공학박사학위논문

Traffic Offloading by User-to-User Opportunistic Sharing in Mobile Social Networks

모바일 소셜 네트워크에서 사용자 간 기회적인 공유 기반 트래픽 오프로딩

2013년 8월

서울대학교 대학원
전기컴퓨터공학부
왕 효 비

Traffic Offloading by User-to-User Opportunistic Sharing in Mobile Social Networks

지도교수 최 양 희

이 논문을 공학박사 학위논문으로 제출함
2013년 5월

서울대학교 대학원
전기컴퓨터공학부
왕 효 비

왕 효 비의 박사 학위논문을 인준함
2013년 6월

위 원 장	김 종 권	(인)
부위원장	최 양 희	(인)
위 원	권 태 경	(인)
위 원	최 성 현	(인)
위 원	이 병 준	(인)

Abstract

Traffic Offloading by User-to-User Opportunistic Sharing in Mobile Social Networks

Xiaofei Wang
Department of Computer Science & Engineering
The Graduate School
Seoul National University

The fast increasing traffic demand becomes a serious concern of mobile network operators. To solve this traffic explosion problem, there have been efforts to offload the traffic from cellular links to local short-range communications among mobile users that are moving around and forming mobile social networks. In my thesis, I mainly focus on the user-to-user opportunistic sharing and try to elaborate its effectiveness and efficiency for to offload mobile traffic.

In the first work, I propose the **T**raffic **O**ffloading assisted by **S**ocial network services via opportunistic **S**haring in mobile social networks, **TOSS**. In TOSS, initially a subset of mobile users are selected as initial seeds depending on their content spreading impact in online social network services (SNSs) and their mobility patterns in offline mobile social networks

(MSNs). Then users share the content via opportunistic local connectivity (e.g., Bluetooth, Wi-Fi Direct) with each other. Due to the distinct access patterns of individual SNS users, TOSS further exploits the user-dependent access delay between the content generation time and each user's access time for the purpose of traffic offloading. I model and analyze process of the traffic offloading and content spreading by taking into account various options in linking SNS and MSN data sets. The trace-driven evaluation shows that TOSS can reduce up to 86.5% of the cellular traffic while satisfying the access delay requirements of all users.

In the second work, I focus on the analytical research on **Push-Share** framework for content disseminating in mobile networks. One content is firstly pushed the to a subset of subscribers via cellular links, and mobile users spread the content via opportunistic local connectivity. I theoretically model and analyze how the content can be disseminated, where handovers are modeled based on the multi-compartment model. I also formulate the mathematical optimization framework, by which the trade-off between the dissemination delay and the energy cost is explored.

Based on the measurement study, trace-driven analysis, theoretical modeling and system optimization in above papers, the traffic offloading by user-to-user opportunistic sharing in mobile social networks is proved to be effective and efficient. Additionally, further discussions on the practical deployment, future vision, and open issues are discussed as well.

Keywords : Traffic Offloading, Opportunistic Sharing, Device-to-Device

(D2D), User-to-User (U2U), Mobile Social Networks, Online Social Networks

Student Number : 2008-30709

Contents

List of Figures

List of Algorithms

List of Tables

Chapter 1

Introduction

Due to the fast development of mobile communication technologies, more and more users tend to download content on mobile devices, for example reading articles and watching videos on phones and tablets. The ever increasing traffic load becomes a serious concern of mobile network operators (MNOs) [1], but studies [1] [2] [3] [4] point out that much of the traffic load is due to the duplicated download of the same popular files. For instance, top 10% of videos in YouTube account for nearly 80% of all the views [4]. Therefore, how to effectively reduce the duplicated download via cellular link by **offloading** the traffic via other networking connectivities becomes a hot research topic.

Recently, by adopting the concept of "peer-to-peer communication" from the BitTorrent in wired Internet into mobile environment [5], there have been many studies to exploit the people-to-people (user-to-user) opportunistic sharing during intermittent meetings of mobile users for traffic offloading in mobile social networks (MSNs), which is a special form of the Delay Tolerant Network (DTN) with more consideration of the social relationship of network users [6] [7] [8] [9] [10] [11] [12] [13]. Also a MSN/DTN can be considered based on the opportunistic network [14] [15]. In MSNs, users are able to discover the adjacent neighbors [16] and set up temporary local connectivities, e.g., Bluetooth, Wi-Fi Direct, Near-

Field-Communication (NFC) [17], and Device-to-Device (D2D) [18] [19], for sharing delay-tolerant content with each other. Especially the D2D is now under very hot discussion, since it is under detailed design in 3GPP as an underlay to LTE-Advanced networks [19], by which users can use operator authorized spectrum for direct communication without the support of infrastructure.

For such kind of user-to-user sharing, some users need to carry the content at the beginning. It is advocated that by selecting an appropriate initial set of **seeds** the peak traffic load can be reduced by 20% to 50% [11]. The study in [12] also proves that content dissemination with a small number of initial seeds can guarantee the delay requirements of all users while reducing a substantial amount of cellular traffic. However there are still several important issues in related research which are not fully elaborated, such as:

- *How to know, or how to predict the dissemination delay of each user for each content?* Recent studies [12] [13] [20] [21], assume the same dissemination deadline of the same content for all users; however users indeed have various delay requirements [22].

- *How to design the seeding strategy to minimize the cellular traffic while satisfying the delay requirements of all users?* Strategies of selecting initial seeds are discussed in prior work [10] [11] [23], but most of them focus on user mobility while ignoring the practical social relationships among users.

- *How to make mobile users share content efficiently with others?* Studies in [10] [12] [13] assume people will always exchange content gratuitously. But in reality, people mostly share information by "word-of-mouth" propagation [24] [25], and it is able to exploit social relationship among users.

To solve the above issues, I seek to exploit the relationship between the offline MSNs and online Social Network Services (SNSs). It is discovered that there is a dramatic rise in the number of mobile users who participate in the online SNSs, e.g., Facebook [26], Twitter [27], Tumblr [28], Sina Weibo [29] and so on, where more and more content is recommended and spread rapidly and widely [25] [30]. By investigating related measurements and modeling studies of the MSNs and SNSs, I discovered the following key points, which can be utilized for content dissemination:

- In online SNSs, the access pattern of each user can be measured, statistically modeled and thus predicted. That is, we can analyze the **access delay** between the content generation time and the user access time [31], which is per-user dependent mainly due to people's different life styles [22] [25] [32]. We can disseminate the content of interest to users considering their different delay sensitivities (requirements).

- In online SNSs, a user's influence, or **spreading impact**, to other users, can be modeled based on the analysis of social behavior histories, for example the forwarding probability.

- In offline MSNs, the mobility patterns of users can be measured and modeled [12] [20] [33] [34] [35], and hence a different offline **mobil-**

ity impact of each user to disseminate the content to others can be derived.

- User relationships and interests in online SNSs have significant **homophily** and **locality** properties (to be detailed in Ch. 2), which is similar to those of offline MSNs [24] [36] [37]. Users are mostly both clustered by geographical regions and interests, which can be exploited for traffic offloading.

Therefore, I am motivated to propose a Mobile ***T***raffic ***O***ffloading framework by ***S***NS-Based opportunistic ***S***haring in MSNs, **TOSS**. TOSS pushes the content object to a properly selected group of seed users, who will opportunistically meet and share the content with others, depending on their spreading impact in the SNS and their mobility impact in the MSN. TOSS further exploits the user-dependent access delay between the content generation time and each user's access time for traffic offloading purposes. From trace-driven evaluation and model-based analysis, TOSS lessens the cellular traffic up to 86.5% while still satisfying the delay requirements of all users. To the best of our knowledge, this is the first study that seeks to combine online SNSs with offline MSNs for traffic offloading considering user access patterns. Furthermore we propose a analytical framework, named **Push-Share** based on TOSS, which extends the offloading scenario to multi-cell environment. We theoretically model and analyze how the content can be pushed to a set of users and then shared to other users, where handovers are modeled based on the multi-compartment model. We also formulate the mathematical optimization framework, by which the trade-off between the dissemination delay and the energy cost is explored.

The advantages of offloading the cellular traffic by the opportunistic user-to-user sharing have been discussed in prior studies [10] [21] [38] [39] [40]. Furthermore, I compare pushing and sharing with other well-known strategies of content dissemination:

- **Pull-based Unicast:** In the traditional pull-based delivery, the file of interest may be downloaded via cellular links as many times as the number of subscribers [3] [4]. Meanwhile, our proposed model leverages the social meets of users, to offload the redundant downloads from the cellular links to local connectivities.

- **Broadcast/Multicast:** When multiple users (in the same cell) wish to receive the same content, broadcasting (or multicasting) would be efficient. However, for broadcasting, the lowest bit rate is normally used to cover all the mobile users in its cell, which reduces the efficiency substantially. And yet, the reliability of the content delivery is still difficult to achieve. There is also a security issue since non-subscriber users can also receive the content.

The rest of the thesis is organized as follows. After reviewing the related work in Ch. 2, I discuss the first study, TOSS framework, in Ch. 3. The framework details are in Sec. 3.1, and related optimization issues are discussed in Sec. 3.2. The trace-driven evaluation and the numerical analysis are shown in Sec. 3.3 and Sec. 3.4, respectively, followed by concluding remarks in Sec. 3.5. The second study, Push-Share, is introduced in Ch. 4. First I introduce the framework details in Sec. 4.1, as well as the system model in Sec. 4.2. Then I detail the content dissemination process in a single cell and multiple cells in Secs. 4.3 and 4.4, respectively. I discuss how to optimize the system parameters in Sec. 4.5. Numerical results are shown in Sec. 4.6, followed by concluding remarks in Sec. 4.7. By the end of the thesis, I make some conclusions in Ch. 5

Chapter 2

Related Work

2.1 Opportunistic Sharing in DTNs/MSNs

The epidemic content delivery in DTNs/MSNs has been extensively studied recently. Zhang et al. [20] have developed a differentiation-based model to study the delay of epidemic content delivery. For the purpose of energy conservation, Li et al. [13] also have designed an energy-efficient opportunistic content delivery framework in DTNs. The scalability and optimality of content dissemination by exploiting user-to-user contacts has been modeled as a social welfare maximization problem in [10]. Similarly, [21] has solved the maximization of traffic offloading utility in DTNs as a knapsack problem. Regarding the slow start and long completion time of the epidemic delivery, strategic pushing is studied to expedite the dissemination in [11]. Whitebeck et al. [9] also demonstrated the effectiveness of opportunistic offloading strategies based on practical mobility traces. While the above studies were limited to single cell environments, Wang et al. extend the pushing and sharing model into multi-cell cellular network environments in [12].

Accelerating the content dissemination by leveraging users' social relationships becomes a more popular research topic recently. BUBBLE Rap [41] utilizes social grouping characteristics for content dissemination. And [23] offloads up to 73.66% mobile traffic through social participation in the

MSN based on selection of the optimal initial seed users. [42] proposes to assign interest tags to the users and content objects to identify their preferences of content, and then utilizes users local centrality for efficient content sharing in DTNs. Similarly, ContentPlace [43] utilizes social central betweenness of mobile users to optimize the mobile content sharing. The SimBet [44] routing scheme in DTNs is also based on the analysis of user similarity due to the clustering effect and thus the calculation of user centrality. The similarity concept is also utilized by [45] and [46], both in which user encounter history is explored for getting the friendship similarity for delegation forwarding in the DTNs/MSNs. Therefore we are also motivated to extend the epidemic sharing in MSNs by considering the real social relationships in SNSs. Furthermore, Bao et al. carried real tests in Manhattan and identified the sharing-based offloading can reduce 30% to 70% mobile traffic [47]. In this thesis, security or privacy problems are not considered, but related studies such as [48] [49] and [50] can be referred.

