Modeling and Analysis of Content Distribution in Disruption Tolerant Networks

Tuan-Minh Pham

Thèse de Doctorat
UPMC Sorbonne Universités

Spécialité

Informatique

présentée par

M. Tuan Minh Pham

pour obtenir le grade de

Docteur de l'UPMC Sorbonne Universités

Modélisation et Analyse de la Distribution de Contenus dans un Réseau DTN

Soutenue le 14 décembre 2011 devant le jury composé de

Philippe NAIN	**Rapporteur**	**Directeur de Recherche, INRIA**
Kavé SALAMATIAN	**Rapporteur**	**Professeur, Université de Savoie**
Mostafa AMMAR	**Examinateur**	**Professeur, Georgia Institute of Technology**
Mario GERLA	**Examinateur**	**Professeur, University of California, Los Angeles**
Sebastien TIXEUIL	**Examinateur**	**Professeur, UPMC Sorbonne Universités**
Serge FDIDA	**Directeur de thèse**	**Professeur, UPMC Sorbonne Universités**

Doctor of Science Thesis
UPMC Sorbonne Universités

Specialization

COMPUTER SCIENCE

presented by

Mr. Tuan Minh PHAM

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF SCIENCE of the UPMC Sorbonne Universités

Modeling and Analysis of Content Distribution in Disruption Tolerant Networks

Committee in charge:

Philippe NAIN	**Reviewer**	**Research Director, INRIA**
Kavé SALAMATIAN	**Reviewer**	**Professor, Université de Savoie**
Mostafa AMMAR	**Examiner**	**Professor, Georgia Institute of Technology**
Mario GERLA	**Examiner**	**Professor, University of California, Los Angeles**
Sebastien TIXEUIL	**Examiner**	**Professor, UPMC Sorbonne Universités**
Serge FDIDA	**Supervisor**	**Professor, UPMC Sorbonne Universités**

Acknowledgement

I want to thank my teacher at Hanoi University of Science and Technology, Thuc Hai Nguyen, and my teachers and colleagues at Hanoi National University of Education, Cam Ha Ho, The Loc Nguyen, Thi Tinh Nguyen and Tho Hoan Pham, for encouraging me to pursue my studies at the University Pierre et Marie Curie.

I express my deep gratitude to my advisor Serge Fdida who welcomed me in 2008 for starting my studies at LIP6. He raised my interest in the subject area of this thesis, guided and supported me in every step of this thesis along these years. I thank Serge for his vision and invaluable lessons that he taught me to research, and for being patient with my questions. It is my great fortune to have pursued my Ph.D. studies under his guidance at LIP6.

I am truthfully grateful to the committee members, Philippe Nain, Kavé Salamatian, Mostafa Ammar, Mario Gerla and Sebastien Tixeuil, who were willing to take responsibility for assessing this dissertation. I also thank the Vietnam Ministry of Education and Training, and the French National Research Agency for their financial support throughout my thesis.

I would like to thank the researchers whom I had the pleasure of meeting at LIP6 and at various events over the past several years, for the open discussions and their valuable comments on part of this research. While any list will inevitably be incomplete, I would like to mention here Mostafa Ammar, Bruno Baynat, Alexandre Brandwajn, Vania Conan, Marcelo Dias de Amorim, Mario Gerla, Jeremie Leguay, Naceur Malouch, Viet Hung Nguyen, and Kavé Salamatian. I am grateful to Martin May and Thi Mai Trang Nguyen for their precious advice on the midterm review of my thesis. I extremely thank Kanchana Kanchanasut for critical reading of the manuscript. Special thanks to Thi Kim Ngan Nguyen and Prométhée Spathis for their corrections on my French writing.

I have been lucky to be a member of the NPA team where I have enjoyed discussing, chatting with my friends during memorable times. In particular, thanks to my friends: Raul Adrian Gorcitz, Asim Ali, Panayotis Antoniadis, Thomas Begin, Nadjet Belblidia, Mohamed Diallo, Mubashir Husain Rehmani, Yesid Jarma Alviz, Adisorn Lertsinsrubtavee, Sophia MacKeith, Mohamed el Mehdi Bezahaf, Hoang Anh Phan, Tiphaine Phe-Neau, Thomas Silverston, Alexandru Tatar, John Whitbeck, Diana Zeaiter Joumblatt, and all the administrative staffs who support great working facilities at LIP6.

My life as a PhD student would not have been enjoyed without my friends from Paris.

Thanks to my friends, Nhu Cuong Tran and Anh Binh Tran, for their help on my initial days in Paris. I hope many of them will not mind for not being individually mentioned here, but their support, company, and sentiments are quite precious.

Most of all, I am sincerely grateful to my parents for their boundless love and tremendous sacrifices to ensure that I had a good education. I am in memory of my maternal grandmother who always believed me and gave me words of encouragement. I dedicate this dissertation to them with my greatest respect.

Résumé

Cette thèse étudie la faisabilité de la diffusion de contenu sur un réseau tolérant aux délais (DTN) dans une zone urbaine. L'application cible est la distribution de la version électronique d'un journal dans une grande ville. Bien que des contraintes de temps ne s'appliquent pas de manière stricte, il est tout de même attendu que la diffusion d'information se fasse dans un délai raisonnable. Deux métriques de performance sont considérés : le délai de message et le temps de propagation. Le délai de message est le délai nécessaire pour transmettre un contenu à partir d'un nœud mobile à un autre nœud, tandis que le temps de propagation est le délai nécessaire pour diffuser un contenu sur un ensemble de nœuds du réseau.

Premièrement, notre objectif est de mesurer de manière analytique les performances d'un environnement DTN simple lorsqu'un contenu est distribué exclusivement à travers des contacts entre les nœuds mobiles. Nos contributions résultent de la prise en compte de la probabilité d'intérêt/acceptation dans l'expression en forme fermée et l'expression asymptotique du délai moyen de message. La probabilité d'intérêt/acceptation représente la probabilité qu'un contenu soit accepté par un nœud qui manifeste de l'intérêt pour ce contenu lors d'un contact. L'expression asymptotique permet de déterminer des moyens efficaces d'améliorer le délai moyen de messages dans une zone où la densité des nœuds mobiles est basse ou haute. Nous montrons aussi une relation entre le délai moyen de message et le temps moyen de propagation dans de tels environnements.

Deuxièmement, si le délai est jugé excessif, nous suggérons un déploiement de kiosques de données dans l'environnement afin d'améliorer les performances de la diffusion de contenu. Les kiosques de données sont des dispositifs simples qui reçoivent les contenus directement depuis leur source, le plus souvent en utilisant des réseaux filaires ou cellulaires. Un des problèmes posés pour concevoir efficacement un tel réseau est le nombre de kiosques de données qu'il faut déployer pour satisfaire aux objectifs de performance. Pour répondre à ce problème, nous déterminons les valeurs de la borne supérieure et de la borne inférieure du nombre de kiosques de données nécessaires pour distribuer le contenu dans une zone géographique en optimisant un temps moyen de propagation pris comme objectif. Nous montrons aussi une propriété importante que ces bornes varient linéairement avec le taux de contact entre un nœud mobile et un kiosque de données.

Enfin, nous considérons le problème de l'emplacement optimal des kiosques de données dans un scénario plus réaliste où les utilisateurs se déplacent en utilisant un système de transport (comme le métro ou le train de banlieue) qui relie plusieurs régions. Nous analysons le choix des stations de métro où installer un kiosque de données pour optimiser le temps moyen de propagation. Les résultats de l'analyse confirment que l'emplacement optimal des kiosques de données n'est pas seulement influencé par les caractéristiques d'une région, mais aussi par le nombre d'utilisateurs mobiles qui recevront le contenu. Nous validons nos résultats analytiques par des simulations en prenant différents modèles de mobilité ainsi que des données de mobilité résultant de mesures réelles.

Mots-clés :

DTN hybride, Délai, Distribution de contenu, Modélisation, Analyse de performance.

Abstract

This dissertation studies the practicality of content distribution over a Delay Tolerant Network (DTN) in an urban area. The target application is the distribution of the electronic version of a newspaper in a large city. Although strict time constraints do not apply, spreading the information should be achieved within a reasonable delay. Two performance metrics, the spreading time and the message delay, are considered. The message delay is the delay required to transmit content from a mobile node to another node, while the spreading time is the delay needed for the content to spread over a part of the network.

Firstly, our goal is to increase our understanding of the performance of a simple DTN environment when content is distributed solely through inter-contact of mobile nodes. We contribute both the closed-form expression and the asymptotic expression of the expected message delay to the literature when considering the probability of interest/acceptance for a given piece of content at each contact. The asymptotic expression provides the insights on the efficient ways for improving the expected message delay in the case of an area with low or high density of mobile nodes. We also show a relationship between the expected message delay and the expected spreading time in such environment.

Secondly, if the delay is found to be excessive, we suggest the deployment of some data kiosks in the environment to better support the dissemination of content. Data kiosks are simple devices that receive content directly from the source, usually using wired or cellular networks. A key issue when designing efficiently such network is to determine the number of data kiosks required to satisfy a performance target. We investigate both an upper bound and a lower bound of the number of data kiosks to distribute the content over a geographical area within an expected spreading time objective. We also show the important property that those bounds scale linearly with the contact rates between a mobile node and a data kiosk.

Finally, we consider the question of the optimal locations of data kiosks in a more realistic scenario where users move along a transportation system (like a subway or suburban train) that connects several regions. We provide an analysis used to decide which subway stop should host a data kiosk to optimize the spreading time. These findings support the view that the optimal locations of data kiosks are influenced not only by the conditions of a region but also by the target number of mobile users that will receive the contents. Analytical results are validated by simulations under a number of mobility models and real

datasets.

Key Words:

Hybrid DTN, Delay, News dissemination, Modeling, Performance analysis.

Table of Contents

Chapter 1

Introduction

This dissertation studies performance analysis and modeling of content distribution in Delay/Disruption Tolerant Networks (DTN) with an emphasis on a solution for satisfying a performance objective. An example of applications is to distribute the electronic version of a newspaper in an urban area through opportunistic contacts between people. For such applications, although strict time constraints do not apply, spreading the information should be achieved within a reasonable delay.

We begin this chapter by showing the application contexts of DTNs including the application context of this dissertation. Next, we show that due to challenging characteristics, building a content distribution system with a performance objective has become a central issue for deploying applications in DTNs (Sections 1.2-1.3). We then outline our solutions to content distribution with delay constraints (Sections 1.4-1.5). Finally, we summarize the contributions and give a road map to the rest of the dissertation (Section 1.6-1.7).

1.1 Application Contexts

Delay and disruption-tolerant networking is an approach to computer network architecture which focuses primarily on the issues in networks where instantaneous end-to-end paths are not available most of the time due to a lack of continuous connectivity [1, 2]. Examples of such networks are interplanetary networks where deep-space communications face extremely high delay, or terrestrial mobile networks where some of these networks may become unexpectedly partitioned due to node mobility. Applications of DTNs have arisen in those contexts and have been studied actively by the network community.

One of the earliest applications of DTNs is file delivery in the context of interplanetary networks [3]. The purpose is to make data communication between Earth and Mars, where orbiters around the planet act as communication relays, in support of Mars exploration. The TCP/IP suite [4] works well for the Internet but it is problematic in long-delay environments. The main reasons that make TCP failed to work for Earth/Mars communications are that TCP uses end-to-end signaling in data communications and adapts to changes in the state of networks. In particular, a TCP connection is disconnected if no data is exchanged for a value of timeout that is specified by TCP; hence, TCP is not suitable for high delay environments. Furthermore, TCP cannot work well even if we assume that the timeout can be assigned to a very high value although it is problematic in practice. Specifically, when a receiver detects an event of segment loss, it signals the event back to the sender. In the responds to this signal, the sender cut its rate in half [5]. While this characteristic of TCP works well in Internet, it prohibits the high utilization of link capacity in interplanetary networks where the propagation delay is extremely high. The DTN architecture has arisen in such contexts.

The sparseness and mobility of sensor nodes in many underwater networks increase the need for DTNs in the environments. While sensor nodes in terrestrial sensor networks usually are dense and stationary, they generally are sparse and mobile in underwater sensor networks due to the fact that underwater sensor nodes are more expensive than terrestrial sensor nodes and deployment areas in ocean environments are often large. In a sparse and mobile network, DTNs can be used to make data communications between nodes by exploiting the opportunistic contacts between nodes. Hence, DTNs are applicable in underwater networks where many potential applications have been explored such as seismic monitoring of oil fields, detection of chemical leaks, and undersea sensing of biological phenomena [6–8].

In rural areas in developing regions, network connectivity is lacking because wireless carriers do not offer coverage or the distances to wireless access points are long. In such contexts, DTNs are applicable to improve the coverage of networks for distributing information to and between people. In this case, content can be provided by ferrying data. Specifically, when means of transport such as buses, and motorbikes meet data sources, they retrieve and store data; then they deliver to users whenever they come in range with the users. [9–11].

In recent years, the mobile data surge has given rise to the use of DTNs for content distribution. Moving people who carry handheld devices with wireless capacity can exchange data during opportunistic contacts between them. Such exchanges can bring benefits to both users and service providers. By using device-to-device connections, users are free of charge for receiving data instead of being charged for cellular access. Service providers can support more subscribers at a lower cost. Many potential applications have been proposed in those contexts [12–16]. In such applications, the dissemination of content exploits the opportunistic contacts between people. Here the performance of DTNs has been shown to be one of the central questions before any application could be really deployed.

Free daily newspapers have been introduced in many countries worldwide for over 10 years now. Their distribution channel uses points of presence (Kiosks) often located at the entrance to the main transportation systems, such as metro or suburban train stations. They also have the presence on the web and exploit user's contributions and social networks. In addition, the widespread deployment of handheld devices provides the opportunity to use the opportunistic contacts between them for content distribution instead of being charged for cellular access. Information conveyed by free daily newspapers does not need to be instantaneous. However, owners of these newspapers expect that the information will reach the reader within a time window related to the period he/she will spend commuting from their home to their place of work as this is the best time to capture their attention. The aim of this dissertation is to understand and improve the performance of such a system with the existence and nonexistence of data kiosks. The contribution of the dissertation is not restricted to the distribution of news as many other applications also exhibit similar expectations, but it provides a practical use case.

Application areas of DTNs are pervasive. The explosive growth of mobile data and the advantages that DTNs offer have given rise to the application of DTNs to content distribution. Therefore, a central question is how efficient can a DTN support content distribution. In addition, we should explore how to build a content distribution system using DTNs with a proper performance objective. The importance and challenges of the above mentioned issues will be discussed in the next two sections.

1.2 Challenges of Content Distribution in DTNs

Delay and disruption-tolerant networking remains a very difficult problem despite its potential in many application contexts that range from deep-space networks to vehicular networks and further to underwater sensor networks [3, 8, 17]. Content distribution in DTNs should map the application requirements and the performance that the system can achieve. This problem is challenging for several fundamental reasons:

- Unlike traditional networks where connections are continuous, the hop-by-hop connection between DTN nodes is opportunistic and the end-to-end connection is not available most of the time. Mobile nodes use a store-carry-and-forward paradigm to forward content towards the destinations. This characteristic of DTNs makes the delivery of content with a proper performance more difficult.

- In DTNs, the network topology is highly dynamic due to the high mobility of DTN nodes. An edge of the network is established when there is a contact between mobile nodes. When mobile nodes are wireless devices carried by humans, it is required that we understand people contacts for analyzing the content distribution. In addition, mobile users are not always willing to store and carry any content for other users. Due to these reasons, understanding the performance of the content distribution in DTNs is difficult.

- The content distribution in DTNs depends on many variables such as the transmission ranges of mobile nodes, the movement of mobile nodes, the limited storage and energy, the number of sources, the number of destination, the density of mobile nodes, the location of sources, and the content size. On the one hand, we must make enough simplifying assumptions to allow us to solve a mathematical model. On the other hand, however, assumptions should be acceptable in practice. Due to the diversity of input information, analyzing factors that have the strongest influence on the performance of content distribution in considered contexts in DTNs and modeling these factors are challenging.

- The heterogeneity of variables such as contacts make the problem more complex. Integrating the behavior of individual node in an analytical model for content distribution in DTNs is nearly impossible. How to approximately characterize the heterogeneity in appropriate contexts is difficult.

- In a DTN where mobile nodes are wireless devices carried by humans, people can walk in a small region and usually use means of transport to move between regions. While content distribution highly depends on the movement of the people, parameterizing these factors is difficult.

Understanding the performance of content distribution in DTNs, and finding the number and the locations of infrastructure nodes for optimizing the performance are important to explore DTN performance to satisfy application constraints. There exists no satisfactory solution to the content distribution with a proper performance objective. This dissertation tackles the problems.

1.3 State of the Art

Delay and disruption-tolerant networking has a broad application environment where connections are disruptive and end-to-end paths are not available most of the time, such as terrestrial mobile networks, interplanetary networks, mobile sensor networks, underwater networks, vehicular networks and social networking. Each solution draws on different aspects of a particular context of networking scenarios and applications. Solutions for content distribution in the field of delay/disruption-tolerant networking or opportunistic networking can be roughly put into two groups:

- The first group focuses on infrastructure-independent solutions where content is disseminated with no needs of infrastructure. In this approach, DTN nodes exploit opportunistic contacts for delivering content towards final destinations. Many papers study forwarding algorithms that control the content replication and use history information of contacts or knowledge of the network to optimize performance [18–30]. Some work exploits the scheduled mobility to enhance the performance of DTNs. For example, [31] proposed an algorithm to find mobile nodes that can act as message ferries while [32] presented an algorithm to design a route that optimize the delivery delay with a constraint of the traffic demand. The DTNRG research group focuses on addressing the architectural and protocol design principles [2, 33, 34]. Other papers propose and extend application protocols to DTNs such as email, HTTP [15,35]. Most of these work provide promising results based on simulation or experiments. Since system models are valuable to estimate the performance of systems before applications are deployed, a number of studies use analytical tools to analyze performance metrics in several application scenarios in DTNs [36–45]. However, due to many challenging characteristics of DTNs as well as many application contexts, each of the proposed solutions selected some particular aspects that are a main concern in the context of their work. Hence, the need for understanding the performance of DTNs will always remain in different perspectives.

- The second group develops infrastructure-dependent solutions for content distribution. From understanding the performance of DTNs in several networking scenarios, recent work makes an effort to improve the performance of DTNs when the performance of DTNs cannot meet an application requirement. For improving the performance of DTNs, a number of papers use the existing infrastructure of communication [46–49] while others add some special infrastructure nodes [50–55]. The results are efficient for the context that the papers consider, and increase significantly the performance of DTNs. However, it cannot be a general solution to every application context. In addition, for each solution, the information that is used as an input to the algorithm needs to be explicitly discussed and it should be explained how it is obtained in practice. Most importantly, no research has found explicitly the answer for the important question of the number of infrastructure nodes needed to meet a performance objective. Furthermore, while the spreading time should be considered to fulfil the interest of all mobile users, the location of infrastructure nodes that optimize the spreading time is still open.

DTN networking clearly has the potential to support content distribution, and changes the way of communication. However, there are many challenges that must be overcome in order to guarantee wide deployment for applications of this kind of communication. While each application has different performance requirements, and each networking scenario has a performance limit, understanding the performance limit of DTNs for a specific application context is very important. When infrastructure is needed to meet a delay requirement, the development of solutions determining the number and the locations of infrastructure nodes with a performance objective is truly crucial in order to leverage real applications of DTNs in practice. This dissertation develops such solutions.

1.4 Goals of the Dissertation

The main goal of this dissertation is to build a content distribution system, which uses opportunistic contacts to distribute content to people on the move in an urban area, with a proper performance objective. For the problem, we can first analyze the performance of content distribution in DTNs when content is distributed solely through opportunistic contacts of mobile nodes. If the performance is unable to meet a delay requirement, we can deploy some data kiosks to support better dissemination of content.

Specifically, our goals include the following:

- Increase a better understanding of the performance of content distribution in DTNs where opportunistic contacts are used to enable communication. The solution should help us to measure the performance metrics that we concern, to understand the performance impacts of input information, and to derive how to improve the performance of content distribution in DTNs.

- Develop a solution for the deployment problem of the content distribution with the presence of data kiosks when the above solution suggests that some data kiosks are necessary to achieve a proper performance. The solution should help us to decide the number and the locations of data kiosks that optimize the performance of the content distribution system. Since people usually walk in a small region and use means of transport to move from a region to another region, the solution should take into account the movement of mobile nodes between regions.

- Develop the solutions that go towards reaching the accuracy, effectivity, and applicability. For the accuracy of solutions, analytical results should be validated on both synthetic and realistic data. For the effectivity of solutions, information that inputs to the analytical model should be easy to evaluate, and solutions should be obtained in finite time. For the applicability of solutions, we should be able to calculate the solution numerically, and estimate the parameters that input to the analytical model from practical situations.

To achieve these goals, we proceed with the following steps:

1. Develop an analytical solution for the expected message delay and the number of copies of the message while considering the interest/acceptance of mobile nodes for carrying a given piece of content at each contact (Chapter 2).

2. Develop an analytical solution for the number of data kiosks needed to satisfy a requirement of the expected spreading time (Chapter 3).

3. Develop an analytical solution for the optimal locations of data kiosks when extending a homogeneous environment to a heterogeneous environment where mobile nodes move from a region to another region by using a subway line (Chapter 4).

4. Evaluate the analytical solutions under both a number of mobility models and real datasets (Chapter 2-4).

1.5 Overview of the Solutions

In this dissertation we develop analytical solutions that aims at reaching the accuracy, efficiency, and applicability for understanding and improving the performance of content distribution in DTNs with the existence and nonexistence of data kiosks. This section outlines our solutions to the above problem in three main steps.

1.5.1 Evaluation and Improvement of Performance in DTNs

We begin by considering the problem of evaluating the performance of content distribution on DTNs in a situation where no data kiosk exists since a service provider wishes to know whether or not adding some data kiosks is necessary. In a DTN context that we consider, mobile nodes are wireless devices carried by humans and they use a store-carry-and-forward paradigm to deliver content towards the destination. Due to the fact that mobile users are not always willing to carry news in which they are not interested, it is important to consider the interest/acceptance of mobile nodes for carrying a given piece of content in the performance evaluation of content distribution in DTNs.

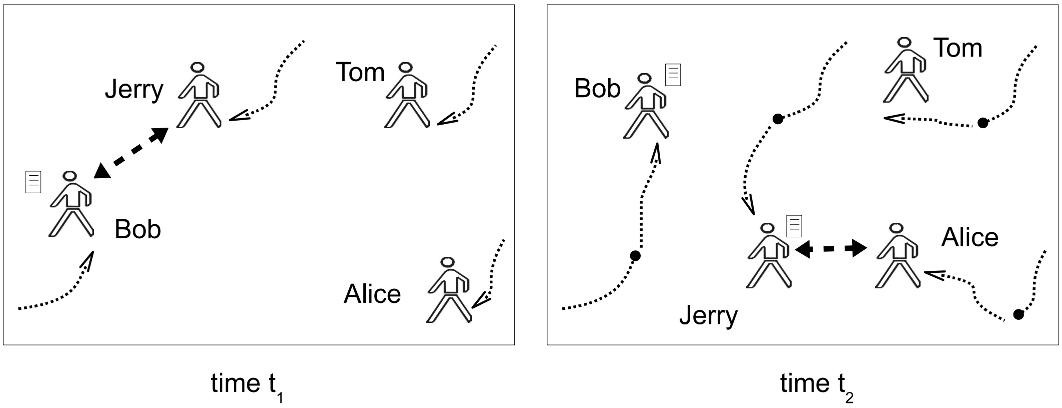


Figure 1.1: Content distribution in infrastructure-independent DTNs.

Fig. 1.1 describes a content distribution scenario that we consider in infrastructure-independent DTNs. Suppose that Bob want to send a single message to Alice at time t_0. Alice, Bob, Tom and Jerry are moving in the area. At time $t_1 > t_0$, Bob meets Jerry. Jerry accepts the message with a probability. Assume that Jerry accepts the message and stores it in his mobile. At time $t_2 > t_1$, Alice runs across Jerry. As a result, Alice receives the message from Jerry. In this content distribution scenario, the time required to deliver the message from Bob to Alice, which we call the message delay that is a common performance metric to measure the networks, is $t_2 - t_0$. Another important performance metric related to resource consumption is the number of copies of the message. The number of copies of a message is the total number of the message that is replicated to mobile nodes including the source and excluding the destination, at the time at which the message is delivered to its destination. In this scenario, when Alice gets the message, the number of copies of the message is 2, which embraces one copy in Bob's mobile and one copy in Jerry's mobile. In this example, Bob is the source, Alice is the destination, and Jerry acts as a relay that stores, carries and forwards the message towards Alice.

We employ a Markov chain to model the dissemination of a given piece of content from a source to a destination in a homogeneous system which is referred to as a network where the contact rates are identical. Heterogeneous contact rates can be approximated by slowly varying contact rates or homogeneous contact rates when we consider an average value of a performance metric [56]. This assumption will be relaxed in some aspects when we consider the placement problem of infrastructure nodes. A major advance of our evaluation is that with our model we take into account the probability of the event that a mobile node wants to receive a given piece of content at each contact. We investigate both closed-form and asymptotic expressions of the expected message delay and the number of copies of the message at the time the message is delivered to its destination. Using the expressions obtained, we analyze the role of interest and contact rates, and determine which one is significant for the improvement of the expected message delay in such a network for the cases of low and high density of mobile nodes. Chapter 2 describes an analytical model for analyzing the message delay and the number of copies of content, and suggestions for improving the performance of content distribution in detail. It also describes simulation results that we conducted on a number of mobility models and real datasets.

1.5.2 Number of Data Kiosks Required to Satisfy a Performance Objective

The solution of the first step provides a quantitative analysis to investigate the message delay of DTNs. It answered the question of how effective a DTN without infrastructure can support the distribution of content. If the message delay is found to be excessive, we extend the problem of content distribution to a situation where some data kiosks exist. Data kiosks are simple devices that receive the content directly from the source, generally using wired or cellular networks. In this situation, the distribution of news uses the opportunistic contacts both between a mobile node and a data kiosk, and between mobile nodes. We are mostly interested in the distribution of content from data kiosks to mobile nodes. Hence, the performance metric that we consider is the spreading time that is the delay required to deliver the content to the last node in a group of nodes, or the time needed for the content to spread over a part of the network. The spreading time would be examined to fulfil the interest of all subscribers. Due to the fact that a spreading time objective for the news dissemination is usually rather short, we suppose that mobile nodes will not be influenced by an incentive scheme, or mobile nodes do not change their interest in the time that we take into consideration. The important question is that given an area and a mobility pattern how many data kiosks one has to invest in order to satisfy a constraint of the expected spreading time.

Fig. 1.2 describes a content distribution scenario that we consider in structure-independent DTNs. Suppose that content is stored to data kiosks at initial time t_0. Alice, Bob, Tom and Jerry are moving in the area. At time $t_1 > t_0$, Bob passes data kiosk 1 and Tom passes data kiosk 2. Hence, Bob and Tom receive content from data kiosks. At time $t_2 > t_1$, Alice encounters Bob; hence, Alice gets content from Bob. In the scenario, the data kiosks that receive content directly from content providers disseminate content to mobile nodes by using opportunistic contacts. The spreading time required to spread content to three nodes is $t_2 - t_0$. Suppose that we want to spread content to three nodes in a time period d. If $d = t_2 - t_0$, a deployment of one data kiosk in the area is not enough to meet a time requirement; however a deployment of three data kiosks is not an optimal investment. Hence, an important question is to analyze the number of data kiosks needed to satisfy a

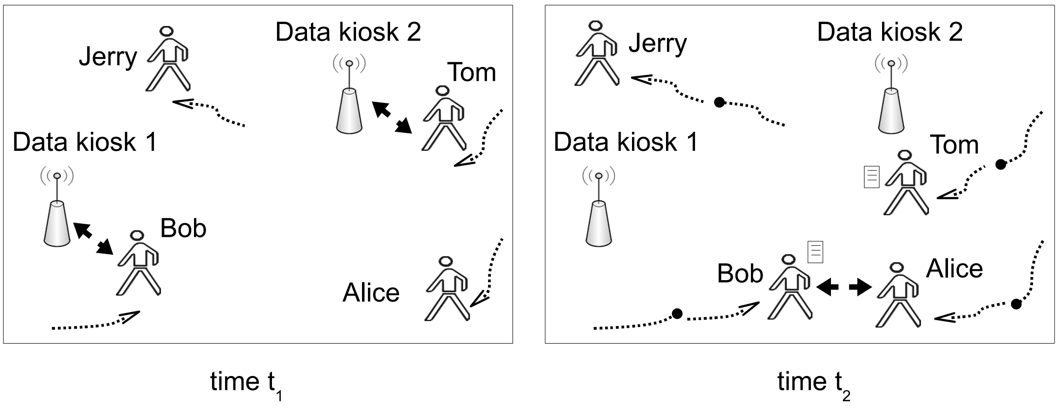


Figure 1.2: Content distribution in infrastructure-dependent DTNs.

constraint of the expected spreading time.

In order to model the content distribution in DTNs with the presence of data kiosks, we represent the number of copies of a given piece of content in the network as an absorbing finite state Markov chain. We obtained the closed-form expression of the expected spreading time represented under the digamma function and show a relationship between the expected spreading time and the expected message delay in a scenario of a homogeneous system. The main advance is that we developed an analysis of the number of data kiosks needed to lower the expected spreading time to a requirement. The results show how the contact rate between a mobile node and a data kiosk influences the quality of service parameters.

In Chapter 3 we provide a detailed description of the analytical model for analyzing the number of data kiosks given an expected spreading time objective and a mobility pattern. We also describes simulation results obtained in the case of the presence of data kiosks on a number of mobility models and real datasets.

