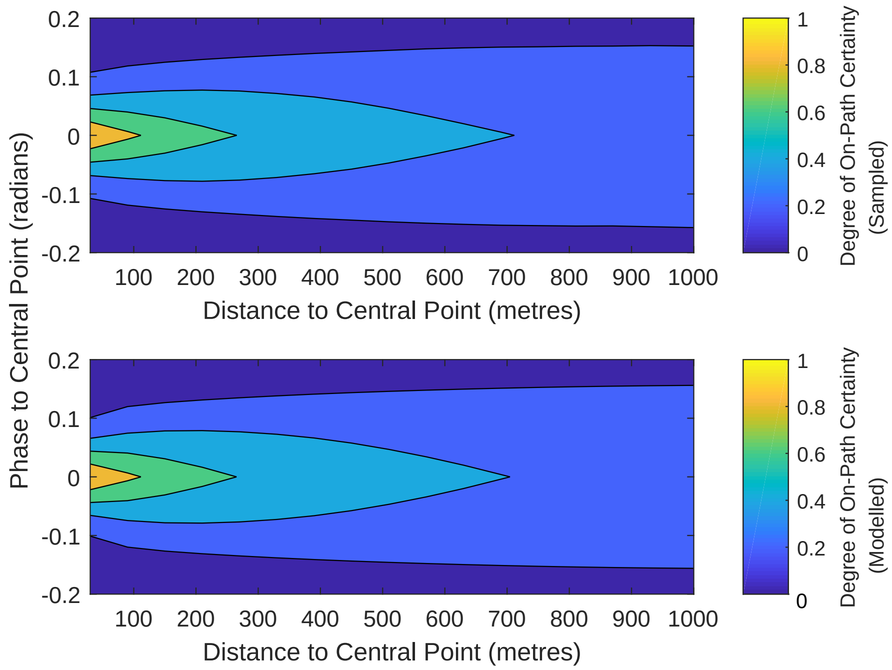


Figure 3.11: Surface Model Evaluation from Contour View

As an illustrative example, the residuals between the fitted surface and experimental distribution (obtained from 50000 sample traces where the SDTA value is 10 degrees) are shown in Figure 3.10. This example shows that the residuals of surface fitting are mostly small relative to the experimental distribution except for minor extreme values on the edges. From the contour map view, the difference resulting from the residuals can be ignored as shown in Figure 3.11.

To evaluate the goodness of surface fitting for different SDTA values (each employing 50000 sample traces), the sum of the relative residuals is considered as the numerical metric. As shown in Figure 3.12, the sum of relative residuals over either the phase dimension or the distance dimension is presented separately. For the sum over the phase dimension, although values are relative large initially, it rapidly decreases and then becomes low and steady. For the sum over the distance dimension, the values increase with absolute phase angle but are mostly within the same order of magnitude.

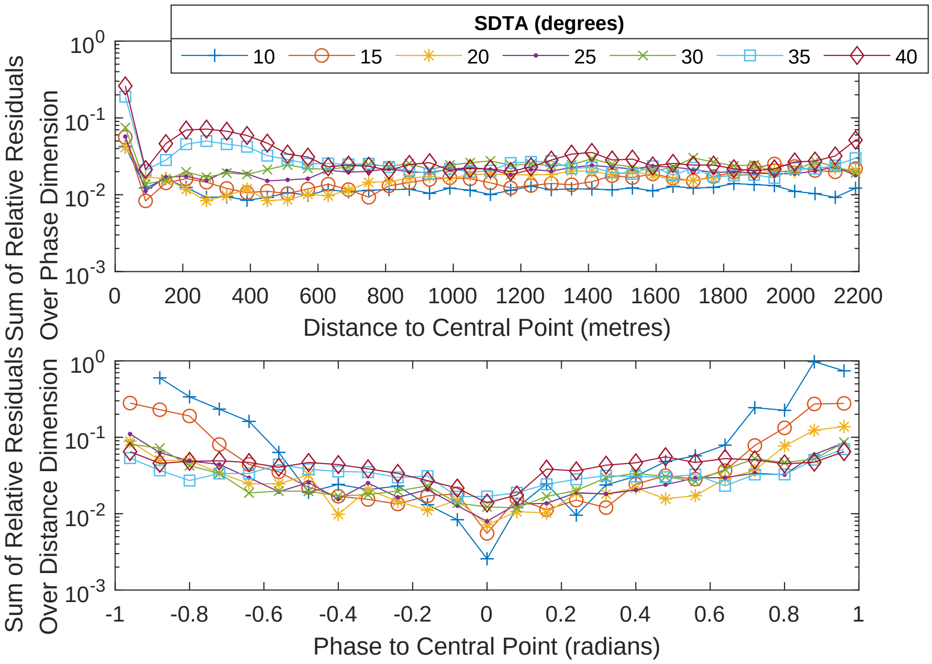


Figure 3.12: Sum of Relative Residuals Over Phase/Distance Dimension

In summary, the analysis results show that random walks with a directional distribution are predictable to a certain extent. However, this analysis does not consider the transmission range of wireless nodes and therefore cannot be directly used for contact prediction. To this end, the contact opportunity is further discussed in the next section where a prediction approach is proposed based on movement estimation.

3.3 Analysis of Contact Opportunity

The aforementioned analysis of directional movement reveals the existence of certain statistical predictability. In this section, the predictability of directional movement is used for modelling contact opportunities, based on the description of the contact model and the estimation of directional movement.

3.3.1 Contact Model

Having described the mobility model, we now turn our attention to the modelling of contact. A traditional model of unit disk [CCJ91] is considered where there is a constant radius of contact range (denoted by R). Let $\mathcal{N}$ denote the whole set of stationary nodes, then a stationary node (denoted by n_j) is contactable by the mobile object if only if

$$d_j\left(i\right) \leqslant R \tag{3.16}$$

where $d_j\left(i\right)$ denotes the distance between n_j and the mobile target as a function of step number i.

The probability that the mobile target can contact a stationary node n_j within finite steps can be described as

$$\mu_\tau\left(n_j\right) = Pr\left[\min\left(d_j\left(i\right)\right) \leqslant R \mid i \leqslant \tau\right] \tag{3.17}$$

where $\tau \in \mathbb{N}$ denotes a bound on the step number and $\min\left(d_j\left(i\right)\right)$ denotes the minimum distance over a number of steps.

Taking equation (3.17) as the membership function, a fuzzy set of the stationary nodes is defined as

$$\mathcal{N}_\tau\left(\eta\right) = \left\{n_j \mid \mu_\tau\left(n_j\right) \geqslant \eta\right\} \tag{3.18}$$

where $\eta \in [0, 1]$ denotes a threshold of contact probability and $\mathcal{N}_\tau\left(\eta\right)$ denotes the subset of stationary nodes considered to be along the movement path of the mobile target given τ and η.

To calculate $d_j\left(i\right)$, a polar coordinate system is considered where r denotes the relative distance from the original position of the mobile target, and φ denotes the relative angle measured with respect to the initial movement direction of the mobile target. Under such a coordinate system, the position coordinates of the mobile target after i

steps (denoted by $\vec{e}_i$) and the position coordinates of n_j (denoted by $\vec{e}_j$) can be respectively described as

$$\vec{e}_i = (r_i, \varphi_i) \tag{3.19}$$

$$\vec{e}_j = (r_j, \varphi_j)\,. \tag{3.20}$$

Note that $\vec{e}_i$ is changing over time step i because it represents the position of the mobile target. Given a set of stationary nodes $\mathcal{N}$ (which have many different positions), $\vec{e}_j$ varies with each $n_j \in \mathcal{N}$. Such notation of $\vec{e}$ intends to show these spatial and temporal variables together in a concise way.

According to the law of cosines in trigonometry,

$$\begin{aligned} d_j\left(i\right) &= \left|\vec{e}_i - \vec{e}_j\right| \\ &= |\vec{e}_i|^2 + |\vec{e}_j|^2 - 2|\vec{e}_i||\vec{e}_j| \cdot \cos\left\langle \vec{e}_i, \vec{e}_j \right\rangle \\ &= r_i^2 + r_j^2 - 2r_i r_j \cdot \cos\left(\varphi_i - \varphi_j\right). \end{aligned} \tag{3.21}$$

The r_j and φ_j are known constants given the original position and the initial movement direction of the mobile target. However, due to the random turning angle as described in equation (3.2), $\vec{e}_i$ is a stochastic process as $\{(r_i, \varphi_i) : i \in \mathbb{N}\}$, where r_i and φ_i are random variables representing the displacement distance and angle observed after i steps, respectively. To quantify the contact potential, this movement process is further analysed in the Section 3.3.2.

3.3.2 Movement Estimation

The prediction of contact is dependent on the estimation of movement, which is central to strategy making. Here, the stochastic properties are further analysed to estimate the movement.

Based on equation (3.3) and (3.4), the displacement distance projected on the axial direction can be given as

$$\begin{aligned} r_i \cdot \sin \varphi_i &= \sum_{\zeta=1}^{i} (\Lambda_i \cdot \sin \omega_i) \\ &= V \cdot T \cdot \sum_{\zeta=1}^{i} (\sin \omega_\zeta) \, . \end{aligned} \tag{3.22}$$

Similarly, the displacement distance projected on the radial direction can be given as

$$r_i \cdot \cos \varphi_i = V \cdot T \cdot \sum_{\zeta=1}^{i} (\cos \omega_\zeta) \, . \tag{3.23}$$

Dividing equation (3.22) by equation (3.23), we obtain

$$\tan \varphi_i = \frac{\sum_{\zeta=1}^{i} (\sin \omega_\zeta)}{\sum_{\zeta=1}^{i} (\cos \omega_\zeta)}. \tag{3.24}$$

Consequently,

$$\varphi_i = \arctan \left(\frac{\sum_{\zeta=1}^{i} \left(\sin \left(\theta_0 + \sum_{\xi=1}^{\zeta} \Theta_\xi \right) \right)}{\sum_{\zeta=1}^{i} \left(\cos \left(\theta_0 + \sum_{\xi=1}^{\zeta} \Theta_\xi \right) \right)} \right) \tag{3.25}$$

where φ_i denotes the displacement angle after i steps.

Due to the complexity of equation (3.25), the probability distribution of φ_i cannot be solved directly (as it is not a linear combination of multiple normal distributions). However, according to the central limit theorem [Fis10], independent and identically distributed (i.i.d.) random variables tend towards the Normal distribution. Therefore, the distribution of the displacement angle after a number of steps is expected to be a bell-shaped curve given as

$$f_{\text{Gauss}}(z) = \alpha \cdot \exp\left(-\left(\frac{z-\beta}{\gamma}\right)^2\right) \tag{3.26}$$

where f_{Gauss} denotes the Gaussian function, z denotes the independent variable, α denotes the amplitude factor, β denotes the centroid factor, and γ denotes the scale factor.

Based on the Monte Carlo method [KBTB14], simulation experiments are conducted to confirm the validity of the distribution and to obtain the model parameters. Given initial parameters[3] (specified in Table 3-B), a set of synthetic traces can be generated as shown in Figure 3.13. A clearer view can be provided by sampling the trace points only at the selected steps as shown in Figure 3.14.

Table 3-B: Synthetic Traces Generation for Monte Carlo Experiments

Parameters	**Values**
Truncation Bound	$\pm\pi$
ς_{th}	4
σ (SDTA)	5,10,15,20,25,30,35,40,45 (degrees)
Movement Speed	1 m/s
Initial Direction	x-axis positive direction
Original Position	central point (0,0)
Step Period	30 seconds
Observation Period	3600 seconds
Number of Traces	10000 for each SDTA value

[3]Settings in this research are based on the normal walking behaviours but scalable to other scenarios.

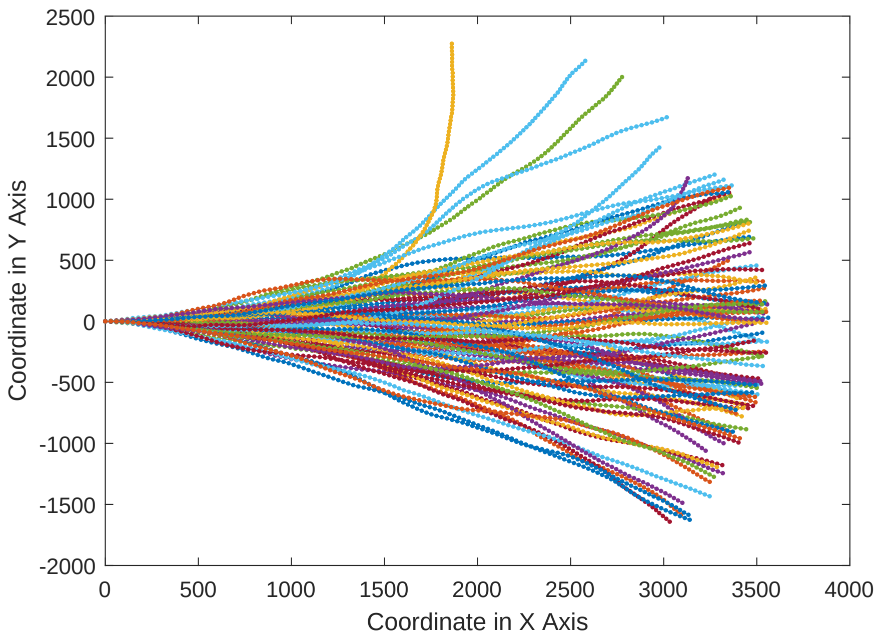


Figure 3.13: 100 Traces Generated by Setting SDTA to 10 Degrees

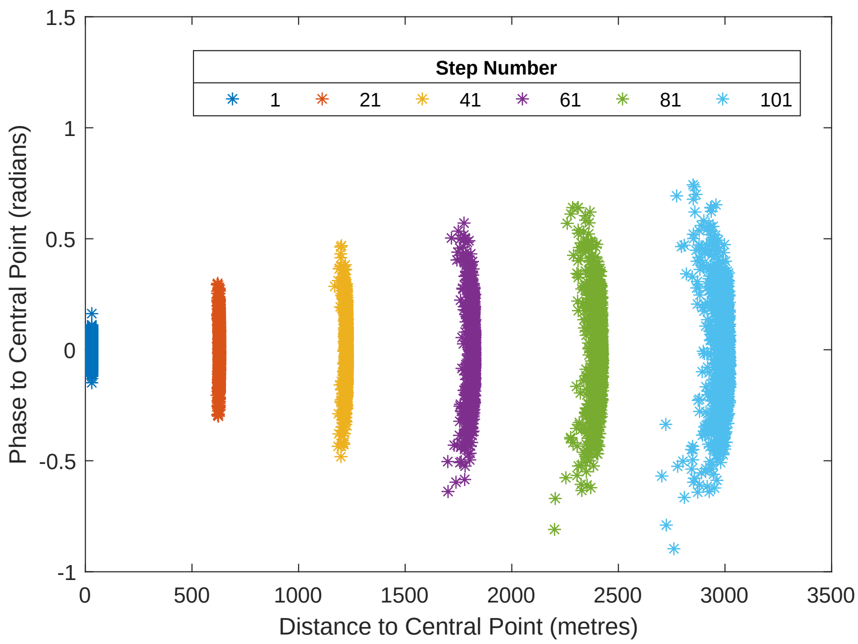


Figure 3.14: Trace Points Sampled at Selected Steps from 1000 Traces (Set SDTA to 10 Degrees)

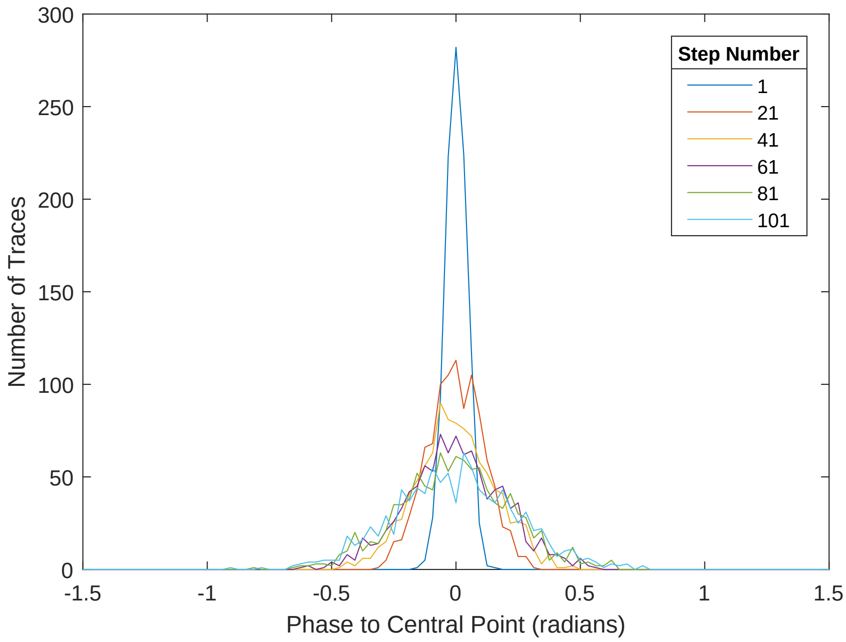


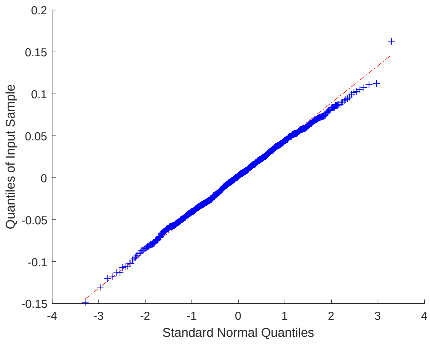
Figure 3.15: Frequency Statistics of Displacement Angle Sampled at Selected Steps from 1000 Traces (Set SDTA to 10 Degrees)

As shown in Figure 3.15, the frequency of a particular displacement angle is counted at each sampled step. The statistical results visually show a trend towards the Normal distribution. To test the normality, the Shapiro-Wilk test [Roy95] is employed and the Shapiro-Wilk null hypothesis is "the sample distribution is normal with unspecified mean and variance." The significance level for the test is set to 0.05 which is a default value. The corresponding p-value can be obtained using Matlab, which is shown in Table 3-C.

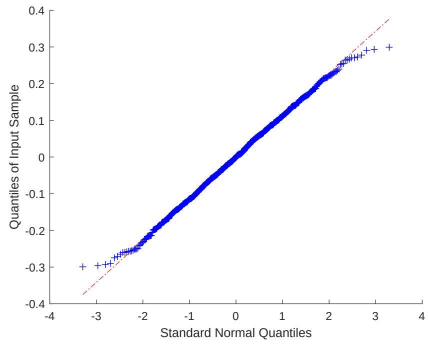
According to Table 3-C, p-values are all larger than the significance level for the test, therefore, the null hypothesis cannot be rejected. That is, the distribution of displacement angles is proved to be Normal distribution. Meanwhile, this normality can also be proved by the following Q-Q plot (Figure 3.16), as the data plot appears linear against Normal distribution.