The sharing in MSNs mainly relies on user-to-user local short-range communication techniques. Among existing user-to-user sharing methods, e.g., Bluetooth, Wi-Fi Direct, and NFC [17], which are based on public short-range communication techniques, the Wi-Fi Direct is becoming popular and popular. For instance, Apple's Airdrop [51] provides convenient user interface for a user to share a content to nearby users with ease. The HyCloud even utilizes cloud computing for enhancing the sharing among mobile users [52]. Recently the Device-to-Device (D2D) communication in the operator authorized spectrum becomes quite hot [18]. Device-to-device (D2D) communication underlaying a 3GPP LTE-Advanced cellular network is studied as an enabler of local services with limited interference impact on the primary cellular network. Based on optimal resource allocation and interference management, D2D communication can increase the total throughput observed in the cell area as studied in [19] [53] and [54]. This will further enhance the development of user-to-user sharing for traffic offloading in emerging mobile networks.

2.2 Mobile Traffic Offloading

There actually have been lots of studies focusing on the mobile traffic offloading by deploying Wi-Fi Access Points (APs). The realistic measurement from Korean Telecom (KT) [55] has pointed out that about 18% to 26% cellular traffic load is offloaded to KT's Wi-Fi APs. Similar studeis are carried in US, such as [56] and [57]. Depending on the density of AP deployment, the Wi-Fi based offloading can have different performance. For example, up to 65% traffic can be offloaded to Wi-Fi APs as practically studied in [58], in the downtown of Seoul, Korea. Y. Im et al. has proposed a cost-aware offloading with the throughput-delay trade-offs for offloading by Wi-Fi APs [59]. The economics of traffic offloading by Wi-Fi APs has been studied in [60] and the work in [61] further analyzes a more complicated offloading economics with the balance between delay and capacity of the network.

Regarding the traffic offloading based on user-to-user opportunistic sharing, how to encourage people to share during moving is thus interesting for researchers to design incentive-based business model, such as pricing study Win-Coupon in [39] to encourage the traffic offloading by DTNs. B. Tang et al. has also studied the benefit-based data caching and forwarding in ad hoc networks [62]. The self-interest-drive incentives for ad dissemination in autonomous MSNs is studied in [63]. Moreover, IPAD is a incentive-based design with conjunctive consideration with privacy [64], and iDEAL [65] is an incentivized cellular offloading based auction game.

2.3 Information/Content Spreading in SNSs

In this thesis, we consider the social relationship of users, so it is necessary to survey related studies on information and content spreading in SNSs. Dozens of years ago, in [66] the people social influence has been researched and identified as a "two-step flow of communication", that is, most people form their opinions under the influence of "opinion leaders", who in turn are influenced by the media source. Also the study in [67] declares that a small number of "opinion leaders" who have strong impact on spreading information perform the key roles to broadcast information by a socially connected network. Currently SNSs have been playing the important role for propagating media content [68]. In SNSs, due to the effect of "word-of-mouth" [24], users can significantly impact the information spreading to other users [69] [70]. Many studies have proposed to use probabilistic modeling to analyze the information/content commenting or re-sharing activities, and thus the information spreading impact among users [31] [32] [68] [70] [71] [72] [73]. Especially, the recommendation from famous people, who have potentially strong impact to others, may accelerate the topic spreading as studied in

[31] [74]. Also [25] indicates that people's historical impact on information sharing can impact and thus enable the accurately forecasting of the future sharing activities. Furthermore [25] points out that there are always some delays of re-sharing behaviors while the spreading impact of each user is accumulated hop by hop. This access delay between the content generation time and the user access time due to people's different life styles has been mentioned in many studies [22] [25] [32]. Researchers can obtain, analyze, and even predict the spreading impact and the access delays of SNS users based on measurement traces [30] [31].

Studies in [24] and [37] report that user relationships and interests in SNSs have significant **homophily** and **locality** characteristics as similar to those in MSNs. Homophily is the tendency of individuals to associate and bond with similar others [75]. The homophily here means online and offline users are both highly clustered by regions and interests, which also is studied as "birds-of-a-feather" in [76]. User homophily significantly impacts the information diffusion in social media. People with similar interests like to share and transfer the interesting information with each other. For instance, if one's friends watched a video, one will watch the video with very high probability. Locality is originally a phenomenon describing the same value, or related storage locations, being frequently accessed [77]. More specifically, in this thesis, the locality means that people who are graphically close may have similar trends of accessing the content and sharing with each other [24]. Even in online SNSs, users may significantly interact with and thus impact others in proximity, which also indicates the locality nature [36]. In other words, users within a short geographical distance have a higher probability of posting the same content than those users who are physically located farther apart. Thus, the locality characteristics of user interests can be utilized to facilitate the traffic load balancing [2] and content delivery [37].

Chapter 3

TOSS: Traffic Offloading Assisted by Social Network Services via Opportunistic Sharing in Mobile Social Networks

3.1 Framework Details

3.1.1 Preliminaries

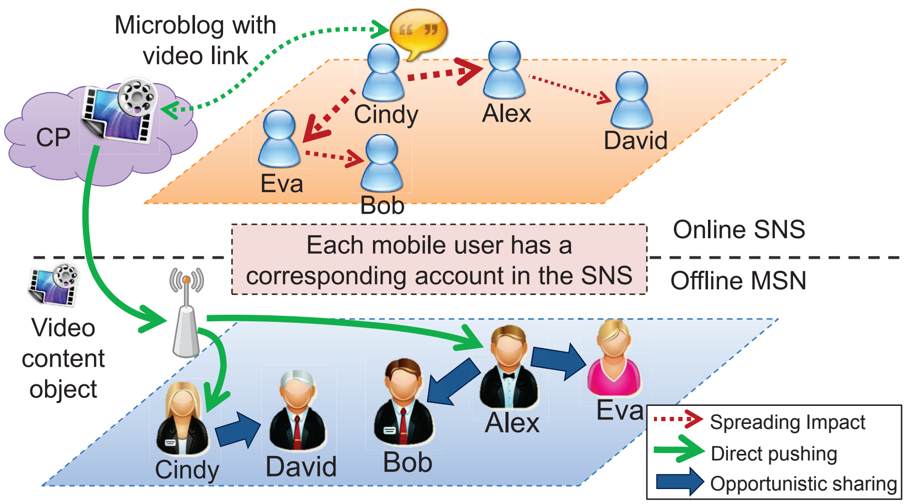


Fig. 3.1. Illustration of the TOSS framework containing the online SNS and the offline MSN

The TOSS framework entails both an online SNS and an offline MSN. Suppose there are total N mobile users, u_i, $i = 1, ..., N$, who have corresponding SNS identities. Because we focus on the content spreading in an online SNS, we use a directional graph to model the SNS[1], e.g., Twitter [28], Sina Weibo [29]. The online SNS can thus be represented by, $G(V, E)$, where V is the set of users in the online SNS, and E is the set of directional edges. If u_j follows u_i, u_j is one **follower** of u_i and u_i is one **followee** of u_j. As we focus on the content spreading, the directional edge (represented by an arrow in Fig. 3.1) is from u_i to u_j, denoted by v_{ij}. That is, u_i has a direct impact to u_j for content spreading. There can be a bidirectional relationship where two users follow each other.

We define the home-site, where a user create and shares content in the SNS platform, as the **microblog**, and we define a short message posted by a user containing the content (or link to the content) as a **micropost**[2], and the content file is called a **content object**. Furthermore, we define the **timeline** of a user in online SNS as the serie of all microposts published by a user in his/her microblog, sorted by time.

At any time, a user may find or create a new interesting article, image, or video, and share it in the SNS as an **initiator**. All his/her followers will then be able to access the content, and some of them will further re-share in their timelines. Making comments will not induce any information spread; thus we only consider the re-share activities. Afterwards, what TOSS seeks

---

[1] TOSS can also work with any SNS based on the bidirectional graph model (e.g., Facebook [26]) since it is a subset of the directional graph model.

[2] It can be a tweet in Twitter or a post in Facebook.

to achieve is that, while the micropost with the content is being spread to other users in the online SNS, the content object will be accessed and delivered among user devices in the offline MSN. Note that the TOSS framework is not confined strictly to the dissemination of one popular content to all the users, but applies to general deliveries of any content to a group of potential recipients with any size.

TOSS defines four factors for user u_i: two for the online SNS, (1) the outgoing spreading impact, $I_i^{S\rightarrow}$, and (2) the incoming spreading impact, $I_i^{S\leftarrow}$, which indicate how important the user is for propagating the micropost (to others or from others); two for the offline MSN, (3) the outgoing mobility impact, $I_i^{M\rightarrow}$, and (4) the incoming mobility impact, $I_i^{M\leftarrow}$, which indicate how important the user is for sharing the content object (to others or from others) via physical encounters. We will discuss how to calculate them in Sec. 3.1.2 and Sec. 3.1.4.

Considering the above factors, TOSS seeks to select a proper subset of users as seeds for pushing the content object directly via cellular links, and to exploit the user-to-user sharing in the offline MSN, while satisfying different access delay requirements of different users. We define a vector $\overrightarrow{p}$ to indicate whether to push the content object to a user via cellular links or not, e.g., $p_i = 1$ means pushing the content object directly to user u_i.

From the illustrated scenario of TOSS in Fig. 3.1, in the online SNS, Cindy shares a video (link) to Eva and Alex, who may in turn share with Bob and David, respectively. Meanwhile, the video content is first downloaded via a cellular link and stored in Cindy's phone. However in the offline MSN, Cindy is geographically distant from other people but David is in proximity.

Although David may not know Cindy, TOSS detects that the $I^{S\rightarrow}$ impact of Cindy to David via Alex is also very strong, and thus lets Cindy share the video with David via a local Wi-Fi connectivity. Furthermore TOSS evaluates the $I^{M\rightarrow}$ impact of Alex, and pushes another copy to him via a cellular link, because Alex is likely to meet Bob and Eva in the offline MSN frequently, and Bob and Eva often access content with some delays. Then the content object will be propagated by local connectivities from Alex to Bob and to Eva at a later time. TOSS reduces 3/5 of the cellular traffic in this example scenario.

3.1.2 Spreading Impact in the Online SNS

We extend the previous probabilistic models [68] [70] [71] [72] [74] to quantify the content spreading impact in the SNS. Hereby we define, the $I^{S\rightarrow}$ factor of user u_i to user u_j, denoted by γ_{ij}, $0 \leq \gamma_{ij} \leq 1$, is the ratio of the number of microposts of u_i that u_j accesses and re-shares to the number of all microposts of u_j in u_j's timeline. Thus for a given object of u_i in the future, And thus γ_{ij} is the probability that u_j will re-share the micropost from u_i [30].

Based on the SNS graph G, we define U_i^h as the set of h-hop upstream neighbors (followees) of user u_i through all possible shortest h-hop paths without a loop, and likewise D_i^h as that of h-hop downstream neighbors (followers). And we use γ_{ij}^h to denote the $I^{S\rightarrow}$ factor from user u_i to u_j by any h-hop path (inversely γ_{ji}^h as the $I^{S\leftarrow}$ factor from user u_j to u_i). From u_j's point of view over a certain period, we need to consider (1) the number of microposts that u_j has created by herself, c_j, (2) the number of re-shared microposts by u_j from u_i, r_{ij}, and (3) the number of re-shared microposts from all h-hop followees, to calculate $I_i^{S\rightarrow}$ as follows:

$$\gamma_{ij}^1 = \frac{r_{ij}}{c_j + \sum_{u_k \in U_j^1} r_{kj}}, \tag{3.1}$$

$$\gamma_{ij}^2 = 1 - \prod_{k \in D_i^1 \cap U_j^1} \left(1 - \gamma_{ik}^1 * \gamma_{kj}^1\right), \tag{3.2}$$

$$\gamma_{ij}^3 = 1 - \prod_{k \in D_i^2 \cap U_j^1} \left(1 - \gamma_{ik}^2 * \gamma_{kj}^1\right), \tag{3.3}$$

$$\gamma_{ij}^h = 1 - \prod_{k \in D_i^{h-1} \cap U_j^1} \left(1 - \gamma_{ik}^{h-1} * \gamma_{kj}^1\right). \tag{3.4}$$

We use γ_{ij}^* to denote the impact from user u_i to user u_j via all possible paths with less than or equal to H hops, computed by:

$$\gamma_{ij}^* = 1 - \prod_{n=1}^{H} \left(1 - \gamma_{ij}^n\right), \tag{3.5}$$

where H is less than or equal to the maximal diameter of the SNS graph G. Then $I_i^{S\rightarrow}$ and $I_i^{S\leftarrow}$ of u_i to and from the whole user base can be respectively calculated by

$$I_i^{S\rightarrow} = \sum_{j=1}^{N} \gamma_{ij}^*, \quad I_i^{S\leftarrow} = \sum_{j=1}^{N} \gamma_{ji}^*. \tag{3.6}$$

Note that it is reported in [25] [76] that the average path length in SNS graphs is about 4.12 and the spreading impact after 3 hops becomes negligible.

