GEOGRAPHICAL MULTICAST DISRUPTION TOLERANT NETWORKING MECHANISM FOR INTERNET OF THINGS

WONG KHANG SIANG

UNIVERSITI SAINS MALAYSIA

2022

GEOGRAPHICAL MULTICAST DISRUPTION TOLERANT NETWORKING MECHANISM FOR INTERNET OF THINGS

by

WONG KHANG SIANG

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

January 2022

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere and utmost gratitude to my supervisor, Associate Professor Dr. Wan Tat Chee, the Deputy Director of National Advanced IPv6 Centre (NAv6), for his valuable insights and guidance throughout my Ph.D. journey that made this thesis possible. I would also like to offer my special thanks to my love Nguyen Lan Anh and my family for their unwavering support and belief in me. Finally, I would like to express my gratitude to Universiti Sains Malaysia for the USM Fellowship funding, as well as the School of Computer Sciences for providing all the facilities in helping me to complete the Ph.D. study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AGML	Average Group Multicast Latency
AODV	Ad Hoc On-Demand Distance Vector
API	Application Programming Interface
BP	Bundle Protocol
BLE	Bluetooth Low Energy
CTMR	Collaborative Time-Stamp based Multicast
DSR	Dynamic Source Routing
DTN	Delay Tolerant Networking
DTNRG	Delay-Tolerant Networking Research Group
EraMobile	Epidemic-based Reliable and Adaptive Multicast for Mobile ad hoc networks
EID	Endpoint Identifier
FIFO	First In First Out
FILO	First In Last Out
GMDTN	Geographical Multicast Disruption Tolerant Networking
GMDTN-R	Geographical Multicast Disruption Tolerant Networking with Retransmission
GPS	Global Positioning System
IEEE	Electrical and Electronics Engineers
ION	Interplanetary Overlay Network

IoT Internet of Things

IP Internet Protocol

IPv4 Internet Protocol version 4

IPv6 Internet Protocol version 6

JPL Jet Propulsion Laboratory

LTP Licklider Transmission Protocol

MANET Mobile Ad Hoc Network

MDM Multiple-Data Multicast

MDR Multicast Delivery Ratio

MOR Multicast Overhead Ratio

MSN Mobile Social Network

NAP Network Access Point

ONE Opportunistic Networking Environment

OTA Over The Air

PAN Personal Area Network

PRoPHET Probabilistic Routing Protocol Using History of Encounters and Transitivity

QGR Q-learning based Gain-aware Routing

RMDTN Reliable Multicast Disruption Tolerant Networking

RMDTN-R Reliable Multicast Disruption Tolerant Networking with Retransmission

RMTP Reliable Multicast Transport Protocol

RTOS Real Time Operating System

SCADA Supervisory Control and Data Acquisition

SDM Single-Data Multicast

SDMR Social Difference Multicast Routing

SPM Social Profile-based Multicast

SPMOR Social Profile-based Multicast - Overhead Reducing

SWN Social Wireless Network

TCP Transmission Control Protocol

TCP/IP Transmission Control Protocol/Internet Protocol

TLMR Two-Level Multicast Routing

TTL Time-To-Live

UAV Unmanned Aerial Vehicle

UDP User Datagram Protocol

URI Uniform Resource Identifier

URL Uniform Resource Locator

USM Universiti Sains Malaysia

VDTN Vehicular Delay Tolerant Network

VLAN Virtual Local Area Network

WANET Wireless Ad Hoc Network

WPAN Wireless Personel Area Network

WSN Wireless Sensor Network

MEKANISMA MULTIKAS BERASASKAN GEOGRAFI DALAM RANGKAIAN TAHAN GANGGUAN UNTUK INTERNET PELBAGAI BENDA

ABSTRAK

Rangkaian Tahan Gangguan (DTN) telah dibangunkan untuk mengatasi masalah sambungan berselang-seli antara nod di kawasan rangkaian tanpa wayar yang lemah dengan menggunakan paradigma simpan-bawa-hantar untuk menghantar mesej ke destinasi. Protokol rangkaian yang sedia ada seperti Protokol Kawalan Penghantaran (TCP) dan Protokol Datagram Pengguna (UDP) adalah tidak sesuai kerana tiada laluan hujung-ke-hujung antara penghantar dan penerima. Perkembangan pesat peranti Internet Pelbagai Benda (IoT) mengupayakan sokongan DTN dalam persekitaran IoT untuk merapatkan jurang komunikasi antara rangkaian sambungan berselang-seli di kawasan terpencil dan senario pasca bencana. Komunikasi kumpulan adalah perkhidmatan yang penting bagi membolehkan pertukaran dan perkongsian maklumat dalam kumpulan dan antara kumpulan dalam rangkaian-rangkaian tersebut. Tambahan pula, sesetengah aplikasi memerlukan sokongan multikas berkebolehpercayaan dalam rangkaian DTN yang kekurangan sumber. Namun begitu, tiada piawaian yang jelas untuk komunikasi kumpulan yang cekap dan berkebolehpercayaan dalam DTN. Komunikasi kumpulan dalam senario pasca bencana yang meliputi kawasan geografi yang luas memberikan cabaran kepada pekerja-pekerja bantuan bencana untuk berkomunikasi dan menyelaraskan misi mencari dan menyelamat. Perkhidmatan penghantaran data berdasarkan kumpulan diperlukan dalam rangkaian DTN dengan sokongan multikas untuk komunikasi di pelbagai kawasan geografi. Dalam rangkaian IoT yang keku-