Table 3-C: Shapiro-Wilk Test Results

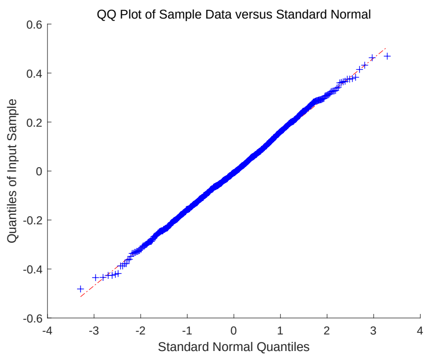
Step Number	1	21	41	61	81	101
p-value	0.3186	0.2290	0.6877	0.8615	0.7095	0.9398



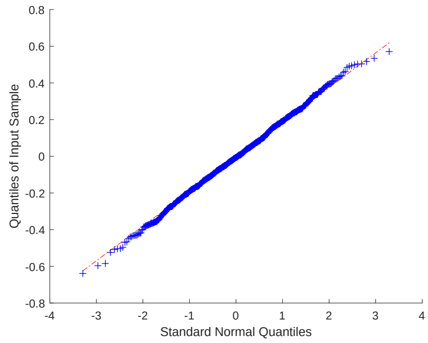
(a) Step Number = 1



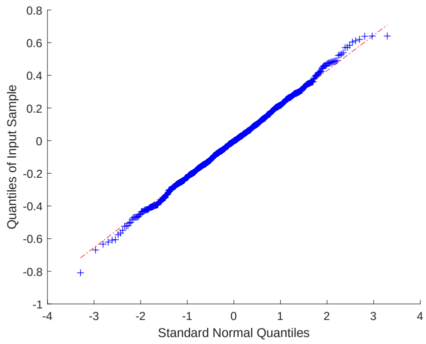
(b) Step Number = 21



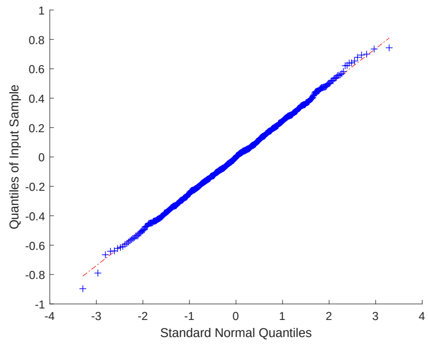
(c) Step Number = 41



(d) Step Number = 61



(e) Step Number = 81



(f) Step Number = 101

Figure 3.16: Q-Q Plot of Sampled Displacement Angles versus Standard Normal

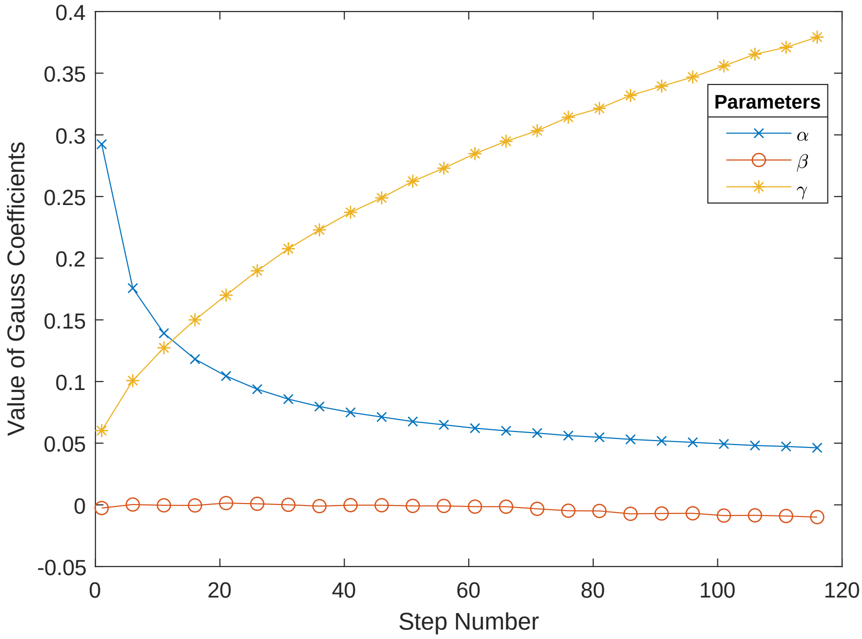


Figure 3.17: Fitted Model Parameters at Different Step Numbers (Set SDTA to 10 Degrees)

By fitting the simulated frequency distribution of each sampled step to the bell-shape curve given in (3.26), functional relationships are reflected between the three model parameters and the step numbers. As shown in Figure 3.17, the amplitude factor decreases and the scale factor increases non-linearly with increasing step number, whilst the centroid factor remains constant. With the 'NonlinearLeastSquares' fitting method and 'Trust-Region' fitting algorithm provided by the MATLAB Curve Fitting Toolbox, the factor relationships are obtained as

$$\begin{cases} \alpha = a_1 \cdot i^{a_2} + a_3 \\ \beta = 0 \\ \gamma = b_1 \cdot i^{b_2} + b_3 \end{cases} \tag{3.27}$$

where α,β,γ are the three factors defined in (3.26), a_1,a_2,a_3,b_1,b_2,b_3 are parametric constants, and i denotes the step number.

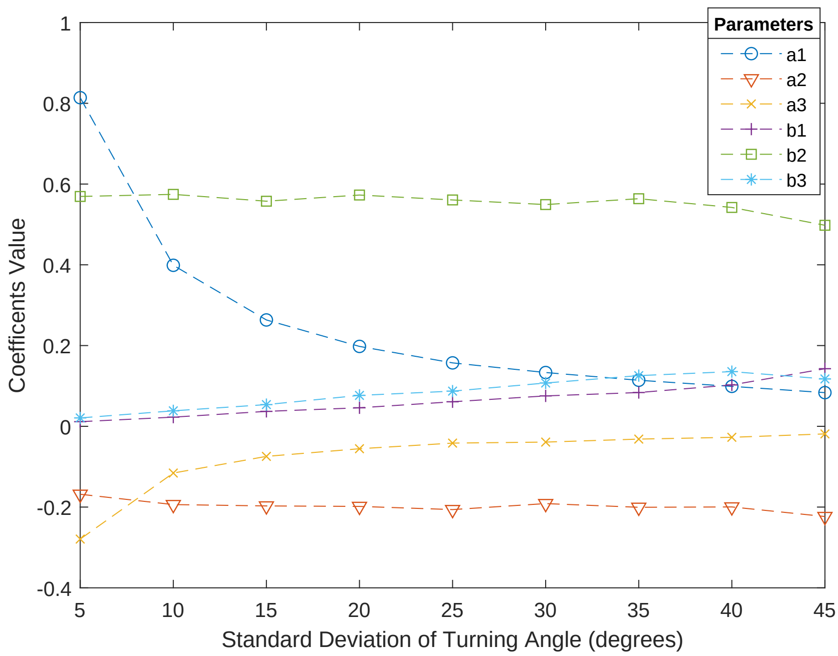


Figure 3.18: Further Fitted Model Parameters at Different SDTA Values

As shown in Figure 3.18, further fitting the parameters obtained in (3.27), the relationships can be found with the SDTA. Using the same fitting technique, the parameter relationships are obtained as

$$\begin{cases} a_1 = p_1 \cdot \sigma^{p_2} \\ a_2 = p_3 \\ a_3 = p_4 \cdot \sigma^{p_5} \\ b_1 = p_6 \cdot \sigma + p_7 \\ b_2 = p_8 \\ b_3 = p_9 \cdot \sigma + p_{10} \end{cases} \tag{3.28}$$

where a_1,a_2,a_3,b_1,b_2,b_3 are parametric constants defined in (3.27), p_1 to p_{10} are fitting parameters, and σ denotes the SDTA.

Based on equations (3.26)(3.27) and (3.28), the curve shape can be defined as a parametric model

$$\begin{cases} f_{\text{shape}}\left(z,\sigma,i\right)=\alpha\left(\sigma,i\right)\cdot\exp\left(-\left(\frac{z-\beta}{\gamma(\sigma,i)}\right)^{2}\right) \\ \alpha\left(\sigma,i\right)=\left(p_1\cdot i^{p_2}\right)\cdot\sigma^{p_3}+\left(p_4\cdot i^{p_5}\right) \\ \gamma\left(\sigma,i\right)=\left(p_6\cdot i+p_7\right)\cdot\sigma^{p_8}+\left(p_9\cdot i+p_{10}\right) \\ \beta=0 \end{cases} \tag{3.29}$$

where f_{shape} denotes the curve shape function, z denotes the independent variable, σ denotes the SDTA, i denotes the step number, and p_1 to p_{10} are parameters (listed in Table 3-D) solved under Table 3-B configurations.

Table 3-D: Solved Model Parameters

Parameters	**Values**	**95% Confidence Bounds**
p_1	0.0675	(0.0662, 0.0687)
p_2	-1.0210	(-1.0290,-1.0120)
p_3	-0.1976	(-0.2087,-0.1865)
p_4	-0.0157	(-0.0177,-0.0137)
p_5	-1.1786	(-1.2360,-1.1210)
p_6	0.1694	(0.1369, 0.2018)
p_7	-0.0092	(-0.0252, 0.0067)
p_8	0.5542	(0.5360, 0.5723)
p_9	0.1627	(0.1211, 0.2042)
p_{10}	0.0137	(-0.0067, 0.0342)

Then the probability distribution function of the displacement angle after certain number of steps (denoted by f_{pdf}) can be obtained as

$$Pr\left[\varphi_i=\delta\right]=f_{\text{pdf}}\left(\delta,\sigma,i\right)=\frac{f_{\text{shape}}\left(\delta,\sigma,i\right)}{\int\limits_{-\pi}^{\pi}f_{\text{shape}}\left(z,\sigma,i\right)dz}. \tag{3.30}$$

As shown in equation (3.30), the probability distribution of displacement angle varies over steps and is dependent on SDTA. This distribution model provides a tool to quantify the directional movement, which will be applied and evaluated in the next section.

3.3.3 Contact Prediction (with Validation)

In Section 3.3.2, a parametric model is proposed to estimate the displacement angle over a number of steps based on the Standard Deviation of the Turning Angles (SDTA). Now we apply this estimation model in a three-phase approach to describe the contact certainty based on the relative distance (denoted by r) from the original position of the mobile target, and the relative angle (denoted by φ) measured with respect to the initial movement direction of the mobile target.

For the first phase, an estimated step number (denoted by i_{est}) is calculated as

$$i_{\mathrm{est}} = \frac{r}{V \cdot T} \tag{3.31}$$

where V denotes the average speed and T denotes the step period.

For the second phase, the range of directional coverage is given as

$$[\varphi_{\mathrm{min}}^{\mathrm{cover}}, \varphi_{\mathrm{max}}^{\mathrm{cover}}] \tag{3.32}$$

$$\varphi_{\mathrm{min}}^{\mathrm{cover}} = \varphi - \sin\left(\frac{R}{r}\right) \tag{3.33}$$

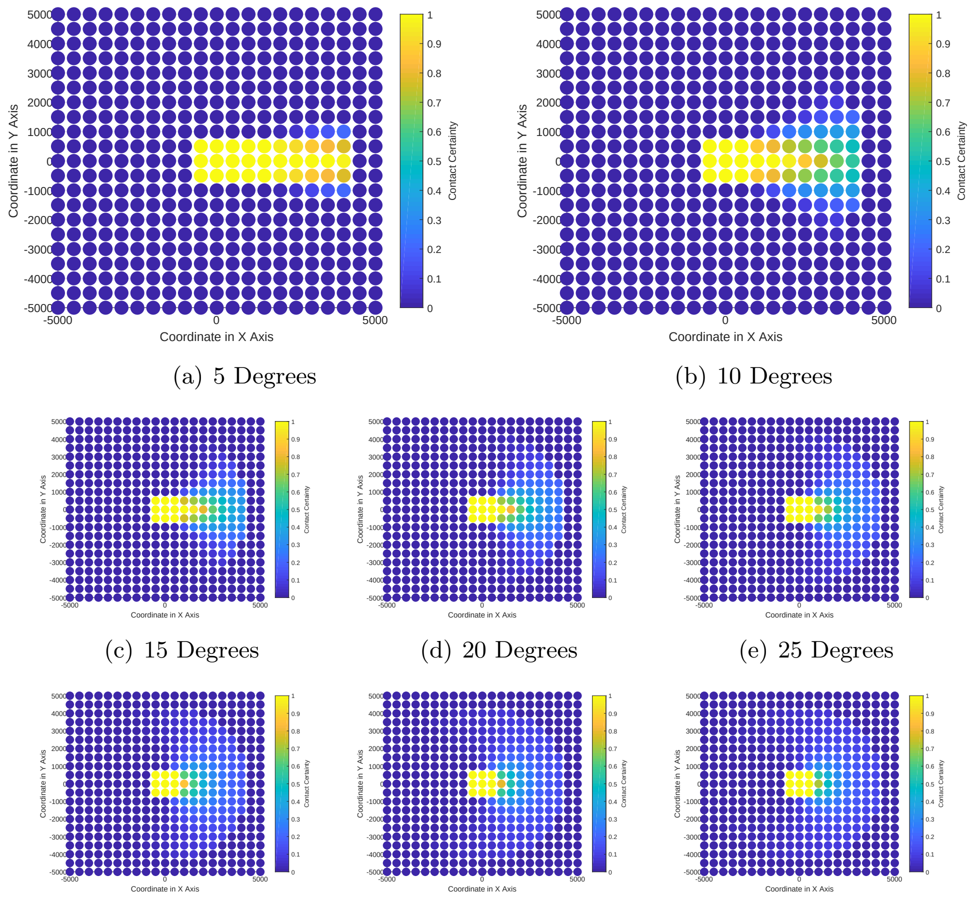
$$\varphi_{\mathrm{max}}^{\mathrm{cover}} = \varphi + \sin\left(\frac{R}{r}\right) \tag{3.34}$$

where $\varphi_{\mathrm{min}}^{\mathrm{cover}}, \varphi_{\mathrm{max}}^{\mathrm{cover}} \in [-\pi, \pi]$ denotes the minimum/maximum coverage angle and R denotes the radius of contact range.

For the third phase, the contact certainty (denoted by μ_{est}) is finally described as

$$\begin{aligned}\mu_{\mathrm{est}} &= Pr\left[\varphi_{\mathrm{max}}^{\mathrm{cover}} \geqslant \varphi_{i_{\mathrm{est}}} \geqslant \varphi_{\mathrm{min}}^{\mathrm{cover}}\right] \\ &= \int_{\varphi_{\mathrm{min}}^{\mathrm{cover}}}^{\varphi_{\mathrm{max}}^{\mathrm{cover}}} f_{\mathrm{pdf}}\left(\delta, \sigma, i_{\mathrm{est}}\right) d\delta \\ &= f_{\sigma}\left(r, \varphi\right)\end{aligned} \tag{3.35}$$

where f_σ denotes the estimation function given σ.



(a) 5 Degrees (b) 10 Degrees

(c) 15 Degrees (d) 20 Degrees (e) 25 Degrees

(f) 30 Degrees (g) 35 Degrees (h) 40 Degrees

Figure 3.19: Spatial Distribution of Contact Certainty Based on the Estimation Model for Different Values of SDTA

Thus, the contact certainty is formulated as a spatial distribution given SDTA. This distribution model describes statistically predicable contact with spatial points, which is later applied for the selection of stationary relay node(s). With configurations specified by Table 3-E, a simulation is firstly conducted to visually demonstrate this distribution model.

Table 3-E: Simulation Configurations for Contact Prediction

Parameters	**Values**
Mobility Configurations	Same with Table 3-B
Model Parameters	Same with Table 3-D
Region Size	10000*10000 metres
Grid Size	500*500 metres
Radius of Contact Range	750 metres

As shown in Figure 3.19, the geographical region is equally divided into multiple grid squares, where the mobile node is originally located at the central grid point with initial movement direction towards the x-axis positive direction. Based on equation (3.35), the contact certainty is estimated for each grid point for a given SDTA. The results show that this model can reflect the decreasing predictability with increasing value of SDTA (i.e. less directional correlation). The varying predictability does not affect the central nine grid points because they are directly contactable and no estimation is required.

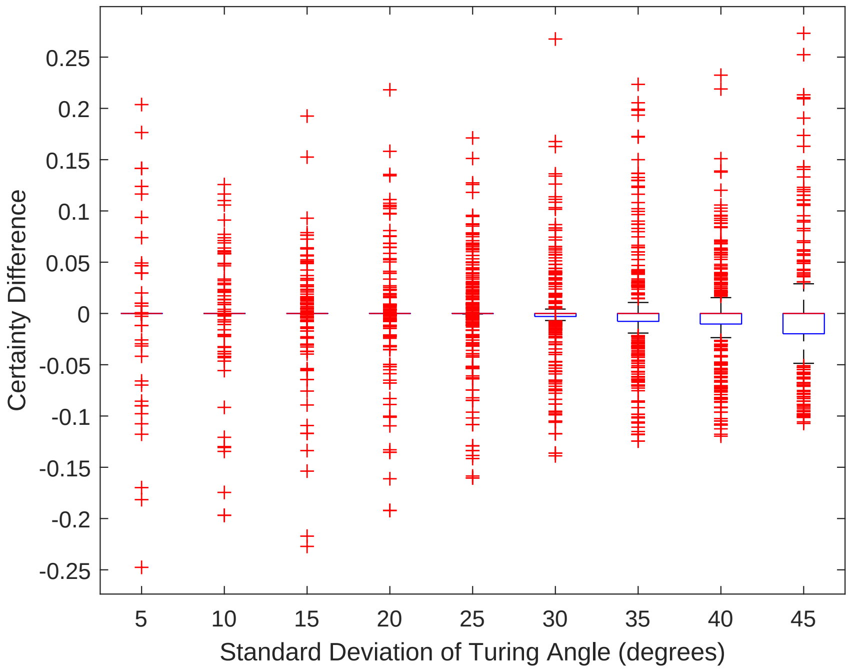


Figure 3.20: Box Plot of Certainty Difference for the Accuracy Validation

To further validate the model accuracy, test traces are generated synthetically (100 traces for each SDTA value with same parameters listed in Table 3-B). The ratio of contactable traces is calculated for each grid point as the tested contact certainty. The certainty difference is defined as the tested certainty minus the estimated certainty, so a positive difference value reveals an underestimate and a negative difference value reveals an overestimate. In Figure 3.20, the certainty difference at each grid points is shown for different SDTA values. Because equation (3.35) is not an accurate calculation of contact probability, either underestimation or overestimation may exist as the results show. However, it can be observed the model can adaptively react to the varying SDTA so that the difference values are mostly bounded by 0.1 and the rare extreme values are roughly bounded by 0.25. As a conclusion, the approximate estimation model can reflect the effect of directional correlation on contact certainty.

3.4 Summary

In this chapter, the directional movement is firstly modelled and investigated on the basis of correlated random walks. A statistical experiment is conducted and the results reveal that predictability of directional movement exists. Given the described contact model, an approach is proposed and shown to be feasible to quantify contact opportunities created by the directional movement. An application of this contact prediction model is provided in Chapter 4 for strategic forwarding.

Chapter 4

Optimisation of Delivery Utility for Efficient Dissemination

Starting with an overview of the scenario description, two types of relay selection are addressed in this chapter. For the case of individual delivery, the relay node is independently selected for each target node. For the case of group delivery, a relay node may serve multiple targets, which provides more potential for optimising network resource utilisation but increases the complexity of the selection problem at the same time.

4.1 Scenario Overview

Consider a SDMSN system where there is one central controller located at the sink node and one local controller located in each sensor node (note that these controllers are logical entities rather than physical nodes). During the system operation, the central controller is responsible for distributing control messages to local controllers regarding instructions for local data forwarding. For control-plane communication, the control policy may change over time and the central controller is required to get the updated control message to all corresponding local controllers.

For the considered scenario, the sink node is stationary; a sensor node can be either a stationary node or a mobile node, where the stationary nodes always stay still while the mobile nodes are able to move. The research problem of control message distribution is formulated as how to efficiently deliver the control message from the sink node (central controller) over a sensor network consisting of stationary nodes and mobile nodes. As such control information dissemination is neither frequent nor delay-sensitive, the opportunistic forwarding is considered to improve the dissemination efficiency.

As shown in Figure 4.1, node 0 is the sink node (i.e. central controller) and nodes 1-9 are sensor nodes (i.e. local controllers), where nodes 1-7 are stationary nodes and nodes 8-9 are mobile nodes. For this example, the control message targets are node 1,5,9, which includes both the stationary node and the mobile node. To focus on the research challenge of efficient delivery to mobile nodes, it is further considered that all message targets are mobile (i.e. sensors attached to or implanted within animals) in the remainder of this chapter.

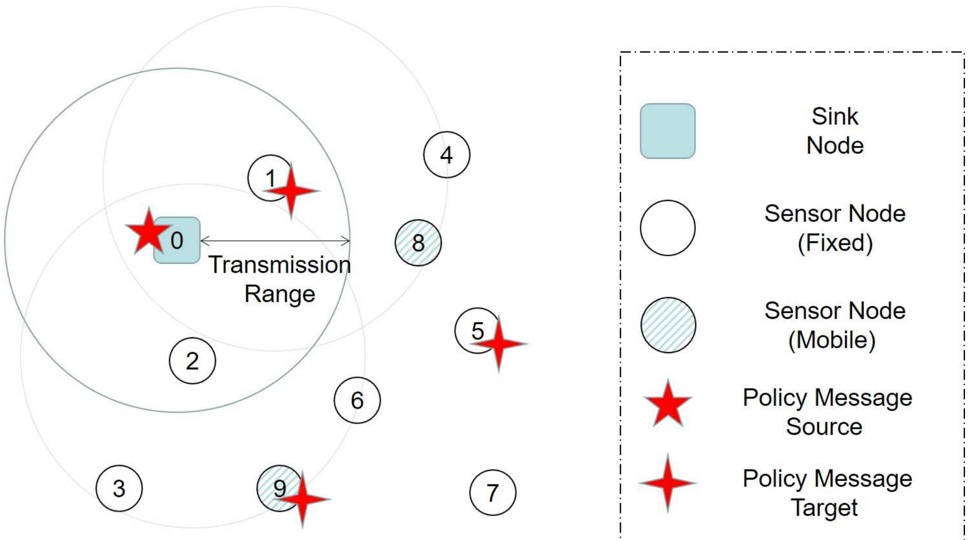


Figure 4.1: Scenario Overview

4.2 Relay Selection for Individual Delivery

For the case of individual delivery, the relay node is independently selected for each target node. The research focus of this case is to identify the relay node with least possible forwarding hops given enough confidence of later contact with the target node. Starting from the scenario model, the research problem is formulated and discussed in this section.