3.1.3 Access Delays of Users in the SNS

3.1.3.1 Access Delay

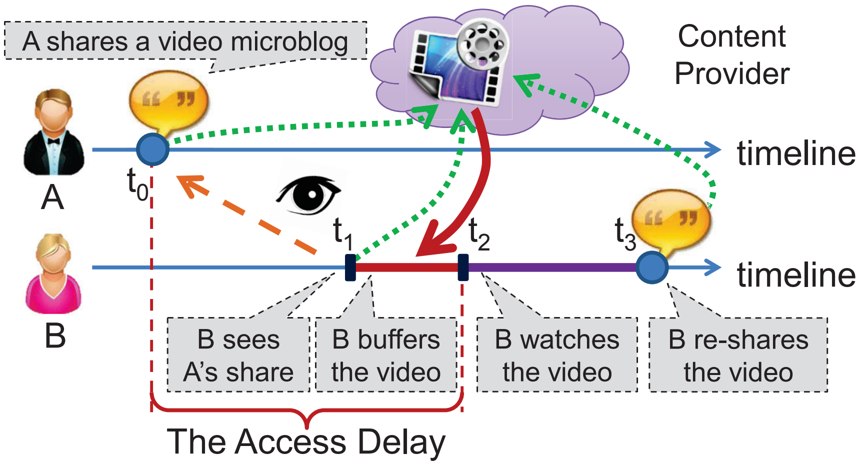


Fig. 3.2. Illustration of the content access delay between A's content generation time and B's access time

Different users have different patterns of accessing content via the online SNS. Some may access the SNS frequently, while others access the SNS at relatively longer intervals. Thus the access delay between the content generation time and user's access time becomes different for each user [22] [25] [32].

As illustrated in Fig. 3.2, user A creates a micropost for an interesting video in the SNS at t_0. One of A's followers, B, happens to see A's micropost after a certain delay at t_1 due to B's personal business. Once B clicks to play it, a buffering delay is needed until t_2; B will re-share the video at t_3 after watching it. In practice, it is hard to obtain t_1 and t_2 data. Thus we consider B's access delay as $t_3 - t_0$, which can be captured from the SNS measurement trace by checking B's re-sharing time from the SNS measurement.

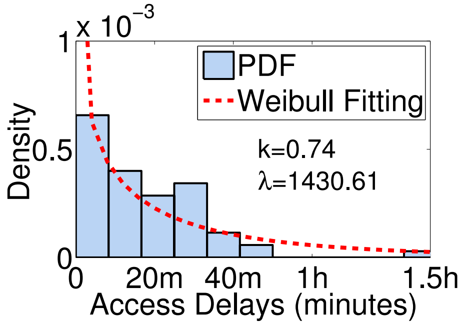
To investigate access delays, we collected the SNS trace data of approximately 2.2 million users from the biggest online SNS in China, Sina Weibo (measurement details will be explained in Ch. 3.3). The access delay is gathered as the time difference between the generation time of the original micropost and the time of re-sharing by a follower.

We pick up three real users from the online SNS trace, and plot their access delays by probability distribution function (PDF) as shown in Fig. 3.3. User u_1 is likely to access the content frequently with short delays. But users u_2 and u_3 have significant delays, on the order of hours and days, respectively. In this regard, we can classify all users into two types: **(a) keen users**, who check microposts frequently, and access content object with short access delays mostly, e.g. u_1; **(b) dull users**, who mostly access the microposts with substantial delays, e.g. u_2 and u_3. Generally, TOSS tends to push content to keen users, but seeks to utilize the opportunistic sharing in the MSN to disseminate the object to dull users.

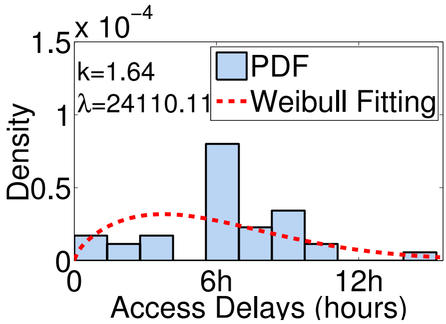
3.1.3.2 Modeling of the Access Delays

We use a PDF to model the access delays of each user, say u_i, in terms of the probability to access the content at t, denoted as $A_i(t)$. As similar to [10], $A_i(t)$ can be considered as the access utility function. If the content object is already obtained locally in the user's device when she has the highest probability to access the content, she will be mostly satisfied.

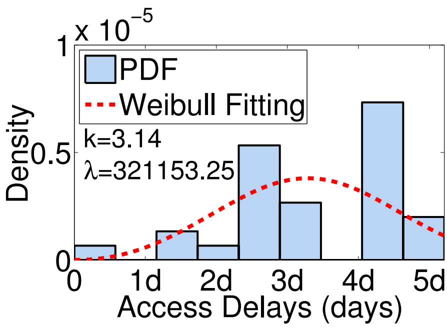
In order to model the various distributions of access delays with different shapes of PDF curves, we choose to use Weibull distribution for fitting,



(a) User u_1



(b) User u_2



(c) User u_3

Fig. 3.3. The access delay distributions of three real users with Weibull fitting

which is commonly used for profiling user behaviors in SNSs [78]:

$$A_i(t, \beta_i, k_i) = \frac{k_i}{\beta_i}\left(\frac{t}{\beta_i}\right)^{k_i-1} e^{-\left(\frac{t}{\beta_i}\right)^{k_i}}, \quad t \geq 0, \tag{3.7}$$

where the fitting parameters β_i and k_i can identify the access pattern of user u_i (note that k_i controls the curve shape). When $k_i \geq 1$, the Weibull fitting curve can represent the distribution of the access delays of keen users; if $k_i < 1$, the Weibull fitting curve has a peak, and thus, can represent the distribution of access delays of dull users. It is measured that (to be discussed in Sec. 3.3) about 2/3 of SNS users are dull ones with large access delays, which is a sufficiently large portion of users that allows TOSS to disseminate the content object by offline opportunistic sharing.

Note that, there can be many different functions for fitting statistical data into functions, but we stick to use Weibull fitting, as a large number of studies use Weibull for fitting human behavior statistics, such as [78] for user behaviors in online SNSs, [79] for user web browsing activities, [80] for user access patterns in Wi-Fi networks, and [81] for traffic flows in online games. In these related papers, the fitting faithfulness is not mentioned, since the parameter k and λ are actually the important ones for analyzing patterns and for reproducing the user behaviors. Also we use MATLAB *wblfit* function [82] to analyze and carry out Weibull fitting on the access delays, and it does not return the value of likelihood (faithfulness), but only returns the k and λ with the maximal likelihood estimation. Therefore although the analysis here has some limitation regarding the faithfulness of the fitting, we believe the Weibull fitting can well reflect the patterns of user access delays.

3.1.4 Mobility Impact in the Offline MSN

It has been studied that mobile users in the offline MSNs (or DTNs), have different mobility patterns [12] [20] [33] [34] [35] , and hence different potentials for sharing content. Thus the mobility impact, I^M, is defined to quantify the capability of a mobile user to share a content object with other users via opportunistic meetings, or say **contacts**, while roaming in the MSN. The temporary connectivity with nearby users mostly relies on active discovery mechanisms; thus we assume all mobile users are synchronized with a low duty cycle for probing as proposed by eDiscovery [16].

Referring to [10] [12] [13] [20] [33] [34] [35] [42] we assume that the inter-contact intervals of any two mobile users follow the exponential distribution. We use λ_{ij} to denote the opportunistic contact rate of user u_i with user u_j. Note that there are many practical methods to measure λ_{ij} values, e.g., centralized measurement by the location management entity in the MNO [83] or by distributed user-to-user exchanges [43]. Note that the contact duration is ignorable in TOSS, because we assume the content delivery is always finished successfully during the contact due to the high bandwidth of local communications (e.g., Wi-Fi) [10] [12] [20] [42].

We adopt the epidemic modeling from [13] [20] to model the opportunistic sharing in TOSS with the continuous time Markov chain. We let $S_i(t)$ be the probability that user u_i may have the content until t, $0 \leq S_i(t) \leq 1$, while $1 - S_i(t)$ is the probability that user u_i has not received the content until t. $S_i(t)$ will be increasing over t while roaming and meeting users in the offline MSN. The increment of $S_i(t)$ within a period Δt, that is $S_i(t + \Delta t) - S_i(t)$, will be calculated in the following procedure.

The probability of user u_i to meet user u_j during Δt, is $1 - e^{-\lambda_{ij}\Delta t}$ due to the exponential decay of inter-contact intervals. The probability that user u_i can get the content from another user u_j via opportunistically meeting, denoted by ε_{ij}, can be calculated by:

$$\varepsilon_{ij} = \left(1 - e^{-\lambda_{ij}\Delta t}\right) \cdot \gamma^{*}_{ji} \cdot S_j\left(t\right), \tag{3.8}$$

where the $I^{S\rightarrow}$ impact factor from u_j to u_i, γ^{*}_{ji}, is considered as both *(i)* the spreading probability that u_j will re-share microposts from u_i and *(ii)* the sharing probability that u_i can obtain the content object from u_j.

Considering the ε_{ij} of u_i from all users, the probably that u_i can get the content from others within Δt is,

$$1 - \prod_{j=1, j\neq i}^{N} (1 - \varepsilon_{ij}). \tag{3.9}$$

Hence based on the probability that u_i has not received the content,

$$S_i\left(t + \Delta t\right) - S_i\left(t\right) = \left(1 - S_i\left(t\right)\right) \cdot \left(1 - \prod_{j=1, j\neq i}^{N} (1 - \varepsilon_{ij})\right). \tag{3.10}$$

Letting $\Delta t \rightarrow 0$, the derivative of $S_i(t)$ will be

$$\begin{aligned} \overset{\bullet}{S_i}\left(t\right) &= \lim_{\Delta t \rightarrow \infty} \frac{S_i(t+\Delta t) - S_i(t)}{\Delta t} \\ &= \lim_{\Delta t \rightarrow \infty} \frac{(1 - S_i(t)) \cdot \left(1 - \prod_{j=1, j\neq i}^{N} \left(1 - \varepsilon_{ij}\right)\right)}{\Delta t} \\ &= \left(1 - S_i\left(t\right)\right) \cdot \sum_{j=1, j\neq i}^{N} \lambda_{ij} \cdot \gamma^{*}_{ji} \cdot S_j\left(t\right) \end{aligned}, \tag{3.11}$$

where initially $S_i\left(0\right) = p_i$ from $\overrightarrow{p}$.

Solving the above matrix of the ordinary differential equation system is complicated. However, we can find a numerical solution easily by approximation with power series [84] [85]. We skip the details of the procedure for getting numerical solutions, since this is trivially straight forward.