rangan sumber, penghantaran data berasaskan kumpulan perlu dipertingkatkan untuk menyokong multikas berkebolehpercayaan untuk penggunaan khas, seperti pengemaskinian konfigurasi kebolehpercayaan. Mekanisme multikas juga perlu dioptimumkan untuk pautan jalur lebar yang berbeza disebabkan rangkaian IoT yang heterogen. Empat ciri telah dikenal pasti sebagai penyelesaian untuk menangani cabaran multikas dalam senario pasca bencana. Pengurusan keahlian kumpulan adalah ciri penting untuk komunikasi kumpulan dalam DTN, manakala pengurusan geohash dijangka meminimumkan overhed multikas dengan membahagikan kawasan pasca bencana kepada beberapa sel geohash. Penghantaran semula multikas membantu meningkatkan nisbah penghantaran untuk memastikan kebolehpercayaan perkhidmatan penghantaran data berasaskan kumpulan, bersama pengagregatan mesej pengakuan yang meminimumkan overhed daripada letupan mesej pengakuan. Menggabungkan kesemua empat ciri tersebut, mekanisme Multikas Berasaskan Geografi Dalam Rangkaian Tahan Gangguan (GMDTN) telah dicadangkan untuk menyokong perkhidmatan penghantaran data berasaskan kumpulan untuk rangkaian yang terdedah kepada gangguan dengan menggunakan informasi keahlian kumpulan dan maklumat geohash dengan kebolehpercayaan yang dipertingkat. Mekanisme ini dioptimumkan lagi untuk pautan lebar jalur yang berbeza untuk mengurangkan overhed mesej pengakuan. GMDTN mencapai 90% nisbah penghantaran mutilcast (MDR) dalam eksperimen tapak uji dengan overhed penghantaran yang lebih rendah berbanding pendekatan konvensional dalam senario Bluetooth. GMDTN juga mencapai lebih daripada 80% nisbah penghantaran multikas dengan hanya nisbah overhed multikas (MOR) 0.5 dalam pautan dengan lebar jalur 5Mbps. Dalam eksperimen simulasi, mekanisme GMDTN menunjukkan prestasi yang lebih baik dari segi nisbah penghantaran multikas dengan lebih daripada 80%

apabila nilai Masa-Untuk-Hidup (TTL) ditetapkan pada separuh daripada masa simulasi. Penyebaran mesej dalam mekanisme GMDTN adalah lebih pantas dengan purata kependaman kumpulan multikas (AGML) terendah bagi senario pasca bencana apabila nod diedarkan secara rawak pada peta simulasi.

GEOGRAPHICAL MULTICAST DISRUPTION TOLERANT NETWORKING MECHANISM FOR INTERNET OF THINGS

ABSTRACT

Disruption Tolerant Networking (DTN) has been developed to overcome the intermittent connection issue between nodes in areas with poor wireless network connectivity by employing a store-carry-forward paradigm to forward messages to the destination. The existing networking protocols such as Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) are not suitable since there may never be an end-to-end path between the sender and the receiver. As the Internet of Things (IoT) devices proliferate, enabling DTN support in IoT environments helps bridge the communication gap between networks with intermittent connectivity such as rural areas and post-disaster scenarios. Group communication is an essential service to enable information exchange and sharing within a group and between groups in such networks. Furthermore, some applications require reliable multicast support over resource-constrained DTN networks. However, there is no well-defined standard for efficient and reliable group communication in DTN. The group communication in a post-disaster scenario that covers a large geographical area presents a more challenging environment for the disaster relief personnel to communicate and coordinate search and rescue missions. A group-based data delivery service is needed in DTN networks with multicast support for communication over multiple geographical areas. In resource-constrained IoT networks, the group-based data delivery needs to be enhanced to provide reliable multicast support for use cases, such as reliable configuration updates. The multicast mechanism also needs to be optimized for different bandwidth links since IoT networks are heterogeneous. Four features are identified as the solutions to address the multicast challenges in post-disaster scenarios. Group membership management is an essential feature for group communication in the DTN, while geohash management is expected to minimize the multicast overhead by dividing the post-disaster area into multiple geohash cells. Multicast retransmission helps to improve the delivery ratio to ensure the reliability of group-based data delivery services, along with acknowledgment aggregation that minimizes the overhead from acknowledgment implosion. Combining all the four features, a novel Geographical Multicast Disruption Tolerant Networking (GMDTN) mechanism supports group-based data delivery services for disruption-prone networks by utilizing group membership, and geohash information with enhanced reliability is proposed. The mechanism is further optimized for different bandwidth links to reduce acknowledgment overheads. GMDTN achieved a 90% multicast delivery ratio (MDR) in the testbed methodologies with reduced transmission overheads compared to the conventional approaches in the Bluetooth scenario. GMDTN also achieved more than 80% multicast delivery ratio with only a 0.5 multicast overhead ratio (MOR) in a link with a bandwidth of 5Mbps. In the simulation experiments, the GMDTN mechanism showed better performance in terms of multicast delivery ratio with more than 80% when the Time-To-Live (TTL) value is fixed at half of the simulation time. The message propagation in the GMDTN mechanism is faster with the lowest average group multicast latency for the post-disaster scenario when the nodes are distributed randomly on the simulation map.

CHAPTER 1

INTRODUCTION

1.1 Background Information

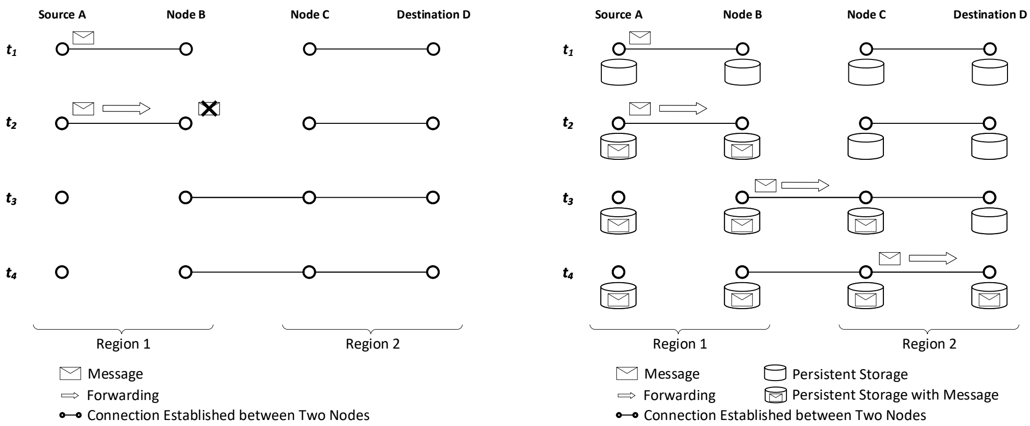
Internet of Things (IoT) devices are being deployed in various environments such as smart cities for environmental monitoring, agricultural lands for precision agriculture, and industrial facilities for Supervisory Control and Data Acquisition (SCADA) data logging. The connectivity between IoT devices is hampered when the end-to-end network path between the source and destinations does not exist, especially in areas with intermittent or poor network connectivity, such as disaster areas. Delay Tolerant Networking (DTN) enables communication in such disruption-prone networks, but multicast support is required in specific group-based data delivery services. The complexity increases when provisioning multicast support in resource-constrained IoT environments. The following subsections present a brief overview of IoT, DTN, and multicast support in DTN.