1.5.3 Locations of Data Kiosks Optimizing the Performance

From the last two steps, we have developed a solution for performance evaluation of content distribution in DTNs and an analysis of a number of data kiosks needed to satisfy an objective of the expected spreading time in a homogeneous environment. However, we have considered the distribution of content only in a homogeneous environment. We should extend the analysis to a heterogeneous environment, namely, where different values are used to describe the contact rates between mobile nodes in distinct sub-areas, as well as the contact rates between a mobile node and a data kiosk. In a heterogeneous environment, where we position a data kiosk influence the performance of content distribution in the system. Hence, two important questions in this situation are: How effective is the distribution of

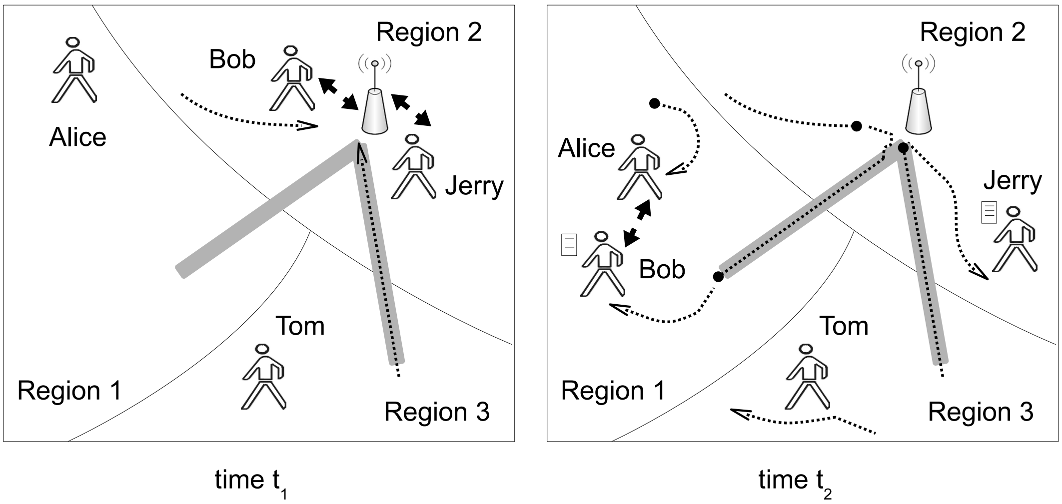


Figure 1.3: Content distribution in infrastructure-dependent DTNs with a subway map.

content in a heterogeneous environment? Where should we place a data kiosk in order to minimize the expected spreading time?

We consider the distribution of content from data kiosks to mobile nodes in a heterogeneous system that is an urban space composed of several regions and covered by a transportation system. In the system, a mobile user moves freely in a region and can move from a region to another region by using a transportation line. There is one subway stop in each region. A data kiosk will be located near a subway stop, which is inspired by the fact that the distribution channel of free daily newspapers is often installed at the entrance to the main transportation systems. In a large area, with the sparseness of data kiosks, selecting the locations of data kiosks will have a strong influence on the service offered.

Fig. 1.3 describes a content distribution scenario that we consider in infrastructure-dependent DTNs where taking into account the movement of the people between regions by using subway lines. In the figure, people carrying WiFi-capable devices are moving in an area composed of three regions. Subway lines connect region 1 with region 2, and region 2 with region 3. A data kiosk is located near the subway stop in region 2. Suppose that content is stored to data kiosks at initial time t_0. At time t_0, Alice and Bob are moving in region 1 and 2 respectively; and Tom and Jerry are moving in region 3. Bob wants to move from region 2 to region 1; hence he goes to the subway stop in region 2. Jerry moves from region 3 to region 2. When Bob and Jerry pass the data kiosk near the subway stop in region 2, Bob and Jerry receive content from data kiosk at t_1. After going out of the subway station in region 1, Bob meets Alice at time $t_2 > t_1$. As a result, Alice gets content from Bob. In this scenario, at time t_2, three people will receive the content if we place the data kiosk near the subway stop in region 2. However, only Jerry will receive the content if we place the data kiosk near the subway stop in region 3. Hence, given a mobility pattern, a number of data kiosks and a subway map, a right decision on the location of data kiosk is very important to optimize the expected spreading time.

We use a system of original differential equations (ODE) to model the distribution of content in such an environment [57]. Using our analytical model, we developed steps to obtain a solution for the locations of data kiosks that optimizes the expected spreading time. In addition, we revealed two general placement strategies of the optimal locations of data kiosks that could make solutions tractable in large systems. These findings support the view that the optimal locations of data kiosks is influenced not only by the conditions of a region but also by the number of mobile users that will receive the content.

In Chapter 4 we elaborate on the analytical model for analyzing the distribution of content in a heterogenous environment and a solution for the locations of data kiosks. We also provide simulation results that we carried out on both a number of mobility models and real datasets to validate the analytical results.

1.6 Contributions of the Dissertation

Most studies in content distribution have been carried out in the case where no infrastructure node exists. Recent works have considered the content distribution in the combination between infrastructure nodes and mobile nodes. However, there is still insufficient assessment of the performance of content distribution in DTNs when mobile nodes are not always willing to carry content. In the combination between infrastructure nodes and mobile nodes, the important questions of the number and the location of infrastructure nodes still remain: How many infrastructure nodes should we provide to meet a performance objective? Where should they be located to optimize the performance? In addition, the spreading time should be considered to fulfil the interest of all mobile nodes. The dissertation focuses on an optimal solution for news distribution using opportunistic contacts between people on the move in an urban area. The most important contribution of the dissertation is a solution for analyzing the performance of content distribution in DTNs while considering the probability of acceptance at each contact, and a solution for the number and the locations of data kiosks for news distribution with a requirement of the expected spreading time. The analytical results that were validated under both a number of mobility models and real datasets provide an accurate and effective evaluation for the number and the locations of data kiosks. The solutions are applicable as the inputs to the analytical model can be evaluated effectively from both synthetic and realistic movements of mobile users. The main contributions of this dissertation are as follows:

- An important contribution of the dissertation is the closed-form and asymptotic expressions of the expected message delay, and the closed-form expression for the distribution of the number of copies. The expressions provide a useful and efficient tool for evaluating numerically the performance of content distribution in DTNs. We used the analytical results to explain the impacts of the transmission range and the probability of interest on the expected message delay and the distribution of the number of copies. The analysis provides a useful suggestion for the improvement in the performance of content distribution in DTNs. In addition, the solution provides an analytical model with a probability of acceptance for the content distribution by using only opportunistic contacts between mobile nodes.

- A major contribution of the dissertation is a solution for the problem of the number of data kiosks. It provides explicitly bounds of the number of data kiosks required to satisfy a performance objective. An important property that is derived from these expressions is that the bounds of the number of data kiosks scale linearly with the contact rates between a mobile node and a data kiosk. Furthermore, the solution provides a general model, where the spreading time from a dynamic number of sources to a dynamic number of destination can be calculated effectively, for news distribution with the presence of data kiosks. Besides, a derivation from the solution is an expression for a relationship between the expected message delay and the expected spreading time in the context of DTNs. The result confirms that the spreading time should be taken into account when one considers to fulfil the interest of all mobile users.

- Another important contribution of the dissertation is the solution for the optimal locations of data kiosks. It provides explicit steps to gain a solution for the locations of data kiosks that optimizes the expected spreading time. The analytical model for the news distribution with the presence of data kiosks is extended to a more realistic scenario where mobile nodes are allowed to move between regions by using means of transport such as a subway line that connects regions. Furthermore, the solution analyzes general trends for placing data kiosks near subway stops in an area composed of several regions, in order to optimize the expected spreading time.

1.7 Structure of the Document

The remainder of the dissertation is structured as follows. Chapter 2 describes analytical results for the expected message delay and the number of copies of the message, and analyzes the impacts of input parameters on the performance of content distribution in

DTNs. Chapter 3 proposes a solution for the number of data kiosks needed to satisfy a constraint of the expected spreading time. Chapter 4 introduces a numerical solution for the placement problem of data kiosks and uses the model to derive two trends of the optimal locations of data kiosks. The last chapter summarizes the dissertation and discusses directions for future research.

The dissertation are organized so that each chapter is relatively self-contained. Chapters 2, 3, and 4 elaborate on the ideas outlined in Section 1.5. The combination of solutions presented in the three chapters provide a comprehensive study that develops from the performance evaluation to a solution for content distribution with a performance objective. The reader can read the Introduction chapter first to quickly have a general view and understand main ideas. The remaining chapters can be read subsequently to complete the picture.

Chapter 2

Evaluation and Improvement of Performance in DTNs

In the previous chapter the benefit of news delivery using the movement of the people was introduced where the people hold handheld devices with Wi-Fi capacity as they can store, carry and forward the news towards the destination. The widespread deployment of handheld devices makes the possibility of using opportunistic contacts for content distribution instead of being charged for cellular access. Since people are not always willing to carry content, it raises important questions above the performance of content distribution in DTNs regarding the acceptance of people for carrying news. Hence, we begin our study by developing an analysis of the performance of content distribution in DTNs while considering a probability of acceptance for a given piece of news at each contact. Specifically, we analyze quantitatively the expected message delay and the number of copies of a given piece of news at the time at which the news is delivered to its destination in a context of infrastructure-independent DTNs where a mobile node accepts news for carrying with a probability at each contact. Chapters 3-4 then extend the analysis described in this chapter to other application contexts (e.g., news distribution with the presence of data kiosks) where it raises several important questions (e.g., the number and the locations of data kiosks).

The chapter is organized as follows. The next section defines the problem of the performance analysis for DTNs in a context of our work. Section 2.2 describes our analytical results for evaluating the performance of content distribution in DTNs and an analysis of the impact of input parameters on the performance metrics. Section 2.3 presents simulations under a number of mobility models and real datasets. Section 2.4 discusses the limitations and extensions of our results and Section 2.5 summarizes the chapter.

2.1 Introduction to the Problem

2.1.1 Problem Statement

In this chapter, we focus on answering the question of how effective a DTN can be without infrastructure. In DTN networks, the store-carry-and-forward paradigm is usually used to enable communication. It is obvious that mobile nodes do not like to store any content in which they are not interested. Base on this statement, we are interested in understanding the performance of a content distribution scheme in which mobile users have an acceptance probability for carrying news at each contact.

Definition 1 (Acceptance Probability). *The acceptance probability denoted by* δ $(0 < \delta \leqslant 1)$ *is the probability of the event that a mobile node accepts a message for carrying at each contact. The probability of interest and the probability of acceptance is referred to the same concept.*

We are interested in understanding two performance metrics that are important for the application of DTNs in the news distribution. The first metric is the message delay that is defined as the time required to transmit a message from a source to a destination. We shall refer to the message as the news or the contents interchangeably since the application target is the distribution of news. The message delay should be considered because it is one of the common performance metrics to measure the networks. For applications that run on top of DTNs, it is not required to deliver the message to its destination instantly, but the message should be delivered to its destination within a reasonable delay. The second metric that we consider is the number of copies of the message at the time at which the message arrives at its destination. The number of copies is related to resource consumption. In the context of our work, the means of carrying news is handheld devices with limited storage. The distribution of news will be deteriorated if mobile nodes do not want to carry news due to their storage full of content. As a result of its important influence on the dissemination of news, the number of copies of the message should be taken into account.

Definition 2 (Message Delay). *The message delay is the time required to transmit a message from a source to a destination.*

Definition 3 (Number of Copies). *The number of copies of a message is the total number of the message that is replicated to mobile nodes including the source and excluding the destination, at the time at which the message is delivered to its destination.*

In our work, we consider the delivery of a single message due to the fact that news content of an event or a story can be completely contained in a message. Indeed, in an online newspaper, the size of news content including both text and images is usually no more than several tens of kilobytes. In practice, a mobile user might be interested in several topics; so there can be multiple news in the network. However, we can consider different news individually. Hence, the problem of multiple news is reduced to the problem of single news.

Due to a number of factors, including the movement of mobile nodes, the unknown behavior of mobile nodes for carrying news, the limited storage and energy of handheld devices, the high heterogeneity of parameters, estimating performance metrics in DTNs is a challenging effort. Integrating a full view of all factors that control the performance of content distribution in DTNs is an impossible task. Hence, we state generally the problem in the following; then we make several proper assumptions in the context of our work, in addition to practical assumptions explained, to focus only on important factors related to our context.

Problem 1 (Expected Message Delay)**.** *Given an area $\mathcal{A}$, a mobility pattern $\mathcal{M}$, a number of mobile nodes n, a transmission range r, find the expected message delay and the number of copies of the message.*

We study a network with n mobile nodes sparsely populated, each with a limited transmission range, moving in a closed area. In the context of the distribution of news, the transmission time for an updated news is small enough to assume that transmission is always successful when two mobile nodes are within each other transmission range; for example, the transmission time for a news 10KB is less than 2ms in Wi-Fi 802.11g with an average throughput 54 Mbit/s [58]. In addition, we suppose that interference from other nodes is negligible in a sparse network.

Contact opportunity plays the decisive role in the performance of content distribution in DTNs. We define three parameters related to contact opportunities:

- The *contact time* is the time interval for which two mobile nodes can communicate when they are within each other's transmission range. This parameter determines the amount of data that the two nodes can exchange at each contact.
- The *inter-contact time* is the time interval between two successive contacts of a pair of mobile nodes. This parameter determines when a contact opportunity occurs. It is an important factor for evaluating the feasibility of opportunistic networking.
- The *contact rate* is the number of contacts between a pair of mobile nodes per a unit of time on average. This parameter relates to how often two nodes meet each other.

Contacts in DTN networks typically are intermittent and classified into the following major types: scheduled contacts, opportunistic contacts, and predicted contacts. In our environment where mobile nodes are wireless devices carried by humans, contacts are unexpected or opportunistic contacts. Inspired by [36], we model contact opportunities between mobile nodes by independent and homogeneous Poisson processes with a rate λ, $\lambda > 0$. When mobile nodes with a transmission range r move according to a mobility model in a square area $\mathcal{A}$, the contact rate can be calculated approximately from an formula. When we have contact traces that describe the movement of mobile nodes, the contact rate can be estimated by using a data fitting method such as the least squares method [59].

2.1.2 Related Work

Many work studied data dissemination schemas or routing protocols in DTNs to optimize performance in DTNs. Preliminary work on routing in DTNs was introduced by Jain et al. [60]. The authors proposed several routing protocols that use knowledge oracles of contacts, queuing, or traffic demand. They concluded that global knowledge may not be required for good performance in many cases. This approach refers as knowledge-based forwarding [9, 61]. Other approaches are based on replication of the message [18, 19, 62], the history information [20, 21], social behavior [22, 23], and probabilistic forwarding [63]. These proposals were validated through simulation on a number of mobility models or several datasets. While results that they provided are promising, results based on analysis is necessary in order to understand several important parameters due to the fact that simulation results are limited to specific settings.

In DTN networks, contact opportunities impact the performance of a variety of applications. For this reason, some existing studies [64–67] have focused on characterizing inter-contact time distributions. Karagiannis et al. observe exponential tail behavior of inter-contact times and find out exponential decay beyond a point [64]. Conan et al. find that the empirical distributions of pairwise inter-contact times tend to be well fitted by log-normal curves, with exponential curves also fitting a significant portion of the distributions [65].

From understanding contact opportunities, recent work makes an effort to analyze performance metrics with different assumptions about content distribution schemes, including the unrestricted multicopy protocols [36], spray-and-wait [68], probabilistic routing [69], k-hop relay [70, 71]. Starting from [36], Hanbali et al. [72] extend the work with lifetime constraints and Zang et al. [41] study variations of the epidemic protocol and recovery schemes. In the first step of our work, this chapter focuses on evaluating the performance of content distribution in DTNs. Our work is different as it considers the probable effect of interest for a given piece of content in estimating the expected message delay and the number of copies of the message in DTNs.

2.2 Analysis of Content Distribution in DTNs

In this section, we explain the analytical model, calculate its parameters, describe and analyze results for the expected message delay and the number of copies of news.

2.2.1 Analytical Model

We extend a stochastic model introduced by Groenevelt et al. [36] to compute the expected message delay and the distribution of the number of copies. In addition to the original model, we assume that nodes have a probability of interest for a given piece of content. Mobile nodes store, carry and forward a message only if they are interested in the message. The model is based on modeling the number of copies in the network as an absorbing finite state Markov chain. The model comprises n states labeled by $\{1, 2, \ldots, n\}$. The Markov chain is in state $i = 1, 2, \ldots, n-1$ when there are i copies of the message in the network including the original message, and it is in state n when the message has been delivered to the destination node.

To complete the model, our task is to compute transition probabilities. Suppose that the Markov chain is in state i, its next state is either state n or state $i+1$. Since we are concerned only about what next state is when the chain departs state i, we may define a random variable Z_i by

$$Z_i = \begin{cases} 1, & \text{if the chain goes to state } n \text{ from state } i, \\ 0, & \text{if the chain goes to state } i+1 \text{ from state } i. \end{cases}$$

Let $T_{i,n}$ be the time at which an infected node (i.e., nodes that have a copy of the message) meets the destination node. $T_{i,i+1}$ is the time at which one of the infected nodes meets one of the healthy nodes (i.e., nodes without a copy of the message except the destination node). Then the probability of the event that changes the chain's state from state i to state n is

$$\begin{aligned} P(Z_i = 1) &= P(T_{i,n} < T_{i,i+1}) \\ &= \int_0^\infty P(x < T_{i,i+1}) dP(T_{i,n} < x). \end{aligned} \tag{2.1}$$

Our task reduces to compute the probability of the event that an infected node meets the destination node and the probability of the event that an infected node meets a healthy node by time t.

The chain is in state i thus there are i infected nodes. Given that g_1, g_2, $\ldots$, and g_i are these nodes. Let Y_{g_1}, Y_{g_2}, $\ldots$, and Y_{g_i} be the time at which the nodes respectively meet the destination node. Contact opportunities between mobile nodes are represented by independent and homogeneous Poisson processes having a rate λ, or the time interval between two successive contacts for each pair of mobile nodes is exponentially distributed with a pairwise contact rate λ. Therefore

$$P(Y_{g_1} \leqslant t) = P(Y_{g_2} \leqslant t) = \ldots = P(Y_{g_i} \leqslant t) = 1 - e^{-\lambda t}.$$

If we are interested in the time until either Y_{g_1}, Y_{g_2}, $\ldots$, or Y_{g_i} occurs, then we are interested in $T_{i,n} = \min(Y_{g_1}, Y_{g_2}, \ldots, Y_{g_i})$. The probability of the event that an infected node meets the destination node by time t is

$$\begin{aligned} P(T_{i,n} \leqslant t) &= 1 - P(\min(Y_{g_1}, Y_{g_2}, \ldots, Y_{g_i}) > t) \\ &= 1 - e^{-\lambda i t}. \end{aligned} \tag{2.2}$$

Given that h_1, h_2, $\ldots$, and h_{n-1-i} are healthy nodes; and $H = \{h_1, h_2, \ldots, h_{n-1-i}\}$, the set of healthy nodes. Let T_{h_v} be the time at which a healthy node $h_v \in H,\ 1 \leqslant v \leqslant n-1-i$ meets an infected node, then the distribution of T_{h_v} is $P(T_{h_v} \leqslant t) = 1 - e^{-\lambda i t}$.

A healthy node h_v is infected with a copy of the message if and only if it meets either g_1, g_2, ..., or g_i; and it would like to receive the message. Let Y_{h_v} be the time at which a node h_v is infected with a copy of the message then the distribution of Y_{h_v} is

$$P\left(Y_{h_v} \leq t\right) = 1 - e^{-\lambda \delta i t}.$$

For explaining the distribution of Y_{h_v}, we image a new process as following. At each time when node h_v is infected with a copy of the message, node h_v deletes the copy immediately (i.e. node h_v changes into a healthy node). The experiment is repeated. Let X_t be the number of events of node h_v being infected by time t. The process $\{X_t\}$ is obtained by thinning the Poisson process with rate λi. Hence, $\{X_t\}$ is a Poisson process having rates $\lambda \delta i$. This Poisson process has interarrival times $X_1 = Y_{h_v}$, X_2, $\ldots$ The random variables of the sequence are independent, and all have the same distribution $P\left(X_1 \leq t\right) = P\left(X_2 \leq t\right) = \ldots = 1 - e^{-\lambda \delta i t}$. This explains the distribution of Y_{h_v}.

If we are interested in the time until either Y_{h_1}, Y_{h_2}, $\ldots$, or $Y_{h_{n-1-i}}$ occurs, then we are interested in $T_{i,i+1} = \min(Y_{h_1}, Y_{h_2}, \ldots, Y_{h_{n-1-i}})$. The probability of the event that one of the healthy nodes is infected with a copy of the message by time t is

$$\begin{aligned} P(T_{i,i+1} \leqslant t) &= 1 - P(\min(Y_{h_1}, Y_{h_2}, \ldots, Y_{h_{n-1-i}}) > t) \\ &= 1 - e^{-\lambda \delta i (n-1-i) t}. \end{aligned} \tag{2.3}$$

Substitute (2.2) and (2.3) into (2.1), and do a little manipulation, we obtain the expression of the probability of going from state i to state n:

$$P\left(Z_i = 1\right) = \frac{1}{1 + \delta\left(n - 1 - i\right)}. \tag{2.4}$$

It is straightforward to show that the probability of going from state i to state $i + 1$ is

$$\begin{aligned} P(Z_i = 0) &= 1 - P(Z_i = 1) \\ &= \frac{\delta\left(n - 1 - i\right)}{1 + \delta\left(n - 1 - i\right)}; \end{aligned} \tag{2.5}$$

and we complete parameters of the transition diagram of the Markov chain for the considered content distribution scheme.

2.2.2 Distribution of the Number of Copies

We investigate the closed-form expression for the distribution of the number of copies of the message at the time the message is delivered to its destination. Let N_2 denote the number of copies of the message at the time the message arrives at its destination. We state the results in the following proposition.

Proposition 1. *Under the content distribution scheme with a probability of interest δ, the distribution of the number of copies $P\{N_2 = i\}$ is given by*

$$P\left(N_2 = i\right) = \frac{1}{1 + \delta\left(n - 1 - i\right)} \prod_{j=1}^{i-1} \frac{\delta\left(n - 1 - j\right)}{1 + \delta\left(n - 1 - j\right)}. \tag{2.6}$$

Proof. Given that the message arrives its destination after $N_2 = i$ copies of the message. The state sequence of Markov chain is 1, 2, ..., $i - 1$, i and n with $i > 1$; 1 and n with $i = 1$. Thus,

$$\begin{aligned} P(N_2 = i) &= P(Z_i = 1) \prod_{j=1}^{i-1} P(Z_j = 0), \\ \textit{where} \quad & \prod_{j=1}^{i-1} P(Z_j = 0) = 1 \text{ with } i = 1. \end{aligned} \tag{2.7}$$

Substituting (2.4) and (2.5) into (2.7), we obtain the expression of the distribution of the number of copies which proves (2.6). □

2.2.3 Expected Message Delay

We investigate both closed-form expressions and asymptotic expressions for the expected message delay when the acceptance probability is small $\delta \to 0$. Let D define as the time needed to send the message from the source to the destination. Proposition 2 presents the analytical results of the expected message delay.

Proposition 2. *Under the content distribution scheme with a probability of interest δ, the expected message delay is given by*

$$E[D] = \frac{1}{\lambda}\sum_{i=1}^{n-1}\left(P\left\{N_2 = i\right\}\sum_{j=1}^{i}\frac{1}{\delta j\left(n-1-j\right)+j}\right) \tag{2.8}$$

$$= \frac{1}{\lambda}\left(1 - \frac{n-2}{2}\delta\right) + o(\delta), \qquad \textit{for } \delta \to 0 \tag{2.9}$$

where N_2 denotes the number of copies of the content at the time it reaches its destination.

Proof. For $\theta \geqslant 0$, let $T^*(\theta) := E[e^{-\theta D}]$ be the Laplace-Stieltjes transforms of the message delay. Then,

$$E[D] = -\left.\frac{dT^*(\theta)}{d\theta}\right|_{\theta=0}. \tag{2.10}$$

We shall compute the Laplace-Stieltjes transforms of the message delay by conditioning on N_2.

$$T^*(\theta) = \sum_{i=1}^{n-1} E[e^{-\theta D}|N_2 = i]P(N_2 = i) \tag{2.11}$$

$$E[e^{-\theta D}|N_2 = i] = E\left[e^{-\theta\sum_{j=1}^{i-1}T_{j,j+1}+T_{i,n}} \,\middle|\, \begin{array}{l} T_{j,j+1} < T_{j,n} \\ j = 1,\ldots,i-1 \\ T_{i,i+1} > T_{i,n} \end{array}\right] \tag{2.12}$$

When in state $i = 1, 2, \ldots, n-1$, the Markov chain can either enter state $i+1$ after a time $T_{i,i+1}$, or enter state n after a time $T_{i,n}$. Let T_i be the sojourn time in state $i = 1, 2, \ldots, n-1$. We have $T_i = \min(T_{i,i+1}, T_{i,n})$ thus,

$$P(T_{i,i+1} < t|T_{i,i+1} < T_{i,n}) = P(T_{i,n} < t|T_{i,i+1} > T_{i,n}) = P(T_i < t). \tag{2.13}$$

From (2.12), (2.13) and the fact that the random variables $\{T_{i,i+1}, T_{i,n}\}_{i=1,2,\ldots,n-1}$ are mutually independent, we obtain

$$E[e^{-\theta D}|N_2 = i] = \prod_{j=1}^{i} E[e^{-\theta T_j}]. \tag{2.14}$$

From (2.14), (2.10) and (2.11), we find

$$E[D] = -\frac{d}{d\theta} \sum_{i=1}^{n-1} \left(P(N_2 = i) \prod_{j=1}^{i} E[e^{-\theta T_j}] \right) \Bigg|_{\theta=0} . \tag{2.15}$$

Our task is now to compute $E[e^{-\theta T_j}]$. First, we use (2.2) and (2.3) to find the cumulative distribution of the sojourn time in state i.

$$\begin{aligned} F(t) &= P(T_i \leqslant t) \\ &= 1 - P(\min(T_{i,i+1}, T_{i,n}) > t) \\ &= 1 - P(T_{i,i+1} > t) P(T_{i,n} > t) \\ &= 1 - e^{-\lambda \delta i (n-1-i) t} e^{-\lambda i t} \\ &= 1 - e^{-\lambda \delta i (n-1-i) t - \lambda i t} \end{aligned}$$

We are ready to compute $E[e^{-\theta T_i}]$.

$$\begin{aligned} E[e^{-\theta T_i}] &= \int_0^{+\infty} e^{-\theta t} dF(t) \\ &= 1 - \theta \int_0^{+\infty} e^{-\theta t} e^{-\lambda \delta i (n-1-i) t - \lambda i t} dt \\ &= \frac{\lambda \delta i\,(n-1-i) + \lambda i}{\theta + \lambda \delta i\,(n-1-i) + \lambda i}. \end{aligned} \tag{2.16}$$

We are almost ready to invoke the closed form of the expected message delay. From (2.15), we find

$$E[D] = -\sum_{i=1}^{n-1} \frac{d}{d\theta} P(N_2 = i) \prod_{j=1}^{i} E[e^{-\theta T_j}] \Bigg|_{\theta=0} .$$

$P(N_2 = i)$ do not contain θ. Thus

$$E[D] = -\sum_{i=1}^{n-1} P(N_2 = i) \times \sum_{j=1}^{i} \left(\prod_{u=1, u \neq j}^{i} E[e^{-\theta T_u}] \Big|_{\theta=0} \right) \frac{d}{d\theta} E[e^{-\theta T_j}] \Bigg|_{\theta=0}$$

Substituting $\theta = 0$ into (2.16) to get $E[e^{-\theta T_i}] \big|_{\theta=0} = 1$. Thus

$$E[D] = -\sum_{i=1}^{n-1} P(N_2 = i) \sum_{j=1}^{i} \frac{d}{d\theta} E[e^{-\theta T_j}] \Bigg|_{\theta=0} \tag{2.17}$$

Our problem reduces to compute $\frac{d}{d\theta} E[e^{-\theta T_j}] \big|_{\theta=0}$.

$$\frac{d}{d\theta}E[e^{-\theta T_j}] = -\frac{\lambda\delta i\,(n-1-j)+\lambda j}{[\theta+\lambda\delta j\,(n-1-j)+\lambda j]^2}$$

Hence, we find

$$\frac{d}{d\theta}E[e^{-\theta T_j}]\bigg|_{\theta=0} = \frac{-1}{\lambda\delta j\,(n-1-j)+\lambda j}.$$

On substituting the above expression into (2.17), we obtain the closed-form expression of the expected message delay (2.8).

Proof of the asymptotic expression. By applying the second order Maclaurin series approximation to a function $f(x)=(1+ax)^{-1}$, we find

$$\frac{1}{1+\delta\,(n-1-j)} = 1-(n-1-j)\,\delta+o\,(\delta)\,, \qquad \text{for } \delta\to 0 \tag{2.18}$$

and so

$$\begin{aligned}\frac{\delta\,(n-1-j)}{1+\delta\,(n-1-j)} &= 1-\frac{1}{1+\delta\,(n-1-j)}\\ &= (n-1-j)\,\delta+o\,(\delta) &(2.19)\\ \frac{1}{\delta j\,(n-1-j)+j} &= \frac{1}{j}-\frac{n-1-j}{j}\delta+o\,(\delta)\,. &(2.20)\end{aligned}$$

From the closed-form of the expected message delay (2.8) and (2.6), we find

$$\lambda E[D] = \sum_{i=1}^{n-1}\left(\begin{array}{l}\dfrac{1}{1+\delta\,(n-1-i)}\\ \times\displaystyle\prod_{j=1}^{i-1}\frac{\delta\,(n-1-j)}{1+\delta\,(n-1-j)}\\ \times\displaystyle\sum_{j=1}^{i}\frac{1}{\delta j\,(n-1-j)+j}\end{array}\right). \tag{2.21}$$

On substituting (2.18), (2.19), and (2.20) into (2.21), we get

$$\lambda E[D] = \sum_{i=1}^{n-1}\left(\begin{array}{l}(1-(n-1-i)\,\delta+o\,(\delta))\\ \times\displaystyle\prod_{j=1}^{i-1}((n-1-j)\,\delta+o\,(\delta))\\ \times\left(\displaystyle\sum_{j=1}^{i}\frac{1}{j}-\sum_{j=1}^{i}\frac{n-1-j}{j}\delta+o\,(\delta)\right)\end{array}\right).$$

It is easy to show that the second factor of the product is asymptotically smaller than $o(\delta)$ for $i > 2$

$$\prod_{j=1}^{i-1} \left((n-1-j)\,\delta + o(\delta) \right) = o(\delta)\,, \qquad \text{for } i > 2$$

then the products of the three factor are asymptotically smaller than $o(\delta)$ for $i > 2$. Therefore, they are missed in the asymptotic representation of the equation to terms of order $o(\delta)$. We find

$$\begin{aligned}
\lambda E[D] &= (1-(n-3)\,\delta + o(\delta))\,((n-2)\,\delta + o(\delta)) \\
&\quad \times \left(\sum_{j=1}^{2} \frac{1}{j} - \sum_{j=1}^{2} \frac{n-1-j}{j}\delta + o(\delta) \right) \\
&\quad + (1-(n-2)\,\delta + o(\delta))\,(1-(n-2)\,\delta + o(\delta)) \\
&= \left(\frac{3\,(n-2)}{2}\delta + o(\delta) \right) + (1 - 2\,(n-2)\,\delta + o(\delta)) \\
&= 1 - \frac{n-2}{2}\delta + o(\delta)\,.
\end{aligned}$$

By changing λ of the equation, we get the asymptotic equation of the expected message delay (2.8). □

We consider a specific case when all n mobile nodes are always interested in the news, or $\delta = 1$. Substitute $\delta = 1$ into (2.6) to get $P\{N_2 = i\} = \frac{1}{n-i} \prod_{j=1}^{i-1} \frac{n-j-1}{n-j} = \frac{1}{n-1}$. We then substitute the expression of $P\{N_2 = i\}$ and $\delta = 1$ into (2.8) to obtain $E[D] = \frac{1}{\lambda} \sum_{i=1}^{n-1} \left(\frac{1}{n-1} \sum_{j=1}^{i} \frac{1}{j(n-j)} \right)$. Some tedious manipulations yield

$$E[D] = \frac{1}{\lambda\,(n-1)} \sum_{i=1}^{n-1} \frac{1}{i}. \tag{2.22}$$

This is the expected message delay when all n mobile nodes are always interested in the news. The expression that is a specific case of our results when δ is assigned to 1 was obtained by Groenevelt et al. [36].