4.2.1 Scenario Model

Based on predictable contact modelled in Chapter 3, strategies can be made for efficient forwarding. Let m_k denote the task of delivering a control message from the sink node to a mobile target node, and a whole set of tasks is denoted by $\mathcal{M}$. Because each delivery task can have its own target, the subscript k is used in later notations to specify the relationship to a certain delivery task m_k. In this section, we assume these delivery tasks are independent, which represents a generalised scenario that can be either a delivery to a whole group of targets or separate deliveries to a single target. The strategy of relay sharing for group delivery, as in [CYY16], is discussed in the next section. Considering delivery tasks of different importance is beyond the scope of this thesis.

For the control message with delay tolerance, predictable contact can be considered in order to reduce the number of transmissions. Let B_k denote the delay budget of m_k, then the corresponding maximum allowed step number (denoted by λ_k) can be calculated as

$$\lambda_k = \lfloor \frac{B_k}{T} \rfloor. \tag{4.1}$$

Thus, for example, if the delay budget is 1300 seconds and the length of a step period is 200 seconds then there are 6 permitted steps.

As with the delivery scheme considered in [CYY16], packets are firstly sent to the selected stationary relay nodes by multi-hop forwarding for later last-hop delivery. Then the delivery is accomplished by the stationary relay node when the mobile target node is directly contactable. Based on the fuzzy set defined by equation (3.18), qualified candidate stationary nodes for m_k can be defined as

$$\mathcal{N}_k(\eta) = \{n_j \mid \mu_{\lambda_k}(n_j) \geqslant \eta\} \tag{4.2}$$

where n_j denotes a target node, μ_{λ_k} denotes the contact probability within step bound λ_k, and η denotes a threshold of contact probability.

Note that η (i.e. a threshold of contact probability) plays an important role in the opportunistic forwarding. More precisely, η defines a minimum acceptable contact probability. The lower value of η, the more tolerance of delivery failure is reflected (e.g. due to less importance of delivery or possible backup schemes), and therefore the opportunistic forwarding is more encouraged for optimisations.

As the stationary nodes form a fixed network topology, we assume knowledge of forwarding hops from the sink node to any stationary node n_j is available at the sink node, which is denoted by H_j. For each delivery, the stationary node with the least forwarding hops is selected from the candidates and $h_k(\eta)$ denotes the minimum hop number for m_k given η.

4.2.2 Problem Formulation

Without considering retransmission and duplication issues, energy consumed due to transmissions is primarily[1] proportional to the number of forwarding hops. As the final hop from the relay to the target should be counted as well, the forwarding cost for m_k is defined as

$$c_k(\eta) = 1 + h_k(\eta) \tag{4.3}$$

[1] This is a simplification made on the basis that the link quality is assumed to be similar for wireless nodes. It is a common way to simplify the application scenario for the research of wireless multi-hop communications, for example in [NF16] and [CYY16]. Such assumption may not stand if nodes/obstacles are very unevenly distributed, which is out-of-scope for this thesis.

where c_k denotes the forwarding cost and the plus one counts for the expected last-hop delivery.

As defined by equation (4.2), any stationary node selected from $\mathcal{N}_k(\eta)$ guarantees the probability η for on-time contact with the mobile target. Ideally, such contact is considered to be a successful delivery and therefore the satisfactory degree of delivery is expected to be

$$s_k(\eta) = \eta \tag{4.4}$$

where $s_k(\eta)$ denotes the expectation of delivery satisfactory degree given η.

Finally, the delivery utility is considered as a metric to assess the efficiency of the completion of the whole set of delivery tasks. As the ratio of overall satisfactory degree to the overall forwarding cost, the delivery utility (denoted by u) is defined as

$$\begin{aligned} u(\eta) &= \frac{\sum\limits_{m_k \in \mathcal{M}} s_k(\eta)}{\sum\limits_{m_k \in \mathcal{M}} c_k(\eta)} = \frac{M \cdot \eta}{M + \sum\limits_{m_k \in \mathcal{M}} h_k(\eta)} \\ &= \frac{\eta}{1 + h_{\text{avg}}(\eta)} \leqslant \frac{\eta}{1 + h_{\text{min}}(\eta)} \end{aligned} \tag{4.5}$$

where $u(\eta)$ denotes the achievable utility given η, M denotes the total number of delivery tasks, $h_{\text{avg}}(\eta)$ denotes the average forwarding hops over all the deliveries, and $h_{\text{min}}(\eta)$ denotes the minimum forwarding hops for a single delivery.

The utility value reflects the satisfactory degree contributed by each forwarding hop for the whole set of delivery tasks. Note that the delivery utility is not simply proportional to the threshold value η, because it is also dependent on the average forwarding hops (which is very likely to be reduced with decreasing threshold value due to the opportunistic forwarding). Although direct optimisation of delivery utility is impossible, as it requires analysis on a case-by-case basis, equation (4.5) shows that the maximum utility value is bounded by the minimum forwarding hops for a single delivery. Consequently, it is possible to realise a higher achievable utility by adjusting η, as a lower value of required threshold can provide more candidates and therefore the minimum forwarding hop number tends to decrease. To explore this possibility, further analysis and simulation (with validation) are conducted.

4.2.3 Simulation and Discussion

In Section 4.2.2, mathematical analysis of the strategy model reveals the possibility of higher achievable utility by adjusting the threshold of contact certainty in stationary relay node selection. We now further explore its practical effect in a simulation scenario (configurations specified by Table 4-A).

Table 4-A: Simulation Configurations for Relay Selection

Parameters	**Values**
Network Area	10000*10000 metres
Transmission Range	750 metres
Number of Stationary Nodes	21*21
Deployment of Stationary Nodes	uniformly distributed
Sink Node Location	bottom right corner
Routing Algorithm	position-based greedy
Original Position	random within network area
Initial Direction	random within $[-\pi, \pi]$
Number of Delivery Tasks	1000

As shown in Figure 4.2, a rectangular region is considered as the network area where a certain number of stationary nodes (denoted by blue squares) are uniformly deployed. Single-hop connectivities (denoted by blue dashed lines) exist between neighbouring stationary nodes, which form a grid network topology. One of the stationary nodes (located at the bottom right corner in our case) serves as the sink node which connects to the outside of the network area. A certain number of delivery tasks are to be accomplished, where all the deliveries come from the sink node and each goes to a mobile target node. The original position (denoted by a blue circle) and initial movement direction (denoted by a blue solid line) of the mobile target node are randomly and independently decided for each delivery task.

To accomplish a delivery task, one stationary node is selected from a set of candidates (denoted by black dots) as the last-hop relay. The selected stationary relay node (denoted by a black asterisk) is responsible for the direct delivery when the mobile tar-

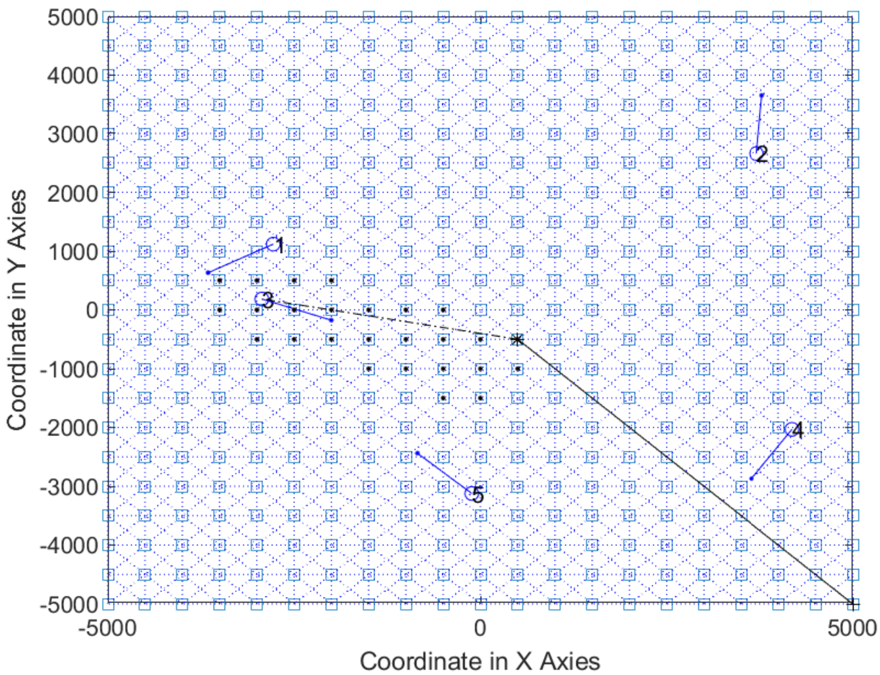


Figure 4.2: Illustrated Delivery Solution for the Third of Five Tasks (Based on the Linear Trajectory Model by Setting Delay Budget to 3000 Seconds)

get node is contactable. Based on the network topology formed by stationary nodes, the sink node connects to the selected stationary relay node by a multi-hop forwarding route. Position-based unicast routing is considered for packet forwarding due to its low computational effort and communication overhead [SMZ07]. In Figure 4.2, the black solid line shows the multi-hop forwarding route and the black dash-dot line indicates the expected contact with the mobile target node. By selecting the stationary node with the fewest possible forwarding hops as the relay node, the transmission times can be reduced which is beneficial for prolonging network lifetime. Thus, the remaining problem is to list candidates from stationary nodes based on the mobility information.

In Figure 4.3 and 4.4, we show the limitation of a linear trajectory model and how our proposed fuzzy path model overcomes it. To evaluate the performance of the relay selection scheme, synthetic traces are generated (one trace for each delivery task) and the trace generation is repeated for different SDTA values (referring to Table 3-B). For

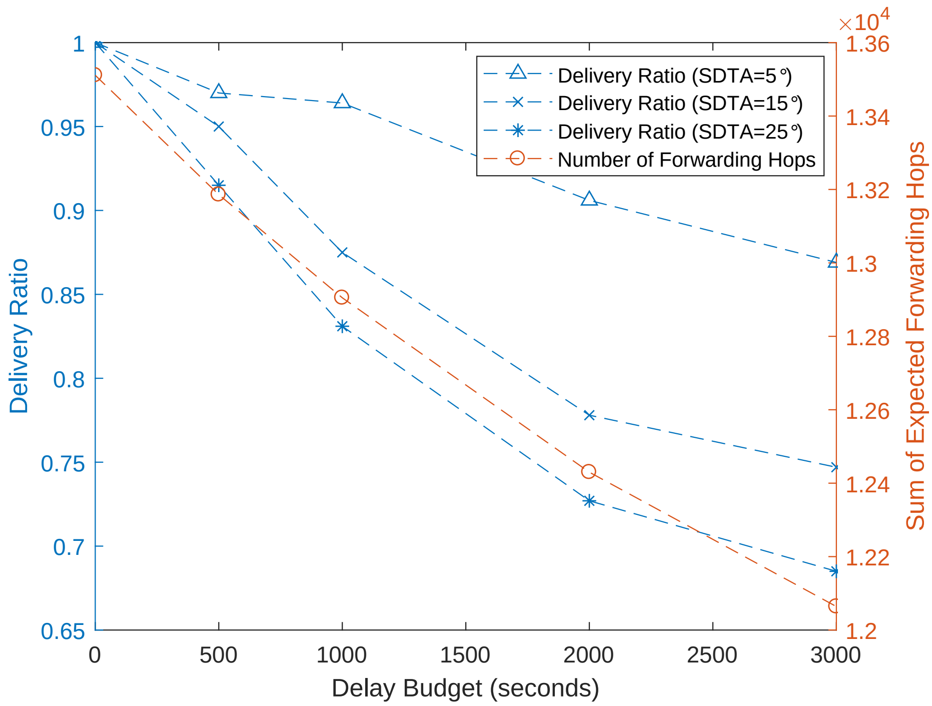
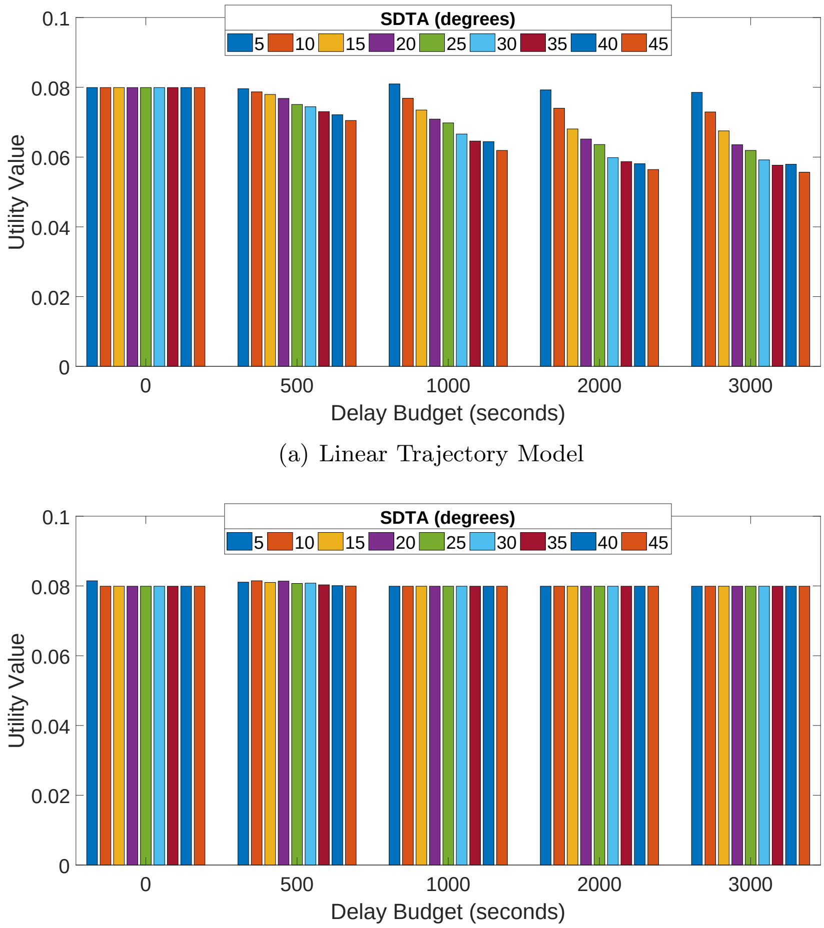


Figure 4.3: Effect of Delay Budget on Linear Trajectory Model

delivery without delay tolerance (i.e. the delay budget equals zero in Figure 4.3), the stationary relay node is selected from the immediate neighbouring nodes identified given the latest position of the mobile node. With increasing budget of delivery delay (500, 1000, 2000, and 3000 seconds in Figure 4.3), more stationary nodes can be listed as candidates by predicting contact based on the linear trajectory model and therefore the number of forwarding hops can be reduced. However, the delivery ratio[2] will decrease at the same time, which worsens with higher SDTA because of the lower persistence of directional movement. Consequently, the delivery utility (as defined in equation (4.5)) will decrease with increasing SDTA especially when the delay budget is large, which is shown in Figure 4.4(a). In contrast, the achievable utility is nearly constant over varying SDTA even for a large delay budget (as shown in Figure 4.4(b)), because the proposed fuzzy path model can actively adapt to different degrees of directional correlation.

In Figure 4.5, more details of the comparison are given for a relative large delay

[2]The percentage of delivery tasks which are accomplished within delay budget.



(a) Linear Trajectory Model

(b) Fuzzy Path Model

Figure 4.4: Comparison of Delivery Utility for Different Values of Delay Budget and SDTA

budget (3000 seconds in this case). For the scheme without prediction, only immediate neighbouring nodes can be selected as the relay and therefore the utility is constant and does not change with SDTA, which overlooks the opportunity of exploiting node movement. For the scheme based on a linear trajectory model, the utility decreases with increasing SDTA due to lower delivery ratio caused by more randomness in the movement. Due to the large delay budget, the linear trajectory model is no longer accurate for contact prediction so the achieved utility is always worse than the scheme without prediction. With the proposed fuzzy path model, higher delivery utility can be achieved because of its dynamic reaction to the directional correlation. As shown in Figure 4.5, the achievable delivery ratio can be adjusted flexibly based on the threshold

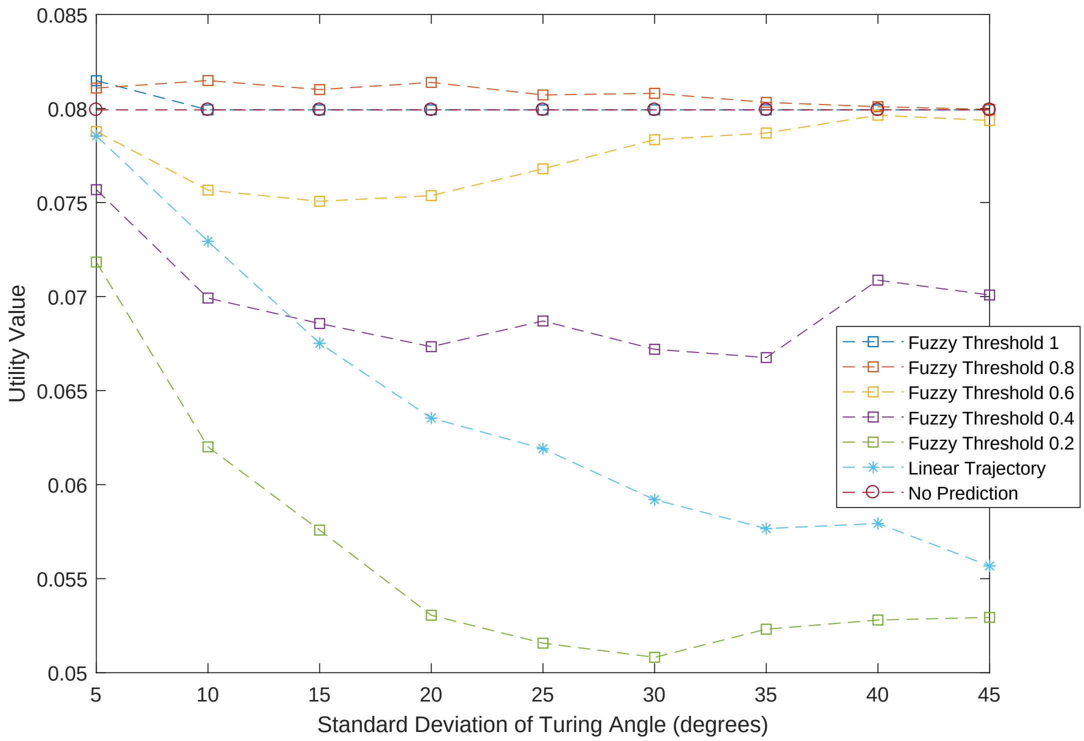


Figure 4.5: Comparison of Delivery Utility for Different Values of SDTA (Setting Delay Budget to 3000 Seconds)

of certainty degree. By setting a higher threshold, the utility achieved by the fuzzy path model approximates the scheme without prediction, due to lower allowed uncertainty. The results show that the delivery utility achieved by the proposed fuzzy path model outperforms the linear trajectory model when the threshold value is no less than 0.6. For the performance at the threshold value of 0.6, the delivery utility increases after a certain SDTA value (15 degrees in this case) because fewer candidates can be identified. With a relative high threshold (no less than 0.8 in this case), the performance is always no worse than the scheme without prediction. Even when the threshold value equals 1.0, a higher utility can still be achieved when the SDTA is small (5 degrees in this case) as directivity of movement is strong enough to be exploited. Besides, it is shown that the mobility with a relative small value of SDTA (less than 30 degrees in this case) the directional movement can be exploited to facilitate packet forwarding.

In summary, our proposal advances the state-of-the-art because the directional correlation of target node movement is considered, allowing node mobility to be dynamically exploited for the optimal selection of the stationary relay node. Simulation results show that higher delivery utility can be achieved by the proposed fuzzy path model compared with a forwarding scheme without contact prediction or one based on linear trajectory prediction. In conclusion, the directional movement is exploitable for the dissemination of delay-tolerant information especially when moderate uncertainty is allowed. Further research can be conducted to devise appropriate forwarding strategies for different scenarios. Above contents have been published in the journal paper "Movement-Aware Relay Selection for Delay-Tolerant Information Dissemination in Wildlife Tracking and Monitoring Applications".