1.1.1 Internet of Things

The IoT is a heterogeneous network comprised of smart objects with embedded technology capable of information sharing and responding to changes in the environment without the need for human-to-human or human-to-computer interaction (Höller et al., 2014). With the popularity of IoT, conventional devices like computers and smartphones are connected to the Internet. However, the inter-connectivity has also been extended to include embedded devices ranging from small sensors, actuators, and ordinary everyday objects to sophisticated industrial automation machines. Every "Thing" in IoT is given a unique identifier to communicate with other "Things" and accomplish certain functionality. The proliferation of IoT devices has enabled diverse applications and solutions across domains such as smart environments, remote monitoring, healthcare, industrial automation, and so on. Wireless Sensor Network (WSN) and Mobile Ad Hoc Network (MANET) serve as the key enabling technologies for the application of the IoT paradigm in urban scenarios (Bellavista et al., 2013). Both WSN and MANET work considerably well in urban environments with stable network connectivity despite the node mobility and limited resources (Bruzgiene et al., 2014; Singh et al., 2014). In other words, each of the connected nodes assumes consistently available end-to-end network links between source and destination using the underlying conventional networking protocols such as Transmission Control Protocol (TCP) and User Datagram Protocol (UDP). Hence, the concept of IoT is feasible on the foundation of well-established telecommunication infrastructure in the cities and towns to ensure continuous network connectivity. However, the heterogeneity and mobility in IoT hinder reliable communication between the nodes (Elsaadany & Aboulhassan, 2019). Their performances suffer from multiple constraints such as limited energy, storage, and bandwidth that cause disruption (Zahoor & Mir, 2018). Furthermore, when IoT devices are deployed over vast geographical areas, and to make matters worse, places like rural areas or countrysides usually lack of telecommunication infrastructures to provide lasting network connection and result in partitioned networks without end-to-end connectivity between them. The conventional Transmission Control Protocol/Internet Protocol (TCP/IP) is no longer suitable for areas with intermittent or poor wireless network connectivity.

1.1.2 Disruption Tolerant Networking

DTN is an alternative solution developed to overcome the intermittent connection problem in the aforementioned challenged networks. It employs the Store-Carry-Forward message switching paradigm by overlaying a Bundle Protocol (BP) on top of the Transport Layer (Scott & Burleigh, 2007). DTN enables data forwarding in the disrupted network without end-to-end connectivity due to the mobility and distribution of nodes across a wide geographical area. The communication of IoT devices in the absence of an end-to-end communication path is achievable by introducing data forwarding nodes, known as "data mules", equipped with short-range or long-range wireless communication links in DTN to bridge the gaps between partitioned networks. DTN is one kind of MANET with intermittent connectivity, also known as sparse MANET. The ability to bridge partitioned networks will expand the reach of IoT for use in vast areas for many IoT applications, where existing networking protocols such as TCP and UDP are not suitable.



(a) TCP/IP Packet Switching

(b) DTN Message Switching

Figure 1.1 Difference of Message Transfer between TCP/IP and DTN.

Figure 1.1 shows the difference in message transfer between TCP/IP and DTN.

When using the conventional TCP/IP, an end-to-end communication path is established between the source and destination to perform data exchange over the network. However, the end-to-end communication path between IoT devices does not exist due to node sparsity and mobility in the network. As shown in Figure 1.1a, there is no end-to-end path to destination D when source A creates the Message at time t_1. At time t_2, source A forwards the message to node B. However, the message is dropped due to the lack of an end-to-end path. Node B moves from Region 1 to Region 2 at t_3. Even though node B established a connection with node C up to t_4, TCP retransmission is not feasible because node B has moved away from the communication range of source A, where the conventional Internet Protocol (IP) fail under such condition. In a DTN-enabled IoT environment in Figure 1.1b, source A generates the message as usual and stores it as a bundle in its persistent storage at t_1. At t_2, it forwards the message bundle to node B, and node B also stores the bundle in its storage. Node B carries the bundle when moving from Region 1 to Region 2 until it encounters a node C and establishes a communication link for the transmission of the message bundle at t_3. Although Node B is disconnected from source A, it can still forward the message originated from source A thanks to the DTN store-carry-forward message switching capability by storing the message bundle in its persistent storage. Finally, node C forwards the message bundle to destination D at t_4. The DTN architecture enables the communication between IoT devices without an end-to-end communication path, where the traditional TCP/IP is not feasible.

1.1.3 Multicast Support in DTN

The communication between source and destination without an end-to-end network link in DTN is possible with the implementation of BP. Group communication is an essential service to enable information exchange and sharing within a group in such networks. Many DTN applications require group communication in their operations to distribute data to a group of members. However, the multicast approaches proposed for the well-connected networks such as the Internet and MANET are not suitable due to the intermittent nature of DTN. Those approaches always assume end-to-end paths between the source and destinations are available. For example, Reliable Multicast Transport Protocol (RMTP) requires continuous network connectivity, and Epidemic-based Reliable and Adaptive Multicast for Mobile ad hoc networks (EraMobile) only works in MANETs with temporary network topology establishment (Genç & Özkasap, 2007). One primitive approach to achieve group communication in DTN is by generating unicast messages for each member of the group. However, the drawback of this simple approach is the high consumption of network resources which leads to poor network performance, especially in resource-constrained DTN (Zhao et al., 2005b).