Given a pushing vector $\overrightarrow{p}$, we can calculate how long it will take for any user u_i to obtain the content by the inverse function of $S_i(t)$ with $S_i(t) = 1$, defined as the **content obtaining delay** of u_i, denoted by t_i^*:

$$t_i^* = S_i^{-1}\left(\{\gamma_{ji}^*\}, \{\lambda_{ij}\}, \overrightarrow{p}\right) \ \ , \ \ j = 1, ..., N, j \neq i, \tag{3.12}$$

where $\left\{\gamma_{ji}^*\right\}$ is the series of $I^{S\leftarrow}$ factors from all other users to u_i in the SNS, and $\{\lambda_{ij}\}$ is the series of meeting rates of user u_i to all other users in the MSN. Note that TOSS mainly seeks the optimal $\overrightarrow{p}$ to match the content obtaining delays of all users with their access delay PDFs.

$I_i^{M\rightarrow}$ is actually the same as $I_i^{M\leftarrow}$ since $\lambda_{ij} = \lambda_{ji}$ for any u_i and u_j due to the symmetric nature of contacts. Hereby, we define the I^M factor for u_i as,

$$I_i^{M\rightarrow} = I_i^{M\leftarrow} = \lambda_i^* = \sum_{j=1}^{N} \lambda_{ij}. \tag{3.13}$$

And then we will only use I^M to denote the mobility impact. We can use approximation methods, e.g., the Newton Method, to get the numerical result of the inverse function of $S_i(t)$.

Note that above content obtaining delay is the expected delay that user can obtain a content based on opportunistic sharing while moving, which is an objective factor depending on the mobility traces by given an initial pushing vector. It is different from previously mentioned content access delay, which is a subjective factor depending on user behaviors (life styles);

TOSS fits the access delays of users by Weibull function, which converts the subjective access delays into objective probability distribution function, and then uses it for indicating user's delay QoS requirement. So TOSS is just right seeking for a perfect match between these two.

3.2 System Optimization

By evaluating I^S (both incoming and outgoing) and I^M values of all users, how to choose proper set of seeds for initial pushing, $\overrightarrow{p}$, to get the content obtaining delay t^* for each user in order to maximize the sum of the access utilities (access probabilities) for all users becomes the objective of TOSS.

$$
\begin{aligned}
\underset{\overrightarrow{p}}{\text{Maximize}}: \quad & \sum_{i=1}^{N} A_i\left(t_i^*, \beta_i, k_i\right) \\
& = \sum_{i=1}^{N} A_i\left(S_i^{-1}\left(\left\{\gamma_{ji}^*\right\}, \left\{\lambda_{ij}\right\}, \overrightarrow{p}\right), \beta_i, k_i\right) \\
& (j = 1, ..., N, j \neq i) \\
\text{Subject to}: \quad & |\overrightarrow{p}| \leq C \quad ,
\end{aligned}
\tag{3.14}
$$

where the number of initial pushing seeds, C, is a constraint controlled by the MNO, and we call $\sum A_i(t)$ the total access utility function of the whole user base.

This problem is similar to the social welfare maximization problem, discussed in [10]. Solving the above optimization problem analytically is hard, since all related equations are not in closed-form. With power series approximations, we can find the maximum values by general numerical methods. Also we can even tune and find the needed C by given a target total access utility value. One of the key remaining future work will be the reduction of the complexity of the equations and thus the optimization problem.

We design a heuristic algorithm to find the near-optimal solution $\overrightarrow{p}$ for maximizing $\sum A_i(t)$ numerically, based on the hill-climbing method [86], as shown in Algorithm 3.1. Initially we select the top C users from all users sorted by I^M in descending order ($I^{S\rightarrow}$ or $I^{S\leftarrow}$ works similarly) and iteratively exchange the p_i and p_j values of any two users u_i and u_j if the larger $\sum A_i(t)$ can be obtained, until the increment of $\sum A_i(t)$ is smaller than a specified threshold. Note that the above modeling and the heuristic algorithm are calculated in MATLAB [82].

3.3 Trace-Driven Measurement

To evaluate the effectiveness of TOSS framework, we need SNS trace data to quantify the spreading impact factors and access delays, as well as MSN trace data to analyze the mobility impact. However, in public, there is no available trace data that contains both the SNS and the MSN activities. Thus, we choose to take separate measurements, and combine them by some mapping strategies, which will be explained in Sec. 3.4.1.

3.3.1 Measurement of the Online SNS

We select the most popular online SNS in China, Sina Weibo, and keep track of 2,223,294 users for four weeks during July, 2012. We collected 37,267,512 microposts generated (and partially re-shared) by the users, and further obtained the list of all the re-sharing activities for each micropost.

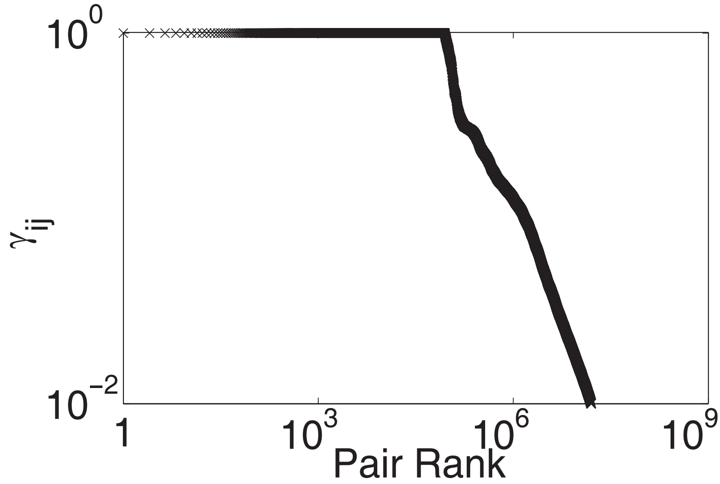
Algorithm 3.1 A Hill-climbing algorithm to seek near-optimal initial pushing seeds

```
// Initializing p⃗
for all i = 1 → N do
    p_i=0; v_i = λ*_i, γ*_i, or random;
end for
Sort v_i by Descent Order (↓);
for all i = 1 → C do
    p_i=1;
end for
A_sum = Σ_{i=1}^{N} A_i( S_i^{-1}( {γ*_ji}, {λ_ij}, p_i ), β_i, k_i ),
      (j = 1...N, j ≠ i);
// Hill-Climbing
repeat
    flag=true;
    for all i = 1 → N do
        for all j = i+1 → N do
            if (flag==true) AND (p_i + p_j == 1) then
                Exchange(p_i, p_j);
                A'_sum = Σ_{i=1}^{N} A_i( S_i^{-1}( γ⃗*_i, λ⃗_i, p_i ), β_i, k_i ),
                      (j = 1...N, j ≠ i);
                if A'_sum > A_sum then
                    δ = A'_sum − A_sum; A_sum = A'_sum;
                    flag=false;
                end if
            end if
        end for
    end for
until δ < Threshold
return A_sum, p⃗
```

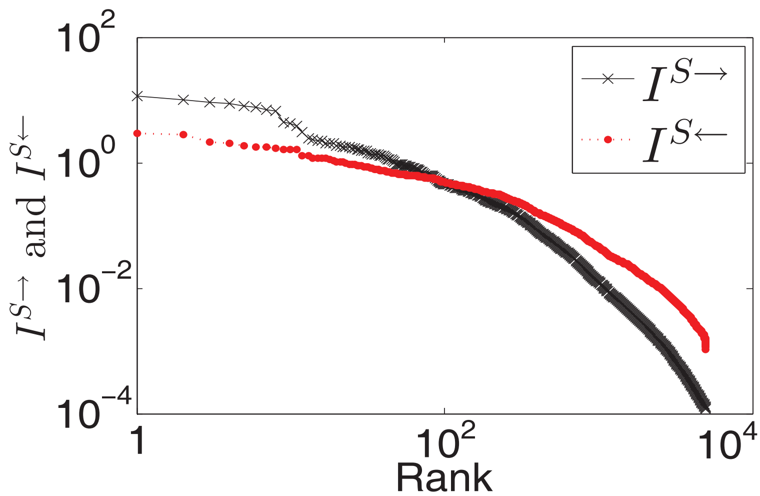
We implemented the data collection software, which starts from 15 famous users of distributing popular video clips, and expands the user base from their followers. Capturing the next hop followers is carried out iteratively. The captured data includes details of owner's account profile, all microposts with timestamps of the owner, all comments and reposts with timestamps, as well as the profile of the users that make comments and reposts to the owner. Note that there are some robots in Sina Weibo, which always re-share some microposts of famous people with extremely short delays, and thus we exclude users with no followers, no followees, or no self-created microposts. How to precisely exclude all the robots in the SNS trace is out of the scope of this thesis, and there are many related studies for reference such as [87] [88] and [89]. In all, we believe that the 2.2 million user base can reflect the ground-truth of the social impact factors and the access delay statistics.

3.3.1.1 Spreading Impact, γ^*_{ij} and I^S

Recall that γ^*_{ij} is the spreading impact of one user to another user based on the accumulation of user-to-user reposting ratio via any possbile paths, calculated by Eq. (3.5) and I^S is the overall spreading impact of the user to all users in the SNS, calculated by Eq. (3.6). However calculating γ_{ij} for the whole user base takes substantially long time. Thus we choose a sub-graph of 4,311 users by random walking method for evaluation (to be detailed later). And we let $H = 4$ to consider up to 4-hop paths among the users as suggested in [25]. The γ^*_{ij} of each pair of users is sorted and shown in the log-log scale in Fig. 3.4(a), which indicates the strong online spread impact among the socially grouped users. 98,168 pairs have $\gamma \geq 0.95$ in Fig. 3.4(a), which may be due to some strongly connected users, and 47,680 pairs have $\gamma = 1$, which may be due to some remaining robots that we cannot exclude. Fig. 3.4(a) shows the user-to-user impact mostly follows the power-law dis-



(a) Analysis of γ_{ij}



(b) Analysis of I^S

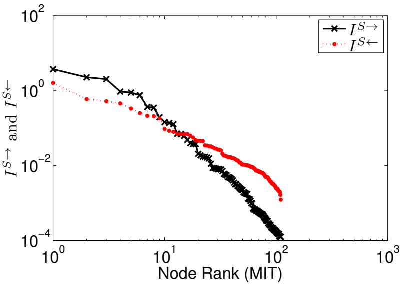
Fig. 3.4. Measurement values of γ_{ij} and I^S (4,311 users)

tribution [90] indicating that a small number of user pairs have very strong impact, while many other users have little impact. The $I^{S\rightarrow}$ and $I^{S\leftarrow}$ values of those users are plotted in Fig. 3.4(b), which also shows that a smaller number of people have significant outgoing impact ($I^{S\rightarrow}$) to the whole SNS, while many users are relatively less impacted by others ($I^{S\leftarrow}$).

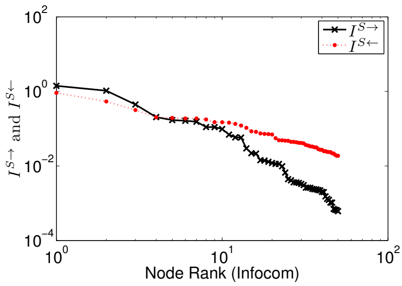
One important issue here is whether the randomly selected sub-graph of the SNS can still reflect the characteristics of the whole SNS user base. There have been some related measurement studies pointing out that: the SNS is a **scale-free** network [25] [91] [92] [93] [94] [95] and [96]. A scale-free network is a complex network whose degree distribution follows the power-law, at least asymptotically, which means in such kind of network, a small number of nodes make dominant impact to the network, while many nodes make very small impact, if we consider the node degree or the spreading impact (re-sharing ratio) as the impact of a node to the network [91] [93] and [97].