We want to note that the expression of the expected message delay that we obtained is different from the expected time needed to transmit news from a mobile node to the first node who receives the news. The later is much easier to compute and does not have much meaning in a practical situation.

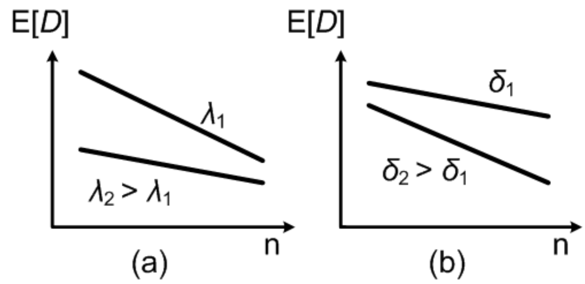


Figure 2.1: Asymptotic representation of the expected message delay.

2.2.4 Impact of Input Parameters on the Performance in DTNs

An important property acquired from the analytical results for the distribution of the number of copies of the message is that the distribution does not depend on the transmission range (Proposition 1). It comes from the fact that the transmission range affects how fast mobile nodes meet each other, or how fast the message arrives at the destination, but it does not affect how the message goes to the destination.

The asymptotic representation of the expected message delay suggests that the role of interest is significant in the case of high node density while the role of contact rates becomes more important in the case of low density of nodes (Proposition 2). Fig. 2.1 plots the asymptotic representation of the expected message delay for two environments with the same area. In Fig. 2.1(a), the probabilities of interests are similar and the contact rates between mobile nodes are λ_1 and λ_2 in the two environments. For a given probability of interest, when the number of mobile nodes or the node density increases, the expected message delay in the case of a large value of the contact rate decreases more slowly than the one in the case of a small value of the contact rate. In Fig. 2.1(b), the contact rates between mobile nodes are similar and the probabilities of interests are δ_1 and δ_2 in the two environments. For a given contact rate, when the number of mobile nodes or the node density increases, the expected delay in the case of a large value of the probability of interest decreases more quickly than the one in the case of a small value of the probability of interest. Consequently, improving the interest is rewarding if the number of mobile nodes is large, while improving contact rates provides better performance if the number of nodes is small. Using the analytical expression of the expected message delay obtained in Proposition 2, we can decide if the performance is acceptable using only contact opportunities between mobile users or if the addition of data kiosks becomes mandatory.

2.3 Evaluation

2.3.1 Simulation under Mobility Models

2.3.1.1 Simulation Setting

We implemented our content distribution scheme using the One simulator [73] with the random waypoint mobility model without pausing [74], the random direction [37] and the random walker mobility model [64] to validate our analytical results. The simulation setting considered in [36] includes n nodes moving in a square of size 4 km $\times$ 4 km, whose initial locations are sampled from the uniform distributions.

In the random waypoint mobility model, a node travels to a destination chosen uniformly in an area with a constant speed. The speed of a mobile node (in km/h) is chosen uniformly in $[v_{\min}, v_{\max}] = [4, 10]$. Upon arrival, the node continues its trip by choosing a new destination and a new speed, independently of all previous destinations and speeds.

In the random direction mobility model, a node selects an initial direction, speed and a finite travel time. The node then travels in the direction at the speed for the duration. Once this time expires, it chooses a new direction, speed and travel time which are independent of all its past directions, speeds and travel times. When a node reaches a boundary it is reflected. In our setting, the speed of a mobile node (in km/h) is chosen uniformly in $[v_{\min}, v_{\max}] = [4, 10]$, a direction is uniformly distributed in $[0, 2\pi)$, and travel time is exponentially distributed with mean 1/4 hour.

In the two-dimensional random walker mobility model, a node moves on a square grid at a constant speed. At crossroads, a node uniformly chooses to go to a next crossroad in the front, in the back, on the left, or on the right. When a node has reached a crossroad at the borders, it will go to the opposite crossroad if the newly chosen crossroad is out of the borders. We set the distance among adjacent crossroads to 80m and a constant speed of 4.8 km/h.

In our simulation, we apply instantaneous transmissions and infinite TTL in order to match the assumptions in our analysis. As mentioned in [75], we discard messages in the initial $1,000$ seconds of simulation time to avoid the high variability in performance results.

2.3.1.2 The Pairwise Contact Rate

The inter-contact time between a pair of mobile nodes is nearly exponentially distributed in common mobility models (such as the random waypoint or random direction). An expression for the pairwise contact rate λ is introduced in [36]:

$$\lambda \approx \frac{2\omega r E[V^*]}{L^2},$$

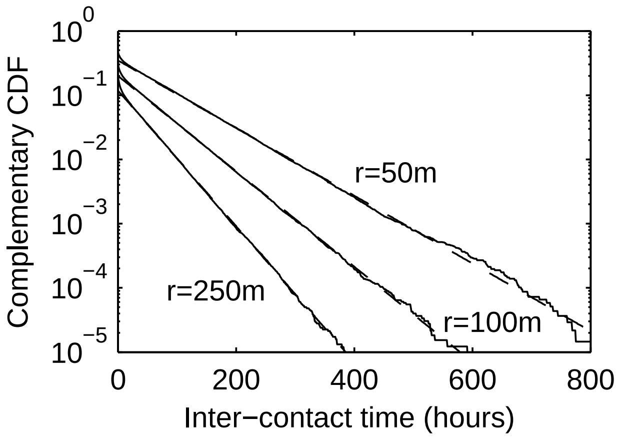


Figure 2.2: CCDF of inter-contact time for random walker.

where r is transmission range of nodes moving on an area $L \times L$ ($r \ll L$), ω is a constant related to mobility models, and $E[V^*]$ is the average relative speed between two nodes.

We find that the contact rates are $\lambda \approx 0.074086$ ($r = 50$ m), $\lambda \approx 0.14817$ ($r = 100$ m), and $\lambda \approx 0.37043$ ($r = 250$ m) for the random waypoint mobility model, $\lambda \approx 0.057731$ ($r = 50$ m), $\lambda \approx 0.11546$ ($r = 100$ m), and $\lambda \approx 0.28866$ ($r = 250$ m) for the random direction mobility model.

Karagiannis et al. proved that the inter-contact time distribution of the random walker mobility model is not exponentially distributed but there exists an exponential tail [64]. A mathematical formula for the exponential tail under the random walker mobility model is, to the best of our knowledge, not known. Therefore we discover its value through simulation. Fig. 2.2 plots the complementary cumulative distribution function (CCDF) of the inter-contact time on log-lin scale. We can see CCDF of an exponential tail (i.e. a straight line on a logarithmic scale on the y-axis) with intensity (i.e. slope) λ. The estimated values of λ are $\lambda \approx 0.0127$ ($r = 50$ m), $\lambda \approx 0.0171$ ($r = 100$ m), and $\lambda \approx 0.0243$ ($r = 250$ m).

2.3.1.3 Simulation Results

We first consider the expected message delay. In a square of size 4 km $\times$ 4 km, for the three probability of interest ($\delta = 1$, 0.4, 0.2), the three transmission ranges ($r = 50$, 100, 250 m), and under the three mobility models, we vary the number of nodes between 5 and 80, and compute the average of the message delay for each n based on $1,000$ observations obtained from simulation. Fig. 2.3 and 2.4 plot the expected message delays as a function of the number of nodes, comparing the theoretical analyses with the simulation results. They confirm the accuracy of the theoretical analysis for these mobility models. Although

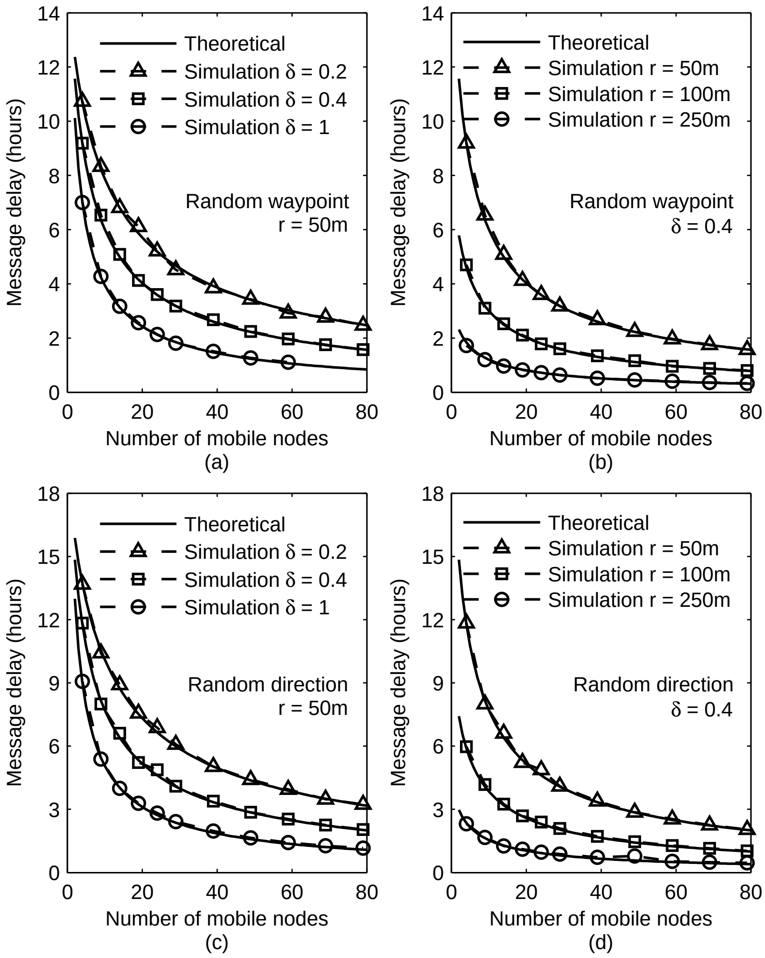


Figure 2.3: Message delay vs. number of nodes under the random waypoint and random direction mobility model.

the inter-contact time is only exponential tail under the random walker mobility model, the theoretical analysis still can predict the delivery delay for the different probability of interest and the different contact rates. The error between simulation and theoretical results increases when the acceptance probability δ is small.

As one can see from Fig. 2.3(a,c) and Fig. 2.4(a,b), the curves of the expected message delay of the different probability of interest converge when the node density is small. They exactly meet at a point where the number of nodes is two. In Fig. 2.3(b,d) and Fig. 2.4(c,d), the curves of the expected message delay of the different contact rates go far away from each others when the number of nodes is small. It agrees with the approximation analysis in which the slope of the curve of the expected message delay decreases when the contact rate increases. From these observations, we see that the improvement of the expected message

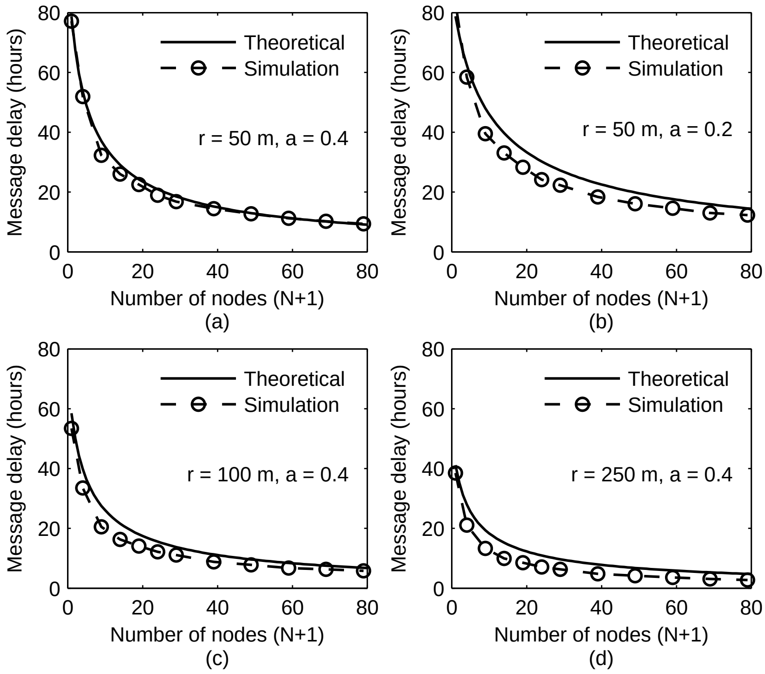


Figure 2.4: Message delay vs. number of nodes under the random walker mobility model.

delay is significant when enhancing the contact rate rather than increasing the acceptance probability, for low-density networks.

For the distribution of the number of duplicate copies of the message in the network at the time at which the message is delivered to the destination node, we have done simulations with $8,000$ runs for the three mobility models with different transmission ranges ($r = 50$, 100, 250 m), different probability of interest ($\delta = 1$, 0.4, 0.2) and 40 nodes. Results depicted in Figs. 2.5-2.6 are compared with our analytical results (solid lines), regarding the distribution of the number of copies, to simulation results (represented by bars). Since the results for the random waypoint mobility model are identical to those of the random direction mobility model, they are not plotted here. The results in the case $\delta = 1$ are not displayed because the content distribution scheme becomes the same as the unrestricted multicopy protocol in this case and we got the same results as obtained by Groenevelt et al. [36]. For the same mobility models, the same probability of interest, and different transmission ranges, we obtained the same distribution of the number of copies. It confirms our analytical results that the distribution of the number of copies does not depend on the transmission range. We observe that the simulation results are in accordance with the analytical results, but the model prediction weakens when transmission ranges increase and when the inter-contact time is not exponential for the case of the random walker

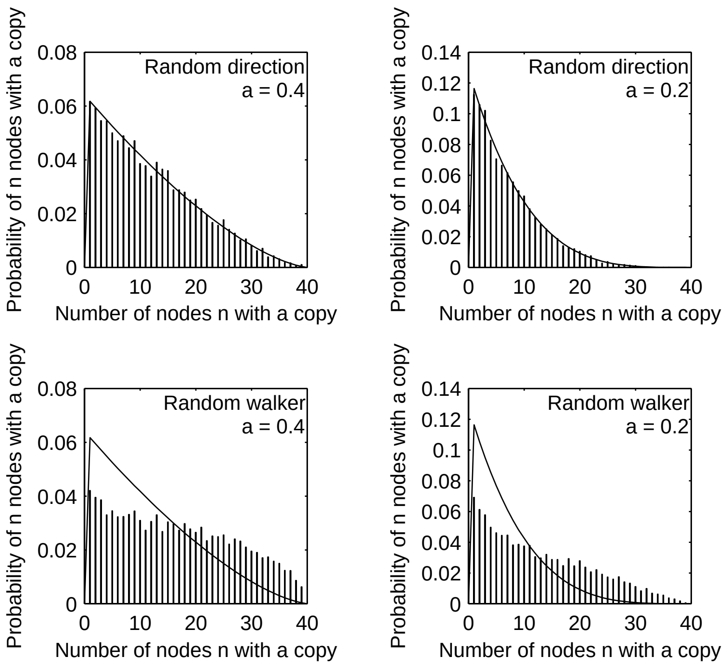


Figure 2.5: Distribution of the number of copies: r=50m.

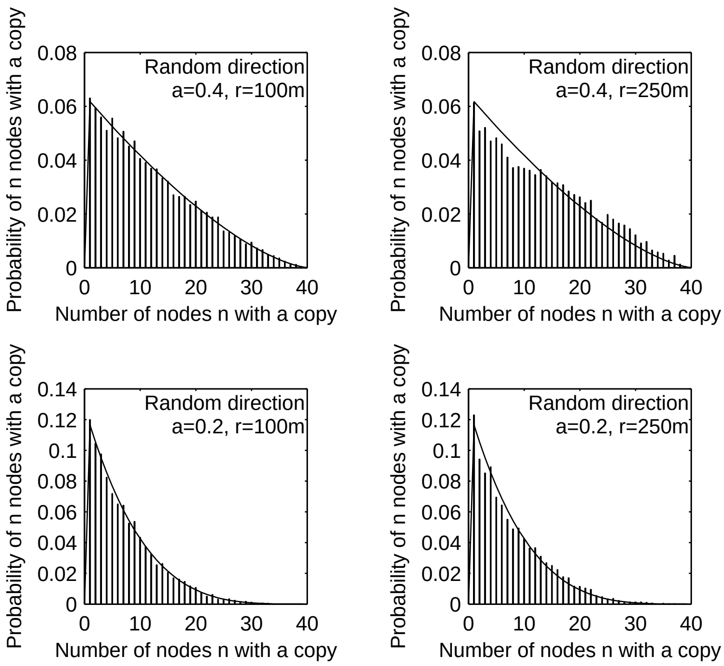
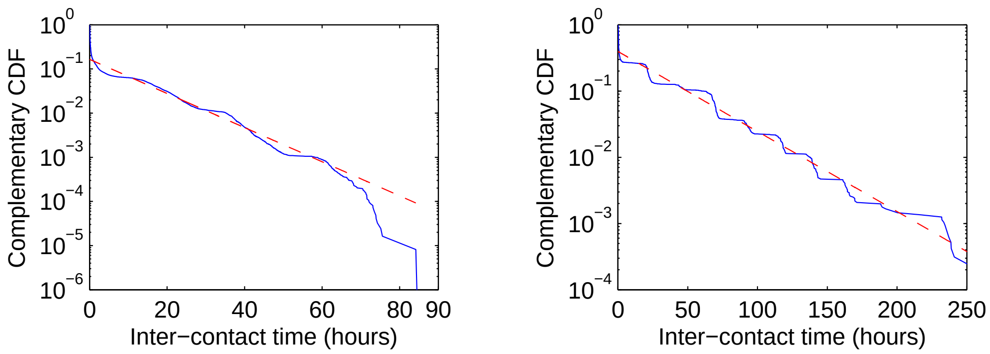


Figure 2.6: Distribution of the number of copies: r=100m, 250m.



(a) Dataset Infocom2006 (b) Dataset Content

Figure 2.7: CCDF of the inter-contact time in dataset Infocom2006 and Content.

mobility model.

2.3.2 Simulation on Datasets

2.3.2.1 Datasets

We use the dataset Infocom2006 and Content in the data repository CRAWDAD for our simulation [76]. While a number of datasets have the advantage of a long duration of experiment, they do not record direct contacts between mobile nodes (i.e. MIT [77], UCSD [78], Dartmouth [79]). In the dataset Infocom2006, the contact trace was collected from 78 mobile devices with ID#21 through #98 and 20 static nodes with ID#1 through #20. In the dataset Content, there are 36 mobile nodes with ID#1 through #36 and 18 static nodes with ID#37 through #54. In our simulation, we only use the contacts between mobile nodes in these datasets.

We analyzed the contact traces to compute the input parameters for our analytical model. Fig. 2.7 plots the CCDF of the inter-contact time between mobile nodes on log-lin scale for both contact traces. We can see CCDF of an exponential tail (i.e. a straight line on a logarithmic scale on the y-axis) with intensity (i.e. slope) λ. The estimated values of λ that we use the method of least squares to compute are $\lambda \approx 0.0890$ for the dataset Infocom2006 and $\lambda \approx 0.0278$ for the dataset Content.

2.3.2.2 Simulation Results

We simulated the distribution of contents when the probability of interest is 0.4. We varied the number of mobile nodes between 3 and 78 nodes in the dataset Infocom2006 and between 3 and 36 nodes in the dataset Content. We repeatedly run simulation and calculated the average value of the message delay from 1,000 runs. For each number

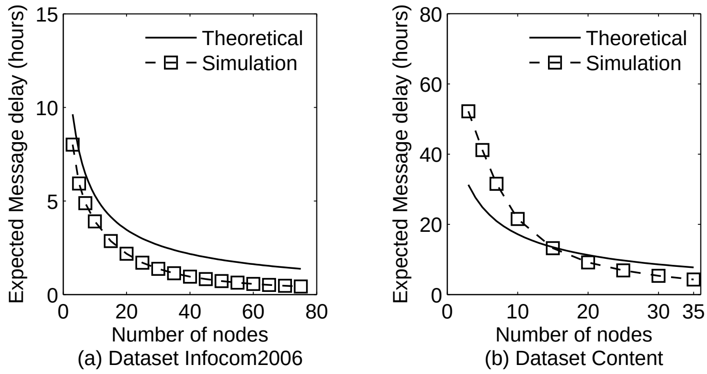


Figure 2.8: Expected message delay on the experiment data.

of mobile nodes, in each run, we chose randomly mobile nodes in the datasets and only considered contacts between these chosen nodes. Fig. 2.8 represents the simulation results compared with the analytical results for the two datasets. With 15 mobile nodes, the expected message delay is around 3 hours in the context of Infocom2006 while it is close to 15 hours in the context of Content. Two contact traces were collected from highly dissimilar environments, as we see that the expected message delay are vastly different in the two environments. However, the analytical model is well adapted to provide approximate predictions for the expected message delay.

2.4 Discussion

We now discuss limitation and extension of our results for evaluating the performance of content distribution in DTNs where considering the probability of interest for a given piece of news at each contact.

First of all, an issue to address is whether we can increase the accuracy of our solutions for estimating the expected message delay and the number of copies of the news. An important parameter in the analytical model is contact rates. Increasing the accuracy of the estimation for contact rates helps to achieve better estimation for the performance metrics. Hence, when mobile nodes move according to a mobility model in which a formula for contact rates is not found, the number of contacts collected in simulations needs to be large enough to ensure the accuracy of the estimation for contact rates. When we extract the movement of mobile nodes from real datasets, the number of contacts in the dataset is required to be large enough as well.

Second, the analytical model is applicable since all information that is input to the model can be determined completely in both synthetic and realistic movements of people.

For a mobility model whose contact rates were characterized, contact rates can be calculated directly to input to the model. For other mobility models and realistic movements, we could estimate approximately contact rates from the contact traces. Nevertheless, the exponential distribution of inter-contact times is only approximately. Hence, examining whether the distribution of inter-contact time can be characterized more precisely and finding a solution for the problem with this distribution are an important area for future research to obtain better analysis of news distribution in DTNs.

Third, our solutions is flexible since they are easy to apply to other opportunistic networks such as vehicular networks where contacts between a car and a infrastructure node and between cars are opportunistic. In spite of that, when we apply the results to other contexts, the assumptions of input information should be validated carefully.

Finally, in an area, the contact rates of different pairs of the people might not be the same. It can be an extension to gain better analysis for the expected message delay and the number of copies of the news. How it can be modeled is a challenge and it deserves for future work.

2.5 Summary

A performance analysis of news distribution in DTNs with an acceptance probability is essentially important for evaluating the delay and resource consumption, and understanding the impacts of input parameters on their performance metrics since a service provider wishes to know whether or not adding some data kiosks is necessary to obtain a proper performance objective. One finds it difficult to understand the performance of content distribution in DTNs while considering the interest/acceptance for a given piece of content at each contact due to the high dynamic of the content distribution scenario. We investigated both the closed-form expression and the asymptotic expression of the expected message delay, and the distribution of the number of copies at the time at which the news is delivered to its destination. Furthermore, using these expressions, we analyzed the role of interests and contact rates and determined which one is significant for the improvement of the expected message delay in such a network for the cases of low and high-density networks. Simulation results on a number of mobility models and real datasets validated the accuracy of analytical results.

In infrastructure-independent DTNs, we find it very difficult to satisfy an performance objective for news distribution in many situations. Specifically, the expected message delay is several hours in many cases. For example, the expected message delay is 2 hours in an ideal case where mobile nodes move according to the random waypoint mobility model, they always accept the message and the node density is 40 mobile nodes in an area $4km^2$

(Fig. 2.3(c)). The expected message delay becomes much worse in a non-ideal case where the node density, the acceptance probability or the mobility of nodes is low. For example, the expected message delay is 20 hours in a case where mobile nodes move according to the random walker mobility model, the acceptance probability is 0.2 and the node density is 40 mobile nodes in an area $4km^2$ (Fig. 2.4(b)). From these observations, it is necessary to extend news distribution in DTNs to a context where some data kiosks are present for obtaining a proper performance. In this context, it raises important questions of the number and the locations of the data kiosks with a performance constraint. We consider the challenging questions in subsequent chapters.

Chapter 3

Number of Data Kiosks

The previous chapter studies performance evaluation of the news distribution in DTNs while considering an acceptance probability for carrying news. The analysis shows the usefulness of opportunistic contacts between mobile nodes as means of disseminating news. However, the performance of news distribution in DTNs is several hours in many cases even if all mobile nodes carry and forward news. Hence, it raises a question of how to strengthen the performance of content distribution in DTNs. Our approach is to deploy some data kiosks in a DTN environment. As discussed in Section 1.2, an important question of content distribution with the presence of data kiosks is to find the number of data kiosks required to satisfy a performance objective. In this chapter, we focus on addressing this challenging question.

The chapter is structured as follows. The next section defines the problem of the number of data kiosks. Section 3.2 describes the analysis of the number of data kiosks in detail. Section 3.3 presents simulations under a number of mobility models and real datasets. Section 3.4 discusses the limitations and extensions of our results and Section 3.5 summarizes the chapter.

3.1 Introduction to the Problem

3.1.1 Problem Statement

We propose the distribution of news in DTNs with the presence of data kiosks in two steps. Data kiosks are simple devices that receive content directly from the source, usually using wired or cellular networks. Such simple devices, which are different from throwboxes [52] that act as relay nodes for passing a message from a mobile node to others, are suitable for the application of news distribution where an infrastructure node is not required to

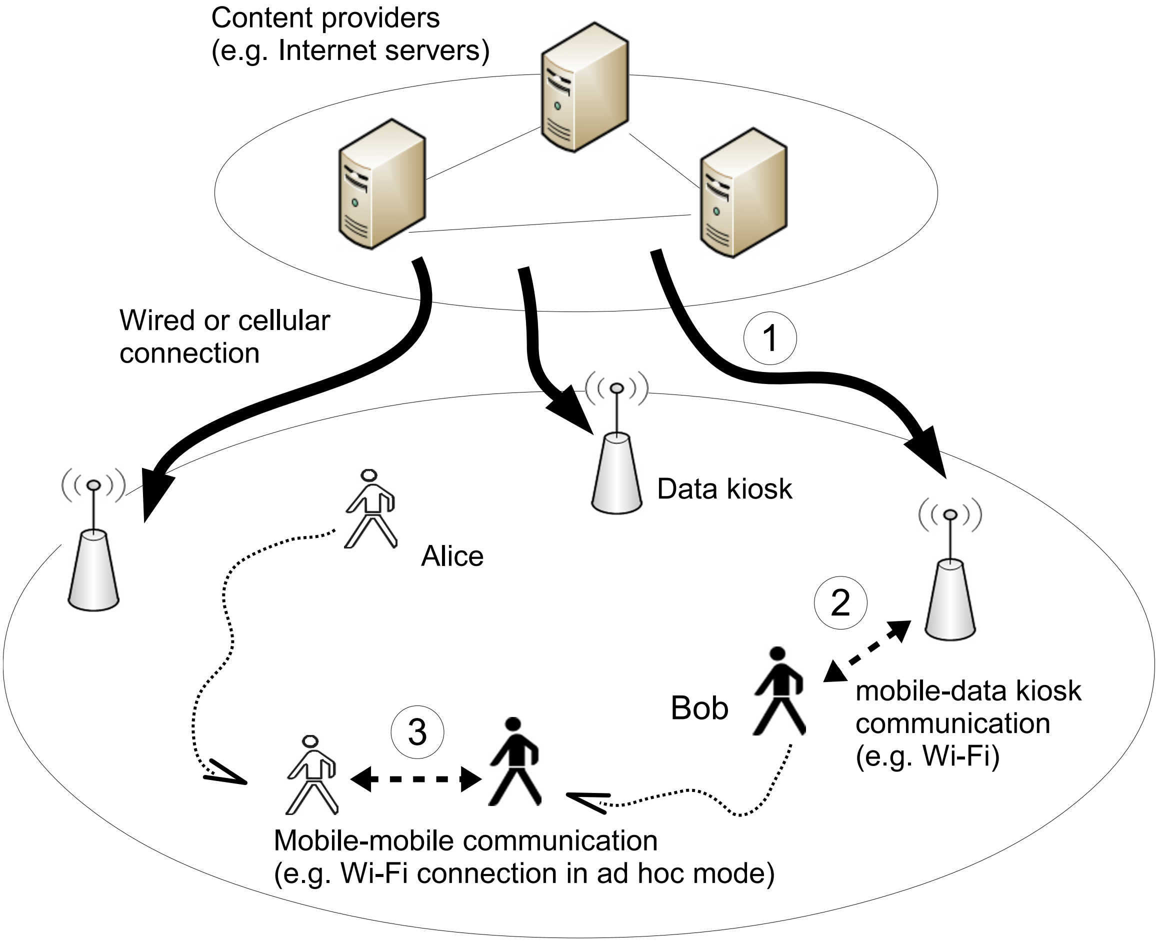


Figure 3.1: A news distribution system.

implement the function of a relay node. Is is due to the fact that news is spread from the data kiosks to mobile nodes, or the data kiosks act as the original sources. News is first updated periodically to data kiosks by using infrastructure based communication. Then, news is disseminated to mobile nodes by using DTN based communication (Fig. 3.1). Such type of system, which we call a hybrid DTN, exploits DTNs for disseminating news while retaining an acceptable performance of system by adding a proper number of data kiosks.

Since the update of news that is from content providers to data kiosks is not an issue of performance in our news distribution system, we focus on providing a guaranteed performance for news dissemination from data kiosks to mobile nodes. Specifically, we assume that news is put in the data kiosks; and we then consider news distribution to mobile nodes by using opportunistic contacts between a data kiosk and a mobile node, and between mobile nodes. A requirement of performance depends on different types of applications. For example, although strict time constraints do not apply for news applications, spreading the news should be achieved within a reasonable delay since people usually expect to receive it within their journey from their home to their office. In addition, the performance analysis in Chapter 2 provides an evidence that the performance of content distribution in DTNs suffers from an area with sparse populations. Therefore, a question is raised as to how many data kiosks does one have to invest in order to satisfy a given set of performance objectives.

In a news distribution system, an important performance metric is the spreading time that is the time required to spread news over a part of a network. Since the main purpose of a news distribution system is to deliver news to all subscribers as soon as possible, the spreading time would be examined to fulfil the interest of all subscribers. So far, however, no research has been found that surveyed the spreading time for such type of distribution systems. Therefore, in the news distribution system, the problem that we focus on is to estimate the number of data kiosks with a constraint of the expected spreading time.