Multicast is a mechanism to support the distribution of data to a group of receivers. The sender sends only a single multicast message addressed to the group rather than duplicating the message for each receiver in the group. For example, several hundred IoT sensors used to monitor water levels in a flood-prone area that need to be reconfigured to handle an upcoming rainy spell could use the help of multicast service. By adopting a multicast service, the group-based data dissemination in DTN applications achieves higher efficiency with lower transmission overheads and resource usage than unicast. Enabling multicast support in DTN faces more challenging problems than MANET, but it was made possible with data mules.

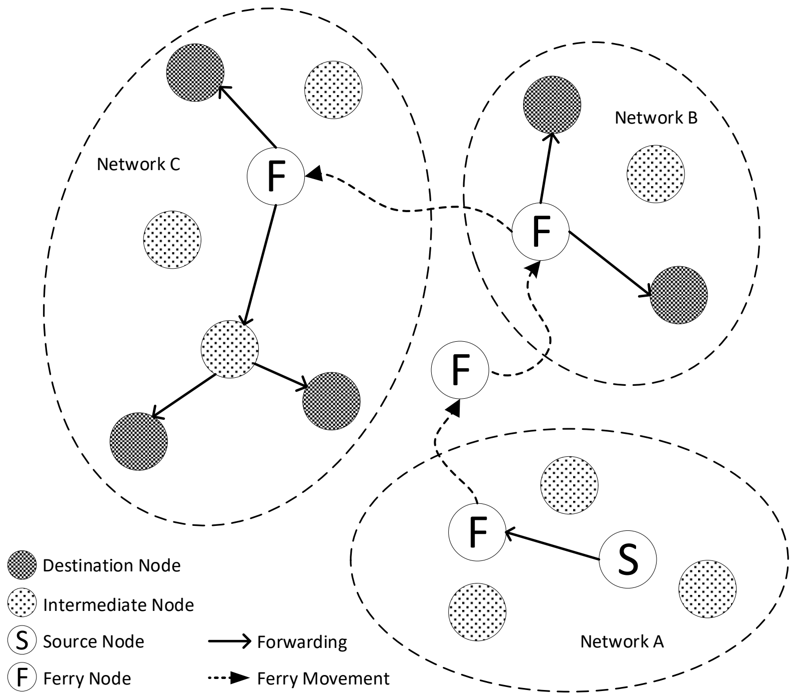


Figure 1.2 Multicast Scenarios Using Data Mule in DTN.

As shown in Figure 1.2, two groups of nodes are distributed sparsely and formed three partitioned networks with nearby nodes. The source node S in network A generates a multicast message for all the members of the temperature group and is forwarded to data mule F. F carries the message and visits network B to forward the message to other group members. When F visits network C, it forwards the message to the members and utilizes non-members nodes to forward the message.

Many DTN multicast routing approaches in the literature attempt to improve the performance through best-effort delivery (Afanasyev et al., 2009; Roy et al., 2017; Santiago et al., 2008, 2009). However, the limitation in IoT environments further increases the complexity of group-based data delivery services, not to mention some applications requiring reliable multicast support to manage and configure deployed devices across multiple geographical areas. Hence, data collection and dissemination become more error-prone in intermittent networks with poor network coverage, and it is tedious to verify that each device was received properly. Reliable multicast is required to ensure that the data is transmitted efficiently to the group of receivers, where the data is received by all intended recipients correctly, and the sender can verify the data delivery. This kind of application imposes more stringent requirements on the DTN routing mechanism when multiple geographical areas are involved.

Since DTN utilizes node mobility to facilitate routing in the area with intermittent network connection, their membership and geographic information play an essential role in making relay selection decisions to improve routing performance. One of the most common localization techniques to obtain the node location is the Global Positioning System (GPS). The GPS coordinate can be converted into a geohash string by geohash encoding. The nodes in the map area can be clustered hierarchically based on the common prefix of their geohashes. Nodes with long common prefixes are near to each other in general. This study proposes a novel geographical multicast mechanism DTN to provide multicast service with good deliverability and reduced transmission overhead. The purpose is to disseminate data to a group of participating nodes in a DTN-based IoT network by utilizing group membership and geographic information with an acknowledgment aggregation mechanism.

1.2 Problem Statements

The emergence of IoT continues with the advancement of telecommunication technologies that rely heavily on network infrastructures such as base stations for 5G networks. Many IoT applications require a continuous network connection to ensure reliable communication (Al-Fuqaha et al., 2015; Safaei et al., 2018). If the network connection is disrupted either through forwarding node failure or network problems, messages would be lost unless the application implements manual retry and retransmission logic to overcome the problem (Shang et al., 2016). More importantly, there may never be an end-to-end path between the sender and the receiver in some usage scenarios. To overcome the intermittent connection problem, DTN has been developed by exploiting data mules to forward data between partitioned networks. In some applications, IoT devices are distributed geographically across vast areas without network infrastructure, and reliable information dissemination to a group of IoT devices using MANETs is not feasible due to the sparsity of nodes. Many proposed routing schemes in the literature focus on Vehicular Delay Tolerant Network (VDTN) that consider only highway scenarios or urban scenarios (Liao & Zhang, 2018; Rahimi & Jabraeil Jamali, 2019), which did not consider the heterogeneity of network (Chourasia et al., 2019). In DTN unicast communication and dissemination, the relay nodes only need to know the location of the destination in order to make forwarding decisions. However, the problem of multicasting in DTN is much more complex, especially when the multicast receivers are spread in vast areas, not just in some targeted areas (Liao & Zhang, 2018). With the advancement of technology, many IoT devices are equipped with GPS (Palma et al., 2012). They can act as data mules utilizing the obtained location information to forward data to a group of IoT devices distributed across multiple geographical areas that share the same interest.