4.3 Relay Selection for Group Delivery

For the case of group delivery, a relay node may serve multiple target nodes. Therefore, this research case focuses on finding optimised relay(s) to reduce overall forwarding while retaining satisfaction of group delivery. In this section, solutions are proposed for the formulated problem with performance evaluation.

4.3.1 Scenario Model

Consider a packet from the sink node to a group of target nodes (denoted by $\mathcal{N}_{\text{tgt}}$). In previous works such as [CYY16], the delivery is simply required to be finished before a maximum acceptable delay is reached. However, for the dissemination of control information, there exists more-or-less flexibility regarding acceptable delay, which brings optimisation opportunities that are overlooked by existing works. To this end, our research describes the delay sensitivity by a two-level budget model:

- The first-level budget (called the main budget) has a soft boundary and any earlier delivery is considered to be a fully satisfied delivery;

- The second-level budget (called the extra budget) has a firm boundary and any later delivery is considered to be a fully unsatisfactory delivery.

- For the delivery delay between these two boundaries (i.e. within the range of the extra budget), it contributes to decreasing satisfaction of delivery.

This two-level budget model describes the relationship between delivery satisfaction and delivery delay, called a delay budget function (denoted by f_{db}). Theoretically, the delay budget function can be any type of monotonically decreasing function. In this thesis, we focus on the case of a sigmoid curve which provides a smooth transition to reflect a realistic tolerance of acceptable delay. Thus, the raw delay budget function is given as

$$f^*_{\mathrm{db}}(\tau) = \frac{1}{1 + e^{\varpi\left(\tau - \frac{\tau_1+\tau_2}{2}\right)}} \tag{4.6}$$

where $\tau \geqslant 0$ denotes the delivery delay[3], $\tau_1 > 0$ denotes the first-level budget boundary, $\tau_2 > \tau_1$ denotes the second-level budget boundary, and ϖ denotes a functional factor.

Given τ_1 and τ_2, the functional factor ϖ can be calculated as

$$\varpi = \frac{2\ln\left((1-\varepsilon)^{-1} - 1\right)}{\tau_1 - \tau_2} \tag{4.7}$$

where ε is a constant that satisfies $\varepsilon > 0$ and $\varepsilon \neq 1$. The smaller ε brings more accuracy because the budget function can have less decrement before τ_1. The value of ε is set to be 0.001 which is sufficiently accurate for our research.

Let the delivery satisfaction be a degree value between 0 and 1, so the budget function is further normalised as

$$f_{\mathrm{db}}(\tau) = \frac{1 + e^{-\varpi\left(\frac{\tau_1+\tau_2}{2}\right)}}{1 + e^{\varpi\left(\tau - \frac{\tau_1+\tau_2}{2}\right)}}. \tag{4.8}$$

[3] As the delivery delay is dominated by the waiting time at the stationary relay, the transmission delay is considered to be negligible in our research, and the case of zero delay represents direct delivery without waiting.

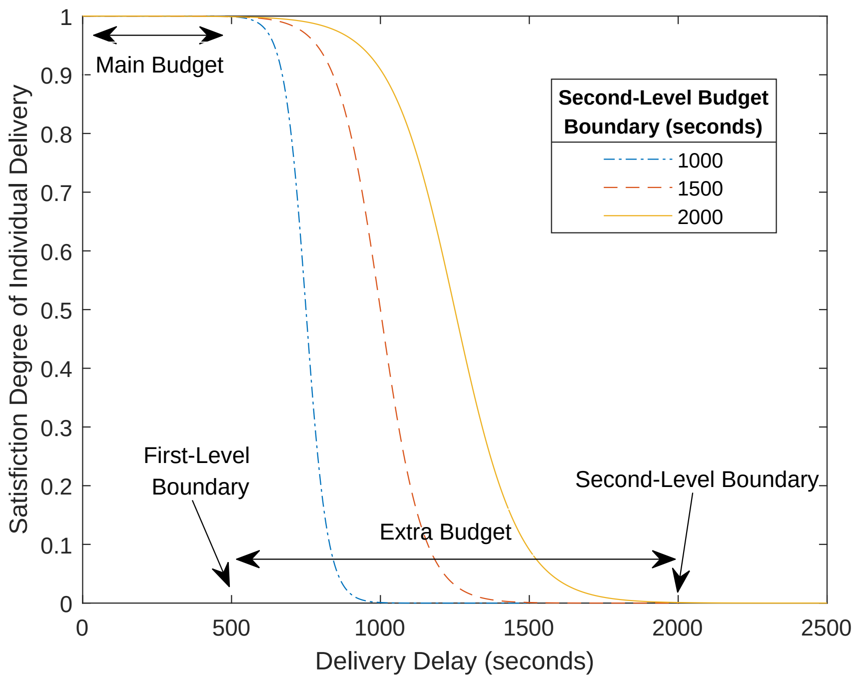


Figure 4.6: Delay Budget Function When the Main Budget is 500 Seconds

In Figure 4.6, the delay budget function is illustrated for a case where the main budget is 500 seconds. This figure shows that the delivery satisfaction starts to decrease from the first-level budget boundary and becomes zero at the second-level budget boundary. With such a transition range, a larger delivery delay is still allowed (but less encouraged) even if the main budget is exceeded, which provides more flexibility in relay selection. For instance, the possibility of 510 seconds delay can be considered for making a possible better offer in other respects (e.g. much fewer forwarding hops). The optimisation opportunity brought by this flexible delivery delay requirement will be further investigated in the remainder of this section.

4.3.2 Problem Formulation with Proposed Solutions

Given the underlying schemes that are already described in Section 4.3.1, the research problem now focuses on the optimisation of the relay strategy. The relay strategy is considered to be a decision (made by the sink node) that determines relay selection and target assignment for achieving a group delivery. Based on knowledge of predictable node contact and tolerable delivery delay, how to identify feasible opportunities for relay sharing is the key issue to be addressed. In this section, the research problem is formulated and solutions are proposed.

4.3.2.1 Problem Formulation

Two optimisation objectives are considered to formulate the research problem: 1) maximise the cost reduction of group delivery; 2) maximise the requirement satisfaction of group delivery. To reflect potential conflicts between these two objectives, this optimisation problem is required to be formally formulated for further analysis.

Let x_{ij} denote a binary indicator of the target assignment decision, such that

$$x_{ij} = \begin{cases} 1 & \text{if relay } n_i \text{ is responsible for target } n_j \\ 0 & \text{otherwise.} \end{cases} \tag{4.9}$$

where $n_i \in \mathcal{N}_{\text{rly}}^{\text{c}}$ denotes a candidate relay node and $n_j \in \mathcal{N}_{\text{tgt}}$ denotes a target node. For the sake of simplicity, it is assumed that all the candidate relay nodes (i.e. $\mathcal{N}_{\text{rly}}^{\text{c}}$ where the superscript c denotes they are candidates to be selected from) are stationary nodes and all the target nodes (i.e. $\mathcal{N}_{\text{tgt}}$) are mobile nodes in the remainder of this section.

Then a relay strategy matrix (denoted by X) can be represented as

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1J} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2J} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{I1} & x_{I2} & x_{I3} & \dots & x_{IJ} \end{bmatrix} = (x_{ij})_{I\times J} \tag{4.10}$$

where I denotes the number of candidate relay nodes, and J denotes the number of delivery target nodes.

For the strategy matrix X, each row represents the relay node selection for a specific target node n_j and each column represents the target node assignment for a specific candidate relay node n_i. The whole set of selected relays can be represented as

$$\mathcal{N}_{\text{rly}} = \{n_i | x_{ij} = 1\} \ , n_j \in \mathcal{N}_{\text{tgt}}. \tag{4.11}$$

Note that multiple target nodes may have the same relay node, which is equivalent to sharing a relay node for a subgroup of targets. The subgroup of targets which are assigned to n_i can be represented as

$$\mathcal{N}_{\text{tgt}|n_i} = \{n_j | x_{ij} = 1\} \ , n_i \in \mathcal{N}^{\text{c}}_{\text{rly}}. \tag{4.12}$$

Under the relay strategy X, let C^*_X denote the overall reduction[4] of forwarding hops and D^*_X denote the overall satisfaction[5] of delivery requirement, i.e.

$$C^*_X = H - \sum_{n_i \in \mathcal{N}_{\text{rly}}} f_{\text{hd}}(n_i) \tag{4.13}$$

$$D^*_X = \sum_{n_j \in \mathcal{N}_{\text{tgt}}} \left(f_{\text{cc}}(\vec{\rho}_{ij}) \cdot f_{\text{db}}(\tau_{ij}) \right) \tag{4.14}$$

where H denotes the sum of least forwarding hops for group relay selection based on immediate delivery (i.e. no movement prediction), f_{db} denotes the delay budget function, and f_{hd} denotes the hop distance (from the sink node). f_{cc} denotes the contact certainty function (defined in equation (3.35)) and $\vec{\rho}_{ij} = (r_{ij}, \varphi_{ij})$ denotes the relative position between stationary node n_i and mobile node n_j. The estimation of the contact delay between n_i and n_j (denoted by τ_{ij}) is based on the minimum required steps.

[4]No multicast tree construction is considered for this thesis but further research could be conducted beyond this basic case.

[5]It is assumed that each individual delivery independently contributes to the group delivery, so the overall satisfaction is a sum of the individual satisfactions.

Let C_X^* and D_X^* be normalised as a degree value between 0 and 1 such that

$$C_X = \frac{H - \sum_{n_i \in \mathcal{N}_{\mathrm{rly}}} f_{\mathrm{hd}}(n_i)}{H} \tag{4.15}$$

$$D_X = \frac{\sum_{n_j \in \mathcal{N}_{\mathrm{tgt}}} \left(f_{\mathrm{cc}}(\vec{\rho}_{ij}) \cdot f_{\mathrm{db}}(\tau_{ij})\right)}{J} \tag{4.16}$$

where J denotes the number of targets, C_X is called the reduction degree of overall cost, and D_X is called the satisfaction degree of group delivery.

Given all the possible relay strategy matrices $X \in \Omega$ (where Ω denotes the whole space of feasible solutions), the problem is to find a suitable strategy considering two optimisation objectives: maximise the reduction degree of overall cost (i.e. C_X); maximise the satisfaction degree of group delivery (i.e. D_X).

4.3.2.2 Group Relay Selection (Design of Evolutionary Approach)

To solve the formulated problem, an approach based on the evolutionary algorihtm is proposed here. As discussed in Section 2.5, the Multi-Objective Genetic Algorithm (MOGA) [DK01] is adopted because it is a standard evolutionary algorithm which is suitable for solving combinatorial optimisation problems (as formulated in Section 4.3.2). Based on the MOGA [DK01] implementation[6], the group relay selection is considered from the following three aspects:

A. Population The main issue to be addressed in the evolutionary approach design is to reduce the solution space for better and faster searching. Without such a reduction, it is hard to obtain a suitable solution in a limited time due to the existence of too many possibilities.

Therefore, instead of using the strategy matrix X, a chromosome is proposed to be

[6]The algorithm follows default settings of the standard implementation provided by the MATLAB Global Optimization Toolbox unless otherwise stated.

a binary selection vector (denoted by $\vec{l}$) as

$$\vec{l} = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_I \end{bmatrix}. \tag{4.17}$$

Then a combination of selected relays can be indicated by

$$\mathcal{N}_{\text{rly}} = \{n_i | l_i = 1\}, n_i \in \mathcal{N}^{\text{c}}_{\text{rly}}. \tag{4.18}$$

Let $\mathcal{L}$ denote a population which consists of a number of chromosomes. An initial population (generation zero) is generated randomly[7] as

$$\mathcal{L}^{(0)} = \left\{\vec{l}_1, \vec{l}_2, \ldots, \vec{l}_L\right\} \tag{4.19}$$

where L denotes the number of chromosomes in the population.

B. Calculation Although the search space can be reduced by using the selection vector $\vec{l}$, a strategy matrix is still required to evaluate the chromosome fitness (i.e. the two optimisation objectives). For generating the strategy matrix from a selection vector, a default assignment matrix is defined as $\hat{X} = (\hat{x}_{ij})_{I \times J}$ such that,

$$\hat{x}_{ij} = \begin{cases} 1 & \text{if } f_{\text{cc}}\left(\vec{\rho}_{ij}\right) \cdot f_{\text{db}}\left(\tau_{ij}\right) > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{4.20}$$

which means that each target node is assigned to all feasible relay candidates.

Given a selection vector $\vec{l}$, a corresponding raw strategy matrix based on the default

[7]For later evaluation, the solutions found by four greedy schemes are included in the initial population to assess the searching performance.

assignment can be generated as

$$X^* = \left(x^*_{ij}\right)_{I\times J} = \vec{l}\cdot \hat{X}. \tag{4.21}$$

Note that the raw strategy matrix brings redundant assignments (i.e. a target node can be assigned to more than one relay node). Although this redundancy can improve the delivery reliability, our research focuses on relay sharing opportunities to reduce the delivery cost and therefore a basic case is considered such that only one target is assigned to one relay (but further research could be conducted beyond this basic case).

Algorithm 1 Raw Strategy Matrix Adjustment and Evaluation

1: **input:** $n_i \in \mathcal{N}_{\text{rly}}^{\eta|n_j}$, $n_j \in \mathcal{N}_{\text{tgt}}$, X^*
2: $(x_{ij})_{I\times J} \leftarrow (0)_{I\times J}$ //initialise an I-by-J matrix of zeros
3: **for** each n_j **do**
4:   $var \leftarrow 0$ //initialise a variable for finding maximum
5:   **for** each n_i **do**
6:     **if** $x^*_{ij} = 1$ **and** $f_{\text{cc}}\left(\vec{\rho}_{ij}\right)\cdot f_{\text{db}}\left(\tau_{ij}\right) > var$ **then**
7:       record n_i as the best relay for n_j
8:       $var \leftarrow f_{\text{cc}}\left(\vec{\rho}_{ij}\right)\cdot f_{\text{db}}\left(\tau_{ij}\right)$ //record a bigger value
9:     **end if**
10:   **end for**
11:   $x_{i'j} \leftarrow 1$, where $n_{i'}$ denotes the recorded best relay for n_j // the target is assigned to only one best relay
12: **end for**
13: $X \leftarrow (x_{ij})_{I\times J}$ // obtain the strategy matrix
14: calculate C_X based on Equation (4.15)
15: calculate D_X based on Equation (4.16)
16: **return** X, C_X, D_X

Based on Algorithm 1, the raw strategy matrix is adjusted to reduce the redundancy and the corresponding fitness values can be calculated. From line 3 to line 12, one best relay node is recorded for each target node. These recorded relay(s) generates the strategy matrix in line 13. In this way, the redundant assignments can be eliminated in the final strategy matrix (which is equivalent to adjusting a raw strategy matrix) and then the fitness values are calculated in line 14 and 15.

C. Iteration Given an initial population and the specified fitness calculation, the algorithm leads later generations to evolve towards the optimal solutions iteratively. The production of a new generation follows the general procedure as:

- select parents from the current population;
- create children by using genetic operators (crossover / mutation) on selected parents;
- calculate fitness using the Algorithm 1 and eliminate inferior chromosomes to retain a fixed size of population.

The iterative mechanism finally terminates when the process reaches a specified generation/time limit or there is sufficient convergence in the solutions. A set of Pareto optimal solutions is provided when the algorithm terminates along with their achieved fitness values.

4.3.2.3 Group Relay Selection (Design of Greedy Approach)

Considering evolutionary algorithms always require iterative computation which can be inefficient, another approach is proposed here based on the greedy algorithm as an alternative.

In [CYY16], a greedy algorithm is proposed for trajectory-based group message delivery. With the relay selection performed by this algorithm, relay nodes can be shared for group delivery so that the overall forwarding hops can be reduced for a given delay budget. As shown in Algorithm 2, the algorithm proposed by [CYY16] solves the problem of group relay selection in two phases:

- For the first phase (from line 7 to line 13), the serving list is identified as a list of targets who can be served by a certain stationary relay.
- For the second phase (from line 14 to line 26), one stationary relay is selected each

round until the serving lists cover all mobile targets.

Finally, a strategy matrix can be obtained in line 27 as the output of the algorithm.

Algorithm 2 Greedy Algorithm for Group Relay Selection

1: **input:** $n_i \in \mathcal{N}_{\text{rly}}^{\eta|n_j}$, $n_j \in \mathcal{N}_{\text{tgt}}$, $\hat{\tau}$
//Initialisation
2: $\mathcal{S}_{\text{ed}} \leftarrow \emptyset$ //initialise an empty set of covered targets
3: $(x_{ij})_{I \times J} \leftarrow (0)_{I \times J}$ //initialise an I-by-J matrix of zeros
4: **for** each pair of n_i and n_j **do**
5: $\quad \mathcal{S}_{ij} \leftarrow \{n_j\}$ //initialise its serving list
6: **end for**
//The first phase: serving list identification
7: **for** each pair of n_i and n_j **do**
8: $\quad$ **for** each $n_{j'} \neq n_j$ **do**
9: $\quad\quad$ **if** $\tau_{ij} \leqslant \tau_{ij'}$ **and** $\tau_{ij'} \leqslant \hat{\tau}$ **then**
10: $\quad\quad\quad \mathcal{S}_{ij} = \mathcal{S}_{ij} \cup n_{j'}$ //add this target node to the list
11: $\quad\quad$ **end if**
12: $\quad$ **end for**
13: **end for**
//The second phase: group relay selection
14: **while** $|\mathcal{S}_{ed}| < |\mathcal{N}_{tgt}|$ **do**
15: $\quad var \leftarrow \infty$ //initialise a variable for finding minimum
16: $\quad \mathcal{S}_{\text{new}} \leftarrow \emptyset$ //an empty list to record newly covered targets
17: $\quad$ **for** each pair of n_i and n_j **do**
18: $\quad\quad \mathcal{S}_{\text{diff}} \leftarrow ((\mathcal{S}_{ij} \cup \mathcal{S}_{\text{ed}}) - \mathcal{S}_{\text{ed}})$ //find newly covered targets
19: $\quad\quad$ **if** $f_{\text{mc}}(n_i, \mathcal{S}_{\text{diff}}) < var$ **then**
20: $\quad\quad\quad$ record n_i as the relay to be selected this round
21: $\quad\quad\quad var \leftarrow f_{\text{mc}}(n_i, \mathcal{S}_{\text{diff}})$, $\mathcal{S}_{\text{new}} \leftarrow \mathcal{S}_{\text{diff}}$
22: $\quad\quad$ **end if**
23: $\quad$ **end for**
24: $\quad \mathcal{S}_{\text{ed}} \leftarrow \mathcal{S}_{\text{ed}} \cup \mathcal{S}_{\text{new}}$ //update the set of covered targets
25: $\quad x_{i'j} \leftarrow 1$ for $n_j \in \mathcal{S}_{\text{new}}$, where $n_{i'}$ denotes the recorded relay to be selected this round
26: **end while**
27: $X \leftarrow (x_{ij})_{I \times J}$ // obtain the strategy matrix
28: **return** X

Our algorithm design bears the same greedy algorithm principle as in [CYY16]. However, it should be noted that the algorithm in [CYY16] is not directly applicable to our research scenario. Firstly, accurate trajectory knowledge assumed by [CYY16] is unavailable so the serving list cannot be explicitly known. In addition, an exact budget boundary may not exist (due to the existence of the extra budget) and this situation is not considered by [CYY16]. To this end, our research provides two variants of the greedy algorithm:

- The method of Delay Boundary based Relay Selection (DBRS) is proposed for the scenario where mobile nodes randomly move with directional correlation. The DBRS is considered as the performance benchmark because it has no awareness of the extra delay budget as with the method proposed in [CYY16].
- A new method of Satisfaction Degree based Relay Selection (SDRS) is proposed to consider the extra delay budget in group relay selection. Based on the design of DBRS, the SDRS has additional awareness so better performance can be expected.

A. DBRS Design The main issue to be addressed in the DBRS design is the identification of the serving list in the first phase of the problem solution. Without accurate trajectory knowledge, it becomes hard to identify a list of targets who can be served by a certain stationary relay.

To overcome the difficulty brought about by the randomness of node movement, a threshold (denoted by η) of contact certainty is proposed to qualify a set of candidate stationary relays as

$$\mathcal{N}_{\text{rly}}^{\eta|n_j} = \{n_i | f_{\text{cc}}\left(\vec{\rho}_{ij}\right) \geqslant \eta\} \ , n_i \in \mathcal{N}_{\text{rly}}^{\text{c}}. \tag{4.22}$$

With the qualified candidates, a serving list (denoted by $\mathcal{S}_{ij}$) can be found for each pair of Stationary Relay n_i and Mobile Target n_j (called the SRMT pair) according to the condition that

$$\mathcal{S}_{ij} = \left\{n_{j'} | \tau_{ij} \leqslant \tau_{ij'} \leqslant \hat{\tau}\right\} \ n_j \in \mathcal{N}_{\text{tgt}}, n_{j'} \in \mathcal{N}_{\text{tgt}} \tag{4.23}$$

where $\hat{\tau}$ denotes a budget boundary to identify a serving list.