As researched in [93] [94] and [97], due to the nature characteristics of scale-free complex networks, no matter we choose any sub-graph from the whole network graph (with not too small size) by random walking or by random-sampling, similar characteristics (power-law distribution of node strength) can be still obtained.

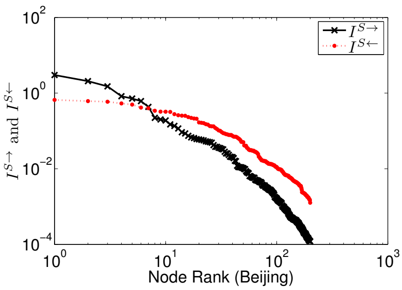
Furthermore, we take a check on whether the sub-graphs that we abstracted from the online SNS graph corresponding to the mobility traces (to be detailed in Ch. 3.3.2) can be suitable to still keep the characteristics. Obviously the number of nodes in each SNS sub-graph is the same as the number of nodes in the corresponding mobility trace, and for each trace we carry out sub-graph sampling for five times, and then make average value. We draw the log-log plots for the spreading impact factor of the nodes from sampled sub-graphs as shown in Fig. 3.5. All of the figures are able to re-



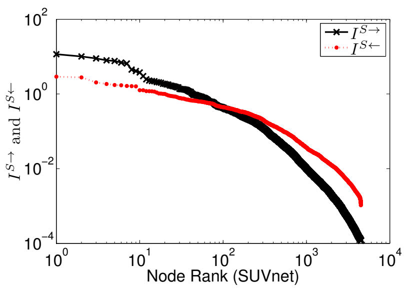
(a) MIT (100 users)



(b) Infocom (41 users)



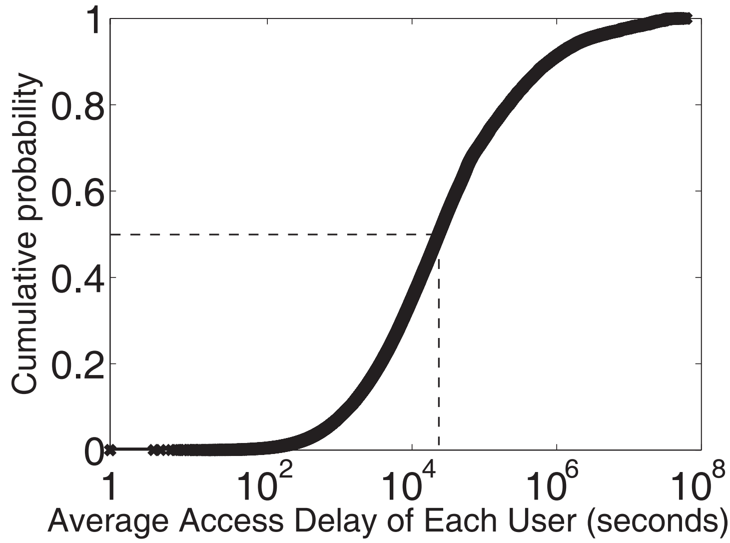
(c) Beijing (182 users)



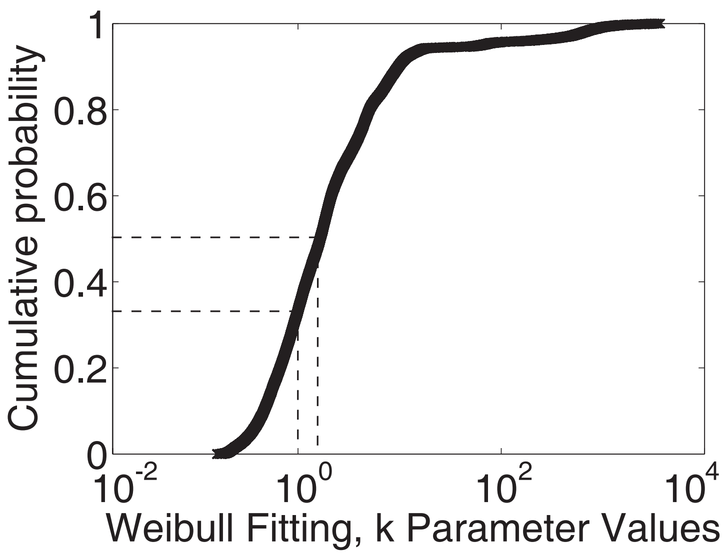
(d) SUVnet (4,311 users)

Fig. 3.5. Measurement values of I^S for sub-graphs sampled from the SNS graph with different sizes corresponding to the mobility traces

flect the asymptotical power-law trend, that is, a very small number of nodes impact the network significantly, but most of the nodes have weak impact. They have quite similar trends to the curves as shown in Fig. 3.4 (b), So conclusively, all of the sub-graphs with different sizes can still represent the SNS, and it will be an acceptable methodology to map the SNS sub-graphs to the mobility traces. Note that in the following part, the online spreading impact factor is normalized and then applied.



(a) Average access delays



(b) Weibull parameter, *k*

Fig. 3.6. Access delays and fitting parameters

3.3.1.2 Access Delay Distribution of u_i, $A_i(t)$

Measurement results of the access delays on the whole user base, $A_i(t)$ from Eq. (3.7), are shown in Fig. 3.6. From the cumulative distribution function (CDF) of the average of all the access delays of each user in Fig. 3.6(a), half of the users have the average access delay larger than 23,880 seconds, which is about 6 hours and 38 minutes. Taking a closer look, we find (1) 3.67% of users have the average access delay less than 10 minutes, (2) 20.38% of users have the delay smaller than 1 hour, and (3) 26.79% of users access the SNS with average delay larger than 1 day. Furthermore, we calculate the Weibull fitting parameters of all users, and the CDF of the shape parameter k of all users is shown in Fig. 3.6(b), which indicates that 32.63% of users have $k < 1$, who are likely to be keen users, while 67.37% of users can be classified as dull users. Therefore, we verify that a substantial number of users access the SNSs with sufficiently large delays, which we can exploit for offline opportunistic sharing.

3.3.2 Measurement of Offline MSNs, λ_{ij} and I^M

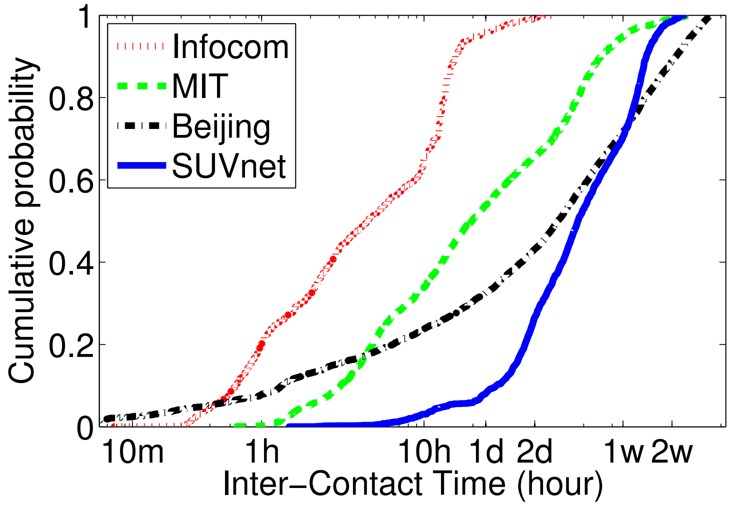
We choose four mobility traces, MIT [98], Infocom [99], Beijing [100], and SUVnet [101], in order to evaluate the performance of TOSS. These traces record either direct contacts among users carrying mobile devices or GPS-coordinates of each user's mobile route, and the traces details are shown in Table 3.1. The four traces differ in their scales, durations, and mobility patterns; The MIT and the Infocom traces are collected by normal people, but the Bejing and the SUVnet traces are collected by vehicles. The Beijing

Table. 3.1. Mobility traces

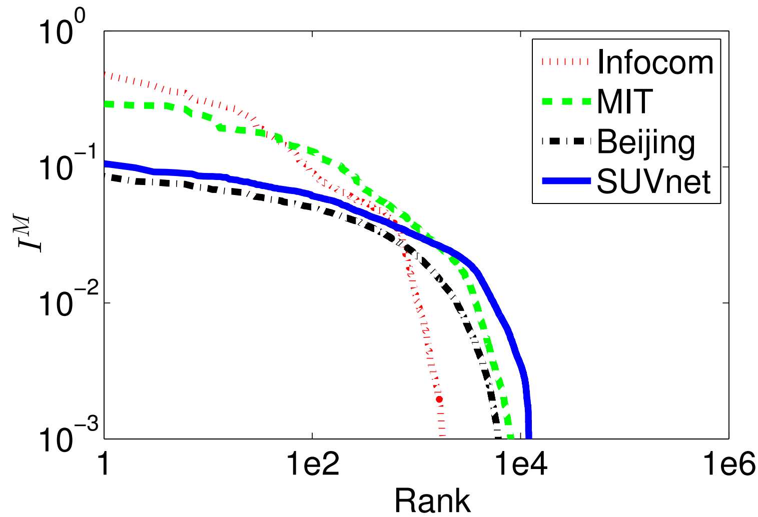
Trace	Link	Users	Days	Contacts	Avg.λ
MIT[98]	Bluetooth	100	246	54,667	0.01532
Infocom[99]	Bluetooth	41	4	22,459	0.14167
Beijing[100]	/	182	150	8,894	0.00023
SUVnet[101]	/	4,311	30	169,762	0.00131

and the SUVnet traces have no record of contacts, but only GPS coordinates with time. We assume two users have a contact once they are within a sufficiently small distance (20 meters) during a short interval (20s).

Recall that the λ_{ij} is the inter-contact rate of two users, which indicates the mobility impact between them. And the I_i^M is the overal mobility impact factor of a user to the whole user base in the MSN base on Eq. (3.13) We analyze the traces and obtain the inter-contact intervals $(1/\lambda_{ij})$ of all user pairs, as shown in Fig. 3.7(a). The Infocom trace has the highest contact rate because users are at a conference spot, and thus have high contact rates. The MIT trace also has high contact rate since users are friends within the campus. The Beijing and the SUVnet traces have large inter-contact intervals because they have relatively low frequency of GPS records and large user base, which is considered as sparse user density. I^M values of all users of the traces (values smaller than 0.001 are ignored) are plotted in Fig. 3.7(b), which indicates the similar trends of the traces as discussed above. Users in the Infocom trace have the highest potentials to obtain the content by sharing, but users in the Beijing trace have the weakest potentials.