One of the possible scenarios is the group communication in post-disaster response for different parties when continuous network connections are no longer available due to the extensive damage on the network infrastructures (Maharjan et al., 2019). There are multiple groups of relief and rescue teams responsible for different operations in the affected area, such as police, army, fireman, and medical team. Depending on the scale of the disaster, the affected area might be divided into different sites, and different parties work in their team at different sites. They all carry some sort of communication device when they move along the affected areas and are mobile in nature. However, information exchange is required between different sites in order to provide immediate rescue and emergency resources. For example, vehicles and drones can be exploited as data mules for information dissemination when they travel along the affected area to transport supplies and conduct surveillance. A data mules will forward multicast messages to other data mules with more contact with non-ferry nodes in different areas to expand the reachability, or forward the messages to non-ferry nodes who have contacted with most data mules in a particular area.

Based on the target scenario in the post-disaster area, the problem statements are identified as follows:

1. The need for a group-based data delivery service in disruption-prone networks with multicast support over multiple geographical areas: DTN introduces data mules to overcome the network partitioning and bridge the partition networks to enable communication without an end-to-end network path between source

and destinations. However, group membership information is insufficient for efficient data dissemination to multicast groups in such networks with intermittent connectivity. Different working groups in the disaster area are distributed sparsely but dense in their cluster. For example, the availability of rescue resources in a post-disaster area is significant to different groups, and multicast support is required to disseminate the resource information in the intermittent networks. The multicast algorithm needs to consider group membership information and other factors to provide efficient data sharing to a group of intended receivers over multiple affected areas. The group-based data delivery service should help a sender forward multicast messages to at least 90% of the intended recipient with lower transmission overheads incurred during multicast forwarding.

2. The need for a reliable group-based data delivery service in the disruption-prone resource-constrained IoT environment: IoT network consists of heterogeneous devices that include battery-powered devices with low storage capacity and short communication range. The best effort delivery by flooding the network with multiple copies of a message is no longer suitable due to the limitation of the devices. For example, the device storage will soon be filled up and requires dropping the older messages before they get populated to other geographical areas. The propagation of information on rescue resources in the disaster scenario will be hampered by intermittent connectivity and limited resources in such an environment. The incorrectness and failure in information sharing might affect the operation afterward, which causes inaccuracy in communication. There is a need for a DTN multicast mechanism to provide reliable multicast support in

resource-constrained IoT to support reliable group-based data delivery services such as correct updating of device configuration in disaster scenarios.

3. The need for a group-based data delivery mechanism optimized for different bandwidth links: The lack of reliable wireless network connectivity in areas without proper network coverage would make data dissemination error-prone and tedious to verify whether the multicast messages have been delivered properly to all intended recipients. Enabling delivery acknowledgment (ACK) to confirm message reception by every group member will result in ACK implosion. Each sender's multicast message results in a flood of ACKs, one from each receiver. This leads to significant network overheads, especially in network links with low bandwidth. Furthermore, the sender will be overwhelmed when each receiver acknowledges each multicast message. There is a need to optimize the support of group-based data delivery services in different bandwidth links to reduce the transmission overheads resulting from the ACK implosion from the IoT receivers when confirming the messages have been received correctly.

1.3 Research Objectives

The objective of the thesis is to design and develop a routing mechanism that can ensure efficient and reliable data dissemination for IoT devices and applications in use cases where TCP/IP is not suitable due to intermittent network connectivity.

The research objectives are:

1. To define a mechanism for group-based data delivery services for disruption-

prone networks targeting a 90% multicast delivery ratio with reduced transmission overhead compared to the conventional approaches.

2. To enhance group-based data delivery services for supporting reliable multicast over disruption-prone resource-constrained IoT networks by improving the multicast delivery ratio with a trade-off in terms of transmission overheads.

3. To optimize group-based data delivery services in different bandwidth links by proposing a mechanism to reduce acknowledgment overheads.

1.4 Expected Contributions

The expected contributions are as follows:

1. The work expects to fill the research gap in provisioning group-based data delivery service in a disruption-prone IoT network by defining a novel DTN multicast routing mechanism targeting a 90% multicast delivery ratio with a reduction in multicast transmission overheads compared to conventional approaches.

2. The enhancement of DTN multicast routing mechanism with reliable multicast support in group-based data delivery over disruption-prone resource-constrained IoT networks by improving multicast delivery ratio with a trade-off in terms of transmission overheads.

3. The DTN multicast routing mechanism will be optimized for group-based data delivery in different bandwidth links by reducing transmission overheads caused by the flooding of acknowledgment messages.

1.5 Scope and Limitations

The focus of this research is to propose and design a novel geographical multicast DTN mechanism for the disruption-prone IoT environment, which considers the issues of multicast routing in a sporadic and resource-constrained environment. A wide range of IoT applications across various fields have different requirements, and there is no multifaceted solution for routing issues in DTN. The followings are the scope and limitations of this research:

- This study considers only the routing of multicast bundles in the bundle layer, although the proposed mechanism should work with lower-layer protocols using respective protocol-specific convergence layers. For example, the proposed solution does not emphasize any specific media access or wireless technology. However, the proposed mechanism in the testbed methodology is setup using the TCP convergence layer and works under Ethernet, Wi-Fi, and Bluetooth setup.

- This thesis does not consider buffer management, and the default First In First Out (FIFO) policy applies to BP bundle queuing and dropping. Energy management is another major topic to be investigated and considered as the future works of this research.

- While multicast security and group key management are complex issues in DTN and IoT networks to ensure data integrity and privacy (Menesidou et al., 2017), data encryption inflicts more significant security overheads (Tselikis et al., 2013). However, it is out of the scope of this thesis.

- The recovery of unreachable nodes is not covered in this research because the unreachable nodes are likely to fail, or they are never in the predefined route of the data mules. Such conditions require manual intervention, which is not a part of the proposed mechanism.