Then, the second phase of problem solution starts from an empty set of covered targets. After each evaluation round, one stationary relay will be selected and the number of covered targets will increase (added from corresponding serving list). The procedure of relay selection ends when the selected relays can cover all mobile targets.

Given the current set of covered targets, each SRMT pair can be evaluated by a cost metric (the hop distance averaged by the number of newly covered targets). The calculation of this cost metric is same as the cost calculation proposed by [CYY16] as

$$f_{\text{mc}}^{\text{DB}}(n_i, \mathcal{S}_{\text{diff}}) = \frac{f_{\text{hd}}(n_i)}{|\mathcal{S}_{\text{diff}}|} \tag{4.24}$$

where $f_{\text{mc}}^{\text{DB}}$ denotes the metric calculation used in DBRS, $f_{\text{hd}}(n_i)$ denotes the hop distance from the sink node to n_i, $\mathcal{S}_{\text{diff}}$ denotes a list of newly covered targets (obtained from line 18 in Algorithm 2), and $|\mathcal{S}_{\text{diff}}|$ denotes the size of this list.

Note that this DBRS scheme relies on the assumption that all deliveries before a given delay $\hat{\tau}$ contribute the same degree of delivery satisfaction. However, in our research scenario, the satisfaction degree gradually decreases after the first-level budget boundary is exceeded. Without this awareness in the algorithm, optimisation opportunities can be overlooked and therefore the SDRS design is also proposed.

B. SDRS Design The main issue to be addressed in the SDRS design is the consideration of the extra delay budget in the second phase of the problem solution. During the first phase, it is similar to our DBRS design except that the budget boundary is fixed to the second-level budget boundary (i.e. $\hat{\tau} = \tau_2$) to provide all possible candidate stationary relays.

After the first phase, all the contactable targets within the second-level budget boundary are included in the serving list. Thus the cost metric should reflect the difference in delivery satisfaction. Instead of treating all the newly served targets the same, the hop distance is averaged by the sum of their satisfaction degree as

$$f_{\mathrm{mc}}^{\mathrm{SD}}(n_i, \mathcal{S}_{\mathrm{diff}}) = \frac{f_{\mathrm{hd}}(n_i)}{\sum_{n_j \in \mathcal{S}_{\mathrm{diff}}} f_{\mathrm{db}}(\tau_{ij})} \tag{4.25}$$

where $f_{\mathrm{mc}}^{\mathrm{SD}}$ denotes the metric calculation used in SDRS and f_{db} denotes the delay budget function.

Compared with DBRS, a more generalised form of cost calculation is provided by SDRS. Because the delivery satisfaction is considered in relay selection, SDRS can make improved decisions and therefore better delivery performance can be achieved. In the next section, DBRS and SDRS will be compared and evaluated, together with an evolutionary approach.

4.3.3 Simulation and Evaluation

In Section 4.3.2, the research problem is formulated and solutions are proposed to select relay node(s) for group delivery. Now, we further evaluate the performance of proposed solutions in different simulation scenarios.

4.3.3.1 Scenario Description

For WSN applications such as habitat monitoring [CEE$^+$01], it is quite common that low-cost wireless nodes are densely deployed over a geographical area. Figure 4.7 shows a scenario where a number of stationary nodes (denoted by squares) are deployed within a rectangular region. A base scenario is defined for our simulations where stationary nodes are uniformly distributed over 10000*10000 metre area as a 21*21 grid.

It is assumed that at a certain time point during network operations (treated as the current time), a packet is planned to be delivered to a group of mobile nodes (denoted by circles) which are randomly located within the base scenario area. These mobile nodes represent the sensors deployed on moving objects such as animals in the application

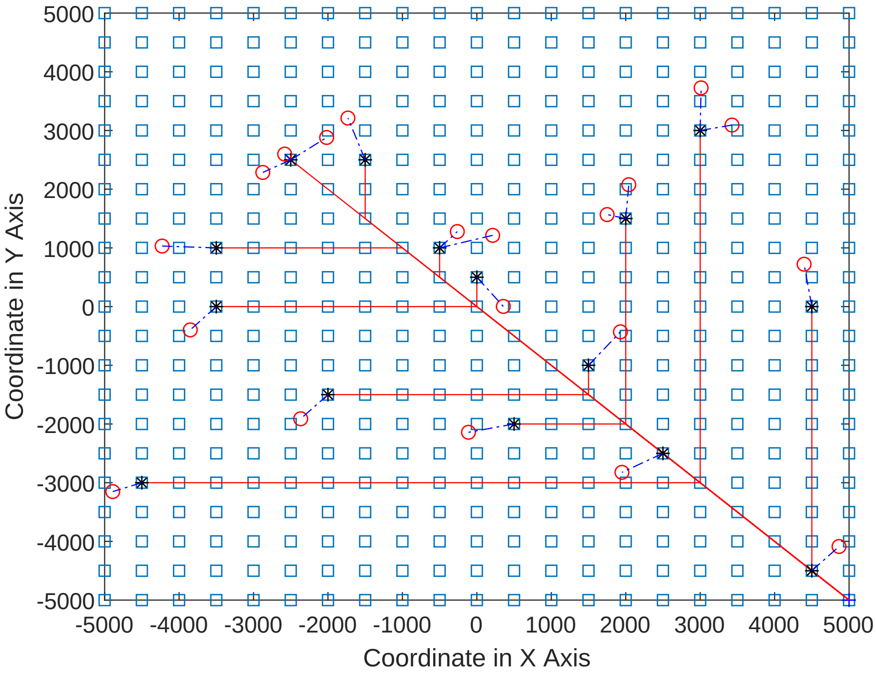


Figure 4.7: Example Simulation Scenario (Relay Sharing without Movement Prediction)

scenario. For the sake of simplicity, it is assumed that the packet source (i.e. the sink node) is the stationary node at the bottom right corner.

In our research, radio coverage is considered to be limited (e.g. subject to device size and/or surrounding environment such as forest or underwater) and therefore the packet delivery mainly relies on multi-hop forwarding. It is assumed that the stationary nodes can form a grid topology to perform forwarding. Considering the base scenario used for our simulations, the communication radius of wireless nodes is set to be 750m (which is a feasible range referring to [SMZ07]) so that each stationary node is able to contact its neighbour nodes.

It is typically possible to achieve group packet delivery with fewer hops by strategic forwarding. Instead of immediate delivery to each target, the packet can be forwarded (via routes shown as solid lines) to the selected relay nodes (denoted by asterisks) and wait to be delivered when the target comes within the relay node coverage (indicated by dash-dot lines). By exploiting predictable contact and delay budget, the opportunity of relay sharing can be brought to further reduce the overall forwarding hops. For the instance shown in Figure 4.8, only ten relay nodes are used for twenty targets (instead of fifteen relay nodes used in Figure 4.7) and therefore the overall forwarding hops can be reduced, by exploiting movement prediction in relay selection.

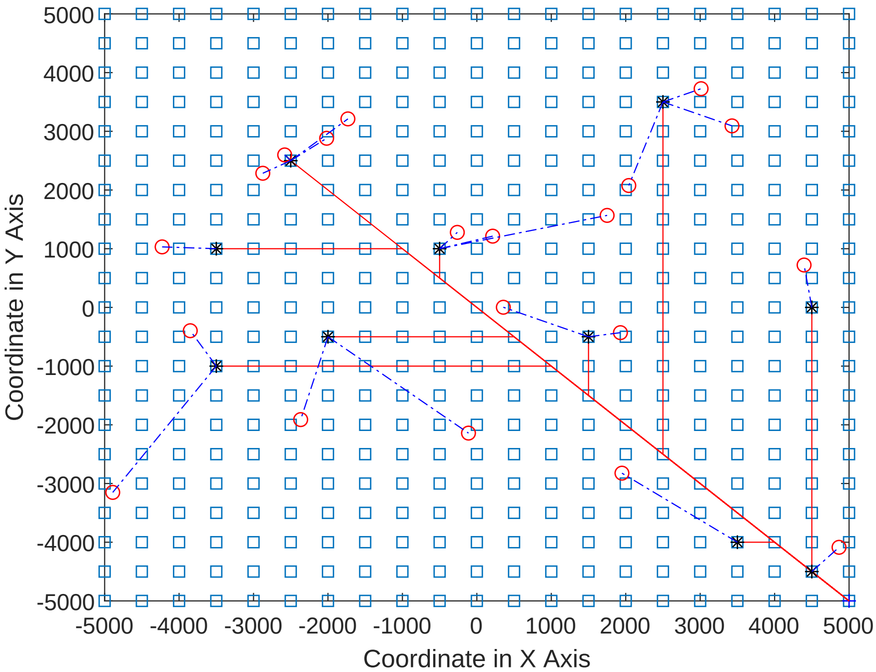


Figure 4.8: Example Simulation Scenario (Relay Sharing with Movement Prediction)

Due to the contact uncertainty caused by movement randomness, proper selection of relay nodes is necessary to ensure delivery satisfaction. For the example simulation scenario (as shown in Figure 4.7), the trade-off between two objectives (defined in Equation (4.15) and (4.16)) can be reflected by the plot of the Pareto front (as shown in Figure 4.9) which is obtained by MOGA (i.e. the evolutionary approach). Figure 4.9 shows that the cost reduction can firstly be improved without much loss of delivery satisfaction but then the delivery satisfaction will degrade rapidly with little further increase in the cost reduction. This means a further reduction of the delivery cost (beyond a certain point) notably impacts on the delivery satisfaction and the two objectives are then indeed conflicted.

To further evaluate the overall delivery performance, the weighted sum of the two objectives is calculated and shown in Figure 4.10. For an easier understanding, it is worth noting that Figure 4.10 transforms the two value of objectives into one weighted sum value. So the two-dimensional plot shown by Figure 4.9 becomes multiple one-dimensional plots in Figure 4.10, given different weights of the cost reduction (the weight of the delivery satisfaction can be known at the same time as the sum of weights given as 1). Figure 4.10 shows the limit of overall delivery performance is a curve and there is a concave region around the value where the two weights are similar. This concave region indicates the two objectives cannot be simultaneously achieved, which results in a performance limit. It is also shown that the concave region is biased to the higher weight for cost reduction, which means it is harder to achieve cost reduction than the delivery satisfaction.

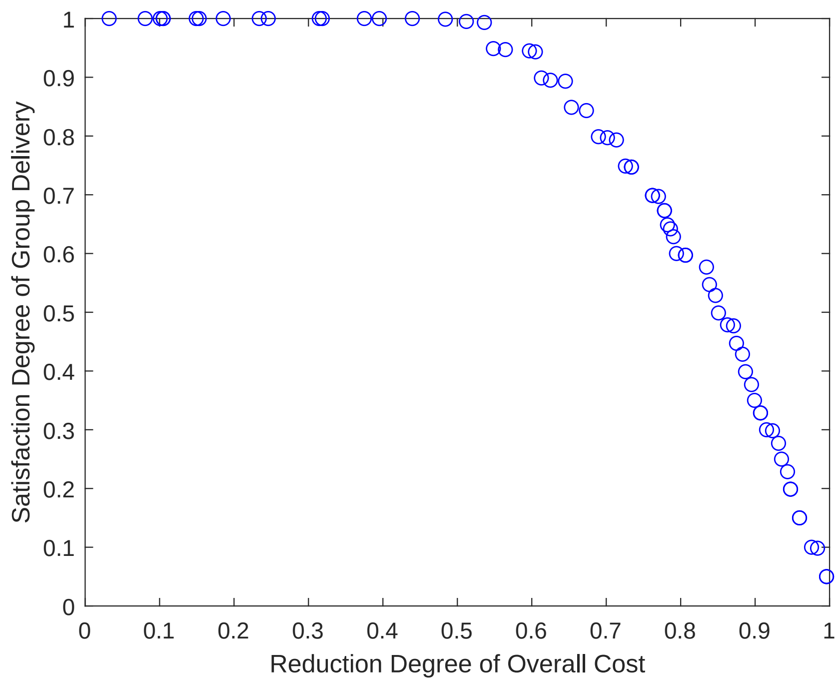


Figure 4.9: Trade-Off between Cost Reduction and Delivery Satisfaction Shown by the Pareto Front

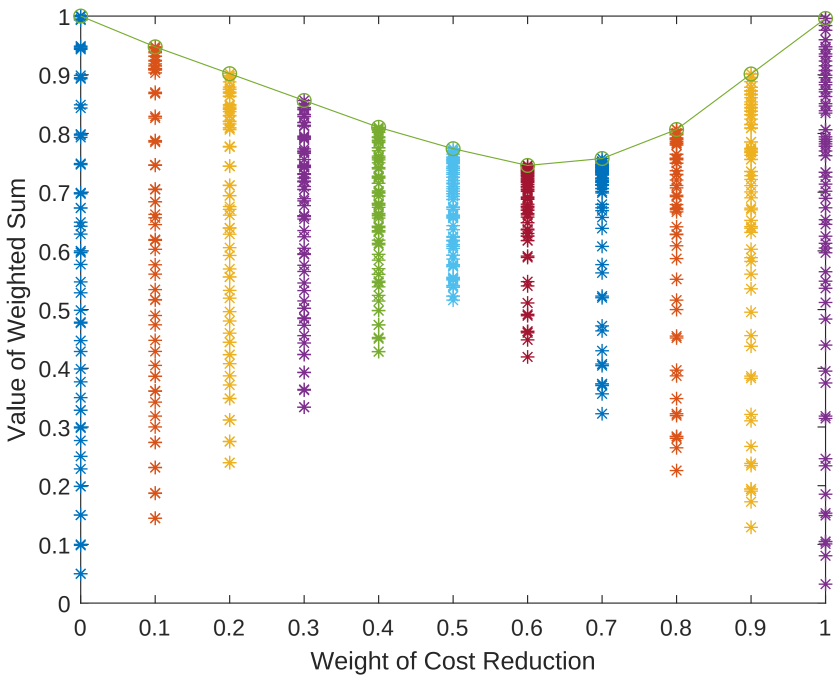


Figure 4.10: Weighted Sum of the Pareto Front over Different Weight Values

4.3.3.2 Performance Analysis

To evaluate the delivery performance, simulation results are obtained from random scenarios for analysis. As described in Section 4.3.3.1, the scenario randomness comes from the mobile targets (i.e. the initial location and direction are randomly chosen for each simulation run). The obtained results are the average of multiple simulation runs (50 random seeds used).

The two objectives defined in Section 4.3.2 are considered as the performance metrics. As there is a trade-off between cost reduction and delivery satisfaction, these two metrics are considered together for evaluation. Furthermore, the weighted sum of the two objectives is introduced as an additional metric to reflect the overall performance given the objective preference.

The following analysis is divided into three parts. For the first part, simulation settings are investigated. For the second part, the two greedy schemes (i.e. DBRS and

SDRS) are compared. For the third part, the performance of greedy schemes is evaluated relative to the evolutionary approach.

A. Simulation Settings This part investigates how the simulation is influenced by three related factors, namely the certainty threshold η, the Standard Deviation of the Turning Angles (SDTA) σ, and the target node number J. The extra budget is set to 0 for this part so there is no difference between DBRS and SDRS. The main budget is the control variable and the default value for other settings are $\eta = 0.9$, $\sigma = 5$, and $J = 20$ (obtained from our investigations).

Certainty Threshold With different values of certainty threshold, Figure 4.11 shows that a lower threshold does not improve the cost reduction (i.e. the C_X defined in Equation (4.15)) much but sightly degrades the delivery satisfaction (i.e. the D_X defined in Equation (4.16)). Therefore the default value of certainty threshold is set to 0.9 for the remaining simulations unless otherwise stated.

Target Node Number With different numbers of target nodes, Figure 4.12 shows that the cost reduction increases with number of targets. Given the number of target nodes, higher reduction degree can be achieved by a larger delay budget, because there are more opportunities of relay sharing. The satisfaction degree is relatively steady but reflects a similar trend, due to the trade-off between the cost reduction and delivery satisfaction. As our research focus is on group delivery, a relative large group is considered and the default value of target number is set to 20 for the remaining simulations unless otherwise stated.

SDTA Figure 4.13 shows that different values of the main budget can lead to different delivery performance especially when SDTA is relatively small. With increasing SDTA, the influence of the main budget decreases as the movement is less predictable for

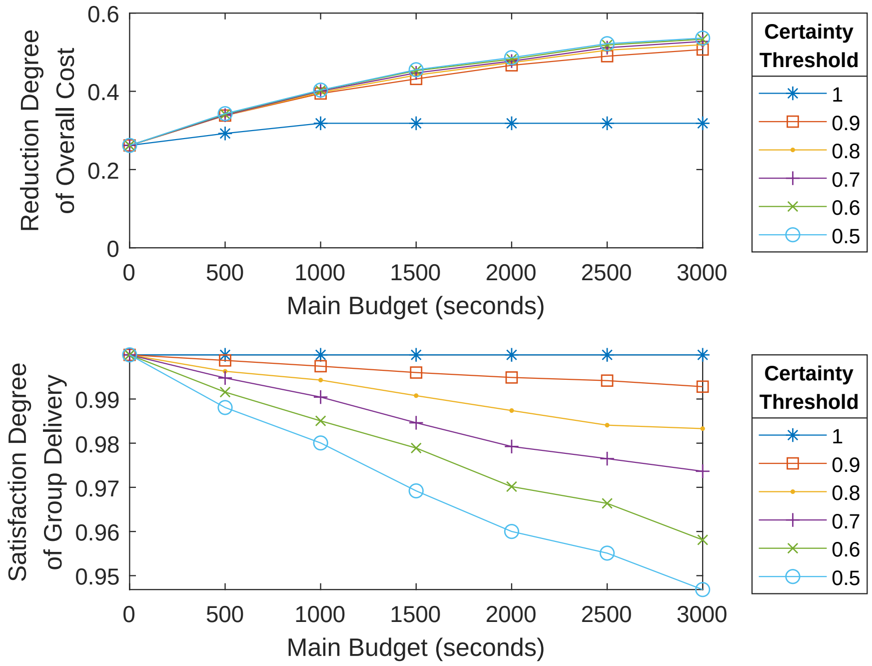


Figure 4.11: Investigation of Certainty Threshold (No Extra Budget)

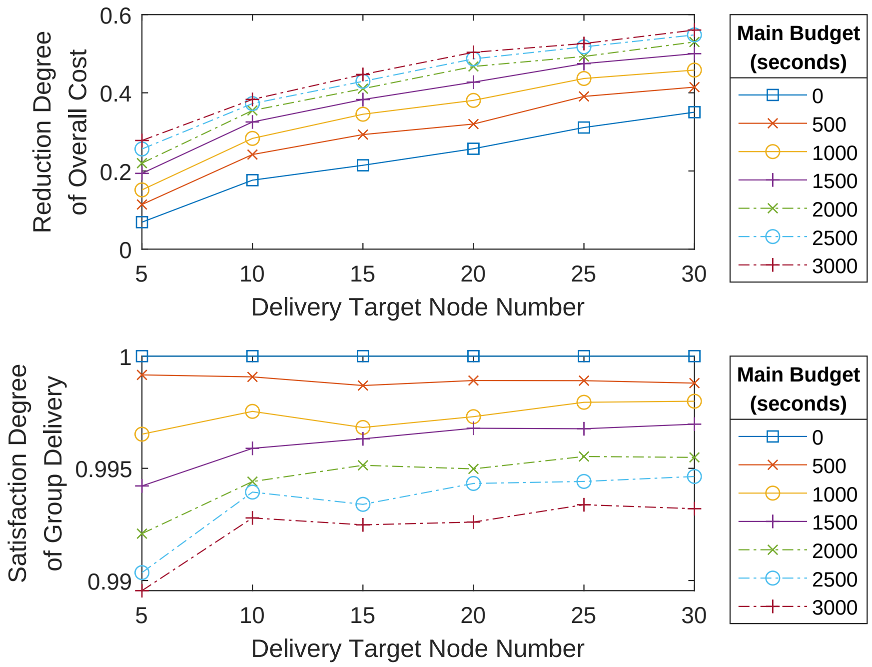


Figure 4.12: Investigation of Target Node Number (No Extra Budget)

exploitation. The concave region shown in the delivery satisfaction results from no more opportunistic forwarding after a certain SDTA value, as fewer contact opportunities can be identified by the prediction model.

As our research focuses on directional movement with strong correlations, a small SDTA is considered and the default value of SDTA is set to be 5 degrees for the remaining simulations unless otherwise stated.