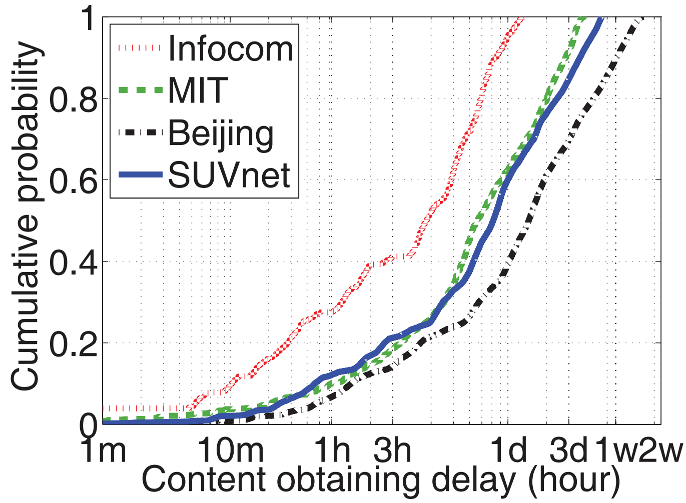


(a) Analysis of $1/\lambda_{ij}$

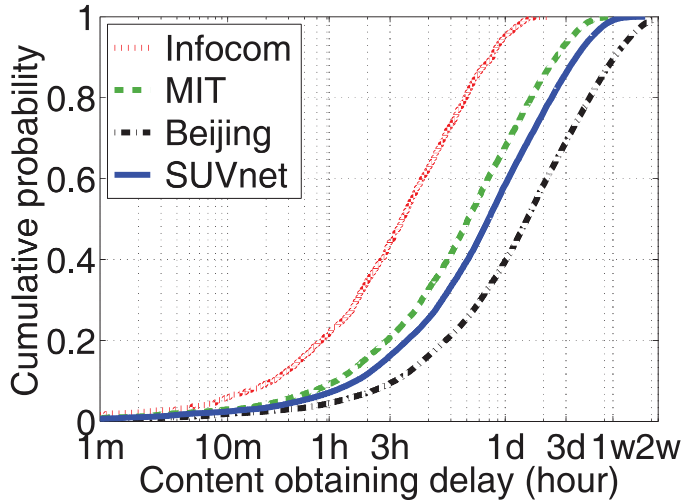


(b) Analysis of I^M

Fig. 3.7. Measurement values of λ_{ij} and I^M



(a) Real Traces



(b) Trace-based Mobility Modeling

Fig. 3.8. Measurement and modeling results of content obtaining delays by 1 random pushing

3.3.3 Content Obtaining Delays, t_i^*

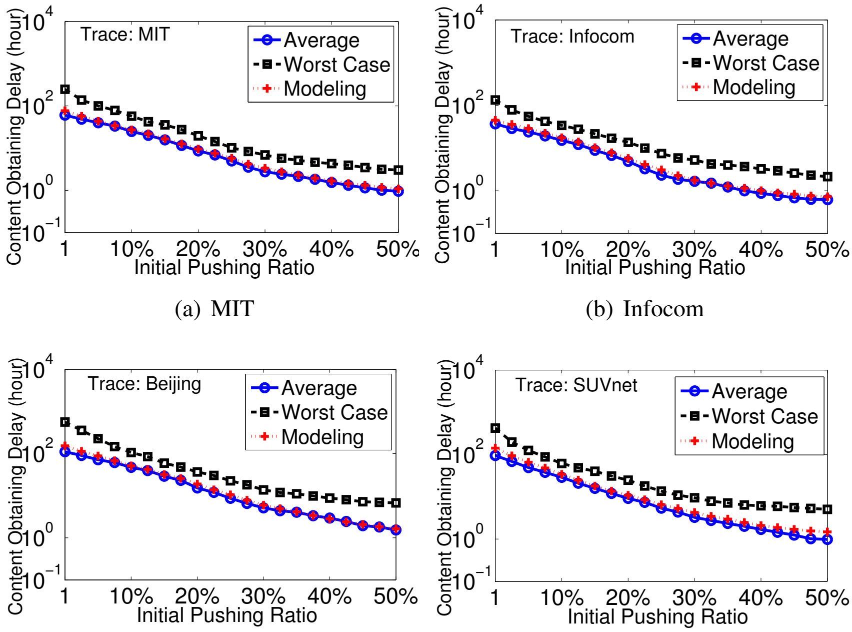
We investigate the content obtaining delays, t_i^* from Eq. 3.12, of all users by just 1 random initial pushing (averaging 20 runs with different random

Table. 3.2. Correlation regression analysis between the traces and modeling

Trace	Pearson Correlation	Significance
MIT[98]	0.973	0.000
Infocom[99]	0.979	0.000
Beijing[100]	0.976	0.000
SUVnet[101]	0.968	0.000

seeds) for the four traces. Made a program to go through the mobility trace entry by entry, to simulate the content propagation and obtain results by the end. And then, we use the empirical λ values of all pairs extracted from the traces and import to the modeling derived in Sec. 3.1.4, and calculate the obtaining delays by MATLAB [82] and Mathematica [102]. From the CDFs in Fig. 3.8(a), the Infocom trace has the smallest obtaining delays mostly within 1 day, while the Beijing trace shows the longest delays even up to 10 days. The model with practical λ values in Fig. 3.8(b) shows the similar performance to the real traces.

In order to precisely verify the accuracy of our modeling to the real traces, from the two figures, Fig. 3.8(a) for the real traces, and Fig. 3.8(b) for the modeling, we carry out the bivariate correlation regression analysis on them, in order to get the Pearson correlation coefficients, by using SPSS [103]. As shown in Table. 3.2, the results of the correlation coefficients between the real traces and by the modeling are in the range of 0.973 to 0.979, which means the simulation and modeling can fit perfectly with a sufficiently high accuracy.



(a) MIT (b) Infocom

(c) Beijing Taxi (d) SUVnet

Fig. 3.9. Performance impact of C on content obtaining delay t^*

3.3.4 How C Impacts the Obtaining Delay, t_i^*

We further investigate how the number of initial seeds, C, impacts the content obtaining delays as shown in Fig. 3.9, where Y-axis shows the average value of the obtaining delays of all users in log scale. Note that we will mostly focus on the initial pushing ratio, which is the ratio of the number of seeds to the total number of users in each trace. We start with pushing to 1 random user (i.e., $C = 1$), until randomly pushing to 50% of all users in each trace, and make the average from 20 runs. Note that we do not consider the I^S impact yet. As more users are selected as initial seeds, the average delay decreases significantly, but there are still some users with large obtaining delays even we push to 50% of the users.

3.4 Performance Evaluation

We now consider how the spreading and mobility impact factors (I^S and I^M) affect the total access utility function ($\sum A_i(t)$) to evaluate TOSS framework.

3.4.1 How C Impacts the Total Access Utility, $\sum A_i(t)$

Due to the lack of a trace that contains the activities of the same users in both online SNSs and offline MSNs, we consider three choices for mapping SNS users to MSN users in each of the four mobility traces: (1) **random**: SNS users are randomly mapped to MSN users; (2) **h-h**: both SNS and MSN users are sorted in descending order of $I^{S\rightarrow}$ and I^M respectively, and then are mapped correspondingly; (3) **h-l**: both users are sorted as similar to **h-h**, but an SNS user with high $I^{S\rightarrow}$ is mapped to an MSN user with low I^M. Since the number of SNS users is much larger than that of MSN users in each trace, we pick accounts from the SNS trace by random-walk sampling to match the number of MSN users in each trace.

Regarding the methodology of mapping a sub-graph of online SNS by random-walk sampling to the offline MSN graph, we carry out following discussion: It is already studied that when we consider the mobility impact (meeting rate) of two users as their vector strength, and the overal mobility impact of one user (sum of all mobility impact to all other users) as the node strength, the MSN can be also classified as a scale-free network [99] [104]

[105] [106] [107] [108] and [109]. That is in the MSN, a small number of users are always moving quickly and meet many components, while many of the users are relatively stable to meet limited number of other users.

So regarding each mobility trace with different amount of mobile users, as we discussed in Sec. 3.3.1.1, we take random-walk-based sampling to obtain the subgraphs from the SNS trace with corresponding number of user accounts, and then map one SNS account to one mobile user by above mapping choices. Note that the online spreading impact factor is normalized and then applied. Conclusively, it is a reasonable methodology to map between online and offline traces in the case of lacking such a trace with both information. To seek or carry out such a measurement study to track both the online SNS activities and offline MSN activities for a group of people is one important future work.

Note that, actually due to the locality nature of human-beings, TOSS framework will still perform well even facing to the scenarios with a very large user base (e.g., a city, or even a country); although people move and travel sometimes, they still meet most of friends in most cases, which is the **clustered effect** for a group of users, which will not be impacted by the whole user base. In another word, people are constrained by our life style and location due to the inherent nature of **locality**, as studied in [109] [110].

To select the users who will be initial seeds, $\overrightarrow{p}$, constrained by the allowed total number of seeds, C, we consider the following five pushing strategies based on the impact factors:

- **p-λ**: we sort users by I^M ($\sum \lambda_i^*$) in descending order and choose the top C ones (similar to [10]);

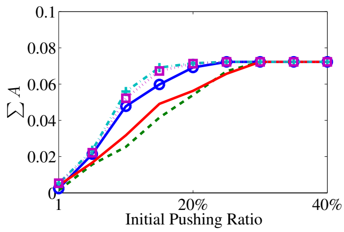
- **p-$\gamma^{\rightarrow}$**: we sort users by $I^{S\rightarrow}$ ($\sum\gamma^*_{ij}$) in descending order and choose the top C ones (similar to [41] [43]);
- **p-$\gamma^{\leftarrow}$**: we sort users by $I^{S\leftarrow}$ ($\sum\gamma^*_{ji}$) in descending order and choose the top C ones;
- **p-$\lambda * \gamma^{\rightarrow}$**: we sort users by $I^M * I^{S\rightarrow}$ conjunctively in descending order and choose the top C ones;
- **p-$\lambda * \gamma^{\leftarrow}$**: we sort users by $I^M * I^{S\leftarrow}$ conjunctively in descending order and choose the top C ones;

There are many **viral marketing** methods to evaluate a SNS user's strength regarding information spreading, for example we can easily qualify by node degree including outgoing degree (number of followees) and incoming degree (number of followers). Note that here the arrow direction is the "following/followed" relationship, reverse to the spreading direction. Furthermore the PageRank algorithm [111] [112] is also comprehensively used for SNS analysis, which is a link analysis algorithm of Google by assigning a numerical weighting to each element of a hyperlinked set of nodes, with the purpose of "measuring" its relative importance. We apply the general PageRank algorithm on the selected SNSN subgraphs and obtain the PageRank scores. We also consider a random pushing and the heuristic algorithm, and hence we have five more initial pushing strategies based on the graphs:

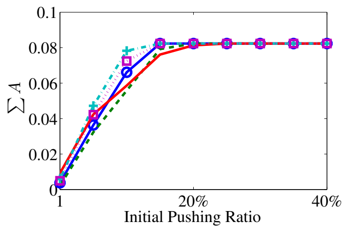
- **p-R**: we randomly choose C users;

- **p-$D^{\rightarrow}$**: we sort users by outgoing node degree in descending order and choose C users;
- **p-$D^{\leftarrow}$**: we sort users by incoming node degree in descending order and choose C users;
- **p-Pr**: we sort users by PageRank score in descending order and choose top C users;
- **p-H**: we run the hill-climbing heuristic algorithm to obtain the near-optimal pushing vector.