1.6 Organization

The organization of this thesis is as follows: Chapter 1 provides an overview of the research background, including the presentation of the problem statements, objectives, and scope of the study. Chapter 2 outlines the current state of multicast routing approaches in DTN and identifies the essential features in proposing a DTN multicast routing mechanism aligned to the research objectives. Chapter 3 presents the proposed solution in designing a geographical multicast DTN mechanism for group-based data delivery service in the disruption-prone network and the overview of experimental methodologies to evaluate the proposed mechanism. The methodologies and setup for testbed and simulation experiments are described in Chapters 4 and 5, presenting results and discussion before ending the chapters. Chapter 6 revisits the research objective and outlines contributions together with future work related to the study area.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

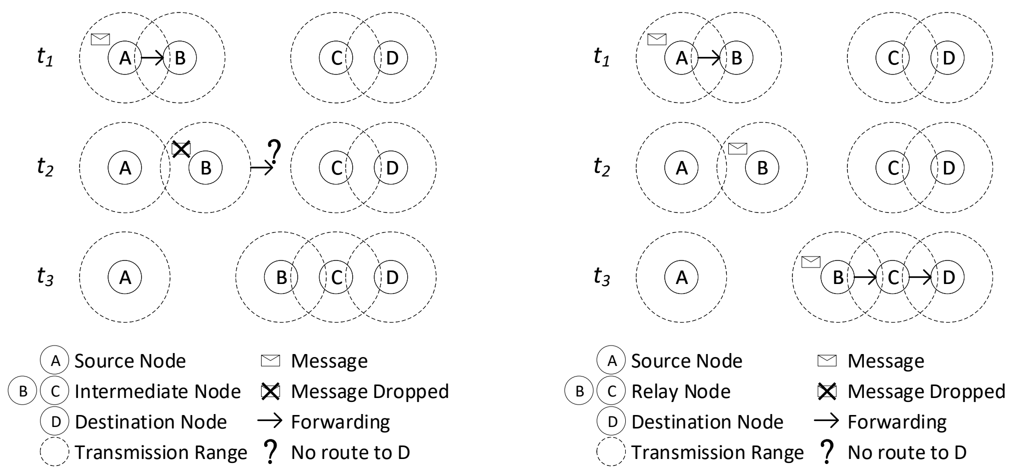
This chapter begins with a background of the study on the difference of DTN from the conventional MANET in Section 2.2, along with the enabling disruption tolerant solution to address intermittent connectivity issues in the IoT context. Section 2.3 covers the related work of the research. The multicast challenges in DTN are first outlined, followed by the state of the art of multicast routing in DTN. A taxonomy is presented to classify the existing multicast routing schemes into four major categories: social-aware, buffer-efficient, energy-aware and geographical. Next, the requirement and constraints to enable multicast support in the resource-constrained IoT network are identified, and the existing DTN bundle protocol implementations are reviewed. Section 2.3 ends with the explanation of geographical and Jaccard index features in provisioning reliable multicast support in DTN-enabled IoT networks. Section 2.4 presents a brief discussion on addressing the research problems with the insights obtained from the literature and identifies the essential features for the proposed solution.

2.2 Background

This section first presents the fundamental differences between DTN and MANET and explains why conventional TCP/IP failed to work in the absence of end-to-end network paths. The significance of DTN is then discussed, followed by an overview of enabling DTN solutions in the IoT context.

2.2.1 DTN and MANET

Originated from the Interplanetary Internet for deep-space communication (Burleigh et al., 2003), the concept of DTN has been adapted to address communication issues in disruption-prone terrestrial networks with various technologies such as MANET (Kang & Chung, 2020; Krug et al., 2018). DTN is rooted in the MANET and has less stringent requirements compared to MANET. Both MANET and DTN comprise mobile nodes in an infrastructure-less environment. The mobility of nodes causes frequent network partition in both networks. The MANET nodes communicate by using the existing network protocols such as TCP and UDP, which assumes an end-to-end path is always available.



(a) Message Passing using TCP/IP

(b) Message Passing using DTN

Figure 2.1 Difference of Message Passing between TCP/IP and DTN. (Redi & Ramanathan, 2011; Vieira et al., 2013)

MANET allows routing between moving nodes and can handle quick changes in the network topology but cannot handle missing end-to-end connections where the TCP/IP fails to work properly as shown in Figure 2.1a. At time t_1, node A creates a message with node D as destination and forwards it to node B within its transmission range. Node B moves away from node A at time t_2 but does not have any other node to forward the message. Therefore, node B dropped the message since there is no route to node D at the moment. When node B is in contact with node C at time t_3, no message is available to be sent. However, an end-to-end communication path is established between nodes B and D. The routing in MANET is best-effort delivery. Packets are dropped when end-to-end connectivity is not available.

On the other hand, DTN implements a bundle layer between the transport and application layers to support data delivery using store-carry-forward mechanisms in the absence of end-to-end network paths. Due to the opportunistic nature of DTN, there is a meager chance that an end-to-end network link exists between a pair of nodes at any given time. In Figure 2.1b, the message created by node A is forwarded to node B at t_1, and node B stores the message as a bundle in its persistent storage. At t_2, node B moves away from node A while carrying the bundle. When node B is in contact with node C at time t_3, it forwards the bundle to node C. Finally, the bundle is delivered to node D. DTN allows intermediate nodes to store the bundle until they encounter the destination node or another relay node that can forward it.

Figure 2.2 shows the flow of bundle from source to destination for DTN protocol stack, where the bundle layer sits on top of the transport layer of the network stack. Assumed that Source A never meets Relay C and Destination D, it can only chance is to forward message bundles to Relay B. Relay B will store the messages bundles into the buffer and carry them until it encounters Relay C. Relay C further stores, carries, and forwards the message bundles to the Destination D. There is no end-to-

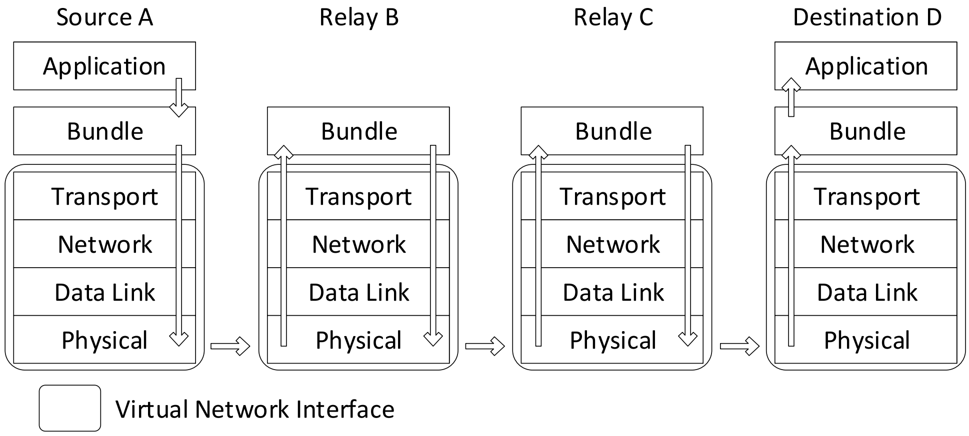


Figure 2.2 The Flow of Bundle from Source to Destination.