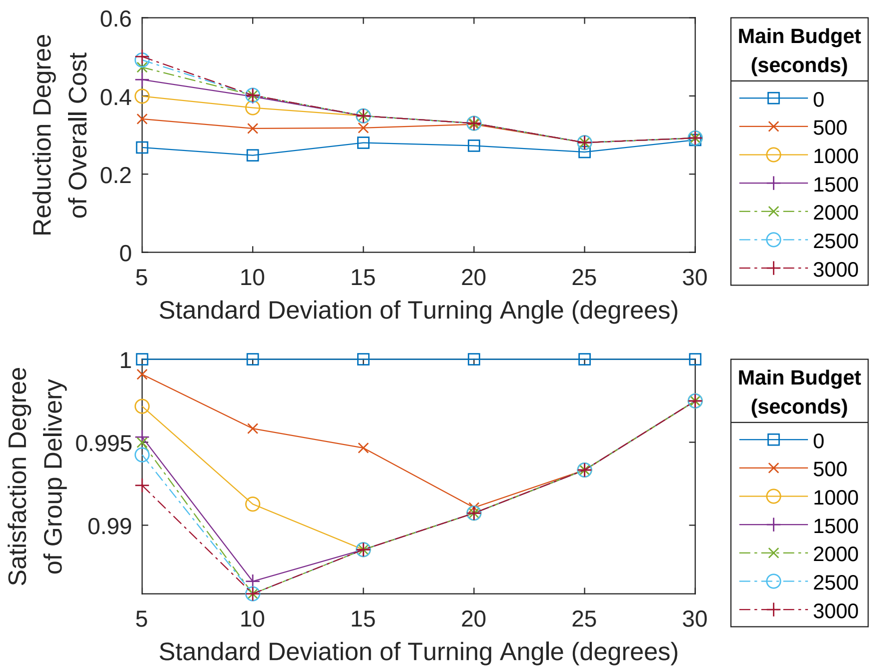


Figure 4.13: Investigation of SDTA (No Extra Budget)

B. Performance Comparison of Greedy Schemes This section provides a performance comparison of relay selection based on different greedy schemes. For the design of DBRS, three kinds of budget boundary are considered: no delay budget (DBRS-N), first-level budget boundary (DBRS-F), and second-level budget boundary (DBRS-S). Consequently, there are four schemes to be compared: DBRS-N, DBRS-F, DBRS-S, and SDRS. As mentioned in Section 4.3.2.3, SDRS is the proposed scheme with extra budget awareness and is highlighted for comparison focusing on the benefit of the extra budget term. As discussed in the first part of the analysis, the following default settings are adopted for the factors: $\eta = 0.9$, $\sigma = 5$, and $J = 20$.

Zero Main Budget In Figure 4.14, the DBRS-N and DBRS-F schemes have the same performance as when the main budget is set to zero. Furthermore, the performance of DBRS-N and DBRS-F do not vary with extra budget due to none-awareness of this factor. With increasing value of extra budget, DBRS-S and SDRS both have an increasing cost reduction degree and decreasing[8] delivery satisfaction degree. Although more cost reduction can be achieved by DBRS-S, SDRS can have a much lower impact on the satisfaction given a similar cost reduction degree as the delay budget function is considered in relay selection.

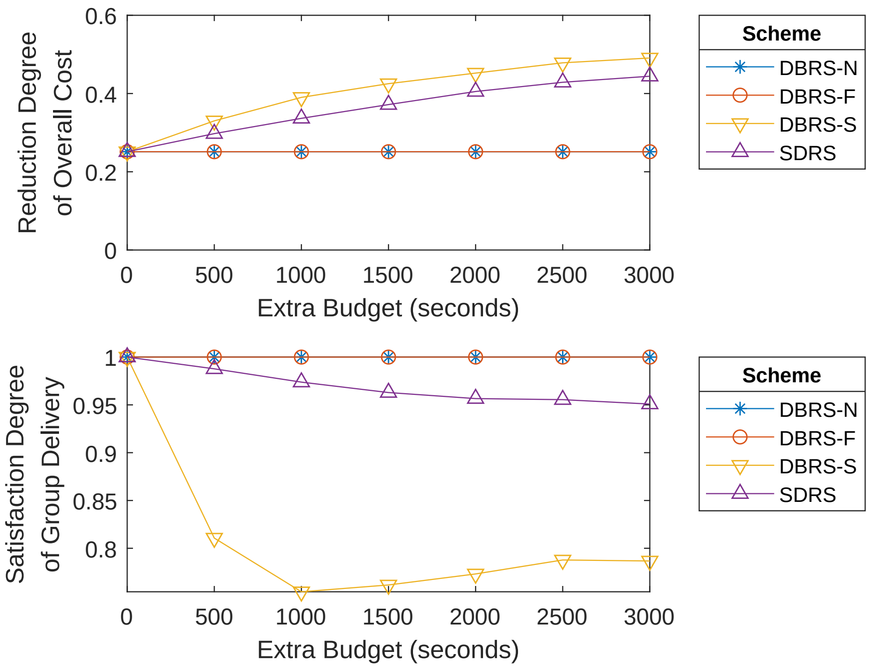


Figure 4.14: Comparison of Greedy Schemes for Zero Main Budget

[8] For the DBRS-S, the delivery satisfaction degree slightly increases when the extra budget is larger than certain value. The reason is that the delay budget function (i.e. delivery requirements) is changed with the changing value of extra budget. Then the same delivery delay (if within the range of extra budget) contributes more delivery satisfaction and therefore the overall satisfaction is increased. Because the delivery delay tends to be larger with a higher value of main budget, this effect becomes even more significant in Figure 4.15, 4.16, 4.17, and SDRS is also affected in Figure 4.17.

Small Main Budget For Figure 4.15, a small main budget is provided (set to 500 seconds). DBRS-F can achieve a higher cost reduction degree than DBRS-N because the main budget brings more relay candidates into consideration and more opportunity for relay sharing. Due to the strong directional correlation reflected by the small SDTA in the simulation settings, a small main budget rarely introduces uncertainty and therefore the satisfaction of DBRS-F is almost the same as for DBRS-N. In addition, the performance of DBRS-N and DBRS-F do not vary with extra budget due to their lack of awareness of this factor. DBRS-S and SDRS are similar in terms of their performance in Figure 4.14, because a small value of main budget does not introduce much difference compared with a zero main budget.

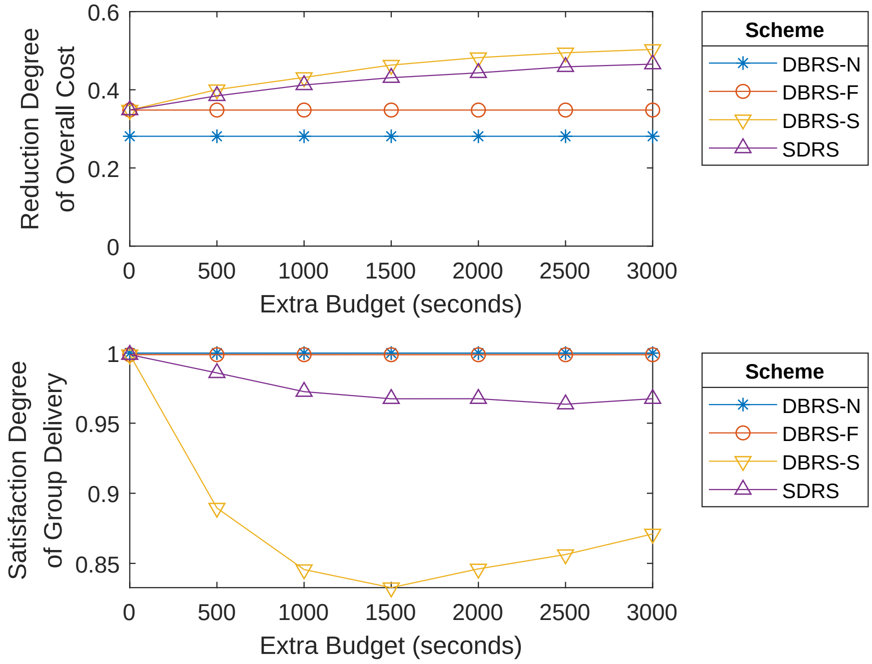


Figure 4.15: Comparison of Greedy Schemes for Small Main Budget (500s)

Medium Main Budget For Figure 4.16, the main budget is introduced as 1000 seconds so the gap between DBRS-N and DBRS-F (over the cost reduction degree) is increased. Meanwhile, the delivery satisfaction of DBRS-F shows a slight degradation because the medium main budget brings more uncertainty. Although more cost reduction can still be achieved by DBRS-S and SDRS, the reduction gain brought about by the extra budget becomes smaller when a medium main budget is adopted, which reflects less opportunities for exploitation.

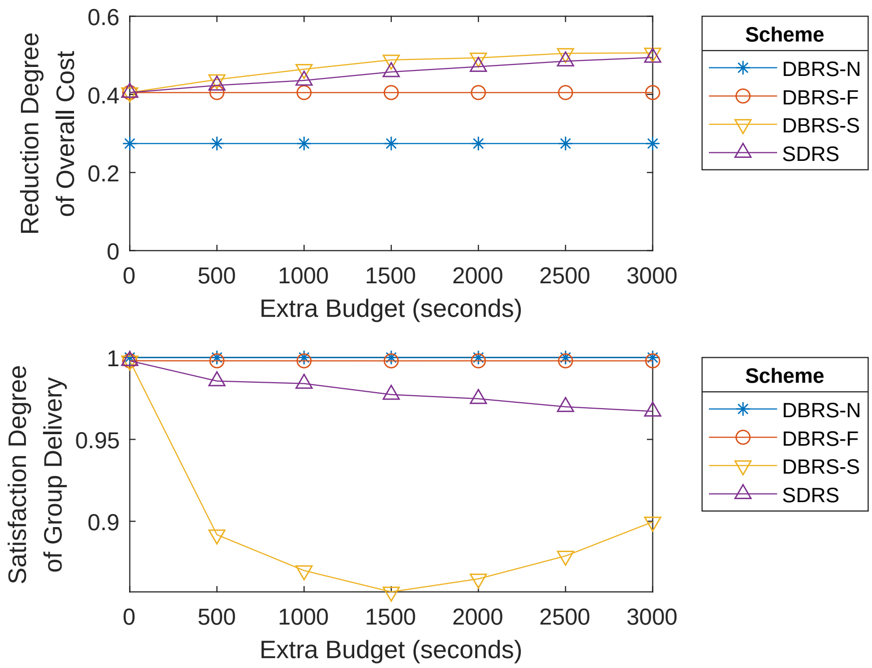


Figure 4.16: Comparison of Greedy Schemes for Medium Main Budget (1000s)

Large Main Budget As shown in Figure 4.17, DBRS-F can nearly achieve the same cost reduction as DBRS-S and SDRS when the main budget is much larger (2000 seconds in this case). However, this brings a more obvious difference between DBRS-N and DBRS-F in terms of delivery satisfaction. The extra budget can hardly bring additional benefit as the directional movement has been exploited by the very large main budget.

In summary, SDRS is not suitable for situations where the main budget is already very large. Otherwise the SDRS seems to bring efficiency gains at reasonable cost. To confirm this hypothesis, the four schemes are further evaluated based on the solutions obtained from the evolutionary approach.

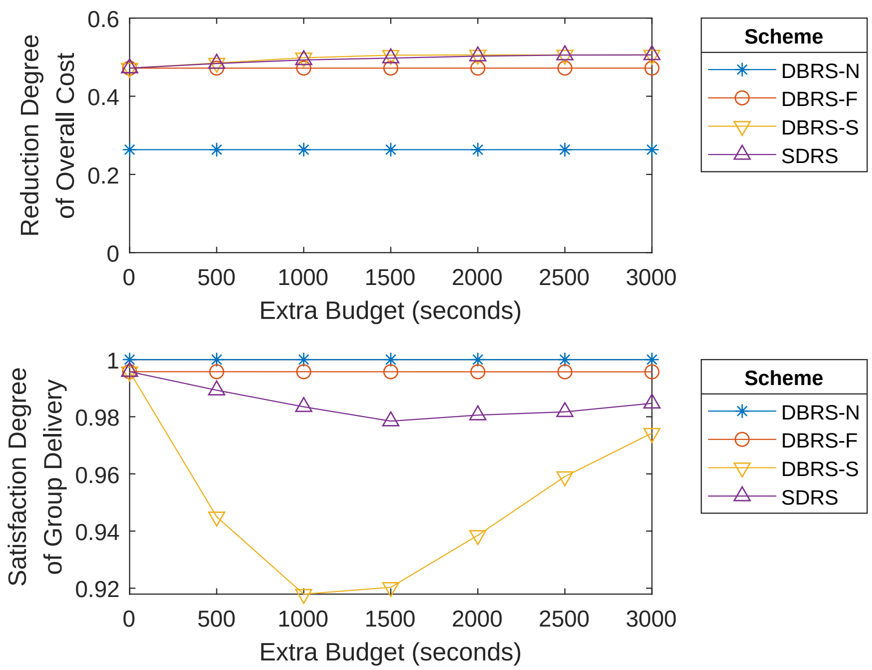


Figure 4.17: Comparison of Greedy Schemes for Large Main Budget (2000s)

C. Performance Evaluation Based on the Evolutionary Algorithm In this section, the weighted sum of two objectives is considered as the performance metric. As shown in Figure 4.10, the limit of the achievable weighted sum can be approximated by the evolutionary algorithm (i.e. the MOGA implementation as referred to in Section 4.3.2.2). From the obtained limit, the four schemes based on the greedy algorithm are evaluated given different weight values for the two objectives. Based on previous analysis and discussion, a small main budget (500 seconds in our case) is considered in these simulations and the extra budget is the control variable. Default settings are adopted for other factors ($\eta = 0.9$, $\sigma = 5$, $J = 20$) and the weight of the cost reduction is varied from 0 to 1 in 0.1 steps (the weight of the delivery satisfaction can be decided at the same time as the sum of weights given as 1).

Figure 4.18 shows how the weighted sum varies with the weight value. Referring to the previous discussion of Figure 4.10, the convex curve obtained from the MOGA reflects a performance limit of the weighted sum and therefore this MOGA curve is considered as the optimum. DBRS-S is far from the optimum when the weight is zero, as it does not achieve a good delivery satisfaction. However, it improves with increasing weight and finally outperforms the other three greedy schemes when the weight is one. The remaining three greedy schemes are near the performance limit when the weight is small. With increasing weight value, DBRS-N is the first to move away from the optimum. DBRS-F is closer because the main budget is considered for the cost reduction. SDRS is closest to the limit as it can exploit more opportunities for cost reduction without notably influencing the delivery satisfaction.

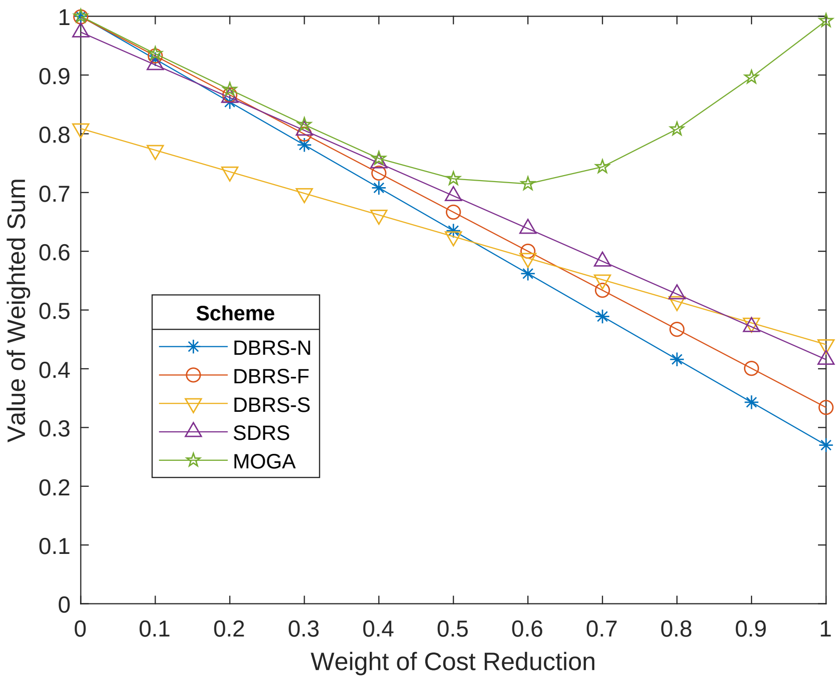


Figure 4.18: Evaluation of Weighted Sum over Different Weight Values (When Main Budget is 500s and Extra Budget is 1000s)

Figure 4.19 further demonstrates the difference from the performance limit (i.e. the MOGA curve) for the four greedy schemes. This figure shows that when the weight value is small (less than 0.3 in this case), the greedy schemes generally have good performance (i.e. a small difference from the optimum) except DBRS-S which considers the second-level budget boundary and therefore introduces more uncertainty. With increasing weight to the cost reduction factor (especially when it is larger than 0.7 in this case), relay selection based on the greedy algorithm becomes no longer suitable. It is worth noting that SDRS outperforms both DBRS-F and DBRS-N when the weight value is larger than 0.3, which confirms that our proposed SDRS scheme can find a better solution when the delivery efficiency becomes similar or more important relative to the delivery satisfaction.

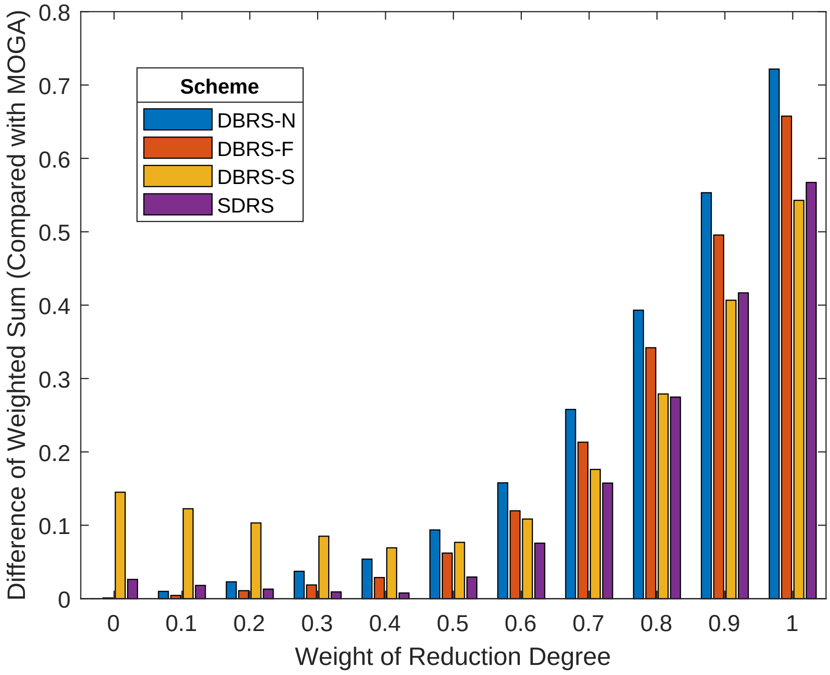


Figure 4.19: Difference of Weighted Sum over Different Weight Values (averaged for extra budget varying from 0s to 3000s in 500s steps where main budget is 500s)

Figure 4.20 provides another view of how the extra budget affects the difference from the performance limit. It is shown that neither DBRS-N and DBRS-S yield good performance when there is an extra budget. With increasing extra budget, the difference of DBRS-F increases while the difference of SDRS decreases due to awareness of extra delay budget. In conclusion, the proposed SDRS scheme can exploit the extra budget and therefore bring better performance.