Definition 4 (Spreading Time)**.** *The spreading time is the time required to spread a message over a part of a network.*

In this system, for a given piece of news, the performance of news distribution depends on several input parameters such as a coverage area, the movement of the people, an incentive scheme for carrying news, a number of subscribers, a target of the dissemination, a transmission range of the handheld devices, and the sizes of the news. The first step for estimating the number of data kiosks is to model the news distribution with the presence of data kiosks. The complexity of modeling the news distribution is measured as a function of multiple parameters of the input. Since it is not possible to find a universal model of the news distribution for all dependent variables, we consider the problem under reasonable assumptions in a proper context. As discussed in Section 2.1.1, in the context of the distribution of news, transmission is assumed to be always successful when two mobile nodes are within each other transmission range. *Due to the fact that a spreading time objective for the news distribution is usually not long, we suppose that mobile nodes receive little influence by an incentive scheme, or mobile nodes do not change their interest in the time that we take into consideration. Hence, we assume that the acceptance probability δ is not significant.* In the context of our problem, the performance of the news distribution are influenced by the following important parameters: a coverage area, the movement of the people, a number of subscribers, a target of the expected spreading time, a transmission range of the handheld devices.

Deployment of a news dissemination system require solving the following fundamental constrained problem:

Problem 2 (Number of Data Kiosks)**.** *Given an area $\mathcal{A}$, a mobility pattern $\mathcal{M}$, a number of subscribers m, a target number of infected nodes b, and a transmission range r, find a number of data kiosks k satisfying a constraint of expected spreading time d.*

3.1.2 Related Work

Recent work was motivated by improving the performance of DTNs. Polat et al. proposed an algorithm to find mobile nodes that can act as message ferries [31]. Another approach is to add some infrastructure nodes. This solution has been used in the context of ad hoc networks where nodes are fixed [80–82]. In the DTN's framework, there are few works that have considered the content dissemination based on the combination of both infrastructure based communication and DTN based communication. Important work in this domain was undertaken in [52,83,84]. In [52], the paper refers to infrastructure nodes as throwboxes that act as disconnected relays for passing a message from a mobile node to others, disseminating a message towards the destination. The paper proposed a greedy algorithm for different routing strategies in order to maximize the total throughput, and provided several intuitive suggestions for the design and deployment of throwboxes. In [83], the paper proposed an wireless ad hoc podcasting architecture, where the traditional podcasting concept is extended to ad hoc domains or a podcast can be downloaded from both infrastructure nodes and mobile nodes. The authors measured the performance of an ad hoc podcasting system, by implementing the system, to give insightful suggestions for designing the system. In [84], the paper analyzed the performance of the Multicopy Two-hop Routing protocol and the Epidemic Routing protocol in a network which includes both throwboxes and mobile nodes. The authors considered two cases of throwboxes: throwboxes compose a mesh network, and throwboxes are full disconnected. They obtained a closed-form expression for the delivery delay and the number of copies of a message when the message is delivered from a source to a destination. While the throwboxes have to implement the function of relay nodes which can store content items from a mobile node and forward them to other mobile nodes, the data kiosks are not required to carry out this function, or they act as the original sources. Such simple devices are suitable for the application of news dissemination where news is updated periodically to data kiosks, usually using wired or cellular networks. These studies make a new approach for news dissemination where the dissemination use both opportunistic contacts between mobile nodes, and between a mobile node and a infrastructure node. However, there have been no studies which provide the analysis of the number data kiosks needed to satisfy a performance objective. In addition, no research has been found that surveyed the spreading time in a news dissemination system while this performance metric would be examined to fulfil the interest of all subscribers. Our study was designed to fill the gap.

3.2 Analysis of Content Distribution with the Presence of Data Kiosks

3.2.1 Analytical Model

This section describes an analytical model for finding the number of data kiosks with a constraint of the expected spreading time d in the news dissemination system.

News dissemination relies on both opportunistic contacts between mobile nodes, and between a mobile node and a data kiosk. The size of the area, the movements of mobile nodes and a transmission range describe opportunistic contacts that are an important factor for spreading news over the network. For the random waypoint mobility model and the random direction mobility model, opportunistic contacts are parameterized as a function of an area size, parameters of a mobility model and a transmission range [36, 84]. For other mobility models such as the random walker mobility model whose formulas for opportunistic contacts are not known, we parameterize the opportunistic contacts by extracting the contact rates from contact data collected from simulation. For an area where the movement of the people can not be modeled on a mobility model, we apply the same method to parameterize the opportunistic contacts but contact data are gathered from a real experiment. Karagiannis et al. observe exponential tail behavior of inter-contact times and find out exponential decay beyond a point in real datasets [64]. Due to these observations of the inter-contact times distribution, we assume that the pairwise inter-contact times for mobile-mobile (resp. mobile-data kiosk) contacts are approximately represented as exponentially distributed random variables with mean $1/\lambda$, $\lambda > 0$ (resp. $1/\mu$, $\mu > 0$). All these random variables are assumed to be homogeneous and mutually independent.

In addition to the general assumptions in Section 3.1.1, we have two following specific assumptions. News dissemination from all data kiosks to a healthy node is usually faster than the news dissemination from only one infected node to a healthy node. Hence, we consider that the contact rate between a mobile node and all data kiosks is greater than the contact rate between a pair of mobile nodes, or the number of data kiosks is greater than λ/μ. In addition, one usually wants to have a lot of subscribers infected by a given spreading time. Hence, we suppose that the number of infected nodes is at least equal to two nodes by the spreading time.

We represent the number of copies of a given piece of content in the network as an absorbing finite state Markov chain comprising $b + 1$ states labeled by $\{0, 1, \ldots, b\}$. The chain is in state i when there are i infected nodes in the network. If a current state i is not the absorbing state b, the chain will make a transition into its next state $i + 1$ after a time period.

In the next sections, we focus on a solution based on our analytical model for the expected spreading time and the number of data kiosks required to meet a target of a given expected spreading time.

3.2.2 Expected Spreading Time

We work out formulas for the expected spreading time which is represented under the digamma function and show a relationship between the expected spreading time and the expected message delay in the same situation. The following proposition presents the expected spreading time for content dissemination in DTNs with the presence of data kiosks.

Definition 5 (Homogeneous System)**.** *The homogeneous system on DTNs is a network where the contact rates between mobile nodes are identical, as well as the contact rates between a mobile node and a data kiosk if there exist some data kiosks in the network. The homogeneous environment is referred to the same concept.*

Proposition 3. *In a homogeneous system with the presence of data kiosks, the expected spreading time is given by*

$$d = \sum_{i=0}^{b-1} \frac{1}{(k\mu + i\lambda)(m-i)} \tag{3.1}$$

Proof. The spreading time that we are going to compute is the time until absorption of the chain.

A healthy node is infected with the content when it meets either one of k data kiosks or one of i infected nodes. Then, a healthy node is infected with the content at an exponential rate $\lambda i + \mu k$. A detailed proof for this result can be obtained by using the same argument in the proof of (2.2) in Section 2.2.1.

The chain leaves state i when one of $m - i$ healthy nodes is infected with the content. The event that a healthy node is infected occurs at an exponential rate $\lambda i + \mu k$, then the event that the chain will make a transition into its next state occurs at an exponential rate $(\lambda i + \mu k)(m - i)$.

Let T_i, $0 \leq i \leq b - 1$, be the time interval in which the chain stays in state i. Since T_i has an exponential distribution with parameter $(k\mu + i\lambda)(m - i)$, we have

$$E[T_i] = \frac{1}{(k\mu + i\lambda)(m-i)}. \tag{3.2}$$

Let T denote the spreading time, i.e. the time until there are b infected nodes in the network. Then,

$$T = \sum_{i=0}^{b-1} T_i. \tag{3.3}$$

We are ready to provide an expression for the expected spreading time d:

$$d = E[T] = E\left[\sum_{i=0}^{b-1} T_i\right] = \sum_{i=0}^{b-1} E[T_i] = \sum_{i=0}^{b-1} \frac{1}{(k\mu + i\lambda)(m-i)},$$

which proves Proposition 3. □

Corollary 1. *In a homogeneous system where no data kiosk exists and when all n mobile nodes are always interested in the news, the expected spreading time needed to disseminate the news from a source to all other mobile nodes is given by*

$$d = \frac{2}{\lambda n} \sum_{i=1}^{n-1} \frac{1}{i} \tag{3.4}$$

Proof. We consider the expected spreading time in a scenario of the news dissemination from a data kiosk to all other mobile nodes in a homogeneous environment. In this environment, there is only one data kiosk, all mobile nodes want to receive the news, and the contact rate between the data kiosk and a mobile node is equal to the contact rate between mobile nodes. Let n_1 be the total number of mobile nodes. We have $k = 1$, $b = m = n_1$, and $\mu = \lambda$.

Substitute $k = 1$, $b = n_1$, and $\mu = \lambda$ into (3.1) to get

$$\begin{aligned} d &= \sum_{i=0}^{n_1-1} \frac{1}{\lambda(1+i)(n_1-i)} \\ &= \frac{1}{\lambda} \sum_{i=1}^{n_1} \frac{1}{i(n_1-i+1)} \\ &= \frac{1}{\lambda(n_1+1)} \sum_{i=1}^{n_1} \left(\frac{1}{i} + \frac{1}{n_1-i+1}\right) \\ &= \frac{2}{\lambda(n_1+1)} \sum_{i=1}^{n_1} \frac{1}{i} \end{aligned}$$

The expected spreading time in the above scenario is equivalent to the one needed to disseminate the news from a source to all other mobile nodes in a homogeneous environment where no data kiosk exists and when $n = n_1 + 1$ mobile nodes are always interested in the news. Therefore, we obtain

$$d = \frac{2}{\lambda n} \sum_{i=1}^{n-1} \frac{1}{i},$$

which proves Corollary 1. □

In a homogeneous system where no data kiosk exists and when all n mobile nodes are always interested in the news, from expression (2.22), the expected message delay is $E[D] = \frac{1}{\lambda(n-1)} \sum_{i=1}^{n-1} \frac{1}{i}$. Using the expression of the expected message delay and (3.4), we obtain a relationship between the expected message delay and the expected spreading time.

Corollary 2. *In a homogeneous system where no data kiosk exists and when all n mobile nodes are always interested in the news, the relationship between the expected message delay and the expected spreading time is given by*

$$\frac{d}{E[D]} = \frac{2\,(n-1)}{n} \tag{3.5}$$

Corollary 2 exhibits a view on the relationship between the expected message delay and the expected spreading time. In a homogeneous environment where no data kiosk exists and when nodes are always interested in the news, a derivation from the formula is that the expected spreading time is approximately twice the expected message delay when the number of mobile nodes is large.

3.2.3 Number of Data Kiosks with a Constraint of an Expected Spreading Time

We now compute the number of data kiosks to meet an objective for the expected spreading time d. Unfortunately, it is not possible to find a closed-form expression for the number of data kiosks in the general case and we will therefore provide bounds. Lemma 1 to 4 will give the foundation to derive these bounds. We present details of the proofs for the lemmas in Appendix.

Lemma 1. *An expression of the expected spreading time represented under the digamma function is*

$$d = \frac{\psi\,(u+b) - \psi\,(u) + \varsigma}{\lambda u + \lambda m} \tag{3.6}$$

where $\varsigma = \sum_{i=0}^{b-1} \frac{1}{m-i}$, $u = \frac{\mu}{\lambda} k$, $\psi\,(x)$ *is the digamma function of* x.

The digamma function is defined as the logarithmic derivative of the gamma function [85]. For a positive real number x, $\psi(x) = \frac{d}{dx} \ln \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$. $\Gamma(x)$ is the gamma function: $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t}\, \mathrm{d}t$.

Lemma 2. *With $u \geq 1$ and $b \geq 2$, inequalities of $\psi(u+b) - \psi(u)$ are*

$$\ln\left(1 + \frac{\tilde{b}_\gamma + 0.5}{u + e^{-\gamma} - 1}\right) < \psi(u+b) - \psi(u) < \ln\left(1 + \frac{b_\gamma + 1}{u - 0.5}\right) \tag{3.7}$$

where $b_\gamma = b + e^{-\gamma} - 1.5$, $\tilde{b}_\gamma = b - e^{-\gamma}$.

Lemma 3. *A solution set of the inequality $\lambda du + \lambda dm - \varsigma < \ln\left(1 + \frac{b_\gamma+1}{u-0.5}\right)$ is*

$$\{u \in \mathbf{R} | u \geq 1, u < u_2\} \tag{3.8}$$

where

$$u_2 = 0.5 + \frac{H + \sqrt{H^2 + 4\lambda d (b_\gamma - 0.5)(H+1)}}{2\lambda d}, H = e^{-(0.5+m)\lambda d - \varsigma} - 1.$$

Lemma 4. *A solution set of the inequality $\lambda du + \lambda dm - \varsigma > \ln\left(1 + \frac{\tilde{b}_\gamma+0.5}{u+e^{-\gamma}-1}\right)$ is*

$$\{u \in \mathbf{R} | u \geq 1, u > u_4\} \tag{3.9}$$

where

$$u_4 = \frac{-(b_\gamma + 2m)}{4} + \frac{2\varsigma + \sqrt{[\lambda d (b_\gamma - 2m) + 2\varsigma]^2 + 8\lambda d \left(2\tilde{b}_\gamma + 1\right)}}{4\lambda d}. \tag{3.10}$$

We are almost ready to derive bounds of the number of data kiosks which is introduced in the following proposition.

Proposition 4. *If the contact rate between a mobile node and all data kiosks is greater than the contact rate between a pair of mobile nodes, or $k\mu \geq \lambda$, then an upper bound and a lower bound of the number of data kiosks to ensure that b subscribers will receive the content by the expected spreading time d are given by $\lfloor L \rfloor < k < \lceil U \rceil$,*

$$U = \frac{\lambda}{2\mu} + \frac{H + \sqrt{H^2 + 4\lambda d (b_\gamma + 1)(H+1)}}{2\mu d}, \tag{3.11}$$

$$L = \frac{-\lambda (b_\gamma + 2m)}{4\mu} + \frac{2\varsigma + \sqrt{[\lambda d (b_\gamma - 2m) + 2\varsigma]^2 + 8\lambda d \left(2\tilde{b}_\gamma + 1\right)}}{4\mu d} \tag{3.12}$$

where $H = e^{-(0.5+m)\lambda d - \varsigma} - 1$, $\varsigma = \sum_{i=0}^{b-1} \frac{1}{m-i}$, $b_\gamma = b + e^{-\gamma} - 1.5$, $\tilde{b}_\gamma = b - e^{-\gamma}$, γ is Euler-Mascheroni constant $\gamma \approx 0.57721$, e is Euler's number $e \approx 2.71828$.

Proof. Using Lemma 1, we find $\lambda du + \lambda dm - \varsigma = \psi(u+b) - \psi(u)$.

Because the number of data kiosks is larger than λ/μ, it follows that $u \geq 1$. Suppose that the number of infected nodes is greater than or equal to two by the spreading time, $b \geq 2$. Hence, using Lemma 2, we get $\ln\left(1 + \frac{\tilde{b}_\gamma + 0.5}{u + e^{-\gamma} - 1}\right) < \lambda du + \lambda dm - \varsigma < \ln\left(1 + \frac{b_\gamma + 1}{u - 0.5}\right)$. From Lemma 3 and Lemma 4, we obtain a solution for the inequalities. Substituting $u = \frac{k\mu}{\lambda}$ into the solution, we find an upper bound and a lower bound of the number of data kiosks $\lfloor L \rfloor < k < \lceil U \rceil$, where U is represented by (3.11) and L is described by (3.12), which demonstrates Proposition 4.

□

In the formulae, if $b = m$, or we want to disseminate a content item to all subscribers, the computed number of data kiosks ensures that a subscriber will receive the content item before a given expected spreading time d on average. The following corollary is derived directly from Proposition 4.

Corollary 3. *Suppose that the contact rate between a data kiosk and a mobile node in two distinct environments A and B is μ_A and μ_B respectively. Let U_A and L_A (resp. U_B and L_B) be an upper bound and a lower bound of the number of data kiosks in the environment A (resp. B). If we keep the same requirement for the expected spreading time, then a relationship between the bounds of the number of data kiosks and the contact rates in these environments is as follows:*

$$\frac{U_B}{U_A} = \frac{L_B}{L_A} = \frac{\mu_A}{\mu_B}.$$

Corollary 3 shows that the bounds of the number of data kiosks scale linearly with the contact rates between a mobile node and a data kiosk.

3.3 Evaluation

In this section, we validate the theoretical results through simulations under three mobility models and datasets, and discuss the application of the results in some specific scenarios. For a given context and a delay objective, we can state if the delay is acceptable to deploy the service. If negative, we provide the number of data kiosks that should be added in a hybrid DTN environment to meet the requirement.

3.3.1 Simulation under Mobility Models

3.3.1.1 Simulation Settings

We simulated the content dissemination with the presence of data kiosks by extending the Opportunistic Network Environment simulator [73]. Three mobility models that we

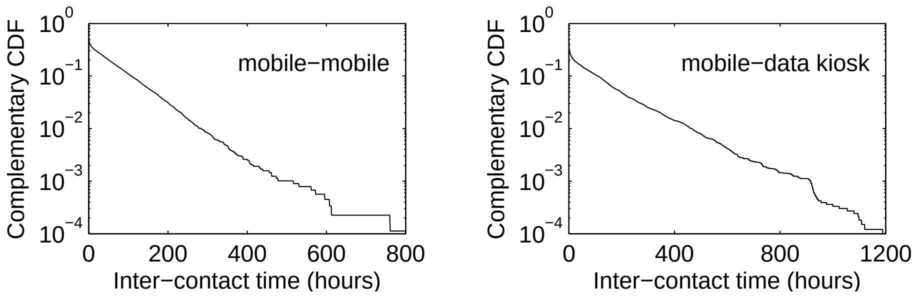


Figure 3.2: CCDF of inter-contact time for the random walker mobility model.

described in Section 2.3.1.1 were considered in the simulation. Mobile nodes move according to the mobility models under consideration and data kiosks are uniformly located in a square of size 4 km $\times$ 4 km. Radio ranges of nodes are 50 meters.

For the random waypoint and random direction mobility model, formulae for λ and μ were introduced [36, 84]. Using the parameter settings of our simulation, we computed the values for λ and μ and found that the contact rates are $\lambda \approx 0.0741$ and $\mu \approx 0.0409$ for the random waypoint mobility model, $\lambda \approx 0.0577$ and $\mu \approx 0.0437$ for the random direction mobility model.

A mathematical formula for the contact rates under the random walker mobility model is, to the best of our knowledge, not known. Therefore we obtained its value by fitting an exponential distribution to the inter-contact time data collected from simulations, using the least squares method [59]. Fig. 3.2 plots the CCDF of the inter-contact time on log-lin scale. Estimates of contact rates for this last mobility model are $\lambda \approx 0.0127$ and $\mu \approx 0.0070$.

3.3.1.2 Simulation Results

We ran simulations for different number of subscribers ($m = 40, 60$). The number of data kiosks was varied between 2 to 20. For each setting, we ran $1,000$ observations and computed the expected spreading time by which the content is delivered to 90 percent of subscribers. Then, we used analytical expressions (3.11) and (3.12) to calculate an upper bound and a lower bound of the number of data kiosks for a given value of the average spreading time that we found from our simulations. For each spreading time, we compared the estimated values of the number of data kiosks provided thanks to the theoretical solution or by simulation. Fig. 3.3 depicts the results under the three mobility models.

Results in Fig. 3.3 show the impact of the number of data kiosks on the expected spreading time and the accurate estimation of the model. As we see in the figure, the slope of the curve of the number of data kiosks increases when the expected spreading time decreases. This implies that we have to pay a high cost for the same delay reduction

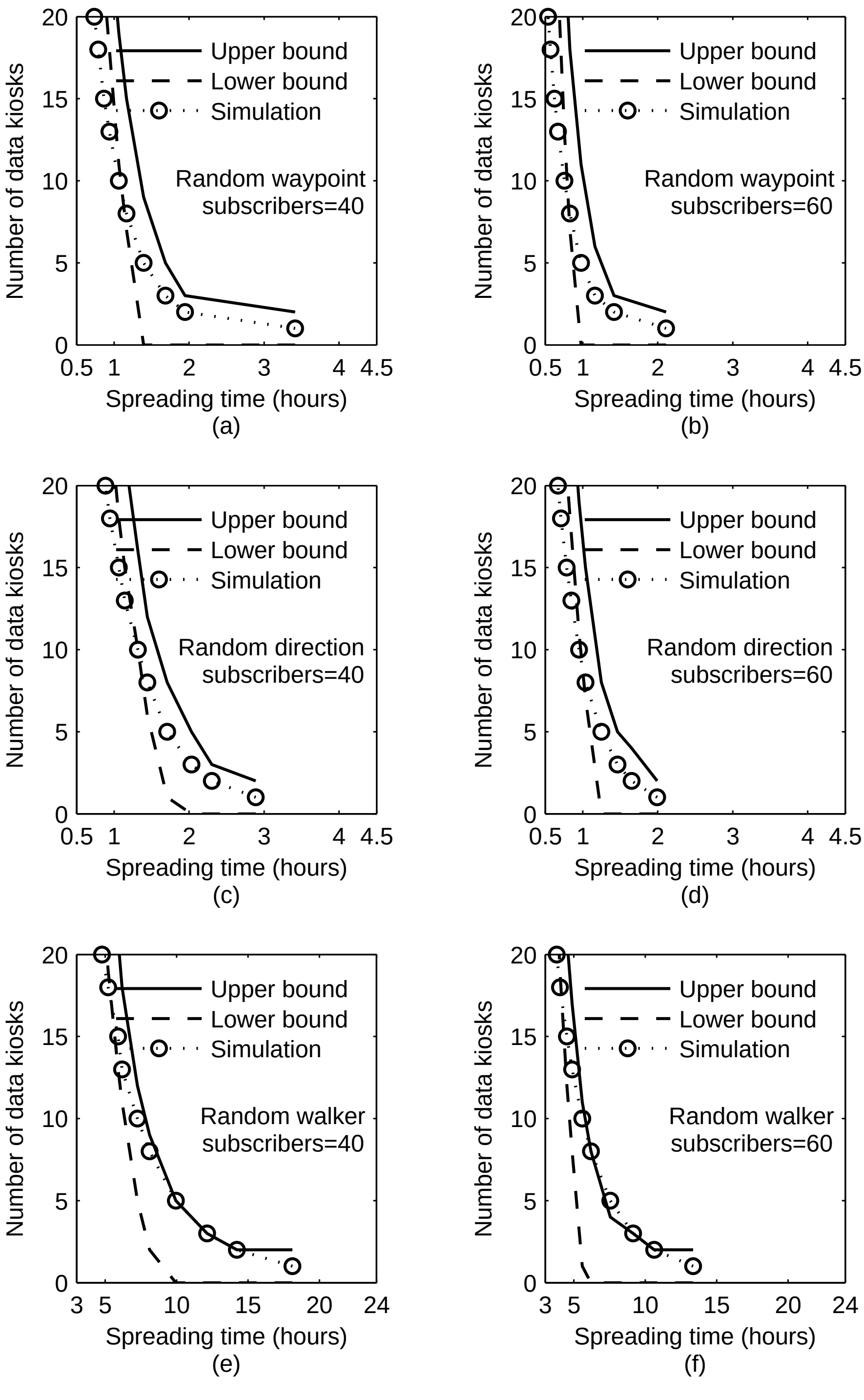


Figure 3.3: Number of data kiosks vs. spreading time.

when the expected spreading time is small. For example, in order to decrease the expected spreading time from 3.5 hours to 2 hours we only need to add 1 data kiosk, but in order to to decrease the expected spreading time from 1.5 hour to 1 hour we need to add at least 5 data kiosks. We also observe that analytical results can approximately predict the number of data kiosks, but the lower bounds deteriorate as the expected spreading time gets very small which is certainly not meaningful.

Fig. 3.4 demonstrates how the bounds approximately estimate the optimal value of the number of data kiosks in the theoretical aspect. For each optimal value of the number of data kiosks, we computed the expected spreading time by using Proposition 3. Then, for each obtained value of the expected spreading time, we estimated the bounds of the number of data kiosks by using Proposition 4. We plotted both the bounds and the optimal value in the figure. For the setting of other input parameters in the formulae, we used the values of λ and μ computed for the random direction mobility model in Section 3.3.1.1: $\lambda = 0.0577$, $\mu = 0.0437$. We set the number of mobile nodes m to 30 and the target number of infected nodes b to 27. The figure demonstrates that the lower bound is tight for the small value of the expected spreading time and the upper bound is tight for the large value of the expected spreading time.

We give an example for the application of our results with ideal parameters. We first consider the scenario without data kiosk. The urban area is a square of size 4 km $\times$ 4 km. There are 40 mobile nodes moving under the random waypoint mobility model. The transmission range is 50 meters, the probability of interest is 1, as every node is interested in the content. Under this scenario, the expected message delay is roughly 2 hours (as in Fig. 2.3(a)). If we wish to spread the content to 90 percent of the subscribers within 1 hour, it is difficult to satisfy this time requirement in a network without data kiosk. We now consider an hybrid environment. We keep these settings and find the number of data kiosks that we should add to meet the requirement. As we see in Fig. 3.3(a), with 10 data kiosks we can lessen the expected spreading time to 1 hour. We also see that the spreading time is improved effectively with some data kiosks. For example, with 5 data kiosks the expected spreading time can decrease to 2 hours in the random waypoint mobility model (as in Fig. 3.3(a)), and approximately to less than 8 hours in the random walker mobility model (see Fig. 3.3(e,f)).

In summary, the simulation results confirm that the analytical models can predict the expected message delay and the number of data kiosks to meet a delay requirement. The lower bound of the number of data kiosks gets close to the simulation results when the expected spreading time is small and the upper bound is accurate when the expected spreading time is large (Fig. 3.4).

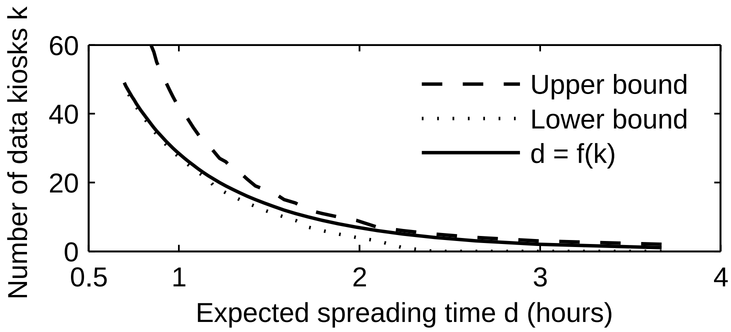


Figure 3.4: Bounds of the number of data kiosks and the function of expected spreading time.

3.3.2 Simulation on Datasets

In addition to the simulation using mobility models, we also carried out simulation based on real datasets to validate the analytical results. Our studies show how the model estimates the number of data kiosks when the input information of node mobility is the contact traces and explain the factors that affect the accuracy of the estimations.

We chose dataset Infocom2006 from the data repository CRAWDAD [76] for our simulation due to several reasons. We have found that dataset Infocom2006 is one of the few datasets that provide contacts directly between mobile devices and also contacts between the mobile users and several fixed nodes [86]. In addition, the dataset supplies a proper number of static nodes for simulation. Although the dataset does not have a long duration of experiment, the number of contacts recorded in the dataset is rich enough for the purpose of simulation.

We simulated the dissemination of content until 90 percent of the mobile nodes received the content. We ran the simulations for 60 mobile nodes that were chosen randomly from the mobile nodes with ID#21-#98, and with different number of data kiosks chosen randomly from static nodes with ID#1-#17. In the dataset, the time zero is the time when the experimental devices were configured. In the simulation, in order to remove the waiting time between the time of configuration and the time of content dissemination actually starting, we choose the start time as the time of the first contact in the datasets. The expected spreading time is the average value of the spreading time from $1,000$ runs.

Fig. 3.5 represents the simulation results and the prediction of the model for the number of data kiosks required to satisfy an objective of the expected spreading time. The model can estimate approximately the number of data kiosks although there are several obstacles that make an estimation more difficult on datasets. One obstacle is that the dataset can omit some contacts due to the sampling rate that is not high enough to detect short

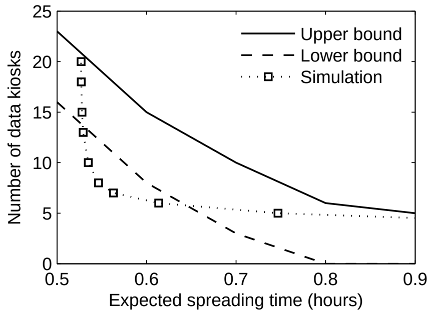


Figure 3.5: Number of data kiosks vs. expected spreading time for dataset Infocom2006.

contact events. Besides, the contact rates in our model are assumed to be homogeneous, but the contact rates in a real dataset can be heterogeneous. In addition, the exponential distribution of inter-contact times is only approximate. In spite of several approximations of input parameters above, the results show that the model still predict approximately the number of data kiosks required to satisfy a given constraint of the expected spreading time.

3.4 Discussion

We now discuss limitations and extensions of our results for estimating the number of data kiosks required for a given expected spreading time constraint in addition to the discussion of the accuracy, applicability, and flexibility that we have analyzed in Chapter 2.

The first issue to address is whether we can increase the accuracy of our solutions for estimating the number of data kiosks. In the previous chapter, we discussed how to improve the accuracy of contact rates between mobile nodes in Section 2.4. We also have the same discussion of the contact rates between a mobile node and a data kiosk for increasing the accuracy of the estimation. The discussion is that when we evaluate the contact rate from contact traces, the number of contacts should be large enough to capture well the value of contact rates.

Second, we have considered the problem of finding the number of data kiosks with a contraint of the average value of the spreading time. However, one might be interested in ensuring a constraint of the exact value of the spreading time. Due to the fact that contacts between a mobile node and a data kiosk, and between mobile nodes are opportunistic, this question is very hard to answer. An idea to approach the problem is to find conditions by which the event that the spreading time equals an objective value happens almost surely, or the event happens with probability one. However, whether we could find an appropriate solution is still a question.

Finally, Corollary 3 shows that the bounds of the number of data kiosks scale linearly with the contact rates between a mobile node and a data kiosk. In a homogeneous environment, the locations of the data kiosks do not influence the performance of the system. However, in a heterogeneous environment, the corollary raises the question of the best location for the data kiosks to maximize the contact opportunity between data kiosks and mobile users, or to minimize the delay. This practical question will be considered in Chapter 4.