end communication path between Source A and Destination D. However, the message bundles are still able to reach the destination. Network partition in MANET will result in packet loss, whereas relays in DTN store the message in their buffer until there is an appropriate node to relay the message. Furthermore, the BP is independent of the lower layers' difference that allows heterogeneous devices to work together in the network. Ito et al. (2013) discussed the features of transmission methods for MANET and DTN and compared their performances in terms of hop count and transmission success probability.

2.2.2 DTN in IoT

The proliferation of smart devices further pushes forward the research endeavor in enabling DTN in the new IoT paradigm. Some domains in IoT experience similar challenges of frequent network disruption and long propagation delay when there is a lack of well-established communication infrastructures. The intermittent connectivity and delayed communication worsen in the IoT networks, especially with heterogeneous resource-constrained and high mobility devices.

In recent years, various research efforts have been made in the literature to enable DTN solutions in opportunistic IoT networks. Abdellaoui Alaoui et al. (2019) attempted to address the heterogeneity issue in IoT by combining the advantages of flooding and forwarding strategy to increase deliverability with lower latency and energy used. Cuka et al. (2019) introduced two fuzzy-based systems to select the best candidates from a group of IoT devices in DTN to carry out a task based on different metrics such as storage size, waiting time, remaining energy, and security of the IoT devices. Maharjan et al. (2019) proposed a multi-dimensional Cloud-Fog-IoT integrated platform with DTN-enabled communication in disaster scenarios. The contents of centralized cloud services for disaster management are distributed to a large number of fog servers in the incident areas, where mobile users are able to query information from the fog servers using IoT devices such as smartphones. However, the authors only considered unicast schemes such as Probabilistic Routing Protocol Using History of Encounters and Transitivity (PRoPHET) and Epidemic in their simulations.

Yaacoub et al. (2020) exploited data mules such as buses to transmit IoT mobile health data of patients in remote villages with limited network connectivity to the nearest hospital for remote monitoring and diagnosis. The collected data is fragmented into multiple parts, and each part is encrypted separately with secure keys shared between the hospital and patients. Each encrypted part is routed to the hospital through different mules for resembling and further processing. This end-to-end secure DTN scheme only focuses on the unicast data transmission from patients to the sink with extra security processing, which is not suitable for multicasting services in resource-constrained networks.

Benhamida et al. (2017) presented the relationship between DTN performance metrics and IoT characteristics and how their dependency affects the design of DTN solutions to address IoT challenges. They suggested that IoT characteristics should outweigh in designing DTN solutions when similar performance metrics are considered. Bounsiar et al. (2019) extended the work from Benhamida et al. (2017) and presented a comprehensive review of the research works that enable DTN solutions to address disruptive and delayed communication challenges in the IoT environment. Their taxonomy includes DTN solutions based on the BP implementations (e.g., IBR-DTN, μDTN, and nanoDTN for resource-constrained devices) and the design of routing approaches in resource-constrained IoT networks such as using scheduling policies (Mao et al., 2019) or adopting butter management schemes (Zguira et al., 2018).

2.3 Related Work

This section first outlines the multicast challenges in DTN, followed by a taxonomy of DTN multicast routing classification in the literature. The taxonomy has classified the surveyed multicast schemes into four major categories, and each category reviews related works according to their characteristics.

2.3.1 Multicast Challenges in DTN

The lack of network infrastructures such as wired and wireless infrastructure-based networks in DTN greatly impacts the delivery ratio of message bundles in the networks. Infrastructure-less MANETs compute end-to-end routing paths, but DTN is not able to do the same. Existing MANET protocols such as Ad Hoc On-Demand Distance Vector (AODV) and Dynamic Source Routing (DSR) cannot be applied in DTN because the construction of a fully connected graph to forward message is required. Otherwise, it will result in packet loss. The nodes in DTN store the message bundles until they find an appropriate relay node to forward the bundles towards the destination. Relay node selection directly affects the routing performance in terms of delivery ratio and latency. A good routing algorithm has to be investigated to achieve low delivery overhead. Due to the DTN characteristics of intermittent connectivity and the long communication delay, it is challenging to develop an optimal forwarding strategy for routing in DTN.

2.3.2 Multicast Routing in DTN

The DTN-enabled IoT applications require multicast routing support to deliver data efficiently to a group of nodes without compromising the overall network performance. There is no multi-faceted solution to address different multicasting issues in various DTN and IoT environments. Most of the researchers proposed problem-specific solutions based on specific scenarios or application areas. In order to gain an overview and insight on multicast routing in DTN, this section presents different multicast routing schemes that are significant and recent in the literature. Figure 2.3 shows the taxonomy of multicast routing in DTN. Four main groups of multicast routing schemes are studied: Social-aware, buffer-efficient, energy-aware, and geographical routing. Unicast routing in DTN is not in the scope of this thesis since comprehensive surveys and reviews have been done in the literature (Cao & Sun, 2013; Cc et al., 2016).