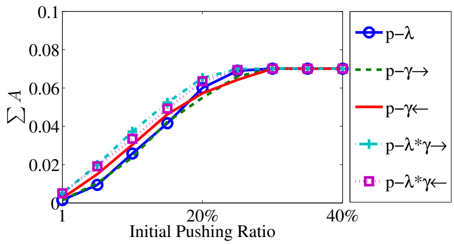
We investigate how $\overrightarrow{p}$ under the 10 pushing strategies impacts the total access utility of all users, $\sum A_i(t)$, with only the MIT trace as shown in Fig. 3.10, and we skip to show the results of other traces since they show very similar trends. The percentage in the figures is C devided by the number of users in each trace. We can see that as the number of initial seeds increases, $\sum A_i(t)$ increases and converges to the maximum. In all cases **p-H** converges to the maximum the fastest, while **p-$\lambda * \gamma^{\rightarrow}$** and **p-$\lambda * \gamma^{\leftarrow}$** as well as **p-$Pr$** perform very well. **p-R** always performs the worst, but **p-$D^{\rightarrow}$** and **p-$D^{\leftarrow}$** also performs poorly. Note that the maximal value of $\sum A$ is capped in different mapping schemes, which means the total user satisfaction is determined by the scenario user nature. The results of different mapping schemes show marginal differences, because TOSS always chooses the users with strong impact strength, and also the access delays provide large space for sharing. In following parts, we will average the evaluation results across the three mapping schemes to reflect various user behaviors in the SNS and



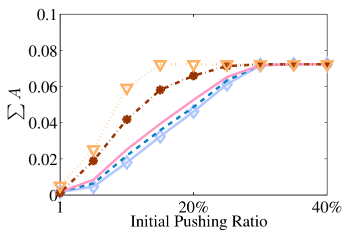
(a) **random - strategies by γ and λ**



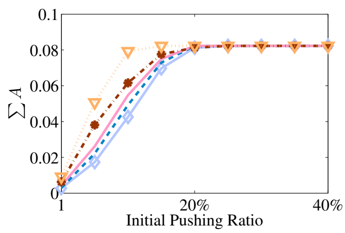
(b) **h-h - strategies by γ and λ**



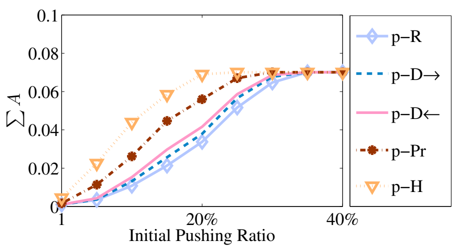
(c) **h-l - strategies by γ and λ**



(d) **random - strategies by graph**



(e) **h-h - strategies by graph**



(f) **h-l - strategies by graph**

Fig. 3.10. As C increases, $\sum A_i(t)$ converges - 3 mappting schemes, 10 pushing strategies - MIT trace as an example

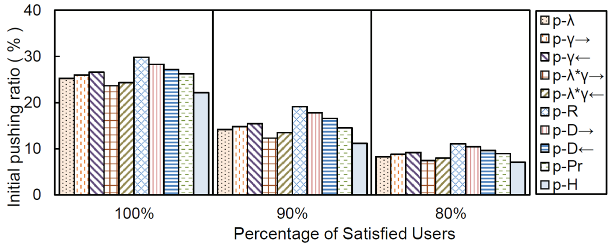
MSN.

3.4.2 Satisfying 100%, 90%, and 80% of Users

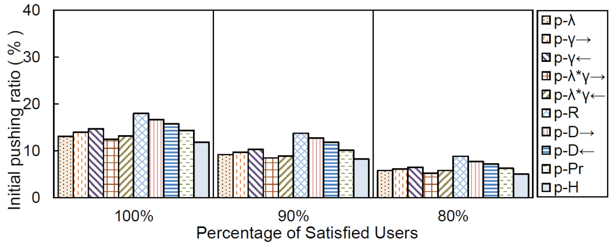
Recall that the access utility function of u_i is $A_i(t)$. A user is **satisfied**, if she can obtain the content when her access probability ($A_i(t)$) approaches its maximum (in the fitting Weibull pdf). If we aim to make 100% of users obtain the content by initial pushing and sharing, substantially large delays may take place for certain users (e.g., a user with low γ and λ values [12]). Therefore, we investigate what percentage of users (initial pushing ratio) should be initial seeds to satisfy the access delay requirements of 100%, 90%, and 80% of users depending on different pushing strategies.

From Sec. 3.2 and Fig. 3.10, $\sum A_i(t)$ is an increasing function of C ($|\overrightarrow{p}|$), and the number of satisfied user is also an increasing function of C. The C value that makes $\sum A_i(t)$ approach its maximum will be the standard number of initial pushing seeds for satisfying 100% of user. We examine how C can be reduced (for higher offloading gains) if we target the satisfaction of 90% and 80% of users.

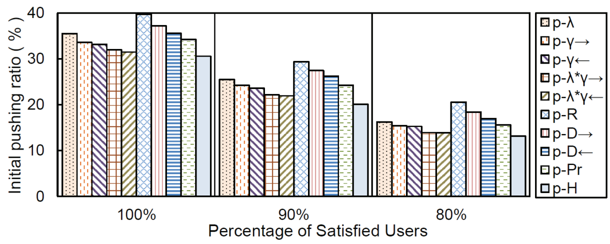
From Fig. 3.11, to satisfy 100% of all users, **p-H** always finds the best initial pushing vector (i.e., the least number of seeds), and **p-R** performs the poorest, while **p-**$D^{\rightarrow}$ and **p-**$D^{\leftarrow}$ also performs poorly, so simply pushing by node degree is not that preferred. In most cases, **p-**$\lambda * \gamma^{\rightarrow}$ and **p-**$\lambda * \gamma^{\leftarrow}$ perform the second best, which implies that we can conjunctively consider the I^S and I^M factor by simple multiplication to achieve near-optimal performance. **p-**Pr achieves not so good performance compared with strategies by impact factors, as it focuses the connections of the network graph but



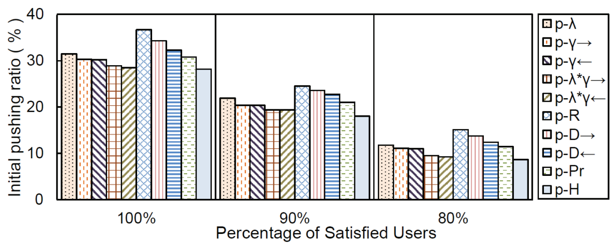
(a) MIT



(b) Infocom



(c) Beijing



(d) SUVnet

Fig. 3.11. Initial pushing ratios to satisfy 100%, 90%, and 80% of all users

ignores the historical spreading impact, while our proposed factors (γ)make better sense. In MIT and Infocom traces, λ-based strategies performance better than γ-base ones, which means the mobility factor decides more on the sharing process when nodes are with high mobility. In Beijing and SUVnet traces, γ-base ones perform better, which means the social factor controls more when nodes are with low mobility. Note that the Infocom trace always has the best performance; only 13.5% initial pushing ratio can satisfy all users by the **p-H**.

When we target to satisfy 90% of all users, the required initial pushing ratio is reduced significantly. With simple pushing strategies, for the MIT and the Infocom traces, only 15.4% and 10.5% of users need to be the initial seeds on average. The number of initial seeds is further dramatically reduced, when satisfying 80% of users. Approximately 10% initial pushing ratio is needed for all traces except the Beijing trace, which requires about 17% initial pushing ratio. The Beijing and SUVnet traces always need relatively larger number of initial seeds due to their low contact rates and large user bases. Also some worse-case users bring ineffectiveness for opportunistic sharing, but it may be better to push the content to them in the beginning, if they have keen access delay requirement, or it will be better to let them to carry out on-demand fetching when they approach the peaks of their access delay PDF.

Generally, **p-H** is about 15-24% better than **p-R**, and 12-16% better than **p-λ** and **p-γ**, and the multiplication of **p-λ** and **p-γ** will be quite a good solution in practical. It is a balance between performance and complexity. The implication is that, if we focus on the best performance, we can run the heuristic algorithm; if we want a balance between complexity and performance, we can evaluate user online spreading impact and offline mobility impact, and choose proper strategy for offloading. **p-R** can still offload certain amount of traffic, which indicates that the sharing-based offloading can work very well in practical actually, because this is mainly due to the potential of the user access delays as discussed in Sec. 3.1.3.1.

3.4.3 On-Demand Delivery

If a user who has not obtained the content (by initial pushing or sharing) until she actually accesses it, we have to deliver it over a cellular link, which is called on-demand delivery. Then the traffic of the content delivered on-demand is not offloaded. We now compare the three target percentages of satisfied users (investigated above) in terms of total offloaded traffic. For example, in the case of 90% of satisfied users, 10% of remaining users (i.e., those who have not received the content) will access the content via cellular links. Table 3.3 shows how much traffic is offloaded from cellular links for the three cases, where the offloaded traffic ratios of the nice pushing strategies are averaged, which are juxtaposed with that of **p-H**. Note that boldfaced numbers are the highest amount of traffic reduction for each trace across the three target satisfaction cases (i.e., 100%, 90% and 80%). When lowering the percentage of satisfied users from 100% to 90% and to 80%, although the initial pushing ratios become reduced, in some cases, the on-demand delivery for abandoned 10% and 20% of users may increase the total cellular traffic instead. In the MIT, Beijing and SUVnet traces, initial pushing for the 90% of users plus on-demand delivery for the 10% of users actually reduces the cellular traffic the most. Overall, TOSS can reduce from 63.8% to 86.5% of the cellular traffic load while satisfying the access delay requirements of all users.

We notice a balance between the traffic reduction due to the initial pushing and the traffic increment by the on-demand delivery, as the satisfaction percentage of users changes. The balance is about how to deal with those worse-case users (with both low online mobility impact and low offline mobility impact). For some of them who have urgent requirement of access delays, TOSS can just push in the beginning, but those who have large access delays will be a burden on selecting the optimal initial pushing seeds by TOSS, as they are hard to reach even by many hops. Instead, it will be better for TOSS to exclude them for a better solution to satisfy a part of other users at the beginning, and then they will carry out on-demand delivery. Note that the Infocom trace can achieve the highest traffic reduction with the target percentage being 100% due to its high contact rates and small user base, and there is very little worse-case users who will not impact the system at all.

Table. 3.3. Percentage (%) of Traffic Reduction With On-Demand Delivery - Simple/Heuristic

Trace	100%	90%	80%
MIT[98]	73.6 / 76.3	**74.6 / 76.9**	70.9 / 72.2
Infocom[99]	**85.3 / 86.5**	79.5 / 80.4	73.4 / 74.1
Beijing[100]	65.3 / 68.4	**65.0 / 68.9**	63.8 / 65.2
SUVnet[101]	68.5 / 70.3	**68.7 / 71.0**	68.3 /70.7

3.5 Conclusion

In this chapter, we proposed the TOSS framework to offload the mobile cellular traffic by leveraging user-to-user local communications, with discussions on the pushing strategies to select the appropriate initial seeds depending on their spreading impact in the online SNS and their mobility impact in the offline MSN. By analyzing the online SNS traces, we learn that a large portion of SNS users have large access delays, which is exploited and utilized for traffic offloading purposes. Trace-driven evaluation reveals that TOSS can reduce from 63.8% to 86.5% of the cellular traffic while guaranteeing the access delay requirements of all users. In particular, users with high mobility impact will play key roles for traffic offloading in scenarios of high user mobility or high user density, and the social spreading impact will then control the content dissemination in scenarios of low user mobility or sparse user density. For worse-case users with both low online and offline impact, it may be better to let them carry out on-demand delivery. Overally, TOSS framework considering both online SNSs and offline MSNs can archieve good performance of disseminating content and offloading traffic efficiently.

Chapter 4

Push-Share: Content Dissemination by Pushing and Sharing in Mobile Cellular Networks - An Analytical Study

4.1 Framework Details

The data explosion problem in mobile cellular networks has become the most critical issue [1]. Mobile network operators (MNOs) seek to mitigate the traffic burden on their cellular links. As the link capacity enhancement in current mobile cellular networks (e.g., 3G and 4G) is unlikely to keep pace with the soaring traffic demand due to limited frequency spectrum, we should investigate this issue from other perspectives.

One of the outstanding trends in the Internet traffic is that increasingly more traffic is attributed to content-oriented applications and services. From this perspective, in addition to the traditional pull-based (request-based) communications, users (or applications) increasingly tend to subscribe to some pushing services from content providers (CPs), and the CPs push the content to subscribers as soon as the content is generated. For instance, the Really Simple Syndication (RSS) is one of the most popular pushing services, by which users can receive the newest photos, documents and video clips. Also YouTube provides some channel-based subscription service to push new and popular videos to users. Many applications in smart phones rely on push mechanisms as well. There are some studies to demonstrate the advantages of push-based models over pull-based models in various contexts (e.g., mission-critical applications [38] and push-to-peer streaming [113]).