3.5 Summary

In this chapter, we focus on a hybrid solution that uses both opportunistic contacts between a mobile node and a data kiosk and between mobile nodes for building a news dissemination system with a performance constraint. In particular, we have described an analysis for the number of data kiosks with a constraint of an expected spreading time, which is an important question in a news dissemination system. The key challenge to the constrained problem is that the news dissemination system using opportunistic contacts is highly dynamic and depends on multiple input information. Our solutions, as represented in Proposition 4 and Corollary 3, provide the bounds for estimating the number of data kiosks. An important property that is derived from these expressions is that the bounds of the number of data kiosks scale linearly with the contact rates between a mobile node and a data kiosk. We also investigated an expression for the expected spreading time and its relationship to the expected message delay. Our analytical model is flexible due to the fact that the information of the movement of mobile nodes that is input to the model can be a number of mobility models as well as datasets. Furthermore, the analysis is extensible to other contexts such as a news dissemination system that combines both car-to-infrastructure and car-to-car communications, by validating and estimating contact rates in the new context. Our simulation results on both a number of mobility models and real datasets confirm the accuracy of the analytical model.

Chapter 4

Locations of Data Kiosks

In the previous two chapters we have studied the problems of performance evaluation of content dissemination using DTNs with a number of data kiosks with respect to an expected spreading time constraint. In this chapter we add two questions. First, we have considered content dissemination problems only for a homogeneous environment. How is the dissemination of content when facing a heterogeneous environment? Since a heterogeneous environment is a more realistic scenario, analysis for the dissemination of the content in a heterogeneous environment should be an integral part of any practical solution to the content dissemination problem.

Second, we have developed a solution for the expected message delay that can be used to decide which parameter should be improved to enhance effectively the expected message delay, as well as an evaluation for a number of data kiosks with an expected spreading time constraint. We have not considered the issue of the locations of data kiosks. It is due to the fact that in a homogeneous environment the locations of data kiosks do not influence the performance of the system. However, in a heterogenous environment, placing a data kiosk in a region where the contact rate is high or in a region where the contact rate is low makes a difference in the spreading time. Thus, the question is: How does the location of a data kiosk influence the spreading time and where should we place a data kiosk to minimize the spreading time?

We develop answers to the above two questions. In the next section we introduce the content distribution related to the locations of data kiosks in a heterogeneous environment. Sections 4.2-4.3 describe our solution. Section 4.4 presents evaluations on both mobility models and real datasets. Section 4.6 summarizes the chapter.

4.1 The Placement Problem in DTNs

In this section we describe the placement problem in a heterogeneous environment in DTNs, and discuss its challenges.

4.1.1 Problem Statement

The problem of the optimal locations of data kiosks in a heterogeneous environment in DTNs has an important role in real-life deployment due to its direct impact on the performance of content distribution. In addition, since the deployment of data kiosks is a long-term decision, one has to address carefully the question of the locations of data kiosks in the planning phase. Successfully addressing this question would greatly aid in both the system performance and deployment cost. In the rest of this section, we first describe the placement problem informally. Then, we describe the parameterization and the formal definition for the problem.

The heterogeneous environment that we consider is an urban space composed of s regions and covered by a transportation system. There is one subway station in the center of a region. A mobile node moves freely in region i for a time period. Then, it travels to region j. It moves firstly to the subway stop in region i, then taking a train to move from the subway stop in region i to the subway stop in region j if there exists a line that connects these subway stops. The behavior of the mobile node in region j is reiterated as same as the one in region i. In the placement problem we assume that we know the number of data kiosks required to satisfy a requirement for the spreading time. Therefore, a data kiosk can be located in one region to be chosen among the s available ones. When we select a region for placing a data kiosk, we will place a data kiosk near the subway station of the region since it is the best position in the region for content dissemination.

Definition 6 (Heterogeneous System). *The heterogeneous system on DTNs is a network where the contact rates between mobile nodes in distinct sub-area are different, as well as the contact rates between a mobile node and a data kiosk in distinct sub-area if there exist some data kiosks in the network. The heterogeneous environment is referred to the same concept.*

Fig. 4.1 describes an example of news dissemination using opportunistic contacts both between a mobile user and a data kiosk, and between mobile users. A data kiosk that stores the news is located near subway stop Pont Neuf. Bob is moving in the region of the Louvre, and he wants to go to the University Pierre et Marie Curie (UPMC). Hence, he moves to subway stop Pont Neuf in the region of the Louvre to take the subway. At the subway station, he passes the data kiosk near the subway stop on the way to enter the

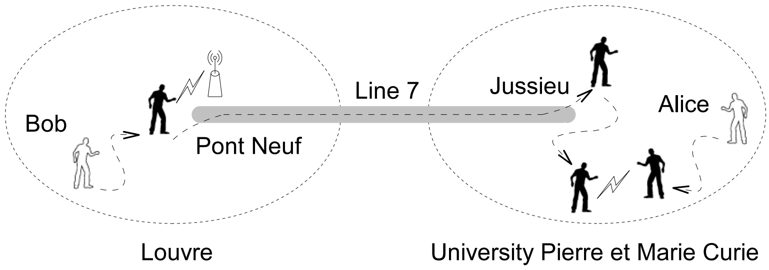


Figure 4.1: News dissemination in an urban area.

subway, and receives news from the data kiosk. After getting off subway stop Jussieu in the region of UPMC, he encounters Alice. As a result, Alice gets the news carried by Bob.

We now describe the information that are input to the problem, and the output that we require solutions to the problem.

Input Information: The specific information that we consider in the problem are:

- *An area map:* An area map provides region areas and subway lines that connect regions. A region area influences the speed of content dissemination in the region. A subway line between regions show that whether the content dissemination in a region has a direct impact on the one in another region.
- *People mobility in each region:* This input allows to parameterize contacts between mobile nodes and between a mobile node and a data kiosk. We use this input to model the content dissemination within a region. This input can refer to mobility models, and datasets.
- *People movement between regions:* We use this type of input to parameterize the influence of content dissemination in one region on the dissemination in another region.
- *A number of data kiosks:* As mentioned, we assume that we have a number of data kiosks that we need to deploy, to cover the area.

Output Solution: We require that a solution to the problem produces a list of regions where should we place a data kiosk to optimize the expected spreading time.

How to parameterize contacts and movement of mobile nodes influences the accuracy of the solution. Motivated by [36, 84], the inter-contact times for mobile-mobile (resp. mobile-data kiosk) contacts in the area of region i can be approximately represented as exponentially distributed random variables with mean $1/\lambda_i$, $\lambda_i > 0$ (resp. $1/\mu_i$, $\mu_i > 0$). The process of mobile nodes going from region i to region j is assumed to be represented by a Poisson process with rate γ_{ij}.

Definition 7 (Moving Rate)**.** *The moving rate γ_{ij} is the number of mobile nodes that move from region i to region j per a unit of time on average. The moving rate γ_{ij} is equal to* 0 *if there is no subway lines connecting region i with region j.*

In order to provide formal definitions for the problem, we introduce the following notations:

- $V = \{i | 1 \leq i \leq s\}$: The set of regions
- $\mathbf{1}_i$: The indicator function, equal to 1 if a data kiosk is located in region i, and equal to 0 otherwise.
- $x_i(t)$: The number of infected nodes in region i at time t
- $n_i(t)$: The number of subscribers in region i at time t

Problem 3 (Locations of Data Kiosks)**.** *Given an area $\mathcal{A}$ composed of s regions, a number of subscribers in region i at the initial time $n_i(0)$, movement of mobile nodes in each region and between regions, a subway map describing the connections between regions, a target number of infected nodes b, a number of data kiosks k, find the location of k data kiosks in an area $\mathcal{A}$ $\{\mathbf{1}_i | 1 \leq i \leq s\}$ in order to minimize the spreading time d, subject to $\sum_{1 \leq i \leq s} \mathbf{1}_i = k$, and $\sum_{1 \leq i \leq s} x_i(d) = b$.*

4.1.2 Challenges of the Placement Problem in DTNs

Addressing the placement problem in DTNs is a challenging effort due to the factors explained in Section 1.2. In addition, our placement problem raises several significant challenges in a heterogeneous environment. The first challenge is how to model the movement of mobile nodes between regions. On the one hand, we must make enough simplifying assumptions to allow us to control a mathematical model. On the other hand, however, assumptions would be applicable to the practice. The Poisson process is a powerful tool used extensively in modeling physical events such as the number of web page requests arriving at a server, the number of telephone calls arriving at a switchboard and the number of raindrops falling over an area. The Poisson process is a good approximation for modeling the movement of mobile nodes between regions due to the fact that the arrival of customers in a queue is well-modeled by the Poisson process [87, 88].

Second, the association between several regions gives rise to the dynamics of system that make more difficult for modeling content dissemination. Many factors influence the dissemination of content: the movement of mobile nodes in a region, the movement frequency of mobile nodes between regions, the location of data kiosk in the subway map, the location of subway station in a region, the location of data kiosk in a region, and the transmission range. In order to reduce the complexity of problem, we try to decrease the dynamic of minor factors by assuming that the subway station is in the center of a region, and a data kiosk is located near the subway station since it the best position to disseminate content.

Finally, the placement problem in DTNs is a challenge because although many techniques have been proposed for placement problems in several contexts, they are not well suited for DTNs due to challenging characteristics that are intermittent connectivity, high mobility, unpredictable movement, and limited capacities. Facility location problems, a related work, concerning optimal placement of facilities in order to optimize performance, have been considered extensively in the fields of operation research. A number of interesting strategies on placing facilities, usually using approximation algorithms, have been proposed [89]. However, facility location problems consider the optimization of cost between location of facilities and a given set of static point sites where costs between them are fixed. In DTNs, the placement problem becomes harder since estimating the performance cost for a given solution is challenging.

In Wireless Local Area Networks (WLAN), placement problems is a process of placing access points in WLANs in order to achieve a performance objective such as the requirements of area coverage, delay, and traffic capacity [90–94]. Several approaches were proposed to find the optimal location of access points in WLANs, such as simplified optimization model [90], a heuristic method Nelder–Mead [91], branch and bound algorithm [92], hierarchical approach [93], and the Genetic Algorithm [94]. Previous studies in the context of WLANs cannot be applied to DTNs due to several challenging characteristics of DTNs. For example, when calculating performance metrics, in WLANs one only needs to consider infrastructure based communication between access points and nodes while in DTNs one needs consider an additional type of communication that is opportunistic hop-by-hop between mobile nodes.

In Wireless Sensor Networks (WSN), the problem of node placement, which has a considerable impact on the effectiveness of a WSN, has been studied in many papers for different aspects. For example, [95, 96] studied the placement of data collectors, which are nodes collecting the data generated by sensors over a multi-hop path for further processing of the data. [97,98] considered the placement of relay nodes, which are added to the network to increase network efficiency by relaying data generated by other sensor nodes to sink nodes. While nodes in WSNs need to be highly connected and not be isolated, nodes in DTNs can stand disruption of connectivity. This distinctive characteristic of DTNs has to be taken into account when considering the placement problem.

In DTN contexts, Chaintreau et al. [99] proposed a greedy solution to optimize the performance measures related to the age distribution for the placement of base stations. The age of a content item is the time elapsed since the content held by this node was emitted by a base station. In [52], the paper finds out the locations of fixed nodes in order to maximize the total throughput in a DTN network by using network flow techniques. Our work differs from previous publications as it provides a model and its solution for locating data kiosks associated with subway stations. This is done to optimize the spreading time required to fulfil the interest of all subscribers

In the rest of the chapter we describe a numerical solution based on an ODE model for the problem. Two trends of the optimal placement of data kiosks were derived from the numerical solution. Simulations were implemented on both mobility models and the datasets, which validated our solution.

4.2 A Numerical Solution for the Placement Problem

We exhaustively search an optimal solution for the locations of data kiosks in all possible candidates for the placement of data kiosks. The important step in finding an optimal solution is first to build systems that describe the dissemination of content for given placement of data kiosks. Then, we solve the systems to obtain the number of infected nodes in regions as a function of the spreading time. By comparing the results of all cases of the placement of data kiosks, we decide which case is optimal.

We first describe the important step, how to model the evolution of content in a heterogeneous environment, for finding an optimal solution. The key idea underlying our approach is to divide an infection process in a large area which is composed of several regions into a system of infection processes in each region. Specifically, in each region we describe the infection process as a function of contacts between nodes in the region, the rates at which infected nodes in the region go out to other regions, the rates at which infected nodes from other regions come into the region, the number of infected nodes and the number of healthy nodes. The relation between infection processes of regions are represented by the moving rates.

We describe the infection process in the area composed of s regions by using an ODE model [57]. The differential equations of the infection process are as follows.

$$\begin{aligned} x_i'(t) &= (x_i(t)\lambda_i + \mathbf{1}_i\mu_i)(n_i(t) - x_i(t)) \\ &\quad + \sum_{j\in V, j\neq i} \gamma_{ji} x_j(t) - \sum_{j\in V, j\neq i} \gamma_{ij} x_i(t), \quad \forall i \in V \\ n_i'(t) &= \sum_{j\in V, j\neq i} \gamma_{ji} n_j(t) - \sum_{j\in V, j\neq i} \gamma_{ij} n_i(t), \quad \forall i \in V \end{aligned} \tag{4.1}$$

Initial conditions:

$$x_i(0) = 0, \quad \forall i \in V,$$
$$\sum_{i \in V} n_i(0) = m,$$

where $n_i(0)$ are constants.

We consider the system at a steady state. At the state, mobile nodes move between regions, but the total number of mobile nodes does not change on average or we have $n_i(t) = n_i(0), \forall i \in V$. Therefore,

$$n_i'(t) = \sum_{j \in V, j \neq i} \gamma_{ji} n_j(t) - \sum_{j \in V, j \neq i} \gamma_{ij} n_i(t) = 0, \quad \forall i \in V.$$

The differential equation of the number of mobile nodes in region s is derived from the equations of the first $(s-1)$ regions; therefore we can withdraw the equation with $i = s$. A system of equations for the number of mobile nodes in regions at a steady state is given by:

$$\begin{aligned} \sum_{j \in V, j \neq i} \gamma_{ji} n_j(t) - \sum_{j \in V, j \neq i} \gamma_{ij} n_i(t) = 0,\ \forall i \in V \setminus \{s\}, \\ \sum_{i \in V} n_i(t) = m. \end{aligned} \tag{4.2}$$

It is a system of s linear equations and s variables. Using Gaussian elimination, we can find a solution for (4.2) to determine the number of mobile nodes in each region at the steady state of (4.1).

We solve (4.1) at the steady state to evaluate the number of infected nodes as a function of the spreading time. The infection process represented by (4.1) is a nonlinear non autonomous system of first-order differential equations. We were not able to find an explicit expression for evaluating the number of infected nodes, but a numerical solution can be exploited by Runge-Kutta methods [100].

To find the optimal locations of data kiosks in a system including k data kiosks and s regions, we have to compute the number of infected nodes as a function of the spreading time in all possible cases of the placement of data kiosks. The total number of possibilities is the number of k-combinations of an s-element set that is equal to the binomial coefficient $C(s,k)$. Hence, a solution for the optimal locations of data kiosks in the system is obtained by solving $C(s,k)$ systems of $(s+k)$-dimensional nonlinear non-autonomous of first-order differential equations.

In summary, we find a solution for the locations of data kiosks that subject to optimize the expected spreading time in an urban area composed of several regions in three main

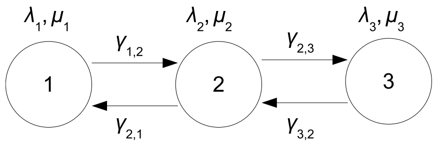


Figure 4.2: The parameters of a system of three regions.

steps: Modeling the dissemination of content for all possible candidates of the placement of data kiosks, solving the system to get the number of the infected nodes as a function of the expected spreading time, and selecting the optimal solution for the locations of data kiosks. In the steps, modeling the dissemination of content is the key step that decides the accuracy of the solution.

4.3 Two Trends of the Locations of Data Kiosks

Current placing solutions (including our solution described in the previous section) have only focused on exact solutions or approximation solutions with different constraints and different optimization objectives. They have not discussed general trends found from their solutions for classes of specific problems. General trends create our profound understanding of the impacts of the locations of data kiosks. In addition, general trends make solutions simplified to face huge size of the placement problem. Our questions are that what is the meaning inside numerical solutions or do solutions have a general rule for classes of problems? In this section, we use the analytical model described in the previous section to explore the impact of the locations of data kiosks with respect to their association to some subway stations while looking for a better service. In particular, we analyze numerically optimal solutions for the locations of data kiosks in a simple case and then derive some expected behavior in more general situations.

4.3.1 Placing One Data Kiosk in a Simple Scenario

We consider a simple scenario with three regions connected by a line of three subway stops (Fig. 4.2). We suppose that the optimal number of data kiosks computed is one and, therefore, we have to place one data kiosk in one of the three regions. We want to study the impact of various parameters such as: the location of a subway station, the contact rate between mobile nodes λ, the contact rate between a mobile node and a data kiosk μ, and the moving rate of people between two stations γ, on the content dissemination by the spreading time.

We first set the same parameters for every region to study the impact of the location of a station on the total number of infected nodes. Note that the location of a station/region

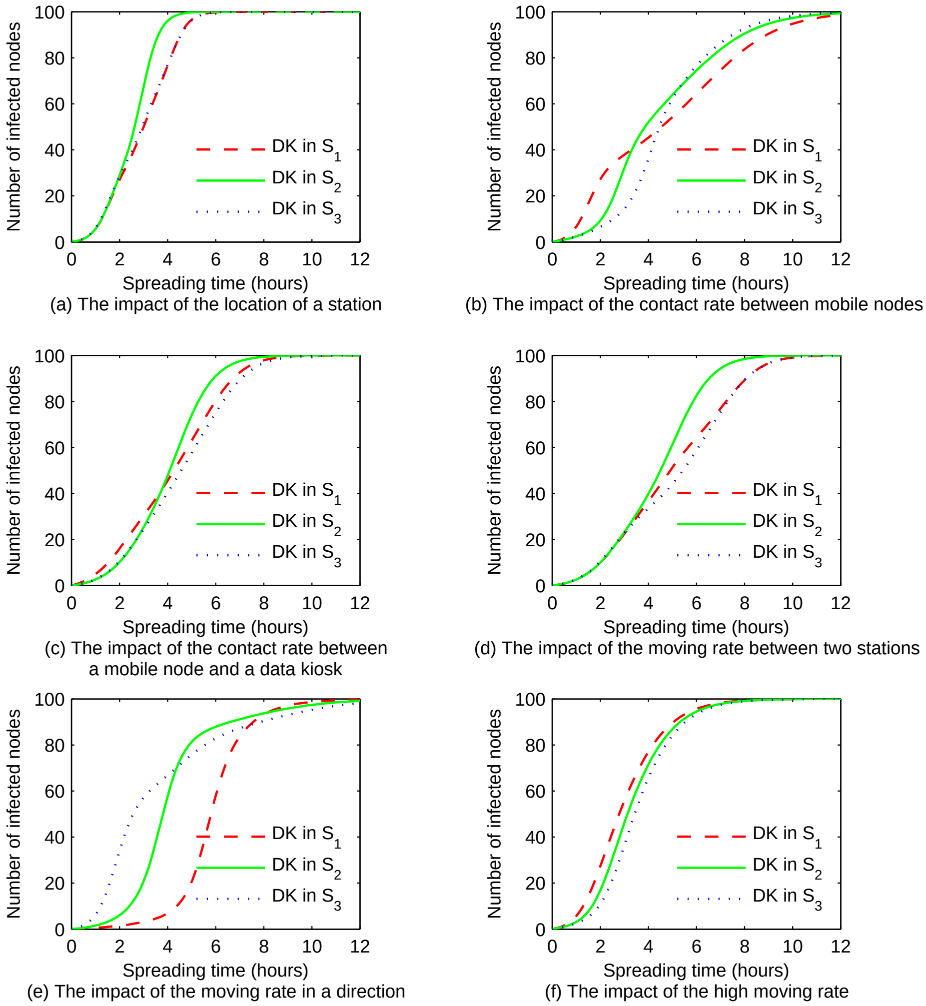


Figure 4.3: Placing one data kiosk in an area of three regions.

that we mention is the relative position of the station/region in the subway map. Settings are $\lambda_1 = \lambda_2 = \lambda_3 = 0.08$, $\mu_1 = \mu_2 = \mu_3 = 0.04$, and $\gamma_{12} = \gamma_{21} = \gamma_{23} = \gamma_{32} = 0.02$. The initial number of mobile nodes in regions is $n_1(0) = 33$, $n_2(0) = 33$, and $n_3(0) = 34$. We plot the total number of infected nodes as a function of the spreading time in Fig. 4.3(a). In the figure and the following figures, the notation "DK in S_i" indicates that we place a data kiosk near the subway station in region i. The results show that the total number of infected nodes increases quickly when a data kiosk is put in a region whose location is central.

Secondly, we set a high value of the contact rate between mobile nodes for a region to examine the impact of the contact rate on the total number of infected nodes. Settings are $\lambda_1 = 0.08, \lambda_2 = \lambda_3 = 0.02$, $\mu_1 = \mu_2 = \mu_3 = 0.05$, and $\gamma_{12} = \gamma_{21} = \gamma_{23} = \gamma_{32} = 0.08$. The initial number of mobile nodes in regions is $n_1(0) = 33$, $n_2(0) = 33$, and $n_3(0) = 34$. Results are depicted in Fig. 4.3(b). In the initial time period, a total number of infected nodes increase quickly when a data kiosk is located in region 1 that has a high value of the contact rate. In the final time, the spreading time that we need to disseminate the content to all nodes is smaller when a data kiosk is put in region 3 that has a low value of the contact rate and do not stay in a central position. If we put a data kiosk in region 2 that has a low value of the contact rate and has a central position, the number of infected nodes has an intermediate value in the first time period and nearly optimal value in the final time. To summarize, if we want to infect the content to all mobile nodes in the area, we should place a data kiosk in a region where the contact rate between mobile nodes is lower. If we want to spread quickly the content to one part of a network, we should place a data kiosk in a region where the contact rate is higher.

Thirdly, we observe the impact of the contact rate between a mobile node and a data kiosk on the total number of infected nodes. Settings are $\lambda_1 = \lambda_2 = \lambda_3 = 0.04$, $\mu_1 = 0.08, \mu_2 = \mu_3 = 0.04$, $\gamma_{12} = \gamma_{21} = \gamma_{23} = \gamma_{32} = 0.04$, $n_1(0) = 33$, $n_2(0) = 33$, and $n_3(0) = 34$. The results represented in Fig. 4.3(c) show a similar impact as the one between mobile nodes, but placing a data kiosk in a central region has advantage in the settings. A central region is a region where its relative location in the subway map is central; for example, region 2 is the central region in the subway map represented in Fig. 4.2.

Next, we set a high value of the moving rate between two stations to study the impact of the moving rate on the total number of infected nodes. Settings are $\lambda_1 = \lambda_2 = \lambda_3 = 0.04$, $\mu_1 = \mu_2 = \mu_3 = 0.04$, $\gamma_{12} = \gamma_{21} = 0.03$, $\gamma_{23} = \gamma_{32} = 0.01$, $n_1(0) = 33$, $n_2(0) = 33$, and $n_3(0) = 34$. The content is spread quickly when we place a data kiosk in a region whose position is central as we see in Fig. 4.3(d).

We also examine the impact of the moving rate on the total of infected nodes when the moving rate is high in only one direction. Settings are $\lambda_1 = \lambda_2 = \lambda_3 = 0.04$, $\mu_1 = \mu_2 = \mu_3 = 0.03$, $\gamma_{12} = \gamma_{23} = 0.04, \gamma_{21} = \gamma_{32} = 0.02$, $n_1(0) = 14$, $n_2(0) = 28$, $n_3(0) = 58$. Fig. 4.3(e) reveals that if we want to spread quickly the content to all mobile nodes, we should add a data kiosk to the initial station in the direction of a high moving rate; and if we want to spread the content to a part of a network, we should support a data kiosk in the last station in the direction.

Finally, when the moving rate is large between stations, the location of the data kiosk has less influence. Settings are $\lambda_1 = 0.08$, $\lambda_2 = \lambda_3 = 0.02$, $\mu_1 = \mu_2 = \mu_3 = 0.05$, $\gamma_{12} = \gamma_{23} = \gamma_{21} = \gamma_{32} = 1$, $n_1(0) = 33$, $n_2(0) = 33$, $n_3(0) = 34$ and displayed in Fig. 4.3(f).

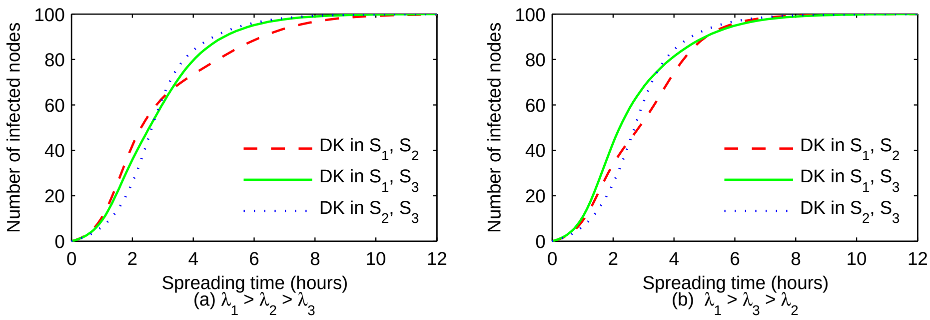


Figure 4.4: Placing two data kiosks in an area of three regions.

These observations conclude two placement strategies for the optimal locations of data kiosks. The first optimization strategy is that we should put a data kiosk to a region whose parameters are rather weak (e.g. its contact rate is low) if we want to minimize the spreading time required to disseminate the content to all subscribers. The second strategy recommends placing a data kiosk in a region whose parameters are good (e.g. its contact rate is high) if the objective is to minimize the spreading time needed to distribute the content to a small set of subscribers. In both strategies, a region whose position is central provides a reasonable trade-off. For a high moving rate between regions, the two strategies have less impact on the optimal locations of data kiosks.

4.3.2 Placing Data Kiosks in a General Scenario

We first examine two strategies, which we investigated, when we put two data kiosks in the scenario of three regions. Suppose that the condition for spreading news in region 1 is the best among the three regions, then there are two possibilities: the first possibility is that the condition of region 2 is better than the one of region 3, and another possibility is that the condition of region 3 is better than the one of region 2. Settings are $\lambda_1 = 0.08$, $\lambda_2 = 0.05$, $\lambda_3 = 0.02$ for the first case, and $\lambda_1 = 0.08$, $\lambda_2 = 0.02$, $\lambda_3 = 0.05$ for the second case. Other parameters are the same as the one in the second case of the analysis of placing one data kiosk. Fig. 4.4 depicts the analytical results.

We inspect two trends by comparing the results in Fig. 4.3(b) when we place one data kiosk and the results in Fig. 4.4 when we place two data kiosks. Suppose that the first data kiosk was located in region 1, we now place a data kiosk to either region 2 or region 3. For the first case, when the contact rate of mobile nodes in region 2 is higher than the one in region 3, placing a data kiosk in region 2 maximizes the number of infected nodes in the initial time period, but placing a data kiosk in region 3 is optimal to infect all mobile nodes. For the second case, when the contact rate of mobile nodes in region 3 is higher

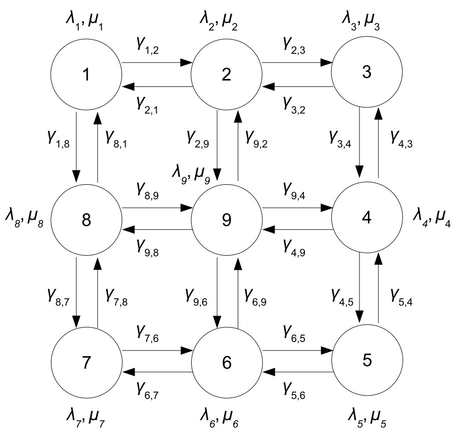


Figure 4.5: The parameters of a system of nine regions.

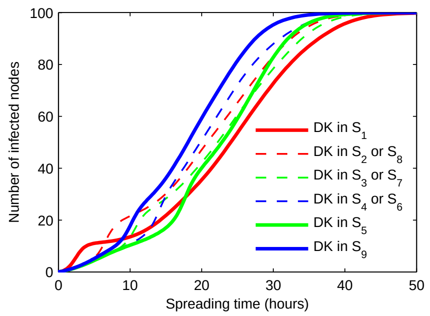


Figure 4.6: Numerical solutions in a case of a system of nine regions.

than the one in region 2, placing a data kiosk in region 3 maximizes the number of infected nodes in the initial time period, but placing a data kiosk in region 2 is optimal to infect all mobile nodes. The results show that when the conditions in each region change, the optimal locations of data kiosks also change according to the two trends.

We obtained the same behavior when the first data kiosk is located in either region 2 or region 3, and we find the location for the second data kiosk. The results confirmed the two placement strategies, which we investigated in Section 4.3.1, for the optimal locations of data kiosks. The first optimization strategy is that we should place a data kiosk in a region whose parameters are rather weak (e.g. its contact rate is low) if we want to minimize the spreading time required to disseminate the content to all subscribers. The second strategy recommends placing a data kiosk in a region whose parameters are good (e.g. its contact rate is high) if the objective is to minimize the spreading time needed to distribute the content to a small set of subscribers.

For a more general topology of the subway system, a grid of nine subway stations (Fig.

Table 4.1: Simulation Settings of Two Regions in an Area $10km^2$

Reg.	The lower left and upper right corners (km)	Subway stop (km)	Moving rate ($1/h$)	Location of DKs (km)
1	(0.0, 0.0), (3.0, 3.0)	(1.6, 1.6)	0.0504	(1.5, 1.5)
2	(4.0, 4.0), (10.0, 10.0)	(6.9, 6.9)	0.0504	(7.0, 7.0)

Table 4.2: Simulation Settings of Two Regions in an Area $20km^2$

Reg.	The lower left and upper right corners (km)	Subway stop (km)	Mov. rate ($1/h$)	Location of DKs (km)
1	(0.0, 0.0), (4.0, 4.0)	(2.1, 2.1)	0.0504	(2.0, 2.0)
2	(14.0, 14.0), (20.0, 20.0)	(16.9, 16.9)	0.0504	(17.0, 17.0)

4.5) where subway station i is located in region i, $i = 1, 2, \ldots 9$, we explore the same expected behavior when putting a data kiosk in the system. The results are depicted in Fig. 4.6 where the contact rate between mobile nodes in region 1 is 0.132, the contact rates between mobile nodes in other regions are 0.033, the contact rates between a mobile node and a data kiosk in each region are 0.041, the moving rates of people between two regions connected by a subway line are 0.003. Region 1 has the most favorable parameters to propagate the news; but when we place a data kiosk to region 1, the news dissemination is fast only in the initial time period then slows down. Putting a data kiosk in region 9 is optimal to scatter the news over the whole area because other regions have the same condition and region 9 has an advantage of a central location in the subway map. We argue that putting a data kiosk in a region whose parameters are good is optimal to spread the news to a small number of mobile users but might not be optimal to spread the news to all mobile users. This is due to the fact that the dissemination of news in other regions where we do not place a data kiosk is too slow. Supporting a data kiosk in a region whose parameters are rather weak is optimal to spread the news over the whole area because we make a balance between regions. Although we do not have a formal proof, we conjecture that these observations will hold in a general topology of the transportation subway map.