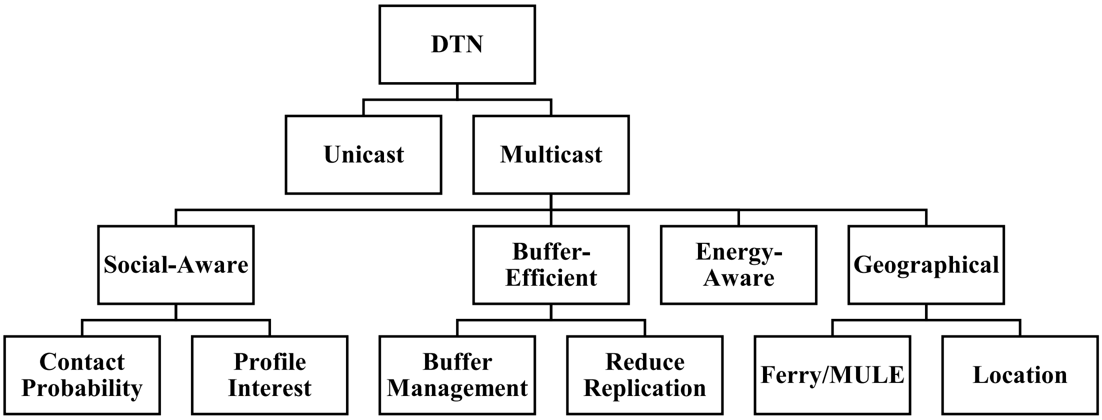


Figure 2.3 Classification of DTN Multicast Routing Strategies.

2.3.2(a) Social-Aware

The DTN nodes usually consist of portable devices carried by people, and they possess social behaviors from their owners, such as visiting certain areas or sharing common interests. These social behaviors facilitate message forwarding in groups of nodes. People with a common interest tend to meet up more frequently, and this interaction pattern speeds up the message propagation in the Social Wireless Network (SWN). For example, the smartphone gathers traffic information along the road when the owner drives to work from home, and the collected information is disseminated to other drivers within the area who are interested in that traffic information. Some social features among mobile users are stable due to their long-term social relationships such as affiliation and language. These are positive social features that are commonly utilized to improve the efficiency of multicast routing in DTN. However, social behaviors might also hinder the routing process when a selfish node wants to conserve energy for more meaningful information to themselves. Table 2.1 compares the performance of social-aware multicast routing in DTN. Most of the works leveraged positive social features to improve the efficiency in multicast data delivery. Epidemic routing is the standard benchmark algorithm to compare the performance thanks to its flooding-based nature leading to high deliverability.

Single-Data Multicast (SDM), Multiple-Data Multicast (MDM) (Gao et al., 2012), Cloud-based Multicasting (Wang et al., 2013) and Two-Level Multicast Routing (TLMR) (Le et al., 2015) utilize social centrality and community for relay selection by calculating the contact probability of a node to the multicast destinations. SDM and MDM minimize the number of forwarding nodes and have lower transmission costs since fewer multicast message replica is duplicated. In cloud-based multicasting, the destination cloud size affects the performance. A small cloud size results in lower latency but higher transmission costs because more relays are used to do multicast forwarding. However, the FIFO message dropping policy drops older messages before they reach the destination cloud. TLMR (Le et al., 2015) forwards multicast messages through the route with lower latency despite more relays and higher delivery costs. Nodes with lower social centrality and community characteristics might not guarantee performance. The Social Difference Multicast Routing (SDMR) (Deng & Chang, 2013) presents two social metrics to calculate the social differences between nodes to expand the message coverage and save resources for nodes with high centrality. The relays carry the message further away and lead to high delivery latency.

Q-learning based Gain-aware Routing (QGR) (Hajiaghajani & Biswas, 2017) implements a coupon dissemination mechanism according to the interest level of a user on specific information. It minimizes the delivery cost with a longer propagation time. Polymorphic epidemic multicast routing (Lorenzo et al., 2013), Multi-CSDO, and Multi-CSDR (Chen et al., 2016) construct interest profiles based on the offline social features. Furthermore, all the social-aware multicast routing strategies do not

Table 2.1 Comparison of Social-Aware DTN Multicast Routing Schemes.

Scheme	Key Solution	Delivery Ratio	Delivery Cost	Delivery Delay	Energy	Mobility	Compared algorithm
SDM	Centrality and Community	Slightly Lower	Low	Slightly Lower	-	Random Way Point	Epidemic
MDM	Destination-Awareness	Slightly Lower	Low	Similar	-	Random Way Point	Epidemic
Cloud-based	Destination Community	Similar	Very Low	Much Higher	-	Real Traces	Epidemic and Small destination size
TLMR	Social-tie Strength	Slightly Lower	Higher	Much Lower	-	Real Trace	Epidemic
SDMR	Similarity and Centrality Difference	Much Lower	Low	High	-	Real Trace	Epidemic
QGR	Consumption Interest and Redemption Probability	-	Low	High	-	Real Trace	SANE
Polymorphic Epidemic	Social Features and Interest	-	High	Low	Consumption Metric	High	SANE
Multi-CSDO Multi-CSDR	Social Features & Interest	Slightly Lower	Low	Moderate	-	Real Traces	Epidemic

consider the energy of the nodes in data forwarding, which is significant in social-based networks with high node mobility.

Multicast routing in DTN can be improved by combining different static and dynamic social metrics to improve the prediction of the mobility patterns of nodes since social behaviors are not only dependent on encounters. Although the combination of different social metrics might improve the routing performance, inappropriate combinations will hamper the overall network performance and increase the complexity in designing a multicast routing strategy with multiple metrics. Furthermore, some selfish nodes are unwilling to spare their resources to facilitate multicast routing. Keykhaie and Rostaei (2017) studied the selfish behavior of nodes and proposed a Congestion Aware and Selfishness Aware Social Routing scheme (CASASR). Li et al. (2011) studied how node selfishness affects multicast routing in DTN and how the multicast group size affects routing performance.

2.3.2(b) Buffer-Efficient

The heterogeneous IoT environments are partly made up of resource-constrained devices with limited storage capabilities, and DTN requires them to store the message bundles until there is an opportunistic link for further forwarding. Multicasting service in DTN causes high buffer overhead because messages are replicated to improve delivery ratio using multi-copy-based routing. Under normal circumstances, obsolete messages are removed from the node when the buffer is full. However, it is difficult to ensure that all destinations have received the message before it is removed. Without proper buffer management, the multicast delivery ratio will be low when a large