In summary, simulation results show that relay selection based on several greedy algorithm variants has been shown to achieve near-optimal performance when the delivery

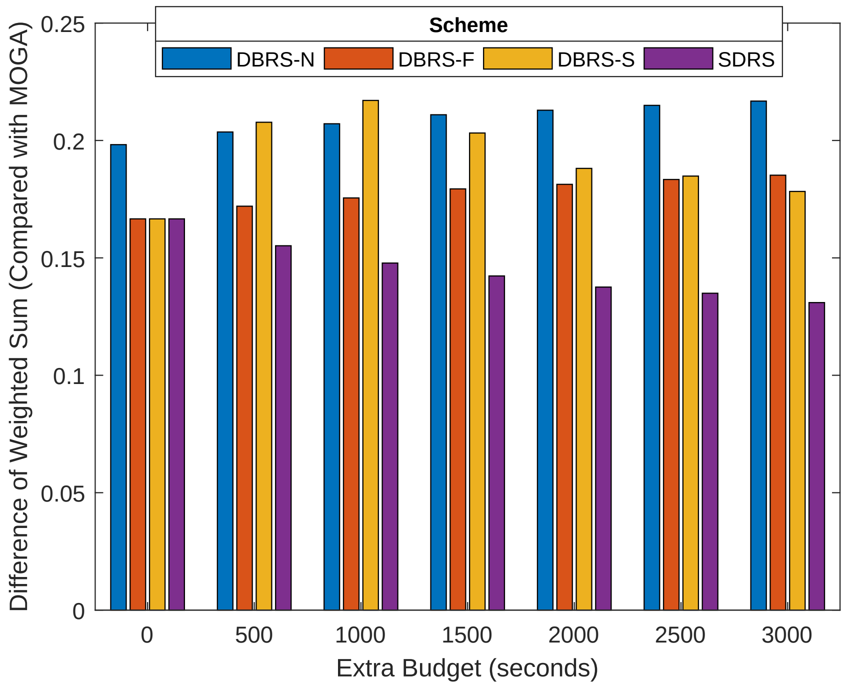


Figure 4.20: Difference of Weighted Sum over Different Extra Budget Values (averaged for weight values varying from 0 to 1 in 0.1 steps where main budget is 500s)

satisfaction is highly desirable. However, the performance of the greedy approach gradually becomes far from optimal with increasing weight given to delivery cost reduction. For cases where delivery efficiency becomes similarly (or more) important than delivery reliability, our proposed SDRS scheme is shown to be capable of finding more suitable solutions, which results in better delivery performance.

Chapter 5

Conclusions and Future Work

In this thesis, strategic packet forwarding is researched to provide more efficient multi-hop communication in software-defined wireless networks. The related literature is reviewed and state-of-the-art dissemination schemes are investigated. A SDMSN system for habitat monitoring is considered as the application scenario. Considering limitations of existing dissemination schemes, the feasible exploitation of directional movement in path-unconstrained mobility is investigated and evaluated. Simulation results show that the proposed scheme outperforms the state-of-the-art because directional correlation of node movement is used to dynamically exploit the delay budget for the optimal selection of relay node(s). In other words, the directional movement is exploited for the dissemination of delay-tolerant information especially under conditions of moderate uncertainty.

Specifically, two kinds of dissemination cases are researched.

- The first case assumes that every delivery is independent and one relay node is selected for each target node. The research focus of this case is to identify the relay node with the least possible forwarding hops given enough confidence of subsequent contact with the target node. A fuzzy path model is therefore proposed for movement prediction by considering SDTA (i.e. the directional correlation). For the simulation scenario given in this thesis, a conclusion can be drawn that mobility with an SDTA value of 30 degrees or lower (i.e. relatively strong directional correlation) can be exploited to improve the efficiency of delay-tolerant dissemination given a certainty threshold value larger than 0.8 (i.e. allowing for moderate uncertainty).

- The second case assumes that a relay can be shared by a (sub)group of target nodes. This research case focuses on finding optimised relay(s) to reduce overall forwarding while retaining sufficient satisfaction of group delivery. Algorithms are therefore proposed for group relay selection where an extra budget of delivery delay is considered to reflect more flexible delay tolerance. When the delivery efficiency is similarly (or more) important relative to the delivery satisfaction, given a small value for the main delay budget, the greedy scheme with an awareness of the extra budget is proven to provide more appropriate node selection than the scheme without such awareness, especially with increasing values of extra budget.

In summary, research issues are addressed in both cases for the scenario where mobile nodes move with directional correlation (i.e. animals in a habitat monitoring system as is the example scenario within this thesis).

Note that this research provides an analytical framework at a simplified fundamental level, which can be extended to more realistic scenarios. Although real-life animals may sometimes move faster/slower (or just stop), they tend to move constantly when there is a certain purpose in mind (e.g. travelling to a point of interest). This research is therefore focused on exploiting such tendencies of directional movement over a certain period. Based on the modelling of directional correlation, general patterns of mobility can be reflected (e.g. a purposeful movement can lead to faster displacement than wanderings), even if fluctuations in movement speed are not considered in our mobility model for the sake of simplicity. With a range of different delay budgets (up to tens of minutes), a specific scenario with walking-speed mobility is investigated in the thesis, which indicates that the exploitation of directional movement can potentially facilitate information dissemination given possible delay tolerance.

Further research can be conducted from following aspects:

- More advanced dissemination strategy could be researched on the basis of hybrid mobility; for example, a scenario involving mobile relay nodes and/or stationary destinations. Particularly for group delivery, new algorithms and mechanisms are required to determine suitable relay nodes and prevent message duplication.

- More factors could be considered to better reflect realistic network conditions such as transmission delay, contact time, residual energy and storage demands. Considering differentiated traffic delivery requirements leads to more complex problem optimisation, which remains unsolved.

- Various mobility information can be taken into consideration, including spatial, temporal, and social constraints and relationships. The "freshness" and accuracy of mobility data can be further studied, along with how this information can be effectively passed to the decision-making process in a timely manner.

- More analysis on real movement traces is required to select and quantify appropriate parameters in the mobility model for more realistic characterisation of animal movements and so forth; for example, the estimation of directional deviation based on recent or historical information.

- Wider use-cases involving path-unconstrained mobility can benefit from this research. For example, the analysis of directional movement can contribute to the development of autonomous underwater/aerial vehicle and robot tracking, especially considering tasks like random searching and target following.

References

[AK04] Ian F Akyildiz and Ismail H Kasimoglu. Wireless sensor and actor networks: research challenges. *Ad hoc networks*, 2(4):351–367, 2004.

[AMC07] Ian F Akyildiz, Tommaso Melodia, and Kaushik R Chowdhury. A survey on wireless multimedia sensor networks. *Computer networks*, 51(4):921–960, 2007.

[APM05] Ian F Akyildiz, Dario Pompili, and Tommaso Melodia. Underwater acoustic sensor networks: research challenges. *Ad hoc networks*, 3(3):257–279, 2005.

[ASBSP14] J. R. A. Amazonas, G. Santos-Boada, and J. Solé-Pareta. A critical review of openflow/sdn-based networks. In *2014 16th International Conference on Transparent Optical Networks (ICTON)*, pages 1–5, July 2014.

[ASSC02] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: A survey. *Comput. Netw.*, 38(4):393–422, March 2002.

[ASV09] Ian F Akyildiz, Zhi Sun, and Mehmet C Vuran. Signal propagation techniques for wireless underground communication networks. *Physical Communication*, 2(3):167–183, 2009.

[AWL15] Ian F. Akyildiz, Pu Wang, and Shih-Chun Lin. Softair: A software defined networking architecture for 5g wireless systems. *Computer Networks*, 85:1 – 18, 2015.

[Bac96] Thomas Back. *Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms.* Oxford university press, 1996.

[BB15] S. Batabyal and P. Bhaumik. Mobility models, traces and impact of mobility on opportunistic routing algorithms: A survey. *IEEE Communications Surveys Tutorials*, 17(3):1679–1707, thirdquarter 2015.

[BCC+17] Armir Bujari, Carlos T Calafate, Juan-Carlos Cano, Pietro Manzoni, Claudio Enrico Palazzi, and Daniele Ronzani. Flying ad-hoc network application scenarios and mobility models. *International Journal of Distributed Sensor Networks*, 13(10):1550147717738192, 2017.

[BCF07] S. Bandyopadhyay, E. J. Coyle, and T. Falck. Stochastic properties of mobility models in mobile ad hoc networks. *IEEE Transactions on Mobile Computing*, 6(11):1218–1229, Nov 2007.

[Ben06] Simon Benhamou. Detecting an orientation component in animal paths when the preferred direction is individual-dependent. *Ecology*, 87(2):518–528, 2006.

[Bet01] Christian Bettstetter. Smooth is better than sharp: A random mobility model for simulation of wireless networks. In *Proceedings of the 4th ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, MSWIM '01, pages 19–27, New York, NY, USA, 2001. ACM.

[BMKL12] Manu Bansal, Jeffrey Mehlman, Sachin Katti, and Philip Levis. Openradio: A programmable wireless dataplane. In *Proceedings of the First Workshop on Hot Topics in Software Defined Networks*, HotSDN '12, pages 109–114, New York, NY, USA, 2012. ACM.

[BS16] H. Byun and J. So. Node scheduling control inspired by epidemic theory for data dissemination in wireless sensor-actuator networks with delay constraints. *IEEE Transactions on Wireless Communications*, 15(3):1794–1807, March 2016.

[BSS+16] Y. Bi, H. Shan, X. S. Shen, N. Wang, and H. Zhao. A multi-hop broadcast protocol for emergency message dissemination in urban vehicular ad hoc networks. *IEEE Transactions on Intelligent Transportation Systems*, 17(3):736–750, March 2016.

[Bye01] John A Byers. Correlated random walk equations of animal dispersal resolved by simulation. *Ecology*, 82(6):1680–1690, 2001.

[BZ16] S. Bharati and W. Zhuang. Crb: Cooperative relay broadcasting for safety applications in vehicular networks. *IEEE Transactions on Vehicular Technology*, 65(12):9542–9553, Dec 2016.

[CCC+10] Long Cheng, Yimin Chen, Canfeng Chen, Jian Ma, Lei Shu, Athanasios V. Vasilakos, and Naixue Xiong. Efficient query-based data collection for mobile wireless monitoring applications. *Comput. J.*, 53(10):1643–1657, December 2010.

[CCJ91] Brent N. Clark, Charles J. Colbourn, and David S. Johnson. Unit disk graphs. *Discrete Math.*, 86(1-3):165–177, January 1991.

[CCLS11] Zhiyuan Cheng, James Caverlee, Kyumin Lee, and Daniel Z. Sui. Exploring millions of footprints in location sharing services. In *ICWSM*, 2011.

[CEE+01] Alberto Cerpa, Jeremy Elson, Deborah Estrin, Lewis Girod, Michael Hamilton, and Jerry Zhao. Habitat monitoring: Application driver for wireless communications technology. *ACM SIGCOMM Computer Communication Review*, 31(2 supplement):20–41, 2001.

[Cen77] Y. Censor. Pareto optimality in multiobjective problems. *Applied Mathematics and Optimization*, 4(1):41–59, 1977.

[CFP+07] Martin Casado, Michael J. Freedman, Justin Pettit, Jianying Luo, Nick McKeown, and Scott Shenker. Ethane: Taking control of the enterprise. *SIGCOMM Comput. Commun. Rev.*, 37(4):1–12, August 2007.

[CLL+17] Zhuangbin Chen, Anfeng Liu, Zhetao Li, Young-June Choi, Hiroo Sekiya, and Jie Li. Energy-efficient broadcasting scheme for smart industrial wireless sensor networks. *Mobile Information Systems*, page 17, 2017.

[CLRS09] Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. *Introduction to algorithms*. MIT press, 2009.

[CLVV+07] Carlos A Coello Coello, Gary B Lamont, David A Van Veldhuizen, et al. *Evolutionary algorithms for solving multi-objective problems*, volume 5. Springer, 2007.

[CPB08] Edward A Codling, Michael J Plank, and Simon Benhamou. Random walk models in biology. *Journal of the Royal Society Interface*, 5(25):813–834, 2008.

[CRMS16] Sobin CC, Vaskar Raychoudhury, Gustavo Marfia, and Ankita Singla. A survey of routing and data dissemination in delay tolerant networks. *Journal of Network and Computer Applications*, 67:128 – 146, 2016.

[CS13] Y. Cao and Z. Sun. Routing in delay/disruption tolerant networks: A taxonomy, survey and challenges. *IEEE Communications Surveys Tutorials*, 15(2):654–677, Second 2013.

[CSW+15] Y. Cao, Z. Sun, N. Wang, M. Riaz, H. Cruickshank, and X. Liu. Geographic-based spray-and-relay (gsar): An efficient routing scheme for dtns. *IEEE Transactions on Vehicular Technology*, 64(4):1548–1564, April 2015.

[CWSC15] Y. Cao, N. Wang, Z. Sun, and H. Cruickshank. A reliable and efficient encounter-based routing framework for delay/disruption tolerant networks. *IEEE Sensors Journal*, 15(7):4004–4018, July 2015.

[CYY16] G. L. Chiou, S. R. Yang, and W. T. Yen. On trajectory-based i2v group message delivery over vehicular ad-hoc networks. *IEEE Transactions on Vehicular Technology*, 65(9):7389–7402, Sept 2016.

[DK01] Kalyanmoy Deb and Deb Kalyanmoy. *Multi-Objective Optimization Using Evolutionary Algorithms*. John Wiley & Sons, Inc., New York, NY, USA, 2001.

[DMR16] B. Das, S. Misra, and U. Roy. Coalition formation for cooperative service-based message sharing in vehicular ad hoc networks. *IEEE Transactions on Parallel and Distributed Systems*, 27(1):144–156, Jan 2016.

[DPSBM13] A. Detti, C. Pisa, S. Salsano, and N. Blefari-Melazzi. Wireless mesh software defined networks (wmsdn). In *2013 IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 89–95, Oct 2013.

[EA12] J. P. Elbers and A. Autenrieth. From static to software-defined optical networks. In *Optical Network Design and Modeling (ONDM), 2012 16th International Conference on*, pages 1–4, April 2012.

[Fis10] H. Fischer. *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*. Sources and Studies in the History of Mathematics and Physical Sciences. Springer New York, 2010.

[FMW16] Klaus-Tycho Förster, Ratul Mahajan, and Roger Wattenhofer. Consistent updates in software defined networks: On dependencies, loop freedom, and blackholes. *Proc. 15th IFIP Networking*, 2016.

[FRZ14] Nick Feamster, Jennifer Rexford, and Ellen Zegura. The road to sdn: An intellectual history of programmable networks. *SIGCOMM Comput. Commun. Rev.*, 44(2):87–98, April 2014.

[FS16] R. A. Freeda and R. N. Sharmila. A review of bulk data dissemination protocols for reprogramming in wsn. In *2016 International Conference on Information Communication and Embedded Systems (ICICES)*, pages 1–4, Feb 2016.

[FSV16] Klaus-Tycho Foerster, Stefan Schmid, and Stefano Vissicchio. Survey of consistent network updates. *arXiv preprint arXiv:1609.02305*, 2016.

[FZZ07] Li Feng, Qinglin Zhao, and Hanwen Zhang. Location management based on distance and direction for pcs networks. *Comput. Netw.*, 51(1):134–152, January 2007.

[GAM14] A. De Gante, M. Aslan, and A. Matrawy. Smart wireless sensor network management based on software-defined networking. In *Communications (QBSC), 2014 27th Biennial Symposium on*, pages 71–75, June 2014.

[GKP+08] Natasha Gude, Teemu Koponen, Justin Pettit, Ben Pfaff, Martín Casado, Nick McKeown, and Scott Shenker. Nox: Towards an operating system for networks. *SIGCOMM Comput. Commun. Rev.*, 38(3):105–110, July 2008.

[GLG16] L. Galluccio, B. Lorenzo, and S. Glisic. Sociality-aided new adaptive infection recovery schemes for multicast dtns. *IEEE Transactions on Vehicular Technology*, 65(5):3360–3376, May 2016.

[GM04] C. Gui and P. Mohapatra. Scalable multicasting in mobile ad hoc networks. In *IEEE INFOCOM 2004*, volume 3, pages 2119–2129 vol.3, March 2004.

[GM07] Chao Gui and Prasant Mohapatra. Overlay multicast for manets using dynamic virtual mesh. *Wirel. Netw.*, 13(1):77–91, January 2007.

[GVL12] Gene H Golub and Charles F Van Loan. *Matrix computations*, volume 3. JHU Press, 2012.

[GZB+16] L. Gou, G. Zhang, D. Bian, W. Zhang, and Z. Xie. Data dissemination in wireless sensor networks with instantly decodable network coding. *Journal of Communications and Networks*, 18(5):846–856, October 2016.

[Haa97] Z. J. Haas. A new routing protocol for the reconfigurable wireless networks. In *Proceedings of ICUPC 97 - 6th International Conference on Universal Personal Communications*, volume 2, pages 562–566 vol.2, Oct 1997.

[HAG16] I. T. Haque and N. Abu-Ghazaleh. Wireless software defined networking: a survey and taxonomy. *IEEE Communications Surveys Tutorials*, PP(99):1–1, 2016.

[HCCP16] J. He, L. Cai, P. Cheng, and J. Pan. Delay minimization for data dissemination in large-scale vanets with buses and taxis. *IEEE Transactions on Mobile Computing*, 15(8):1939–1950, Aug 2016.

[Hil87] Francis Begnaud Hildebrand. *Introduction to numerical analysis.* Courier Corporation, 1987.

[HYG12] Soheil Hassas Yeganeh and Yashar Ganjali. Kandoo: A framework for efficient and scalable offloading of control applications. In *Proceedings of the First Workshop on Hot Topics in Software Defined Networks*, HotSDN '12, pages 19–24, New York, NY, USA, 2012. ACM.

[JC01] Lusheng Ji and M. S. Corson. Differential destination multicast-a manet multicast routing protocol for small groups. In *Proceedings IEEE INFOCOM 2001. Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society (Cat. No.01CH37213)*, volume 2, pages 1192–1201 vol.2, 2001.

[JKK12] K. Jeong, J. Kim, and Y. T. Kim. Qos-aware network operating system for software defined networking with generalized openflows. In *2012 IEEE Network Operations and Management Symposium*, pages 1167–1174, April 2012.

[JKM+13] Sushant Jain, Alok Kumar, Subhasree Mandal, Joon Ong, Leon Poutievski, Arjun Singh, Subbaiah Venkata, Jim Wanderer, Junlan Zhou, Min Zhu, Jon Zolla, Urs Hölzle, Stephen Stuart, and Amin Vahdat. B4: Experience with a globally-deployed software defined wan. *SIGCOMM Comput. Commun. Rev.*, 43(4):3–14, August 2013.

[JP13] R. Jain and S. Paul. Network virtualization and software defined networking for cloud computing: a survey. *IEEE Communications Magazine*, 51(11):24–31, November 2013.

[KBTB14] Dirk P. Kroese, Tim Brereton, Thomas Taimre, and Zdravko I. Botev. Why the monte carlo method is so important today. *Wiley Interdisciplinary Reviews: Computational Statistics*, 6(6):386–392, 2014.

[KBV10] T. Karagiannis, J. Y. Le Boudec, and M. Vojnovic. Power law and exponential decay of intercontact times between mobile devices. *IEEE Transactions on Mobile Computing*, 9(10):1377–1390, Oct 2010.

[KCG+10] Teemu Koponen, Martin Casado, Natasha Gude, Jeremy Stribling, Leon Poutievski, Min Zhu, Rajiv Ramanathan, Yuichiro Iwata, Hiroaki Inoue, Takayuki Hama, and Scott Shenker. Onix: A distributed control platform for large-scale production networks. In *Proceedings of the 9th USENIX Conference on Operating Systems Design and Implementation*, OSDI'10, pages 351–364, Berkeley, CA, USA, 2010. USENIX Association.

[KF13] H. Kim and N. Feamster. Improving network management with software defined networking. *IEEE Communications Magazine*, 51(2):114–119, February 2013.

[KLE10] S. Kim, C. H. Lee, and D. Y. Eun. Superdiffusive behavior of mobile nodes and its impact on routing protocol performance. *IEEE Transactions on Mobile Computing*, 9(2):288–304, Feb 2010.

[KLG14a] I. Ku, Y. Lu, and M. Gerla. Software-defined mobile cloud: Architecture, services and use cases. In *2014 International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 1–6, Aug 2014.

[KLG+14b] I. Ku, Y. Lu, M. Gerla, R. L. Gomes, F. Ongaro, and E. Cerqueira. Towards software-defined vanet: Architecture and services. In *Ad Hoc Networking Workshop (MED-HOC-NET), 2014 13th Annual Mediterranean*, pages 103–110, June 2014.

[KLK16] R. Kim, H. Lim, and B. Krishnamachari. Prefetching-based data dissemination in vehicular cloud systems. *IEEE Transactions on Vehicular Technology*, 65(1):292–306, Jan 2016.

[KRV+15] D. Kreutz, F. M. V. Ramos, P. E. Veríssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig. Software-defined networking: A comprehensive survey. *Proceedings of the IEEE*, 103(1):14–76, Jan 2015.

[KS15] Navneet Kaur and Amarpreet Singh. A survey on data dissemination protocols used in vanets. *International Journal of Computer Applications*, 120(23), 2015.

[KSKD+12] A. Kassler, L. Skorin-Kapov, O. Dobrijevic, M. Matijasevic, and P. Dely. Towards qoe-driven multimedia service negotiation and path optimization with software defined networking. In *Software, Telecommunications and Computer Networks (SoftCOM), 2012 20th International Conference on*, pages 1–5, Sept 2012.