4.4 Evaluation

In this section, we present simulation results to validate the policies of the placement of data kiosks in an area divided into several regions linked by some subway lines. We begin by describing simulation parameters for this context. Then, we show the estimation of analytical parameters in the model. Finally, we compare simulation results and analytical results to examine how the policies adjust to the variation of a number of parameters.

Table 4.3: Theoretical Results of Contact Rates

Area (km^2)	λ $(1/h)$	μ $(1/h)$
3	0.132 269	0.099 553
4	0.074 401	0.055 998
6	0.033 067	0.024 888

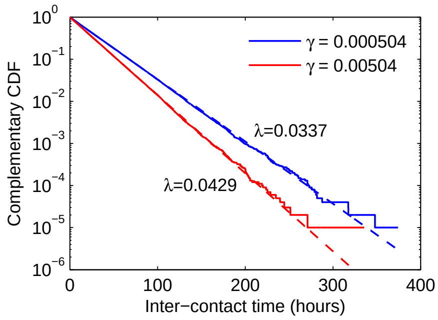


Figure 4.7: CCDF of inter-contact times for different moving rates.

4.4.1 Simulation under Mobility Models

4.4.1.1 Simulation Settings

We simulated a hybrid DTN with a number of regions connected by subway lines. In a region, mobile nodes move according to the random waypoint mobility model that was described in Section 2.3.1.1 and data kiosks are located in the center of a region that is near the subway stop. In a region, a mobile node will move for a time interval that depends on the moving rates. When the time expires, the node will go to the subway stop in the region to travel to another region. We set the moving time between two subway stops to 3 minutes. Table 4.1 and Table 4.2 give other parameters of a scenario that includes two regions connected by a subway line in areas of size 10km^2 and 20km^2. A region is limited by the coordinates of its lower left and upper right corner. A subway stop and a data kiosk are represented by their coordinates in an area. We will place a data kiosk in one of the two regions.

For the random waypoint mobility model, the estimation of λ and μ were introduced [36,84]. Using the formulae and the parameter settings of our simulation, we compute some theoretical values of contact rates in Table 4.3.

We examine the impact of the moving rate on the prediction of the contact rate. We run simulations using the settings in Table 4.1 with different moving rates $\gamma = 0.000504$ and $\gamma = 0.00504$ until we collect $100,000$ contacts between mobile nodes in region 2. We compute the contact rate between mobile nodes in region 2 for each case of the moving

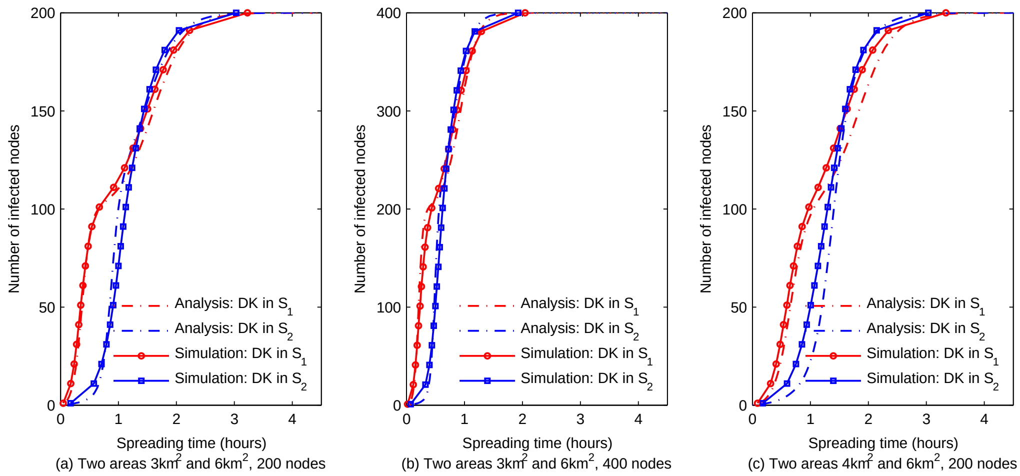


Figure 4.8: Number of infected nodes for different number of mobile nodes and inter-contact times.

rate. Fig. 4.7 plots the complementary cumulative distribution function (CCDF) of the inter-contact time on log-lin scale. Estimate of contact rates for $\gamma = 0.000504$ is $\lambda = 0.0337$ and the one for $\gamma = 0.00504$ is $\lambda = 0.0429$. By comparing the theoretical results in Table 4.3 and the simulation results, we see that the prediction of contact rates is accurate when the moving rate is low. In the case of high moving rates between subway stations, nodes tend to concentrate in a small area around the subway stop since they do not have enough time to go away from the subway stop. That explains why the contact rates are underestimated when the moving rate is high.

4.4.1.2 Simulation Results

We vary the number of mobile nodes, area sizes, and moving rates to verify how two optimal placement strategies of data kiosks scale to different environments. In the first simulation study, we use the settings in Table 4.1 and vary the number of mobile nodes. The results in Fig. 4.8(a,b) confirm that the model accurately describes the growth of the number of infected nodes for different number of mobile nodes. We then check the prediction of the model when we change the value of λ and μ by modifying the areas of regions. Fig. 4.8(a,c) plot the results under two settings of the areas described in Table 4.1 and Table 4.2. We observe that the analytical results can approximately predict the behavior of the system. Finally, we vary the moving rates in Table 4.1 to $\gamma = 0.072$ and $\gamma = 0.144$ while keeping all other parameters to examine their impact on the prediction of the model. The results are depicted in Fig. 4.8(a)-4.9 that compare the theoretical analyses with the simulation results. When the moving rates increase, the prediction known from the analytical model

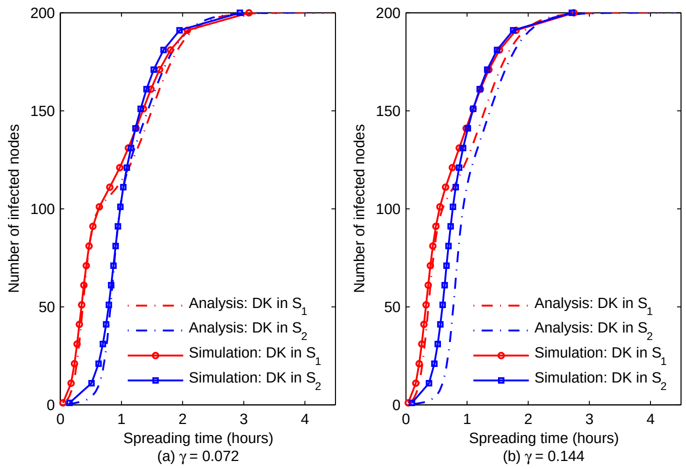


Figure 4.9: Number of infected nodes for different moving rates.

weakens because of the impact of moving rates on the contact rate.

For different number of mobile nodes, contact rates and moving rates, the model, although approximate, can tell us about both the time evolution of the number of infected nodes and two trends of the placement of the data kiosks. When the moving rate is small, the simulation results confirm two trends of minimizing the spreading time. When the moving rates increase, due to a short present time of a node in a region, although the fitting between simulation results and analytical results weakens, the model still can predict the behavior of the system. If the moving rate is high enough, the two trends disappear due to a strong connection between two regions. This occurs because one change in a region can influence quickly on other regions. Hence, with a high moving rate, placing a data kiosk in a region with good parameters is optimal in minimizing the spreading time, regardless of a target number of subscribers that will receive content.

4.4.2 Simulation on Datasets

This section analyzes experimental results on real datasets as a complement to the use of mobility models in simulation. Our studies provide a better understanding of how the model can estimate the quantity of interest upon the approximation of the exponential distribution of inter-contact times with high heterogeneity, and the dependence of the numerous variables in real datasets.

We use dataset Infocom2006 and Content in data repository CRAWDAD for our simulation [76]. While a number of datasets have the advantage of a long duration of experiment, they do not record direct contacts between mobile nodes (i.e. MIT [77], UCSD [78], Dartmouth [79]). We have found three datasets that meet the context of our research: Infocom2006 [86], Content [101] and PMTR [102]. Although PMTR provides a fine-grained dataset, the number of static nodes in the dataset is very few (i.e. 5 static nodes). Dataset Infocom2006 and Content supply a better number of static nodes for our simulation. A limitation of dataset Content is that the number of contacts between static nodes and mobile nodes is not as rich as the one in dataset Infocom2006 [101]. However, two datasets Infocom2006 and Content are the most appropriate datasets for our experiment due to the fact that they provide contacts directly between mobile devices and also contacts between mobile users and several fixed nodes.

In the setting of a wide area that is composed of several regions, dataset Infocom2006 and Content describe the movement of people in region 1 and region 2, respectively. It is due to the fact that we have not found any dataset that records the movement of people in a scenario that fits with the one we consider. The set of moving people in each region is a subset of the people in each dataset. In order to simulate the movement of the people between connected regions, we will remove all contacts of a person from dataset Infocom2006 when the person moves from region 1 to region 2. In region 2, we will add a new person to the set of moving people. The new person is chosen randomly from people who belong to dataset Content but are not in the current set of moving people in region 2. The movement of people in the reverse way is simulated similarly. In this simulation, we ignore any contacts between a person and other people if he moves from his current position to the subway stop for moving to other regions. The approximation is reasonable because the moving time between his current position and the subway stop is normally very short as it is compared to the total time that he spends in each region.

More details of simulation setting are as follows. At the start time, there are 18 mobile nodes in each region. The mobile nodes are chosen randomly from the set of people that belongs to the trace set of the region. A data kiosk is put near a subway stop in the region 1 or the region 2. We assume that there is a subway stop that is near the location of one of the fixed nodes in the dataset. In our simulation, the location of the subway stop in the region 1 is near the location of the node with ID#16 and the one in the region 2 is near the location of the node with ID#54. The moving rates from region 1 to region 2 is $0.504.10^{-3}$ and the one from region 2 to region 1 is $0.504.10^{-3}$. We computed the average number of infected nodes from 100 observations.

The simulation results are depicted in Fig. 4.10. The figure plots the evolution of the average number of infected nodes in each region as a function of the spreading time when we put a data kiosk in the region 1. Note that people in the region 1 meet more frequently than people in the region 2. The result confirms the strategies of the placement of data kiosks, which were revealed in Section 4.3.

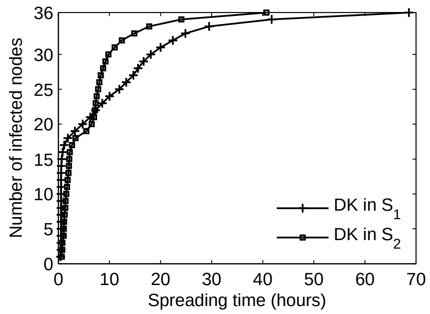


Figure 4.10: The evolution of infected nodes in an area of two regions using real-world traces.

4.5 Discussion

The analytical model and two placement strategies of data kiosks completely answer to the two questions about the content dissemination and the optimal locations of data kiosks in a heterogeneous environment of DTNs. For the heterogeneity of an environment and the approximation of the input parameters, the accuracy of the answers is quite impressive, but it remains natural to ask what are limitations of our model and how to gain higher accuracy.

First, when the size of the problem related to the number of regions and the number of data kiosks is large, the analytical model for the optimal locations of the data kiosks faces a combinatorial explosion since it evaluates the number of infected nodes as a function of the expected spreading time for every possibility of the placements of the data kiosks. However, in practice the number of subway stations in a large city does not reach a size limitation of the problem. For example, in Paris metro system, a huge metro system, the number of stations is 300 stations [103]. Hence, the analytical model for the optimal locations of data kiosks is efficient enough to solve the problem in practice. For a problem with a huge size, this limitation can be solved by applying the two placement strategies of the data kiosks.

Second, the accuracy of the solutions depends on how well the contact rates and moving rates are captured. Hence, the solutions gain higher accuracy if we can improve the evaluation of the contact rates and the moving rates. When the problem applies to an area, we should carry out an experiment in a long time duration for collecting a rich dataset of contacts. An assumption that people move as a mobility model can be used to support our evaluation. In this case, the evaluation of the contact rate when it is aware of a moving rate can be an extension to obtain higher accuracy.

Finally, some realistic aspects could be considered for a comprehensive solution of content dissemination by using opportunistic hop-by-hop transfers between neighboring devices. For example, the problem of neighbor discovery and efficient energy use in ad-hoc mode should be taken into account for practical implementation [104]. In addition, we have extended the dissemination of news from a homogeneous environment to a more realistic environment where mobile nodes can move from one region to another region. However, the model has not considered contacts between people when they are in the train. It can be an extension to gain a fine analysis for the news dissemination in an urban area.

4.6 Summary

In a heterogeneous environment with DTNs, for a given number of data kiosks, the placement of a data kiosk significantly influences the performance of content dissemination as Corollary 3 demonstrates. In this chapter we have described a solution to the problem of finding the locations of data kiosks with respect to their association with a subway map in order to optimize the expected spreading time. The key challenge with the optimization problem is mainly due to DTN characteristics where an end-to-end path between a source and a destination for communication does not exist most of the time. Our solutions, as embodied in the analytical model and the two trends, provide both an exact solution and general trends for locations of data kiosks with respect to their association with a subway map. The results show that the optimal locations of data kiosks are influenced not only by the conditions of a region but also by the number of mobile users that will receive the contents. The solutions are flexible due to the fact that the subway map is not fixed or we can change both a number of subway stations and links between them. In addition, they can easily be applied to new contexts where the means of content dissemination can be different such as vehicles in VANETs, by evaluating contact rates in the contexts. The solutions are also extensible to more complicated scenarios such as a scenario in which an infected node does not keep content forever, by adding appropriate parameters. Our simulation results on both a number of mobility models and datasets demonstrate that the analytical model and the two trends gain accuracy in the prediction of the content dissemination and the locations of data kiosks in the system.

Chapter 5

Conclusion

The widespread deployment of handheld devices with wireless capacity provides the opportunity to use opportunistic contacts between them for content distribution instead of being charged for cellular access. Due to the explosion of data usage in cellular networks, the use of delay tolerant networking for content distribution becomes more important. An application cannot be deployed on a network without performance objectives. Hence, it is important to understand and improve the performance of content distribution in DTNs to meet a certain time requirement. This dissertation has contributed to both the understanding of the performance of content distribution in DTNs and the development of a content distribution system with a performance objective in DTNs. In this chapter we conclude the dissertation by summarizing the key contributions and discussing directions for future research.

5.1 Summary of Contributions

The central goal in our work is to model and analyze DTN supports for content dissemination in an urban area within a certain time constraint. This dissertation makes two major contributions.

The first contribution is that we have formally modeled the content dissemination in DTNs with and without the presence of data kiosks. In these models, we introduce a performance metric for the content dissemination, namely, the spreading time. The metric was considered to evaluate the delivery delay of content to all subscribers. An important consequence of the models is that it suggests approaches to model data dissemination in other contexts for Delay-Disruption Tolerant Networks, including Vehicular Ad-hoc Networks (VANET) and Underwater Acoustic Sensor Networks.

The second major contribution of the dissertation is the evaluation and the improvement of the performance of content distribution in infrastructure-independent DTNs and infrastructure-dependent DTNs, where we contribute the analysis of content distribution with an acceptance probability and the explicit solutions for the number and the locations of data kiosks with an expected spreading time constraint. The key innovations of the solutions are:

- In the context of DTNs without the presence of data kiosks, we brought the necessity of considering the acceptance of mobile users for carrying news to the performance evaluation. The key idea underlying our solution is to model the number of copies in the network as an absorbing finite state Markov chain while considering a probability of acceptance for carrying content at each contact. We investigated both closed-form and asymptotic expressions for evaluating the expected message delay. The closed-form expression explicitly provides a numerical evaluation of the expected message delay. The asymptotic expression suggests the impacts of contact rates and the probability of interest on the expected message delay. Specifically, improving the interest for carrying content is rewarding if the number of mobile nodes is large, while improving contact rates provides better performance if the number of nodes is small. We also show a relationship between the expected message delay and the expected spreading time when all mobile nodes always are interested in the content. Furthermore, we provide the closed-form expression for analyzing the distribution of the number of copies of the content at the time at which the content is delivered to its destination. The proposed model is easily adapted to other contexts of DTNs such as VANETs.

- If the performance cannot meet an expected spreading time requirement, we consider a model where a number of data kiosks are added in order to meet an objective for the expected spreading time. Data kiosks are simple devices that receive content directly from the source, usually using wired or cellular networks. We provided the expressions for evaluating bounds on the number of data kiosks. This result is significant because virtually all previous works have not supplied an explicit evaluation of the number of data kiosks with a constraint of the expected spreading time while this question is very important in practice. For a given constraint of the expected spreading time, we show an important property that the bounds of the number of data kiosks scale linearly with the contact rates between a mobile node and a data kiosk. Again, this solution is highly adaptable to other contexts of DTNs. For example, we can use the solution to evaluate the number of data kiosks in VANETs where the application is to provide traffic information from data kiosks to drivers by using opportunistic contacts between a car and a data kiosk and between cars.

- Finally, in a heterogeneous environment, we developed a solution for the locations of data kiosks that optimizes the expected spreading time. The locations of data kiosks play an important part in the performance of content dissemination for a given number of data kiosks. The idea underlying the solution is to model the growth of the number of infected nodes by using an ODE model. Our solution provides the explicit steps, which are modeling the dissemination of content for all possible candidates of the placement of data kiosks, solving the system to get the number of the infected nodes as a function of the expected spreading time and selecting the optimal solution, to obtain the optimal locations of data kiosks. In addition, we reveal the two decision guidelines on choosing the locations of data kiosks. The first optimization strategy is that we should place a data kiosk in a region whose parameters are rather weak (e.g. its contact rate is low) if we want to minimize the spreading time required to disseminate the content to all subscribers. The second strategy recommends placing a data kiosk in a region whose parameters are good (e.g. its contact rate is high) if the objective is to minimize the spreading time needed to distribute the content to a small set of subscribers. These results are significant because virtually all previous works have not considered the locations of data kiosks that optimize the expected spreading time for content dissemination in a large area composed of several regions.

5.2 Future Research Directions

We have made significant contributions in understanding the performance of content distribution in DTNs, and finding the number and the locations of data kiosks that optimize the performance of content distribution in DTNs, but substantial work remains towards the goal of achieving a comprehensive performance model for content dissemination in DTNs. Future research work on topics covered by this dissertation can proceed along the following directions: (1) Verifying the flexibility of the solutions in other DTN contexts; (2) Appending a subscription model; (3) Considering the usefulness of the contacts; (4) Adding spatial dimensions; (5) Analyzing other important performance measures.

5.2.1 Verifying the Flexibility of the Solutions in Other DTN Contexts

Dissemination of traffic information by using opportunistic contacts between cars would be a potential application. In such contexts, the problem of deciding the number and the locations of data kiosks arises when some data kiosks may be necessary to distribute effectively the traffic information of interest to drivers. Our solutions for the number and the locations of data kiosks consider a DTN context where mobile nodes are wireless devices carried by human. However, the solutions would be applicable to other contexts of DTNs. It will be an interesting direction to examine how flexible the model can be when we apply the solutions to other contexts.

5.2.2 Appending a Subscription Model

An extension that we should consider is to analyze the performance of content distribution when considering a subscription model in a DTN context. Such models are based on an idea of loosely coupled communication between senders and receivers. Content that is published by senders without the need to know receivers is delivered to receivers who express their interest in the content [105]. Most popular subscription models that have been implemented in the Internet are Scribe, Gryphon, Siena, and Hermes [106–109]. Such communication models have been also presented for mobile networks [110–112]. While publish-subscribe communication is considered as a communication service that is effective and scalable for large scale communication systems, how effective is this model when applied to DTNs remains open when mobile users are not always willing to store-and-forward the content.

5.2.3 Considering the Usefulness of the Contacts

Our solutions deal with the dissemination of news; hence, we can assume that the content fits in a message and a transmission is always successful in each contact. However, when dealing with the dissemination of a video clip, we cannot make such assumptions or we have to deal with the usefulness of contacts. When considering the message size and the short contact time, we usually divide the message into several fragments to increase the probability of a successful transmission for a contact. However, too many fragments can decrease the performance of the system. In other words, it takes more time to carry all of them to the destination. Hence, it is important to quantify the number of fragments that a message should be split so that the spreading time is optimized.

5.2.4 Adding Spatial Dimensions

Our solutions consider the evolution of content in the temporal dimension; however, in some cases, it is important to add the spatial dimension. For example, when one is interested in the impacts of physical constraints such as interference, the number of mobile nodes in a location at a given instant has to be considered. Another challenging example is to develop a solution for the number and the locations of data kiosks when taking into account a spatial model for content distribution in DTNs. Point processes and stochastic geometry would be potential tools in such cases [88,113–115]. Despite the importance of this problem in practice, it has not been addressed in the literature.

5.2.5 Analyzing Other Important Performance Measures

We have evaluated the performance of content dissemination in some contexts of DTNs in terms of the expected message delay and the expected spreading time. Although the expected message delay and the expected spreading time are important performance measures for building applications, other metrics such as the message delivery ratio and buffer occupancy are especially important when the number of different content items is very large. This problem is related to forwarding, replication strategies and resource constraints. Finding the number and the locations of data kiosks with constraints of such performance measures is a challenge due to the expansion of the input variable space.

Appendix A

Proofs

A.1 Proof of Lemma 1

Using equation 6.3.16 of the digamma function in [85], we have

$$\psi(x+1) = -\gamma + \sum_{i=1}^{\infty}\left(\frac{1}{i} - \frac{1}{x+i}\right)(x \neq -1, -2, \ldots)$$

with γ is Euler-Mascheroni constant $\gamma \approx 0.57721$.

Since $k \geq \lambda/\mu$, we get $u = \mu k/\lambda \geq 1$. Applying the equation of the digamma function, we find $\psi(u+b) = -\gamma + \sum_{i=1}^{\infty}\left(\frac{1}{i} - \frac{1}{u+b-1+i}\right)$ and $\psi(u) = -\gamma + \sum_{i=1}^{\infty}\left(\frac{1}{i} - \frac{1}{u-1+i}\right)$. Subtracting the later equation from the former, we get

$$\psi(u+b) - \psi(u) = \sum_{i=1}^{b}\frac{1}{(u-1)+i} = \sum_{i=0}^{b-1}\frac{1}{u+i}. \tag{A.1}$$

From Proposition 3, we have $d = \frac{1}{\lambda u + \lambda m}\left(\sum_{i=0}^{b-1}\frac{1}{u+i} + \varsigma\right)$ where $\varsigma = \sum_{i=0}^{b-1}\frac{1}{m-i}$, $u = \frac{\mu}{\lambda}k$. Substituting (A.1) into the equation, we get (3.6) which proves Lemma 1.

A.2 Proof of Lemma 2

From Theorem 2.1 in [116], the digamma function satisfies the following inequalities:

$$\psi(x) \leq \ln\left(x + e^{-\gamma} - 1\right) \qquad \text{for } x \geq 1,$$

$$\psi(x) > \ln(x - 0.5) \qquad \text{for } x > 0.5,$$

where γ is Euler-Mascheroni constant $\gamma \approx 0.57721$ and e is Euler's number $e \approx 2.71828$.

Since $u \geq 1$ and $u+b \geq 1$, we apply the above inequalities of the digamma function:

$$\ln(u+b-0.5) < \psi(u+b) \leq \ln\left(u+b+e^{-\gamma}-1\right)$$
$$-\ln\left(u+e^{-\gamma}-1\right) \leq -\psi(u) < -\ln(u-0.5)$$

Adding the respective sides of the two inequalities, we obtain

$$\psi(u+b)-\psi(u) > \ln\left(\frac{u+b-0.5}{u+e^{-\gamma}-1}\right)$$
$$\psi(u+b)-\psi(u) < \ln\left(\frac{u+b+e^{-\gamma}-1}{u-0.5}\right).$$

Substitution of $b_\gamma = b+e^{-\gamma}-1.5$, $\tilde{b}_\gamma = b-e^{-\gamma}$ into the inequalities generates (3.7) which proves Lemma 2.

A.3 Proof of Lemma 3

From (3.8), we have

$$e^{\lambda dv+c} < 1+\frac{b_\gamma+1}{v} \tag{A.2}$$

with $v = u-0.5, c = 0.5\lambda d + m\lambda d - \varsigma$.

Expanding $f(v) = e^{\lambda dv+c}$ as a Maclaurin series, we find

$$f(v) = e^c + \lambda de^c v + \frac{(\lambda d)^2 e^c}{2!}v^2 + \frac{(\lambda d)^3 e^c}{3!}v^3 + ...$$

Since $u \geq 1$, $v = u-0.5$, and $\lambda d > 0$, all terms of $f(v)$ are positive. Thus, $e^{\lambda dv+c} > e^c + \lambda de^c v$. From this inequality and (A.2), we find

$$\lambda de^c v^2 + (e^c-1)v - (b_\gamma+1) < 0 \quad \text{with } v \geq 0.5. \tag{A.3}$$

We first find where $\lambda de^c v^2 + (e^c-1)v - (b_\gamma+1)$ is equal to zero. Since it is a quadratic equation and $-\lambda d(b_\gamma+1) < 0$, the roots of the equation is

$$v_{2,1} = \frac{e^{-c}-1 \pm \sqrt{(e^{-c}-1)^2 + 4\lambda de^{-c}(b_\gamma+1)}}{2\lambda d}.$$

Since $b \geq 2, b_\gamma = b+e^{-\gamma}-1.5$, and $\lambda d > 0$, we find $v_1 < 0$. Hence, the solution of the quadratic inequality (A.3) is $\{v \in \mathbf{R} | v \geq 0.5, v \leq v_2\}$. Substitution of $H = e^{-(0.5+m)\lambda d-\varsigma} - 1 = e^{-c}-1$ and $v = u-0.5$, we obtain a solution set of the inequality (3.8) which proves the Lemma 3.

A.4 Proof of Lemma 4

Since $\left(\tilde{b}_\gamma + 0.5\right) / \left(u + e^{-\gamma} - 1\right) > 0$, we use logarithm inequalities in [117] to find

$$\ln\left(1 + \frac{\tilde{b}_\gamma + 0.5}{u + e^{-\gamma} - 1}\right) > \frac{\frac{\tilde{b}_\gamma + 0.5}{u + e^{-\gamma} - 1}}{1 + \frac{\tilde{b}_\gamma + 0.5}{2(u + e^{-\gamma} - 1)}} = \frac{2\tilde{b}_\gamma + 1}{2u + b_\gamma}.$$

From the inequality and (3.9), we obtain $\lambda du + \lambda dm - \varsigma > \left(2\tilde{b}_\gamma + 1\right) / \left(2u + b_\gamma\right)$. Since $b \geq 2, b_\gamma = b + e^{-\gamma} - 1.5$, and $u \geq 1$, $2u + b_\gamma$ is positive. Multiplying both sides of the inequality by a positive number $2u + b_\gamma$, we obtain a quadratic inequality of

$$Au^2 + (B + C)\,u + D > 0 \tag{A.4}$$

where $A = 2\lambda d$, $B = \lambda b_\gamma d$, $C = 2\left(\lambda dm - \varsigma\right)$, and $D = b_\gamma\left(\lambda dm - \varsigma\right) - \left(2\tilde{b}_\gamma + 1\right)$. First, replace the inequality symbol with an equal symbol and find the solution of the quadratic equation $Au^2 + (B + C)\,u + D = 0$

$$u_{4,3} = \frac{-\left[B + C\right] \pm \sqrt{\left[B - C\right]^2 + 8\lambda d\left(2\tilde{b}_\gamma + 1\right)}}{4\lambda d}.$$

Substitution of $B + C = \lambda d\left(b_\gamma + 2m\right) - 2\varsigma$ and $B - C = \lambda d\left(b_\gamma - 2m\right) + 2\varsigma$ into the expression of u_4, we get the expression of u_4 represented by (3.10).

Due to $B = \lambda d b_\gamma > 0$, we find that u_3 is always negative. Combining $u_3 \leq 0$ and the condition $u \geq 1$, we find the solution set of the quadratic inequality (A.4) that is $\{u \in \mathbf{R} | u \geq 1, u \geq u_4\}$ which proves Lemma 4.

Annexe B

Résumé de la thèse

The Université Pierre et Marie Curie (Paris VI) requires all PhD theses written in English to include a summary in French. The next pages correspond to this summary.

B.1 Introduction

Cette thèse propose une analyse de performance et une modélisation de la distribution de contenu dans les réseaux tolérants aux délais (DTN) en mettant l'accent sur une solution qui permet de satisfaire le temps de propagation des contenus pris comme objectif de performance. Un exemple d'application est la distribution de la version électronique d'un journal dans une zone urbaine en tirant parti des contacts opportunistes entre utilisateurs. Bien que les contraintes temporelles ne s'appliquent pas de manière stricte dans ce contexte applicatif, la diffusion d'information doit cependant être réalisée dans un délai raisonnable.

Nous commençons ce résumé en introduisant les domaines applicatifs, les défis soulevés par la diffusion de contenu dans le contexte des DTN, les objectifs de cette thèse et l'ensemble des solutions proposées. Nous concluons ce résumé en listant les contributions de cette thèse.

B.1.1 Domaines d'application

Les réseaux tolérants aux délais ou DTN sont des architectures proposées dans le contexte des réseaux où la connectivité entre nœuds n'est pas garantie, pouvant entraîner l'absence de chemins de bout en bout de façon permanente [1,2]. Des exemples de réseaux DTN sont les réseaux interplanétaire où les communications spatiales font face à un délai très élevé ou les réseaux mobiles terrestres dont certains peuvent se retrouver partitionnés de manière inopinée en raison de la mobilité des nœuds. Depuis l'apparition du concept des réseaux DTN, une très large communauté de recherche s'est impliquée dans l'étude de ces réseaux.

Une des premières applications des DTN a été la livraison de fichiers dans le contexte des réseaux interplanétaires [3]. Ces travaux s'inscrivent dans le contexte des futures missions d'exploration de Mars. L'objectif est d'établir une communication de données entre la Terre et la Mars, planète autour de laquelle des orbiteurs seront déployés pour agir comme des relais de communication. La pile TCP/IP [4] dont le fonctionnement s'avère efficace dans le contexte de l'Internet, pose problème dans des environnements où des délais longs caractérisent les échanges entre nœuds. Les raisons qui empêchent TCP de fonctionner dans le contexte de communications entre la Terre et Mars sont liées à l'utilisation d'une signalisation de bout en bout basée sur le retour d'information des récepteurs et aux mécanismes d'adaptation mis en œuvre pour s'adapter aux changements d'état du réseau. En particulier, une connexion TCP est déconnectée en l'absence de données échangées durant une période de temps représentée par la valeur d'un temporisateur spécifiée par le protocole TCP. Le protocole TCP n'est donc pas adapté à des environnements où les délais sont élevés. Bien que le temporisateur d'inactivité soit paramétrable en prenant une valeur reflétant le délai attendu, le fonctionnement de TCP dans un environnement caractérisé par des délais élevés reste problématique. Plus précisément, lorsqu'un récepteur détecte un événement de certains segments perdu, il signale l'événement à l'expéditeur. À la réponse de ce signal, l'expéditeur a coupé de moitié son taux. Quoique cette caractéristique de TCP fonctionne bien dans l'Internet, il interdit la haute utilisation de la capacité de liaison dans les réseaux interplanétaire où le délai de propagation est extrêmement élevé. L'architecture de DTN a surgi dans de tels contextes.

Avec l'utilisation massive de terminaux mobiles dans des réseaux tels que les réseaux GSM où le trafic de données est en perpétuelle augmentation, le paradigme des DTN a été appliqué à la distribution de contenu dans les réseaux mobiles ad-hoc. L'objectif est de tirer partie des capacités de connectivité des terminaux mobiles qui accompagnent les utilisateurs dans leur déplacement pour échanger des données lors des contacts opportunistes entre eux. De tels échanges présentent de nombreux avantages aussi pour les utilisateurs que pour les fournisseurs de services. En utilisant des connexions directes entre terminaux, les utilisateurs peuvent recevoir des données sans avoir à payer les coûts d'accès au réseau cellulaire. Les fournisseurs de services de leur côté peuvent augmenter leur nombre d'abonnés sans que cette augmentation n'induise de coût supplémentaire. De nombreuses applications potentielles ont été proposées dans ce contexte [12–16]. Ces applications exploitent les contacts opportunistes entre utilisateurs pour diffuser des contenus. La question principale soulevée par ces applications concerne les performances des DTN.

Ces dernières années, des journaux gratuits ont vu le jour dans de nombreux pays et sont rapidement devenu les quotidiens les plus lus. Leurs canaux de distribution utilisent des points de présence (ou kiosques) souvent placés aux abords de stations de métro ou des gares de banlieue. Ils sont également disponibles au format électronique sur des sites web où les utilisateurs peuvent participer et contribuer à la manière des réseaux sociaux. Un autre moyen de distribution envisagée exploite les capacités de connectivité des terminaux mobiles dont le nombre ne cesse d'augmenter. Un utilisateur en contact avec un autre peut en profiter pour recevoir un contenu sans que l'accès à un réseau cellulaire ne lui soit facturé. Dans le cas des journaux électroniques, il n'est pas nécessaire que l'information contenue soit transmise instantanément dès sa création. Toutefois, il est attendu que l'information soit reçue par ses lecteurs entre le moment où ces derniers quittent leur domicile et l'instant où ils arrivent au bureau. C'est durant cette période que les utilisateurs sont les plus réceptifs et que leur attention peut être capturée. L'objectif de cette thèse est de comprendre et d'améliorer les performances d'un tel système en étudiant l'apport supplémentaire qui pourra résulter du déploiement de kiosques de données. La contribution de la thèse ne se limite pas à la distribution de journaux électroniques. Cette application donne un scénario d'utilisation pratique à nos travaux, au même titre que d'autres applications caractérisées par des attentes similaires.

Les domaines d'application des DTN sont multiples. La croissance explosive des données mobiles a motivé l'utilisation du paradigme des DTN au contexte de la diffusion de contenu dans les réseaux mobiles ad-hoc. Parmi les défis soulevés par le déploiement d'applications dans les réseaux DTN, celui de l'efficacité des mécanismes mises en œuvre dans un DTN pour supporter la distribution de contenu. De plus, il est important de déterminer les choix de conception nécessaires à la définition d'un système de diffusion de contenu qui permette d'atteindre un objectif de performance en se plaçant dans le contexte d'un réseau DTN. La section suivante présente l'importance des défis soulevés par la conception d'un tel système.

B.1.2 Défis de la diffusion de contenu dans un DTN

Le réseautage tolérants aux délais demeure un problème très difficile, malgré son potentiel dans plusieurs contextes d'applications qui vont des réseaux de communications avec l'espace lointain aux réseaux ad-hoc de véhicules et plus loin aux réseaux sous-marin [3,8,17]. La diffusion de contenu dans un DTN règle le problème de la rencontre entre le demande de quelques services d'application et la performance que le système peut atteindre. Ce problème est difficile pour plusieurs raisons fondamentales :

- Contrairement aux réseaux traditionnels où les connexions sont continues, la connexion hop-by-hop entre des nœuds DTN est opportuniste et la connexion de bout en bout n'est pas disponible la plupart du temps. Les nœuds mobiles utilisent un paradigme store-carry-and-forward pour transmettre le contenu aux destinations. Cette caractéristique de DTN fait la distribution de contenu avec bonne performance d'être plus difficile.

- En DTN, la topologie de réseau est très dynamique en raison de la forte mobilité des nœuds DTN. Une arête du réseau est établie lorsque il y a un contact entre les nœuds mobiles. Lorsque les nœuds mobiles sont des appareils sans fil portée par des humains, il est nécessaire que nous comprenions les contacts des personnes pour analyser la diffusion de contenu. En plus, l'utilisateur de téléphone mobile n'est pas toujours prêt à conserver et transporter tout contenu pour des autres utilisateurs. Par ces raisons, la compréhension de la performance de la diffusion de contenu dans un DTN est difficile.

- La diffusion de contenu dépend beaucoup des variables telles que les portées de transmission des nœuds mobiles, le mouvement des nœuds mobiles, le stockage limitée, l'énergie limitée, le nombre de sources, le nombre de destination, le nombre de nœuds mobiles, l'emplacements des sources, et la taille du contenu. D'une part, nous devons faire des hypothèses assez simplificatrices qui nous permettent de contrôler un modèle mathématique. D'un autre côté, cependant, les hypothèses seraient applicables à la pratique. En raison de la diversité de l'information entrée, l'analyse des facteurs qui ont une influence principale sur la performance du réseau et la modélisation de ces facteurs sont difficiles.

- L'hétérogénéité des variables telles que les contacts fait des affaires plus complexes. Intégrer le comportement des nœuds individuels dans un modèle analytique pour la diffusion de contenu dans un DTN est presque impossible. Comment caractériser approximativement l'hétérogénéité dans des contextes appropriés est difficile.

- Dans un DTN où les nœuds mobiles sont des appareils sans fil portée par des humains, les gens peuvent marcher dans une petite région et utiliser souvent des moyens de transport pour se déplacer entre des régions. Alors que la diffusion de contenu dépend fortement de leur mouvements, le paramétrage de ces facteurs est difficile.

Comprendre la performance de la diffusion de contenu dans un DTN, et trouver le nombre et le emplacement des nœuds d'infrastructures pour optimiser les performances de la diffusion de contenu dans un DTN sont importants pour faire la rencontre entre la demande des services d'application et les performances de DTN. Les problèmes restent difficiles en raison des caractéristiques difficiles. Grâce à aucune solution satisfaisant la diffusion de contenu avec une objectif de performance, cette thèse s'attaque à les problèmes.

B.1.3 Objectifs de la thèse

L'objectif principal de cette thèse est de construire un système de diffusion de contenu, qui utilise des contacts opportunistes pour distribuer le contenu à des personnes déplaçant dans une zone urbaine, avec un objectif de bonne performance. Pour le problème, nous pouvons d'abord analyser la performance de diffusion de contenu dans un DTN lorsque le contenu est distribué uniquement par des contacts opportunistes de quelques nœuds mobiles. Si la performance est incapable de satisfaire à une demande, nous pouvons déployer des kiosques de données pour supporter une meilleure diffusion de contenu.

Plus précisément, nos objectifs sont les suivants :

- Augmenter une meilleure compréhension de la performance de la diffusion de contenu dans un DTN où des contacts opportunistes sont utilisés pour permettre la communication. La solution devrait nous aider à mesurer les indicateurs de performance qui nous concernent, pour comprendre les impacts de performances sur l'information d'entrée, et gagner comment améliorer les performances de la diffusion de contenu dans un DTN.

- Développer une solution pour le problème du déploiement de la diffusion de contenu avec la présence de quelques kiosques de données lorsque la solution ci-dessus suggère que certains kiosques de données sont nécessaires pour parvenir à une bonne performance. La solution devrait nous aider à décider le nombre de kiosques de données et les emplacements des kiosques des données qui permettent d'optimiser la performance du système de diffusion de contenu. Puisque les gens marchent souvent dans une petite région et utilisent des moyens de transport pour se déplacer d'une région à une autre région, la solution a besoin de prendre en compte le mouvement des nœuds mobiles entre des petites régions.

- Développer les solutions qui vont obtenir la précision, l'effectivité et l'applicabilité. Pour la précision des solutions, des résultats d'analyse doivent être validées sur des données synthétiques et réalistes. Pour l'effectivité des solutions, des informations entrant au modèle d'analyse doivent être faciles pour évaluer, et les solutions doivent être obtenues au temps fini. Pour l'applicabilité des solutions, nous pouvons calculer la solution numérique, et estimer les paramètres entrant au modèle analytique à partir de certains situations pratiques.

B.1.4 Aperçu des solutions

Dans cette thèse, nous développons des solutions analytiques qui vont obtenir la précision, l'effectivité et l'applicabilité pour comprendre et améliorer la performance de distribution

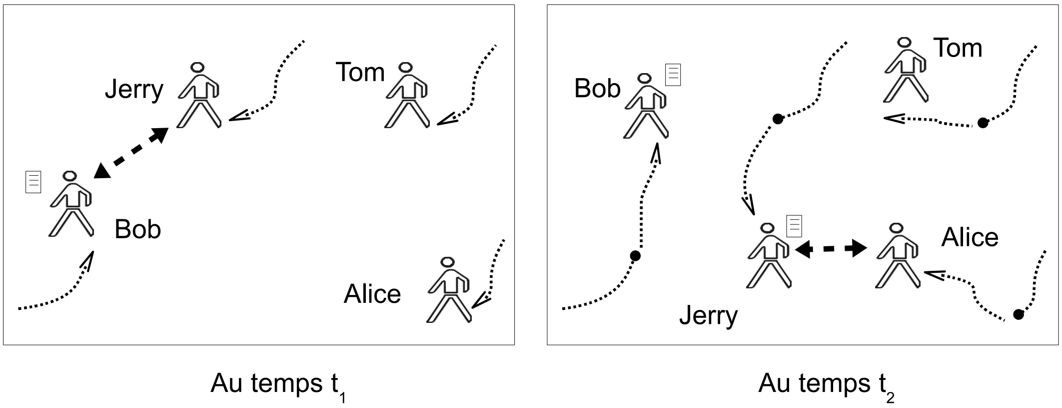


FIGURE B.1 – Diffusion de contenu dans un DTN sans la présence des kiosques de données.

de contenu dans un DTN avec l'existence et l'inexistence de quelques kiosques de données. Cette section résume nos solutions pour ce problème ci-dessus en trois étapes principales.

B.1.4.1 Evaluation et amélioration des performances dans un DTN

D'abord, nous considérons le problème de l'évaluation de performance de distribution de contenu sur un DTN dans une situation où aucune kiosque de données existe puisque un fournisseur de services souhaite savoir l'addition de quelques kiosques de données est nécessaire ou non. Dans un contexte DTN que nous considérons, les nœuds mobiles sont des appareils sans fil portés par des humains et ces personnes utilisent un paradigme store-carry-and-forward pour transmettre le contenu à la destination. En raison du fait que les utilisateurs mobiles ne sont pas toujours contents de porter les informations qui ne leur intéressent pas, il est important de considérer l'intérêt / l'acceptation des nœuds mobiles pour transporter un morceau de contenu donné dans l'évaluation de la performance de distribution de contenu sur un DTN.

La Fig. B.1 décrit un scénario de diffusion de contenu que nous considérons dans des infrastructures indépendantes DTN. Supposons que Bob veut envoyer un seul message à Alice au temps t_0, et que Alice, Bob, Tom et Jerry se déplacent dans la région. Au temps $t_1 > t_0$, Bob rencontre Jerry. Jerry accepte le message avec une probabilité. Supposons que Jerry accepte le message et le stocke dans son mobile. Au temps $t_2 > t_1$, Alice traverse Jerry. En conséquence, Alice reçoit le message de Jerry. Dans ce scénario de diffusion de contenu, le temps nécessaire pour transmettre le message de Bob à Alice, que nous appelons le délai de message qui est une métrique commune pour mesurer les réseaux, est $t_2 - t_0$.

Une autre mesure de métrique importante de la performance concernant à la consommation des ressources est le nombre de copies du message pour le transmettre à sa destination. Le nombre de copies est le nombre total du message qui est reproduit à des noeuds mobiles, y compris la source et l'exclusion de la destination, au moment à laquelle le message est livré à sa destination. Ainsi, quand Alice reçoit le message, le nombre de copies du message est égale à 2 qui embrasse une copie dans le mobile de Bob et une copie dans le mobile de Jerry. Dans cet exemple, Bob est la source, Alice est la destination, et Jerry agit comme un relais qui stocke, porte et transmet le message à Alice.

Nous employons une chaîne de Markov pour modéliser la diffusion d'un morceau de contenu donné d'une source à une destination dans un système homogène qui est considéré comme un réseau où des taux de contact sont identiques. Les taux de contact hétérogène peuvent être approchées par la variabilité lente des taux de contact ou des taux de contact homogène à l'égard d'une valeur moyenne d'un métrique de performance [56]. Cette hypothèse sera détendue dans certains aspects quand nous considérons le problème d'emplacement des nœuds de d'infrastructure. Une avance importante de notre évaluation est que, avec notre modèle, nous prenons en compte la probabilité de l'événement qu'un nœud mobile veut recevoir un contenu donné à chaque contact. Nous étudions à la fois les expressions de forme fermée et les expressions asymptotiques du délai moyen de message et le nombre de copies du message au temps où le message est transmis à sa destination. En utilisant les expressions obtenues, nous analysons le rôle de l'intérêt et des taux de contact, et déterminons celle qui est significative pour améliorer le délai moyen de message dans un tel réseau pour les cas de basse et haute densité des nœuds mobiles. Les résultats d'analyse ont été validés par des simulations sous un certain nombre de modèles de mobilité et de données réelles.

B.1.4.2 Nombre de Kiosques de Données Nécessaires pour Satisfaire un Objectif de Performance

La solution de la première étape fournit une analyse quantitative pour enquêter le délai de message de DTN. Elle a répondu à la question que comment un DTN peut être effectif sans l'infrastructure. Si le délai de message trouvé est excessif, nous étendons le problème de la distribution de contenu à une situation où certains kiosques de données existent. Les kiosques de données sont des dispositifs simples qui reçoivent directement le contenu de la source, en utilisant souvent des réseaux filaires ou des réseaux cellulaires. Dans cette

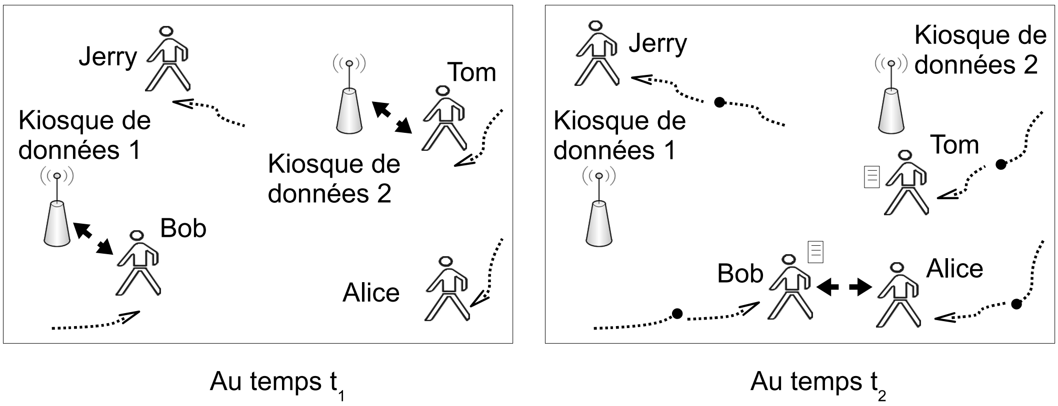


FIGURE B.2 – Diffusion de contenu dans un DTN avec la présence des kiosques de données.

situation, la diffusion des nouvelles utilise des contacts opportunistes entre un nœud mobile et un kiosque de données, et entre les nœuds mobiles. Nous sommes surtout intéressés par la diffusion de contenus des kiosques de données aux nœuds mobiles. Ainsi, la métrique de performance que nous considérons est le temps de propagation qui est le délai nécessaire pour transmettre le contenu au dernier nœud dans un groupe de nœuds, ou le temps nécessaire pour diffuser le contenu sur une partie du réseau. Le temps de propagation serait examinée à remplir l'intérêt de tous les abonnés. En raison du fait que l'objectif du temps de propagation pour la diffusion de nouvelles n'est souvent pas longtemps, nous supposons que les nœuds mobiles reçoivent un peu d'influence par un système d'incitation, ou nœuds mobiles ne changent pas leur intérêt dans le temps considéré. La question importante dans la situation, c'est que, étant donné une région et un motif de mobilité, combien kiosques de données on doit investir pour satisfaire une contrainte du temps moyen de propagation.

La Fig. B.2 décrit un scénario de diffusion de contenu que nous considérons dans un DTN avec la présence de quelques kiosques de données. Supposons que un contenu est mis à certains kiosques de données au temps initial t_0. Alice, Bob, Tom et Jerry se déplacent dans la région. Au temps $t_1 > t_0$, Bob passe le kiosque de données 1 et Tom passe le kiosque de données 2. Ainsi, Bob et Tom reçoivent le contenu à partir des kiosques de données. Au temps $t_2 > t_1$, Alice rencontre Bob, Alice prend donc le contenu à partir de Bob. Dans le scénario, les kiosques de données recevant directement le contenu à partir de quelques fournisseurs de contenu diffusent le contenu à des nœuds mobiles en utilisant des contacts opportunistes. Le temps de propagation nécessaire pour diffuser du contenu à trois noeuds est $t_2 - t_0$. Supposons que nous souhaitons diffuser le contenu à trois nœuds pendant une période de temps d. Si $d = t_2 - t_0$, un déploiement d'un kiosque de données dans la région n'est pas suffisant pour satisfaire le temps demandé, cependant un déploiement de trois kiosques de données n'est pas un investissement optimal. Ainsi, une question importante est d'analyser le nombre de kiosques données nécessaires pour satisfaire une contrainte du temps moyen de propagation.

Afin de modéliser la diffusion de contenu dans un DTN avec la présence de quelques kiosques de données, nous représentons le nombre de copies d'un morceau de contenu donné dans le réseau comme une chaîne de Markov absorbante à espace d'états finis.Nous avons obtenu l'expression de forme fermée du temps moyen de propagation qui est représenté sous la fonction digamma, et montré une relation entre le temps moyen de propagation et le délai moyen de message dans un scénario d'un système homogène. La principale avancée est que nous avons développé une analyse du nombre de kiosques données nécessaires pour réduire le temps moyen de propagation d'une demande. Les résultats montrent comment le taux de contact entre un nœud mobile et un kiosque de données influence quelques paramètres de la qualité de service. Les résultats de la simulation que nous avons simulé la diffusion de contenu dans le contexte de la présence de quelques kiosques de données sur certain modèles de mobilité et sur des données réelles confirment nos résultats analytiques.

B.1.4.3 Emplacement des Kiosques de Données pour Optimiser les Performances

Dans deux dernières étapes, nous avons développé une solution pour l'évaluation de performances de la distribution de contenu dans un DTN et une analyse d'un nombre de kiosques de données nécessaires pour satisfaire un objectif du temps moyen de propagation dans un environnement homogène. Toutefois, nous avons considéré la diffusion de contenu seulement dans un environnement homogène. Nous avons besoin d'étendre l'analyse à un environnement hétérogène, à savoir, où les valeurs différentes sont utilisées pour décrire les taux de contact entre les nœuds mobiles dans sous-zones distinctes, ainsi que les taux de contact entre un nœud mobile et un kiosque de données. Dans un environnement hétérogène, l'emplacement d'un kiosque de données influence la performance de la distribution de contenu dans le système. Ainsi, deux questions importantes dans cette situation sont : Quelle est l'efficacité de la diffusion de contenus dans un environnement hétérogène ? Où devons-nous mettre un kiosque de données pour minimiser le temps moyen de propagation ?

Nous considérons la diffusion de contenu à partir des kiosques de données à quelques

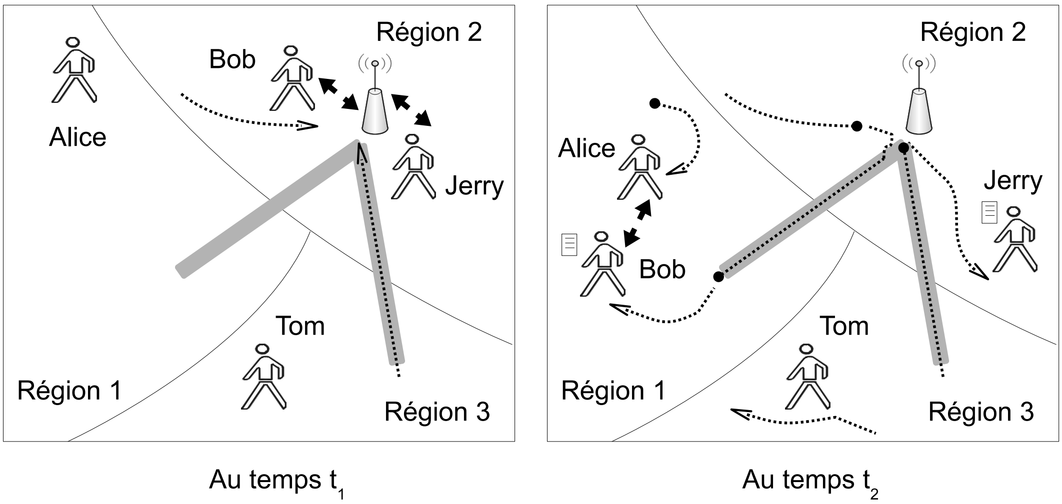


FIGURE B.3 – Diffusion de contenu dans un DTN avec la présence des kiosques de données associé à un plan de métro.

nœuds mobiles dans un système hétérogène qui est un espace urbain composé de certains régions et couvert par un système de transport. Dans le système, un utilisateur mobile se déplace librement dans une région et peut se déplacer d'une région à une autre région en utilisant une ligne de transport. Il y a une station de métro dans chaque région. Un kiosque de données sera situé près d'une station de métro, qui est inspiré par le fait que le canal de distribution des quotidiens gratuits est souvent installé à l'entrée des systèmes de transport principals. Dans une grande région, avec la rareté de kiosques de données, en sélectionnant l'emplacement des kiosques de données influencera fortement le service offert.

La Fig. B.3 décrit un scénario de diffusion de contenu que nous considérons dans un DTN qui dependent de l'infrastructures où nous prenons en compte le mouvement des personnes entre les régions en utilisant des lignes de métro. Dans la figure, les personnes portant des appareils compatibles WiFi se déplacent dans une zone composée de trois régions. Les lignes de métro connectent la région 1 avec la région 2, et la région 2 avec la région 3. Un kiosque de données est situé près de la station de métro dans la région 2. Supposons que le contenu est mis aux kiosques de données au temps initial t_0. Au temps t_0, Alice et Bob se déplacent dans la région 1 et 2 respectivement, et Tom et Jerry se déplacent dans la région 3. Bob veut se déplacer de la région 2 à la région 1, il va donc à la station de métro dans la région 2. Jerry se déplace de la région 3 à la région 2. Quand Bob et Jerry passent le kiosque de données près de la station de métro dans la région 2, ils reçoivent le contenu de kiosque de données 2 au temps t_1. Après sortir de la station de métro dans la région 1, Bob rencontre Alice au temps $t_2 > t_1$. En tant que résultats, Alice reçoit le contenu à partir de Bob. Dans ce scénario, au temps t_2, trois personnes recevront le contenu si on place le kiosque de données près de la station de métro dans la région 2. Cependant, seulement Jerry recevra le contenu si on place le kiosque de données près de la station de métro dans la région 3. Ainsi, étant donné un modèle de mobilité et un nombre kiosques de données, une bonne décision pour l'emplacement des kiosques de données est très important pour optimiser le temps moyen de propagation.

Nous utilisons un modèle ODE pour modéliser la diffusion de contenu dans un tel environnement [57]. En utilisant notre modèle analytique, nous avons développé des étapes explicites pour obtenir une solution exacte pour l'emplacement des kiosques de données qui optimise le temps moyen de propagation. De plus, nous avons révélé deux stratégies de placement généraux de l'emplacement optimal des kiosques de données qui font des solutions simplifiées pour faire face à la taille énorme du problème de placement. Ces résultats renforcent l'idée que l'emplacement optimal des kiosques de données n'est pas seulement influencée par les conditions d'une région, mais aussi par le nombre d'utilisateurs mobiles qui recevront le contenu. Les résultats de simulation que nous avons menée à la fois certains modèles de mobilité et quelques données réelles ont validé les résultats analytiques.

B.2 Contributions de la thèse

Les capacités de connectivité dont disposent les terminaux mobiles de poche très largement déployés à l'heure actuelle, offrent la possibilité de tirer parti des contacts entre terminaux pour distribuer des contenus de manière opportuniste sans avoir à payer l'accès à un réseau cellulaire. L'explosion du nombre des applications mobiles basées autour des contenus montre le fort potentiel des réseaux tolérants aux délais utilisés dans le contexte de la distribution de contenu entre terminaux mobiles. Parmi les applications qu'il est envisagé de déployer dans les réseaux DTN, la plupart ne requiert pas de transmission instantanée des messages. Il est cependant attendu que un message soit livré à sa destination dans un délai raisonnable. Il est donc important de comprendre et d'améliorer les performances de distribution des contenus dans les réseaux DTN de façon à satisfaire les contraintes temporelles demandées par de telles applications. L'objectif principal de notre travail est de modéliser et d'analyser l'apport des DTN dans la diffusion de contenu dans une zone urbaine selon des contraintes de temps connues. Cette thèse apporte deux contributions majeures.

La première contribution est que nous avons formellement modélisés la diffusion de contenu dans un DTN avec et sans la présence de quelques kiosques de données. Dans ces modèles, nous introduisons une métrique de performance qui correspond au temps de propagation des contenus. Cette métrique est prise en compte pour évaluer le délai de livraison d'un contenu à tous les abonnés. Une propriété importante de nos modèles découle du fait qu'ils ne se limitent pas au contexte des DTN. Ils peuvent être utilisés pour modéliser la diffusion de données dans d'autres contextes tels que celui des réseaux véhiculaires (VANET) ou celui des réseaux de capteurs acoustiques sous-marins.

La deuxième contribution majeure de la thèse est l'évaluation et l'amélioration de la performance de la distribution de contenu dans un DTN avec et sans la présence de quelques kiosques de données. En comparaison aux travaux existants, notre analyse de la distribution de contenu présente la particularité de prendre en compte la probabilité d'acceptation d'un contenu et de proposer des solutions explicites permettant d'évaluer le nombre des kiosques de données et de déterminer leur localisation afin de satisfaire un temps de propagation attendu. Les principales innovations de ces contributions sont :

- Dans le contexte des DTN ne comportant pas de kiosques de données, nous évaluons et analysons les performances de la distribution de contenu en fonction de la disposition des utilisateurs à accepter de véhiculer de nouveaux contenus à chaque contact. L'idée sous-jacente à notre solution est de modéliser le nombre de copies dans le réseau sous la forme d'une chaîne de Markov absorbante à espace d'états fini en considérant la probabilité qu'un nœud accepte de véhiculer le contenu à chaque contact. Nous proposons une expression de forme fermée et une expression asymptotique pour évaluer le délai moyen de message. L'expression de forme fermée fournit une évaluation quantitative du délai moyen de message. L'expression asymptotique indique l'impact du taux de contacts et de la probabilité d'intérêt sur le délai moyen de message. Plus particulièrement, l'amélioration de l'intérêt est enrichissante si le nombre de nœuds mobiles est grand, d'autre part l'améliorant les taux de contact offre une meilleure performance si le nombre de nœuds est petit. Nous montrons aussi une relation entre le délai moyen de message et le temps moyen de propagation lorsque l'ensemble des nœuds mobiles s'intéressent toujours au contenu disséminé. De plus, nous fournissons l'expression de forme fermée permettant d'analyser la distribution du nombre de copies du contenu existant dans le réseau lorsque le contenu est reçu par sa destination.

Le modèle proposé est conçu pour les DTN mais peut être facilement transposé à d'autres contextes tel que celui des VANET.

- Si les performances ne peuvent pas satisfaire à une contrainte de temps représentant le temps moyen de propagation attendu, nous considérons un modèle où des kiosques de données sont déployés afin de satisfaire à cette contrainte. Les kiosques de données sont des dispositifs simples qui reçoivent directement le contenu depuis la source via des réseaux filaires ou des réseaux cellulaires. Nous avons produit les expressions permettant d'évaluer les bornes inférieure et supérieure sur le nombre de kiosques de données. Nos résultats sont importants au regard des travaux antérieurs qui n'ont pas fourni d'évaluation explicite concernant le nombre de kiosques de données étant donné la contrainte du temps donnée par le délai moyen de propagation. Ces questions revêtent toute leur importante notamment dans le contexte d'un déploiement réel. Une propriété importante retirée de nos travaux concerne les bornes sur le nombre de kiosques de données qui varie linéairement avec le taux de contacts entre un nœud mobile et un kiosque de données. Comme pour l'ensemble de nos contributions, ces résultats concernant le déploiement de kiosques de données s'appliquent à d'autres domaines que celui des DTN. Nos résultats permettent de déterminer le nombre de kiosques de données dans un réseau véhiculaire de type VANET pour une application qui permet à des kiosques de données de fournir des informations de trafic aux conducteurs en utilisant les contacts opportunistes entre véhicules et kiosques de données ou entre véhicules.