[KSKS98] Ioannis Karatzas, Steven E Shreve, I Karatzas, and Steven E Shreve. *Methods of mathematical finance*, volume 39. Springer, 1998.

[KY11] Slawomir Koziel and Xin-She Yang. *Computational optimization, methods and algorithms*, volume 356. Springer, 2011.

[LC15] Y. Li and M. Chen. Software-defined network function virtualization: A survey. *IEEE Access*, 3:2542–2553, 2015.

[LFT+12] Wei Liu, Xiaotian Fei, Tao Tang, Rangkun Li, Pengjun Wang, Hong Luo, Kun Li, Linyuan Zhang, Beixing Deng, and Huazhong Yang. Design and implementation of a hybrid sensor network for milu deer monitoring. In *Advanced Communication Technology (ICACT), 2012 14th International Conference on*, pages 52–56, Feb 2012.

[LH03] B. Liang and Z. J. Haas. Predictive distance-based mobility management for multidimensional pcs networks. *IEEE/ACM Transactions on Networking*, 11(5):718–732, Oct 2003.

[LJHC16] Y. Li, D. Jin, P. Hui, and S. Chen. Contact-aware data replication in roadside unit aided vehicular delay tolerant networks. *IEEE Transactions on Mobile Computing*, 15(2):306–321, Feb 2016.

[LK16] J. Liu and N. Kato. A markovian analysis for explicit probabilistic stopping-based information propagation in postdisaster ad hoc mobile networks.

IEEE Transactions on Wireless Communications, 15(1):81–90, Jan 2016.

[LMO16] Jonathan Loo, Jaime Lloret Mauri, and Jesús Hamilton Ortiz. *Mobile ad hoc networks: current status and future trends*. CRC Press, 2016.

[LMR12] L. E. Li, Z. M. Mao, and J. Rexford. Toward software-defined cellular networks. In *2012 European Workshop on Software Defined Networking*, pages 7–12, Oct 2012.

[LNJ+16] J. Li, Z. Ning, B. Jedari, F. Xia, I. Lee, and A. Tolba. Geo-social distance-based data dissemination for socially aware networking. *IEEE Access*, 4:1444–1453, 2016.

[LQZ16] L. Li, Y. Qin, and X. Zhong. A novel routing scheme for resource-constraint opportunistic networks: A cooperative multiplayer bargaining game approach. *IEEE Transactions on Vehicular Technology*, 65(8):6547–6561, Aug 2016.

[LTQ12] T. Luo, H. P. Tan, and T. Q. S. Quek. Sensor openflow: Enabling software-defined wireless sensor networks. *IEEE Communications Letters*, 16(11):1896–1899, November 2012.

[LW09] C. Liu and J. Wu. Scalable routing in cyclic mobile networks. *IEEE Transactions on Parallel and Distributed Systems*, 20(9):1325–1338, Sept 2009.

[LW16] Z. Li and P. Wu. Intermittent geocast routing in urban vehicular delay tolerant networks. *Tsinghua Science and Technology*, 21(6):630–642, Dec 2016.

[LWH+12] Dan Levin, Andreas Wundsam, Brandon Heller, Nikhil Handigol, and Anja Feldmann. Logically centralized?: state distribution trade-offs in software defined networks. In *Proceedings of the first workshop on Hot topics in software defined networks*, pages 1–6. ACM, 2012.

[LWL16] Xiao Liu, Tianyi Wei, and Anfeng Liu. Fast program codes dissemination for smart wireless software defined networks. *Scientific Programming*, page 21, 2016.

[LXJ15] Xiaofeng Liu, Ning Xu, and Aimin Jiang. Tortuosity entropy: A measure of spatial complexity of behavioral changes in animal movement. *Journal* *of Theoretical Biology*, 364(Supplement C):197 – 205, 2015.

[MAB+08] Nick McKeown, Tom Anderson, Hari Balakrishnan, Guru Parulkar, Larry Peterson, Jennifer Rexford, Scott Shenker, and Jonathan Turner. Openflow: Enabling innovation in campus networks. *SIGCOMM Comput. Commun. Rev.*, 38(2):69–74, March 2008.

[Mai04] C. Maihofer. A survey of geocast routing protocols. *IEEE Communications Surveys Tutorials*, 6(2):32–42, Second 2004.

[MB17] P. Mach and Z. Becvar. Mobile edge computing: A survey on architecture and computation offloading. *IEEE Communications Surveys Tutorials*, 19(3):1628–1656, thirdquarter 2017.

[MCP+02] Alan Mainwaring, David Culler, Joseph Polastre, Robert Szewczyk, and John Anderson. Wireless sensor networks for habitat monitoring. In *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*, WSNA '02, pages 88–97, New York, NY, USA, 2002. ACM.

[MGT+15] J. Matias, J. Garay, N. Toledo, J. Unzilla, and E. Jacob. Toward an sdn-enabled nfv architecture. *IEEE Communications Magazine*, 53(4):187–193, April 2015.

[Mor12] Annu Mor. Study of different type of data dissemination strategy in vanet. *Research Scholar, Deptt. Of Computer Science Applications, Kurukshetra University, Kurukshetra, Haryana, India*, 2012.

[MS13] A. Mason and J. Sneddon. Automated monitoring of foraging behaviour in free ranging sheep grazing a biodiverse pasture. In *2013 Seventh International Conference on Sensing Technology (ICST)*, pages 46–51, Dec 2013.

[MV15] V. Muthumanikandan and C. Valliyammai. A survey on link failures in software defined networks. In *2015 Seventh International Conference on Advanced Computing (ICoAC)*, pages 1–5, Dec 2015.

[MZSS16] Guowang Miao, Jens Zander, Ki Won Sung, and Slimane Ben Slimane. *Fundamentals of Mobile Data Networks*. Cambridge University Press, 2016.

[NBGT17] V. G. Nguyen, A. Brunstrom, K. J. Grinnemo, and J. Taheri. Sdn/nfv-based mobile packet core network architectures: A survey. *IEEE Communications Surveys Tutorials*, 19(3):1567–1602, thirdquarter 2017.

[NBW09] P. Nouvellet, J. P. Bacon, and D. Waxman. Fundamental insights into the random movement of animals from a single distance related statistic. *The American Naturalist*, 174(4):506–514, 2009. PMID: 19737110.

[NdML16] I. O. Nunes, P. O. S. Vaz de Melo, and A. A. F. Loureiro. Leveraging d2d multihop communication through social group meeting awareness. *IEEE Wireless Communications*, 23(4):12–19, August 2016.

[NF16] M. T. Nuruzzaman and H. W. Ferng. A low energy consumption routing protocol for mobile sensor networks with a path-constrained mobile sink. In *2016 IEEE International Conference on Communications (ICC)*, pages 1–6, May 2016.

[NMN+14] B. A. A. Nunes, M. Mendonca, X. N. Nguyen, K. Obraczka, and T. Turletti. A survey of software-defined networking: Past, present, and future of programmable networks. *IEEE Communications Surveys Tutorials*, 16(3):1617–1634, Third 2014.

[NS15] Nishtha and M. Sood. A survey on issues of concern in software defined networks. In *2015 Third International Conference on Image Information Processing (ICIIP)*, pages 295–300, Dec 2015.

[NTCS99] Sze-Yao Ni, Yu-Chee Tseng, Yuh-Shyan Chen, and Jang-Ping Sheu. The broadcast storm problem in a mobile ad hoc network. In *Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking*, MobiCom '99, pages 151–162, New York, NY, USA, 1999. ACM.

[OCF15] F. Olivier, G. Carlos, and N. Florent. Sdn based architecture for clustered wsn. In *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2015 9th International Conference on*, pages 342–347, July 2015.

[OL13] Akira Okubo and Smon A Levin. *Diffusion and ecological problems: modern perspectives*, volume 14. Springer Science & Business Media, 2013.

[PBD13] Claire M. Postlethwaite, Pieta Brown, and Todd E. Dennis. A new multiscale measure for analysing animal movement data. *Journal of Theoretical Biology*, 317(Supplement C):175 – 185, 2013.

[PiAAB16] J. Pedersen, A. Graell i Amat, I. Andriyanova, and F. Brännström. Distributed storage in mobile wireless networks with device-to-device communication. *IEEE Transactions on Communications*, 64(11):4862–4878, Nov 2016.

[PJT+14] T. Parker, J. Johnson, M. Tummala, J. McEachen, and J. Scrofani. Dynamic state determination of a software-defined network via dual basis representation. In *Signal Processing and Communication Systems (ICSPCS), 2014 8th International Conference on*, pages 1–7, Dec 2014.

[Pos11] Stefan Poslad. *Ubiquitous computing: smart devices, environments and interactions*. John Wiley & Sons, 2011.

[PRR16] Kavita Pandey, Saurabh Kumar Raina, and Ram Shringar Raw. Distance and direction-based location aided multi-hop routing protocol for vehicular ad-hoc networks. *Int. J. Commun. Netw. Distrib. Syst.*, 16(1):71–98, December 2016.

[PS82] Christos H. Papadimitriou and Kenneth Steiglitz. *Combinatorial Optimization: Algorithms and Complexity*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1982.

[PSM+04] Joseph Polastre, Robert Szewczyk, Alan Mainwaring, David Culler, and John Anderson. Wireless sensor networks. chapter Analysis of Wireless Sensor Networks for Habitat Monitoring, pages 399–423. Kluwer Academic Publishers, Norwell, MA, USA, 2004.

[PSM17] S. Pal, B. K. Saha, and S. Misra. Game theoretic analysis of cooperative message forwarding in opportunistic mobile networks. *IEEE Transactions on Cybernetics*, PP(99):1–12, 2017.

[QDG+14] Z. Qin, G. Denker, C. Giannelli, P. Bellavista, and N. Venkatasubramanian. A software defined networking architecture for the internet-of-things. In *2014 IEEE Network Operations and Management Symposium (NOMS)*, pages 1–9, May 2014.

[RMTO11] R. N. B. Rais, M. Mendonca, T. Turletti, and K. Obraczka. Towards truly heterogeneous internets: Bridging infrastructure-based and infrastructure-less networks. In *2011 Third International Conference on Communication Systems and Networks (COMSNETS 2011)*, pages 1–10, Jan 2011.

[Roy95] Patrick Royston. Remark as r94: A remark on algorithm as 181: The w-test for normality. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 44(4):547–551, 1995.

[RRNS14] H. K. Rath, V. Revoori, S. M. Nadaf, and A. Simha. Optimal controller placement in software defined networks (sdn) using a non-zero-sum game. In *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2014 IEEE 15th International Symposium on a*, pages 1–6, June 2014.

[RY13] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 2013.

[SB13] Josef Stoer and Roland Bulirsch. *Introduction to numerical analysis*, volume 12. Springer Science & Business Media, 2013.

[SCC+10] Rob Sherwood, Michael Chan, Adam Covington, Glen Gibb, Mario Flajslik, Nikhil Handigol, Te-Yuan Huang, Peyman Kazemian, Masayoshi Kobayashi, Jad Naous, Srinivasan Seetharaman, David Underhill, Tatsuya Yabe, Kok-Kiong Yap, Yiannis Yiakoumis, Hongyi Zeng, Guido Appenzeller, Ramesh Johari, Nick McKeown, and Guru Parulkar. Carving research slices out of your production networks with openflow. *SIGCOMM Comput. Commun. Rev.*, 40(1):129–130, January 2010.

[SGY17] Y. Sun, J. Guo, and Yuhui Yao. Speed up-greedy perimeter stateless routing protocol for wireless sensor networks (su-gpsr). In *2017 IEEE 18th International Conference on High Performance Switching and Routing (HPSR)*, pages 1–6, June 2017.

[SHNS16] S. Scott-Hayward, S. Natarajan, and S. Sezer. A survey of security in software defined networks. *IEEE Communications Surveys Tutorials*, 18(1):623–654, Firstquarter 2016.

[SK11] U. Sadiq and M. Kumar. Proximol: Proximity and mobility estimation for efficient forwarding in opportunistic networks. In *2011 IEEE Eighth International Conference on Mobile Ad-Hoc and Sensor Systems*, pages 312–321, Oct 2011.

[SKLR16] J. Sung, M. Kim, K. Lim, and J. K. K. Rhee. Efficient cache placement strategy in two-tier wireless content delivery network. *IEEE Transactions on Multimedia*, 18(6):1163–1174, June 2016.

[SMZ07] Kazem Sohraby, Daniel Minoli, and Taieb Znati. *Wireless sensor networks: technology, protocols, and applications*. John Wiley & Sons, 2007.

[SSZM+12] Lalith Suresh, Julius Schulz-Zander, Ruben Merz, Anja Feldmann, and Teresa Vazao. Towards programmable enterprise wlans with odin. In *Proceedings of the First Workshop on Hot Topics in Software Defined Networks*, HotSDN '12, pages 115–120, New York, NY, USA, 2012. ACM.

[SW14] Henry Stark and John W Woods. *Probability and Random Processes with Applications to Signal Processing: International Edition*. Pearson Higher Ed, 2014.

[TG10] Amin Tootoonchian and Yashar Ganjali. Hyperflow: A distributed control plane for openflow. In *Proceedings of the 2010 Internet Network Management Conference on Research on Enterprise Networking*, INM/WREN'10, pages 3–3, Berkeley, CA, USA, 2010. USENIX Association.

[Toh01] C.K. K Toh. *Ad Hoc Wireless Networks: Protocols and Systems*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1st edition, 2001.

[VN11] Steven J. Vaughan-Nichols. Openflow: The next generation of the network? *Computer*, 44(8):13–15, August 2011.

[WBS14] Xiuchao Wu, Kenneth N. Brown, and Cormac J. Sreenan. Data pre-forwarding for opportunistic data collection in wireless sensor networks. *ACM Trans. Sen. Netw.*, 11(1):8:1–8:33, July 2014.

[wCsL16] S. w. Chang and S. s. Lee. A routing protocol for urban vehicular multi-hop data delivery. *Chinese Journal of Electronics*, 25(2):348–356, 2016.

[WCZL16] Tong Wang, Yue Cao, Yongzhe Zhou, and Pengcheng Li. A survey on geographic routing protocols in delay/disruption tolerant networks. *Inter-*

national Journal of Distributed Sensor Networks*, 12(2):3174670, 2016.

[WLSN00] Hsin-i Wu, Bai-Lian Li, Timothy A Springer, and William H Neill. Modelling animal movement as a persistent random walk in two dimensions: expected magnitude of net displacement. *Ecological Modelling*, 132(1):115–124, 2000.

[Wol05] J. Wolberg. *Data Analysis Using the Method of Least Squares: Extracting the Most Information from Experiments*. Springer Berlin Heidelberg, 2005.

[WWH+16] S. Wang, X. Wang, J. Huang, R. Bie, Z. Tian, and F. Zhao. The potential of mobile opportunistic networks for data disseminations. *IEEE Transactions on Vehicular Technology*, 65(2):912–922, Feb 2016.

[XHC17] M. Xing, J. He, and L. Cai. Utility maximization for multimedia data dissemination in large-scale vanets. *IEEE Transactions on Mobile Computing*, 16(4):1188–1198, April 2017.

[XLH16] Yixuan Xu, Anfeng Liu, and Changqin Huang. Delay-aware program codes dissemination scheme in internet of everything. *Mobile Information Systems*, page 18, 2016.

[XWF+15] W. Xia, Y. Wen, C. H. Foh, D. Niyato, and H. Xie. A survey on software-defined networking. *IEEE Communications Surveys Tutorials*, 17(1):27–51, Firstquarter 2015.

[XZD+14] X. Xu, H. Zhang, X. Dai, Y. Hou, X. Tao, and P. Zhang. Sdn based next generation mobile network with service slicing and trials. *China Communications*, 11(2):65–77, Feb 2014.

[YGH+14] F. Yang, V. Gondi, J. O. Hallstrom, K. C. Wang, and G. Eidson. Openflow-based load balancing for wireless mesh infrastructure. In *2014 IEEE 11th Consumer Communications and Networking Conference (CCNC)*, pages 444–449, Jan 2014.

[YKS+10] Kok-Kiong Yap, Masayoshi Kobayashi, Rob Sherwood, Te-Yuan Huang, Michael Chan, Nikhil Handigol, and Nick McKeown. Openroads: Empowering research in mobile networks. *SIGCOMM Comput. Commun. Rev.*, 40(1):125–126, January 2010.

[YKU+09] Kok-Kiong Yap, Masayoshi Kobayashi, David Underhill, Srinivasan Seetharaman, Peyman Kazemian, and Nick McKeown. The stanford openroads deployment. In *Proceedings of the 4th ACM International Workshop on Experimental Evaluation and Characterization*, WINTECH '09, pages 59–66, New York, NY, USA, 2009. ACM.

[YLJ+13] Mao Yang, Yong Li, Depeng Jin, Li Su, Shaowu Ma, and Lieguang Zeng. Openran: A software-defined ran architecture via virtualization. In *Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM*, SIGCOMM '13, pages 549–550, New York, NY, USA, 2013. ACM.

[YLK07] Chia-Mu Yu, Chun-Shien Lu, and Sy-Yen Kuo. Habitual behavior-based opportunistic data forwarding in wildlife tracking. In *2007 4th International Symposium on Wireless Communication Systems*, pages 807–808. IEEE, 2007.

[YNP17] B. Yu, M. Ni, and J. Pan. Performance analysis of message dissemination in icmn under the last-create first-relay scheme. *IEEE Transactions on Vehicular Technology*, 66(1):818–830, Jan 2017.

[YS14] T. C. Yen and C. S. Su. An sdn-based cloud computing architecture and its mathematical model. In *Information Science, Electronics and Electrical Engineering (ISEEE), 2014 International Conference on*, volume 3, pages 1728–1731, April 2014.

[YSK+10] Kok-Kiong Yap, Rob Sherwood, Masayoshi Kobayashi, Te-Yuan Huang, Michael Chan, Nikhil Handigol, Nick McKeown, and Guru Parulkar. Blueprint for introducing innovation into wireless mobile networks. In *Proceedings of the Second ACM SIGCOMM Workshop on Virtualized Infrastructure Systems and Architectures*, VISA '10, pages 25–32, New York, NY, USA, 2010. ACM.

[YYK+11] Yiannis Yiakoumis, Kok-Kiong Yap, Sachin Katti, Guru Parulkar, and Nick McKeown. Slicing home networks. In *Proceedings of the 2Nd ACM SIGCOMM Workshop on Home Networks*, HomeNets '11, pages 1–6, New York,

NY, USA, 2011. ACM.

[YZW16] Q. Yang, B. Zhu, and S. Wu. An architecture of cloud-assisted information dissemination in vehicular networks. *IEEE Access*, 4:2764–2770, 2016.

[ZCZY16] A. Zhang, J. Chen, L. Zhou, and S. Yu. Graph theory-based qoe-driven cooperation stimulation for content dissemination in device-to-device communication. *IEEE Transactions on Emerging Topics in Computing*, 4(4):556–567, Oct 2016.

[ZLW16] Z. Zhang, L. Luo, and L. Wang. D2d multicast retransmission algorithm in mobile cloud based on sinr constraint. *China Communications*, 13(8):41–52, Aug 2016.

[ZLZY16] Q. Zhao, S. C. Liew, S. Zhang, and Y. Yu. Distance-based location management utilizing initial position for mobile communication networks. *IEEE Transactions on Mobile Computing*, 15(1):107–120, Jan 2016.

[ZSJ+17] Y. Zhang, L. Song, C. Jiang, N. H. Tran, Z. Dawy, and Z. Han. A social-aware framework for efficient information dissemination in wireless ad hoc networks. *IEEE Communications Magazine*, 55(1):174–179, January 2017.

[ZW06] M. Zhao and W. Wang. A novel semi-markov smooth mobility model for mobile ad hoc networks. In *IEEE Globecom 2006*, pages 1–5, Nov 2006.

[ZWD+13] X. Zheng, J. Wang, W. Dong, Y. He, and Y. Liu. Survival of the fittest: Data dissemination with selective negotiation in wireless sensor networks. In *2013 IEEE 10th International Conference on Mobile Ad-Hoc and Sensor Systems*, pages 443–451, Oct 2013.

[ZWD+16] X. Zheng, J. Wang, W. Dong, Y. He, and Y. Liu. Bulk data dissemination in wireless sensor networks: Analysis, implications and improvement. *IEEE Transactions on Computers*, 65(5):1428–1439, May 2016.

[ZYL16] X. Zhang, L. Yan, and W. Li. Efficient and reliable abiding geocast based on carrier sets for vehicular ad hoc networks. *IEEE Wireless Communications Letters*, 5(6):660–663, Dec 2016.