- Enfin, nous étudions un plan de déploiement des kiosques de données adapté aux environnements hétérogènes qui optimise le temps moyen de propagation des messages. L'emplacement des kiosques de données joue un rôle important dans les performances de la diffusion de contenu dans un DTN hétérogène où le nombre de kiosques de données est limité. Pour déterminer le plan de déploiement des kiosques, nous modélisons l'accroissement du nombre de nœuds qui reçoivent un contenu en utilisant un modèle ODE. Notre solution fournit les étapes nécessaires à l'obtention du plan de déploiement optimal des kiosques. Dans les étapes, nous estimons le temps moyen de propagation de tous les candidats possibles de la mise en place de kiosques de données. Cette étude nous a permis de définir deux stratégies générales pour déterminer l'emplacement des kiosques de données. La première stratégie consiste à placer un kiosque de données dans des régions dont les propriétés ne favorisent pas l'échange efficace des données en raison par exemple d'un taux de contacts trop faibles. Cette stratégie permet minimiser le temps de propagation nécessaire pour diffuser un contenu à l'ensemble des abonnés. La seconde stratégie s'applique aux régions dont les propriétés facilitent les échanges efficaces de données en raison par exemple d'un taux de contacts élevé. Cette stratégie installe un kiosque dans de telles régions si l'objectif est de minimiser le temps moyen de propagation nécessaire pour distribuer le contenu à un sous-ensemble réduit des abonnés. Nos contributions en matière de placement des kiosques sont en rupture avec les travaux antérieurs qui n'ont pas considéré l'emplacement des kiosques de données comme moyen d'optimiser le temps moyen de propagation pour la diffusion de contenu dans une zone couvrant de large étendue et composée de plusieurs sous-régions hétérogènes.

Appendix C

List of Publications

- T.-M. Pham and S. Fdida, “DTN Support for News Dissemination in an Urban Area,” submitted to the *Computer Networks (Elsevier) Journal*, 2011.
- T.-M. Pham and S. Fdida, “DTN Support for News Dissemination in an Urban Area,” in IFIP NETWORKING 2011 (J. Domingo-Pascual, P. Manzoni, S. Palazzo, A. Pont, and C. Scoglio, eds.), vol. 6640 of *Lecture Notes in Computer Science*, pp. 120-133, Springer, May 2011.
- T.-M. Pham and S. Fdida, “Delay Estimation of a User-Preferred Content Distribution Scheme in Disruption Tolerant Networks,” in *ACM AINTEC 2009: Proceedings of the 5th Asian Internet Engineering Conference*, (Bangkok), pp. 3-10, Nov. 2009.

List of Notations

δ	The acceptance/interest probability
γ_{ij}	The moving rate of the mobile nodes from region i to region j
λ	The contact rate between a pair of mobile nodes
μ	The contact rate between a mobile node and a data kiosk
b	The target number of infected nodes
D	The message delay
d	The expected spreading time
$E[D]$	The expected message delay
k	The number of data kiosks
m	The number of subscribers
n	The number of mobile nodes
r	The transmission range

List of Figures

List of Tables

References

[1] K. Fall, "A delay-tolerant network architecture for challenged internets," in *SIGCOMM 2003: Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, (Karlsruhe), pp. 27–34, Aug. 2003.

[2] V. Cerf, S. Burleigh, A. Hooke, L. Torgerson, R. Durst, K. Scott, K. Fall, and H. Weiss, "RFC 4838, Delay-Tolerant Networking Architecture," 2007.

[3] I. Akyildiz, O. Akan, C. Chen, J. Fang, and W. Su, "The state of the art in interplanetary internet," *IEEE Communications Magazine*, vol. 42, no. 7, pp. 108 – 118, Jul. 2004.

[4] J. F. Kurose and K. W. Ross, *Computer Networking: A Top-Down Approach*. Addison Wesley, 5th ed., 2009.

[5] G. Montenegro, S. Dawkins, M. Kojo, V. Magret, and N. Vaidya, "RFC 2757, Long thin networks," 2000.

[6] I. F. Akyildiz, D. Pompili, and T. Melodia, "Challenges for efficient communication in underwater acoustic sensor networks," *ACM SIGBED Review*, vol. 1, no. 2, pp. 3–8, Jul. 2004.

[7] J. Heidemann, W. Ye, J. Wills, A. Syed, and Y. Li, "Research challenges and applications for underwater sensor networking," in *WCNC 2006: Proceedings of the 2006 IEEE Conference on Wireless Communications and Networking*, vol. 1, (Las Vegas), pp. 228 –235, Apr. 2006.

[8] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: research challenges," *Ad Hoc Networks*, vol. 3, no. 3, pp. 257 – 279, May 2005.

[9] M. Demmer and K. Fall, "Dtlsr: delay tolerant routing for developing regions," in *NSDR 2007: Proceedings of the 2007 ACM SIGCOMM workshop on Networked systems for developing regions*, (Kyoto), pp. 5:1–5:6, Aug. 2007.

[10] X. Zhang, J. Kurose, B. N. Levine, D. Towsley, and H. Zhang, "Study of a bus-based disruption-tolerant network: mobility modeling and impact on routing," in *MobiCom 2007: Proceedings of the 13th ACM International Conference on Mobile Computing and Networking*, (Montréal), pp. 195–206, Sep. 2007.

[11] A. Seth, D. Kroeker, M. Zaharia, S. Guo, and S. Keshav, "Low-cost communication for rural internet kiosks using mechanical backhaul," in *MobiCom 2006: Proceedings of the 12th ACM International Conference on Mobile Computing and Networking*, (Los Angeles), pp. 334–345, Sep. 2006.

[12] T. Islam, "Dt-talkie: Interactive voice messaging for heterogeneous groups in delay-tolerant networks," in *CCNC 2009: Proceedings of the 6th IEEE Conference on Consumer Communications and Networking*, (Las Vegas), pp. 1–2, Jan. 2009.

[13] E. Altaian, P. Nain, and J. C. Bermond, "Distributed storage management of evolving files in delay tolerant ad hoc networks," in *INFOCOM 2009: Proceedings of the 28th Conference on Computer Communications*, (Rio De Janeiro), pp. 1431–1439, Apr. 2009.

[14] A. Balasubramanian, Y. Zhou, W. B. Croft, B. N. Levine, and A. Venkataramani, "Web search from a bus," in *CHANTS 2007: Proceedings of the second ACM workshop on Challenged networks*, (Montréal), Sep. 2007.

[15] T. Hyyryläinen, T. Kärkkäinen, C. Luo, V. Jaspertas, J. Karvo, and J. Ott, "Opportunistic email distribution and access in challenged heterogeneous environments," in *CHANTS 2007: Proceedings of the second ACM workshop on Challenged networks*, (Montréal), pp. 97–100, Aug. 2007.

[16] B. Hull, V. Bychkovsky, Y. Zhang, K. Chen, M. Goraczko, A. Miu, E. Shih, H. Balakrishnan, and S. Madden, "Cartel: a distributed mobile sensor computing system,"

in *SenSys 2006: Proceedings of the 4th ACM International Conference on Embedded Networked Sensor Systems*, (Boulder), pp. 125–138, Oct. 31 - Nov. 3 2006.

[17] M. Gerla and L. Kleinrock, "Vehicular networks and the future of the mobile internet," *Computer Networks*, vol. 55, no. 2, pp. 457–469, Feb. 2011.

[18] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," tech. rep., Duke University, Jul. 2000.

[19] K. Harras, K. Almeroth, and E. Belding-Royer, "Delay tolerant mobile networks (dtmns): Controlled flooding in sparse mobile networks," in *NETWORKING 2005* (R. Boutaba, K. Almeroth, R. Puigjaner, S. Shen, and J. Black, eds.), vol. 3462 of *Lecture Notes in Computer Science*, pp. 41–91, Springer Berlin / Heidelberg, 2005.

[20] J. Burgess, B. Gallagher, D. Jensen, and B. N. Levine, "Maxprop: Routing for vehicle-based disruption-tolerant networks," in *INFOCOM 2006: Proceedings of the 25th IEEE International Conference on Computer Communications*, (Barcelona), pp. 1–11, Apr. 2006.

[21] A. Lindgren, A. Doria, and O. Schelén, "Probabilistic routing in intermittently connected networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 7, no. 3, pp. 19–20, Jul. 2003.

[22] E. M. Daly and M. Haahr, "Social network analysis for routing in disconnected delay-tolerant manets," in *MobiHoc 2007: Proceedings of the 8th ACM International Symposium on Mobile Ad hoc Networking and Computing*, (Montréal), pp. 32–40, Sep. 2007.

[23] P. Hui, J. Crowcroft, and E. Yoneki, "Bubble rap: social-based forwarding in delay tolerant networks," in *MobiHoc 2008: Proceedings of the 9th ACM International Symposium on Mobile Ad hoc Networking and Computing*, (Hong Kong), pp. 241–250, May 2008.

[24] L. McNamara, C. Mascolo, and L. Capra, "Media sharing based on colocation prediction in urban transport," in *MobiCom 2008: Proceedings of the 14th ACM International Conference on Mobile Computing and Networking*, (San Francisco), pp. 58–69, Sep. 2008.

[25] J. Leguay, T. Friedman, and V. Conan, "Dtn routing in a mobility pattern space," in *WDTN 2005: Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, (Philadelphia), pp. 276–283, Aug. 2005.

[26] B. Burns, O. Brock, and B. Levine, "Mv routing and capacity building in disruption tolerant networks," in *INFOCOM 2005: Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications*, vol. 1, (Miami), pp. 398–408, Mar. 2005.

[27] M. Musolesi, S. Hailes, and C. Mascolo, "Adaptive routing for intermittently connected mobile ad hoc networks," in *WoWMoM 2005: Proceedings of the 6th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, (Taormina), pp. 183–189, Jun. 2005.

[28] V. Erramilli, A. Chaintreau, M. Crovella, and C. Diot, "Diversity of forwarding paths in pocket switched networks," in *IMC 2007: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, (San Diego), pp. 161–174, Oct. 2007.

[29] P. Costa, C. Mascolo, M. Musolesi, and G. Picco, "Socially-aware routing for publish-subscribe in delay-tolerant mobile ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 5, pp. 748 –760, Jun. 2008.

[30] A. Mashhadi, S. Ben Mokhtar, and L. Capra, "Habit: Leveraging human mobility and social network for efficient content dissemination in delay tolerant networks," in *WoWMoM 2009: Proceedings of the 10th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, (Kos), pp. 1–6, Jun. 2009.

[31] B. K. Polat, P. Sachdeva, M. H. Ammar, and E. W. Zegura, "Message ferries as generalized dominating sets in intermittently connected mobile networks," in *MobiOpp 2010: Proceedings of the Second International Workshop on Mobile Opportunistic Networking*, (Pisa), pp. 22–31, Feb. 2010.

[32] W. Zhao, M. Ammar, and E. Zegura, "Controlling the mobility of multiple data transport ferries in a delay-tolerant network," in *INFOCOM 2005: Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications*, vol. 2, (Miami), pp. 1407–1418, Mar. 2005.

[33] K. Scott and S. Burleigh, "RFC 5050, Bundle protocol specification," 2007.

[34] S. Burleigh, M. Ramadas, and S. Farrell, “RFC 5325, Licklider transmission protocol - motivation,” 2008.

[35] L. Wood and P. Holliday, “Using HTTP for delivery in Delay-/Disruption-Tolerant Networks, Internet-Draft, draft-wood-dtnrg-http-dtn-delivery-07,” 2011.

[36] R. Groenevelt, P. Nain, and G. Koole, “The message delay in mobile ad hoc networks,” *Performance Evaluation*, vol. 62, no. 1-4, pp. 210–228, Oct. 2005.

[37] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, “Performance analysis of mobility-assisted routing,” in *MobiHoc 2006: Proceedings of the 7th ACM International Symposium on Mobile Ad hoc Networking and Computing*, (Florence), pp. 49–60, May 2006.

[38] T. Spyropoulos, K. Psounis, and C. Raghavendra, “Efficient routing in intermittently connected mobile networks: The single-copy case,” *IEEE/ACM Transactions on Networking*, vol. 16, no. 1, pp. 63 –76, Feb. 2008.

[39] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, “Efficient routing in intermittently connected mobile networks: the multiple-copy case,” *IEEE/ACM Transactions on Networking*, vol. 16, no. 1, pp. 77–90, Feb. 2008.

[40] A. Al Hanbali, P. Nain, and E. Altman, “Performance of ad hoc networks with two-hop relay routing and limited packet lifetime (extended version),” *Performance Evaluation*, vol. 65, no. 6-7, pp. 463–483, Jun. 2008.

[41] X. Zhang, G. Neglia, J. Kurose, and D. Towsley, “Performance modeling of epidemic routing,” *Computer Networks*, vol. 51, no. 10, pp. 2867–2891, Jul. 2007.

[42] S. Ioannidis, A. Chaintreau, and L. Massoulié, “Optimal and scalable distribution of content updates over a mobile social network,” in *INFOCOM 2009: Proceedings of the 28th Conference on Computer Communications*, (Rio De Janeiro), pp. 1422–1430, Apr. 2009.

[43] Y. Wang, H. Wu, and H. Dang, “Analytic study of delay/fault-tolerant mobile sensor networks (dft-msn’s),” in *WoWMoM 2009: Proceedings of the 10th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, (Kos), pp. 1–9, Jun. 2009.

[44] A. Jindal and K. Psounis, "Performance analysis of epidemic routing under contention," in *IWCMC 2006: Proceedings of the 2006 International Conference on Wireless Communications and Mobile Computing*, (Vancouver), pp. 539–544, Jul. 2006.

[45] A. Krifa, C. Baraka, and T. Spyropoulos, "Optimal buffer management policies for delay tolerant networks," in *SECON 2008: Proceedings of the 5th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, (San Francisco), pp. 260–268, Jun. 2008.

[46] K. Harras, M. Wittie, K. Almeroth, and E. Belding, "Paranets: A parallel network architecture for challenged networks," in *HotMobile 2007: Proceedings of the 8th IEEE Workshop on Mobile Computing Systems and Applications*, (Tucson), pp. 73–78, Mar. 2007.

[47] D. Goodman, J. Borras, N. Mandayam, and R. Yates, "Infostations: a new system model for data and messaging services," in *VTC 1997: Proceedings of the 47th IEEE Conference on Vehicular Technology*, vol. 2, (Phoenix), pp. 969–973, May 1997.

[48] M. Liberatore, B. N. Levine, and C. Barakat, "Maximizing transfer opportunities in bluetooth dtns," in *CoNEXT 2006: Proceedings of the 2nd International Conference on Emerging Networking Experiments and Technologies*, (Lisboa), pp. 3:1–3:11, Dec. 2006.

[49] H. Jun, M. H. Ammar, M. D. Corner, and E. W. Zegura, "Hierarchical power management in disruption tolerant networks with traffic-aware optimization," in *CHANTS 2006: Proceedings of the 2006 SIGCOMM workshop on Challenged networks*, (Pisa), pp. 245–252, Sep. 2006.

[50] R. C. Shah, S. Roy, S. Jain, and W. Brunette, "Data mules: modeling and analysis of a three-tier architecture for sparse sensor networks," *Ad Hoc Networks*, vol. 1, no. 2-3, pp. 215–233, May 2003.

[51] Z. J. Haas and T. Small, "A new networking model for biological applications of ad hoc sensor networks," *IEEE/ACM Transactions on Networking*, vol. 14, no. 1, pp. 27–40, Feb. 2006.

[52] Z. Wenrui, C. Yang, A. Mostafa, C. Mark, L. Brian, and Z. Ellen, "Capacity enhancement using throwboxes in dtns," in *MASS 2006: Proceedings of the 3rd IEEE International Conference on Mobile Adhoc and Sensor Systems*, (Vancouver), pp. 31–40, Oct. 2006.

[53] B. Burns, O. Brock, and B. Levine, "Autonomous enhancement of disruption tolerant networks," in *ICRA 2006: Proceedings 2006 IEEE International Conference on Robotics and Automation*, (Orlando), pp. 2105–2110, May 2006.

[54] B. Burns, O. Brock, and B. N. Levine, "Mora routing and capacity building in disruption-tolerant networks," *Ad Hoc Networks*, vol. 6, no. 4, pp. 600 – 620, Jun. 2008.

[55] N. Banerjee, M. D. Corner, and B. N. Levine, "Design and field experimentation of an energy-efficient architecture for dtn throwboxes," *IEEE/ACM Transactions on Networking*, vol. 18, no. 2, pp. 554–567, Apr. 2010.

[56] T. Small and Z. J. Haas, "The shared wireless infostation model: a new ad hoc networking paradigm (or where there is a whale, there is a way)," in *MobiHoc 2003: Proceedings of the 4th ACM international symposium on Mobile Ad hoc Networking and Computing*, (Annapolis), pp. 233–244, Jun. 2003.

[57] T. G. Kurtz, "Solutions of ordinary differential equations as limits of pure jump markov processes," *Journal of Application of Probability*, vol. 7, no. 1, pp. 49–58, Apr. 1970.

[58] D. Vassis, G. Kormentzas, A. Rouskas, and I. Maglogiannis, "The IEEE 802.11g standard for high data rate WLANs," *IEEE Network*, vol. 19, no. 3, pp. 21 – 26, May-Jun. 2005.

[59] J. H. Mathews and K. K. Fink, *Numerical Methods Using Matlab*. Prentice Hall, 4th ed., 2004.

[60] S. Jain, K. Fall, and R. Patra, "Routing in a delay tolerant network," *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 4, pp. 145–158, Aug. 2004.

[61] W. Zhao, M. Ammar, and E. Zegura, "Multicasting in delay tolerant networks: semantic models and routing algorithms," in *WDTN 2005: Proceedings of the 2005*

ACM SIGCOMM workshop on Delay-tolerant networking, (Philadelphia), pp. 268–275, Aug. 2005.

[62] A. Balasubramanian, B. Levine, and A. Venkataramani, "Dtn routing as a resource allocation problem," *ACM SIGCOMM Computer Communication Review*, vol. 37, no. 4, pp. 373–384, Aug. 2007.

[63] T. Matsuda and T. Takine, "(p,q)-epidemic routing for sparsely populated mobile ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 5, pp. 783 –793, Jun. 2008.

[64] T. Karagiannis, J.-Y. Le Boudec, and M. Vojnović, "Power law and exponential decay of inter contact times between mobile devices," in *MobiCom 2007: Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking*, (Montréal), pp. 183–194, Sep. 2007.

[65] V. Conan, J. Leguay, and T. Friedman, "Characterizing pairwise inter-contact patterns in delay tolerant networks," in *Autonomics 2007: Proceedings of the 1st International Conference on Autonomic Computing and Communication Systems*, (Brussels), pp. 1–9, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Sep. 2007.

[66] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot, "Pocket switched networks and human mobility in conference environments," in *WDTN 2005: Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, (Philadelphia), pp. 244–251, Aug. 2005.

[67] E. Yoneki, D. Greenfield, and J. Crowcroft, "Dynamics of inter-meeting time in human contact networks," in *ASONAM 2009: Proceedings of the 2009 International Conference on Advances in Social Network Analysis and Mining*, (Washington), pp. 356–361, IEEE Computer Society, Jul. 2009.

[68] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: an efficient routing scheme for intermittently connected mobile networks," in *WDTN 2005: Proceedings of the 2005 ACM SIGCOMM Workshop on Delay-tolerant Networking*, (Philadelphia), pp. 252–259, Aug. 2005.

[69] A. Lindgren, A. Doria, and O. Schelén, "Probabilistic routing in intermittently connected networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 7, no. 3, pp. 19–20, Jul. 2003.

[70] G. Sharma, R. Mazumdar, and N. B. Shroff, "Delay and capacity trade-offs in mobile ad hoc networks: a global perspective," *IEEE/ACM Transactions on Networking*, vol. 15, no. 5, pp. 981–992, Oct. 2007.

[71] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 477–486, Aug. 2002.

[72] A. Al Hanbali, P. Nain, and E. Altman, "Performance of ad hoc networks with two-hop relay routing and limited packet lifetime (extended version)," *Performance Evaluation*, vol. 65, no. 6-7, pp. 463–483, Jun. 2008.

[73] A. Keränen, J. Ott, and T. Kärkkäinen, "The ONE simulator for DTN protocol evaluation," in *Simutools 2009: Proceedings of the 2nd International Conference on Simulation Tools and Techniques*, (Brussels), pp. 1–10, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Mar. 2009.

[74] W. Navidi, "Stationary distributions for the random waypoint mobility model," *IEEE Transactions on Mobile Computing*, vol. 3, no. 1, pp. 99–108, Jan. 2004.

[75] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Communications and Mobile Computing*, vol. 2, no. 5, pp. 483–502, Aug. 2002.

[76] "A community resource for archiving wireless data at Dartmouth." available at `http://crawdad.cs.dartmouth.edu/`.

[77] N. Eagle, A. Pentland, and D. Lazer, "Inferring social network structure using mobile phone data," *Proceedings of the National Academy of Sciences*, vol. 106, no. 36, pp. 15274–15278, Sep. 2009.

[78] M. McNett and G. M. Voelker, "Access and mobility of wireless pda users," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 9, no. 2, pp. 40–55, Apr. 2005.

[79] T. Henderson, D. Kotz, and I. Abyzov, "The changing usage of a mature campus-wide wireless network," *Computer Networks*, vol. 52, no. 14, pp. 2690–2712, Oct. 2008.

[80] B. Liu, P. Thiran, and D. Towsley, "Capacity of a wireless ad hoc network with infrastructure," in *MobiHoc 2007: Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, (Montréal), pp. 239–246, Sep. 2007.

[81] A. Zemlianov and V. Gustavo de, "Capacity of ad hoc wireless networks with infrastructure support," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 3, pp. 657–667, Mar. 2005.

[82] J. Robinson, M. Singh, R. Swaminathan, and E. Knightly, "Deploying mesh nodes under non-uniform propagation," in *INFOCOM 2010: Proceedings of the 29th Conference on Computer Communications*, (San Diego), pp. 2142–2150, Mar. 2010.

[83] M. May, V. Lenders, G. Karlsson, and C. Wacha, "Wireless opportunistic podcasting: implementation and design tradeoffs," in *CHANTS 2007: Proceedings of the second ACM workshop on Challenged networks*, (Montréal), pp. 75–82, Sep. 2007.

[84] M. Ibrahim, A. Al Hanbali, and P. Nain, "Delay and resource analysis in MANETs in presence of throwboxes," *Performance Evaluation*, vol. 64, no. 9-12, pp. 933–947, Oct. 2007.

[85] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables.* Dover Publications, 1965.

[86] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of human mobility on opportunistic forwarding algorithms," *IEEE Transactions on Mobile Computing*, vol. 6, no. 6, pp. 606–620, Jun. 2007.

[87] S. M. Ross, *Introduction to Probability Models.* Academic Press, 9th ed., 2006.

[88] Haight, *Mathematical theories of traffic flow, Volume 7.* Academic Press, 1963.

[89] Z. Drezner, *Facility Location: A Survey of Applications and Methods.* Springer, 1st ed., 1995.

[90] D. Stamatelos and A. Ephremides, "Spectral efficiency and optimal base placement for indoor wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 4, pp. 651 –661, May 1996.

[91] M. Wright, “Optimization methods for base station placement in wireless applications,” in *VTC 1998: Proceedings of the 48th IEEE Conference on Vehicular Technology*, vol. 1, (Ottawa), pp. 387–391, May 1998.

[92] T. Fruhwirth and P. Brisset, “Placing base stations in wireless indoor communication networks,” *IEEE Intelligent Systems and their Applications*, vol. 15, no. 1, pp. 49 –53, Jan/Feb 2000.

[93] X. Huang, U. Behr, and W. Wiesbeck, “Automatic base station placement and dimensioning for mobile network planning,” in *VTC 2000: Proceedings of the 52nd IEEE Conference on Vehicular Technology*, vol. 4, (Boston), pp. 1544–1549, Sep. 2000.

[94] J. K. Han, B. S. Park, Y. S. Choi, and H. K. Park, “Genetic approach with a new representation for base station placement in mobile communications,” in *VTC 2001 Fall: Proceedings of the 54th IEEE Conference on Vehicular Technology*, vol. 4, (Atlantic City), pp. 2703–2707, Oct. 2001.

[95] Y. Hu, Y. Xue, Q. Li, F. Liu, G. Y. Keung, and B. Li, “The sink node placement and performance implication in mobile sensor networks,” *Mobile Networks and Applications*, vol. 14, no. 2, pp. 230–240, Apr. 2009.

[96] K. Akkaya and M. Younis, “Cola: A coverage and latency aware actor placement for wireless sensor and actor networks,” in *VTC 2006 Fall: Proceedings of the 64th IEEE Conference on Vehicular Technology*, (Montréal), pp. 1–5, Sep. 2006.

[97] Y. Hou, Y. Shi, H. Sherali, and S. Midkiff, “On energy provisioning and relay node placement for wireless sensor networks,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2579 – 2590, Sept. 2005.

[98] A. Cerpa and D. Estrin, “Ascent: adaptive self-configuring sensor networks topologies,” *IEEE Transactions on Mobile Computing*, vol. 3, no. 3, pp. 272 – 285, Jul.-Aug. 2004.

[99] A. Chaintreau, J.-Y. Le Boudec, and N. Ristanovic, “The age of gossip: spatial mean field regime,” in *SIGMETRICS 2009: Proceedings of the 11st International joint Conference on Measurement and Modeling of Computer Systems*, (Seattle), pp. 109–120, Jun. 2009.

[100] A. Iserles, G. Ramaswami, and M. Sofroniou, "Runge-kutta methods for quadratic ordinary differential equations," *BIT Numerical Mathematics*, vol. 38, no. 2, pp. 315–346, Jun. 1998.

[101] J. Leguay, A. Lindgren, J. Scott, T. Friedman, and J. Crowcroft, "Opportunistic content distribution in an urban setting," in *CHANTS 2006: Proceedings of the 2006 SIGCOMM Workshop on Challenged Networks*, (Pisa), pp. 205–212, Sep. 2006.

[102] S. Gaito, E. Pagani, and G. P. Rossi, "Strangers help friends to communicate in opportunistic networks," *Computer Networks*, vol. 55, no. 2, pp. 374–385, Feb. 2011.

[103] Wikipedia, "Paris métro." available at `http://en.wikipedia.org/wiki/Paris_Metro`, 2005.

[104] S. Jung, U. Lee, A. Chang, D.-K. Cho, and M. Gerla, "Bluetorrent: Cooperative content sharing for bluetooth users," *Pervasive and Mobile Computing*, vol. 3, no. 6, pp. 609 – 634, Mar. 2007.

[105] R. Baldoni, M. Contenti, and A. Virgillito, "The evolution of publish/subscribe communication systems," in *Future directions in distributed computing* (A. Schiper, A. A. Shvartsman, H. Weatherspoon, and B. Y. Zhao, eds.), pp. 137–141, Springer-Verlag, 2003.

[106] M. Castro, P. Druschel, A.-M. Kermarrec, and A. Rowstron, "Scribe: a large-scale and decentralized application-level multicast infrastructure," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 8, pp. 1489 – 1499, Oct. 2002.

[107] G. Banavar, T. Chandra, B. Mukherjee, J. Nagarajarao, R. Strom, and D. Sturman, "An efficient multicast protocol for content-based publish-subscribe systems," in *ICDCS 1999: Proceedings of the 19th IEEE International Conference on Distributed Computing Systems*, (Austin), pp. 262–272, 31 May - 04 Jun. 1999.

[108] A. Carzaniga, D. S. Rosenblum, and A. L. Wolf, "Design and evaluation of a wide-area event notification service," *ACM Transactions on Computer Systems*, vol. 19, no. 3, pp. 332–383, Aug. 2001.

[109] P. Pietzuch and J. Bacon, "Hermes: a distributed event-based middleware architecture," in *ICDCS 2002 Workshops: Proceedings of the 22nd International Conference on Distributed Computing Systems Workshops*, (Vienna), pp. 611–618, Jul. 2002.

[110] A. Emmanuelle, A. K. Datta, M. Gradinariu, and G. Simon, "Publish/subscribe scheme for mobile networks," in *POMC 2002: Proceedings of the 2nd ACM international workshop on Principles of mobile computing*, (Toulouse), pp. 74–81, Oct. 2002.

[111] Y. Huang and H. Garcia-Molina, "Publish/subscribe in a mobile environment," *Wireless Networks*, vol. 10, no. 6, pp. 643–652, Nov. 2004.

[112] G. Sollazzo, M. Musolesi, and C. Mascolo, "Taco-dtn: a time-aware content-based dissemination system for delay tolerant networks," in *MobiOpp 2007: Proceedings of the 1st international MobiSys workshop on Mobile opportunistic networking*, (Puerto Rico), pp. 83–90, Jun. 2007.

[113] D. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes, Volume 1*. Springer, 2nd ed., 2002.

[114] P. D. Stoyan, D. W. S. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications*. Wiley, 2nd ed., 1996.

[115] K. Leung, W. Massey, and W. Whitt, "Traffic models for wireless communication networks," *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 8, pp. 1353 –1364, Oct. 1994.

[116] N. Batir, "Some new inequalities for gamma and polygamma functions," *Journal of Inequalities in Pure and Applied Mathematics*, vol. 6, no. 4, no. 4, 2005.

[117] E. R. Love, "64.4 some logarithm inequalities," *The Mathematical Gazette*, vol. 64, no. 427, pp. 55–57, Mar. 1980.

Modeling and Analysis of Content Distribution in Disruption Tolerant Networks

Abstract. This dissertation studies the practicality of content distribution over a Delay Tolerant Network (DTN) in an urban area. The target application is the distribution of the electronic version of a newspaper in a large city. Although strict time constraints do not apply, spreading the information should be achieved within a reasonable delay. Two performance metrics, the spreading time and the message delay, are considered. The message delay is the delay required to transmit content from a mobile node to another node, while the spreading time is the delay needed for the content to spread over a part of the network.

Firstly, our goal is to increase our understanding of the performance of a simple DTN environment when content is distributed solely through inter-contact of mobile nodes. We contribute both the close-form expression and the asymptotic expression of the expected message delay to the literature when considering the probability of interest/acceptance for a given piece of content at each contact. The asymptotic expression provides the insights on the efficient ways for improving the expected message delay in the case of an area with low or high density of mobile nodes. We also show a relationship between the expected message delay and the expected spreading time in such environment.

Secondly, if the delay is found to be excessive, we suggest the deployment of some data kiosks in the environment to better support the dissemination of content. Data kiosks are simple devices that receive content directly from the source, usually using wired or cellular networks. A key issue when designing efficiently such network is to determine the number of data kiosks required to satisfy a performance target. We investigate both an upper bound and a lower bound of the number of data kiosks to distribute the content over a geographical area within an expected spreading time objective. We also show the important property that those bounds scale linearly with the contact rates between a mobile node and a data kiosk.

Finally, we consider the question of the optimal locations of data kiosks in a more realistic scenario where users move along a transportation system (like a subway or suburban train) that connects several regions. We provide an analysis used to decide which subway stop should host a data kiosk to optimize the spreading time. These findings support the view that the optimal locations of data kiosks are influenced not only by the conditions of a region but also by the number of mobile users that will receive the contents. Analytical results are validated by simulations under a number of mobility models and real datasets.

Keywords. Hybrid DTN, Delay, News dissemination, Modeling, Performance analysis.