From delay perspective, users may not always have to instantly access the content of interest as soon as the content is generated. Instead, some delay is tolerable depending on the users' daily lives and the content natures. For instance, a new music video is generated in the morning, but many people may watch it in the evening or even after some days. Also as reported in [3], when people download content files, there is a substantial disparity in the popularity of the files. That is, only a small portion of content files may be downloaded by a large number of users, which results in multiple users downloading the same content multiple times via cellular links redundantly [3] [4]. Therefore, it is attractive to exploit the affordable delivery delay in such a way that users can receive the content via non-cellular links (e.g., Wi-Fi). For instance, if a user who is to be pushed a file learns that another user who already got the file is nearby, they can "share" the file via Wi-Fi ad hoc connectivity.

From the above observations, we propose **Push-Share** framework to use both "pushing" (over cellular links) and "sharing" over Wi-Fi links (or other local short-range communication techniques) for the content dissemination to subscribers, which can reduce the traffic load on cellular links. The content can be of any type, such as news articles, stock information, advertisements, social events, weather forecasts, and video clips (which currently consumes more than a half of the whole mobile traffic [1]).

We simply illustrate how a file is disseminated in Push-Share by pushing and sharing in Fig. 4.1. Once a file (to which users have subscribed) is generated, the CP sends the file to a dissemination server (DS) in the MNO. The DS is in charge of disseminating the file to the subscribed users until its deadline (or the maximum tolerable delivery delay). The DS will deliver the file to the caching spaces of base stations (BSs), each of which then pushes the file to mobile stations (MSs) of the subscribed users via cellular links. Note that only a subset of MSs will receive the file by the pushing. If an MS with the file opportunistically gets in contact with another nearby MS without the file, they will set up a Wi-Fi connectivity to share the file. The opportunistic content delivery by these "social meets" has been extensively studied in the name of delay tolerant networks (DTNs) [10] [20] [33] [34] [114]. We assume that every MS wakes up periodically with a low duty cycle to probe other MSs nearby for content "sharing" purposes referring to study in eDiscovery [16]. For sake of clarity, we call the direct delivery between MSs "sharing", while "pushing" is for the delivery via cellular links.

Therefore, the focus of **Push-Share** framework is on how to coordinate the pushing and the sharing in the dissemination. Also, by using the multi-compartment model, we discuss how the content is disseminated among multiple cells with handovers. We further formulate an optimization framework for the dissemination performance, and explore the trade-off between the energy cost and dissemination delay. To the best of our knowledge,

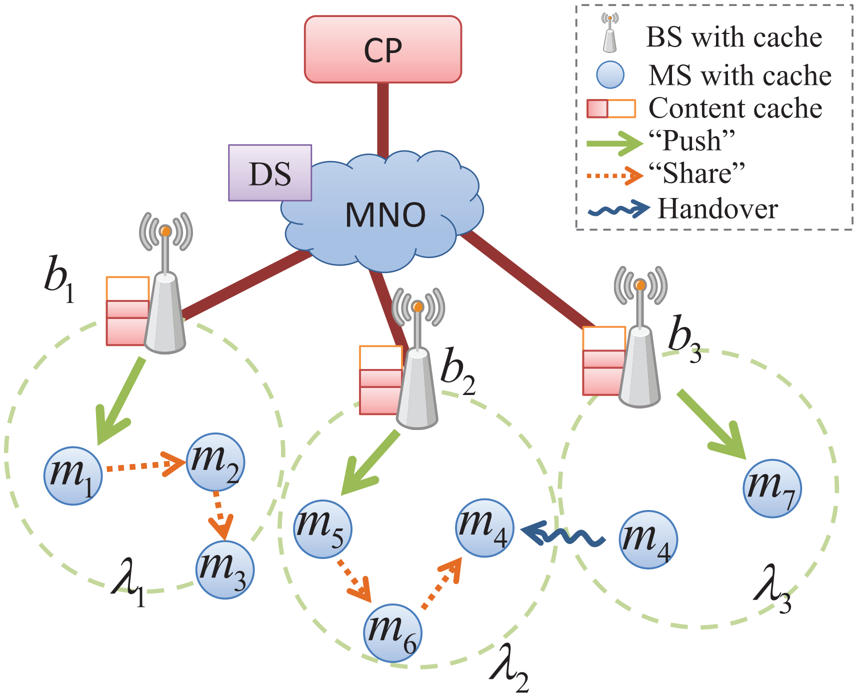


Fig. 4.1. Illustration of content dissemination by pushing via cellular links and by sharing via Wi-Fi links among MSs with handovers

Push-Share framework is the first study to theoretically model and analyze the content dissemination across multiple cells in cellular networks based on pushing and sharing.

4.2 System Model

We illustrate a dissemination scenario in Fig. 4.1, where there is one CP and one MNO with three BSs, b_1, b_2 and b_3. Each BS services multiple MSs in its cell who are interested in the CP's content. For example, m_1, m_2 and m_3 are within the cell of b_1. Hereby, we only focus on those MSs who have

Table. 4.1. Variables and notation of the system model

Variable	Explanation (default value in evaluation)
b_i	BS with id i
n	number of total BSs (20)
M_i	number of MSs in the area of b_i (1000)
m_k	a typical MS with index k
λ	average meeting rate of MSs in the cell (0.00001)
ϕ	energy consumption per delivery via Wi-Fi (1)
Φ	energy consumption per delivery via cellular link (4)
ρ	probing cost per time unit (0.001)
P_{init}	the amount of initial push (50)
P_{final}	the amount of final push (50)
$S(t)$	the function of number of updated MSs in the cell to time t
t^O	dissemination completion time with only pushing
t^*	dissemination completion time with both sharing and pushing
$C(t)$	accumulative cost function of MSs in the cell to time t
C^*	cost to disseminate content to all MSs in the cell
ℓ_{xy}	handover rate of MSs from BS b_x to BS b_y

subscribed to the content from the CP, and thus ignore other ones. Even though a single CP may disseminate multiple files to the MSs periodically or concurrently, we focus a single file for sake of exposition. Also we do not consider MSs who may turn off during the dissemination. The notation and the default values are shown in Table 4.1.

As for "pushing", the CP first delivers a file for a particular group of MSs(its group identifier is needed) to the DS of the MNO. In a cellular network, the location management entity (LME) [83] keeps track of the locations of the MSs. Thus, along with the LME, the DS knows: (i) which MSs have subscribed the content, and (ii) which MSs of the group are serviced by each BS. Then the DS will dispatch the file to all the BSs that service the MSs. Each BS will initially push the content file to some of the MSs in its cell. For instance, BS b_1 will deliver the content to m_1 at the beginning. There can be different strategies regarding which MSs will be pushed first, but this is out of our scope (see [10] and [23] for details). Here by I deploy a random strategy which we will describe later in Sec. 4.3.2.

As for "sharing", MSs will move with a certain mobility model. According to [10] [20] [33] [34] and [114], the intervals between consecutive meets of any pair of MSs, called the inter-contact times (ICTs), are assumed to follow an exponential distribution. Also based on the measurements in [33], we assume that MSs at different places will have different mobility patterns and thus MSs at different BSs will have different mean rates of inter-contacts, denoted as λ_i for BS b_i, (also called meeting rate interchangeably). For instance, a park area will have a longer ICT than a subway station. Each MS periodically probes to check whether there is any nearby MS that holds the content being disseminated. We assume the MSs are synchronized and the probing is triggered with a sufficiently low duty cycle, say, during the first 5ms period in every second. (The energy consumption per time unit due to probing is denoted by ρ.) If there is, two MSs will share the content via ad-hoc Wi-Fi connectivity. For instance, m_1 occasionally meets m_2 and shares the content, and later m_2 meets and shares with m_3. If an MS obtained the content by either pushing or sharing, we say the MS is "updated".

Some MSs do not like to participate in carrying and sharing content with others due to security, privacy or cost issues. We will exclude those MSs from the model. Note that related security and privacy issues in sharing can be handled by some prior work in opportunistic DTN such as [48] [49] and [50]. Also since the focus is to model and analyze how the content can be disseminated across multiple cells in a macro perspective, we assume that the content can be shared successfully via Wi-Fi during the meets with fairly high bit rates.

With the initial pushing and sharing, some MSs may not be able to obtain the content for a long time due to the limitation of the opportunistic sharing. Those MSs will request the final push from the BS, to be detailed later in Sec. 4.3.2.

4.3 Content Dissemination in Single Cell

In this section, we discuss the content dissemination within a single cell. For simplicity, we temporarily assume that for a certain amount of duration, the MSs will stay in a single cell and will not make handovers. We will discuss the case of multiple cells with handovers in Sec. 4.4. As we consider a single BS in this section, we omit the BS's index i.

4.3.1 Content Dissemination by Sharing Only

We first focus on how the content is gradually disseminated to MSs over time t in a single cell by sharing only, where the number of MSs who are to receive the content is denoted by M. Let $S(t)$ be the state of the continuous-time Markov chain system, which indicates the number of MSs that have received the content until time t by sharing. We will obtain $S(t)$ from its derivative based on the similar methodology as used in [13] and [20]; thus we show only the main steps for the sake of simplicity.

Due to the synchronized probing among MSs, an MS, say m_k, will always be able to discover other MSs in the Wi-Fi range. During a short period, say Δt, the probability for m_k to get the content from any MS who already got the content within Δt, denoted by $\theta_{t,t+\Delta t}(m_k)$, can be calculated by

$$\theta_{t,t+\Delta t}(m_k) = 1 - \left(1 - \left(1 - e^{-\lambda \Delta t}\right)\right)^{S(t)}. \tag{4.1}$$

Then summing this probability across all the MSs that have not received the content at time t, the current number of updated MSs after Δt, $S(t+\Delta t)$, can be calculated as

$$S(t+\Delta t) = S(t) + \sum_{k=1}^{M-S(t)} \theta_{t,t+\Delta t}(m_k), \tag{4.2}$$

whose expectation is given by

$$E\left[S\left(t+\Delta t\right)\right]=E\left[S\left(t\right)\right]+\left(M-E\left[S\left(t\right)\right]\right)\cdot E\left[\theta_{t,t+\Delta t}\left(x\right)\right] \tag{4.3}$$

We obtain the derivative of $E\left[S\left(t\right)\right]$ by letting $\Delta t\to 0$,

$$\begin{aligned}
&E\left[\overset{\bullet}{S}\left(t\right)\right]=\lim_{\Delta t\to 0}\tfrac{E\left[S\left(t+\Delta t\right)\right]-E\left[S\left(t\right)\right]}{\Delta t}\\
&=\lim_{\Delta t\to 0}\tfrac{\left(M-E\left[S\left(t\right)\right]\right)\cdot\left(1-\left(1-\left(1-e^{-\lambda\Delta t}\right)\right)^{S\left(t\right)}\right)}{\Delta t}\\
&=\left(M-E\left[S\left(t\right)\right]\right)\left(\lambda\cdot E\left[S\left(t\right)\right]\right).
\end{aligned} \tag{4.4}$$

By solving the above ordinary differential equation (ODE), we finally obtain the function $S(t)$ by

$$S\left(t\right)=\frac{S\left(0\right)Me^{M\lambda t}}{M-S\left(0\right)+S\left(0\right)e^{M\lambda t}}. \tag{4.5}$$

Note that if there is no MS who has the content at the beginning, i.e., $S(0)=0$, $S\left(t\right)$ will always be zero. Therefore, the BS should push the file to at least one MS, i.e., $S(0)=1$, and then the MS with the content will disseminate the file to other MSs by sharing. Thus, $S(t)$ starting with only a single seed will increase by sharing over time as

$$S\left(t\right)=\frac{Me^{M\lambda t}}{M-1+e^{M\lambda t}}. \tag{4.6}$$

From Eq. (4.5) we can calculate the required delay, t_{req}, to disseminate the content to S_{des} MSs ($1 \leq S_{des} \leq M$) by,

$$t_{req} = S^{-1}\left(S_{des}\right) = \frac{\log\left(\frac{S_{des}(M-S(0))}{S(0)(M-S_{des})}\right)}{M\lambda}. \tag{4.7}$$

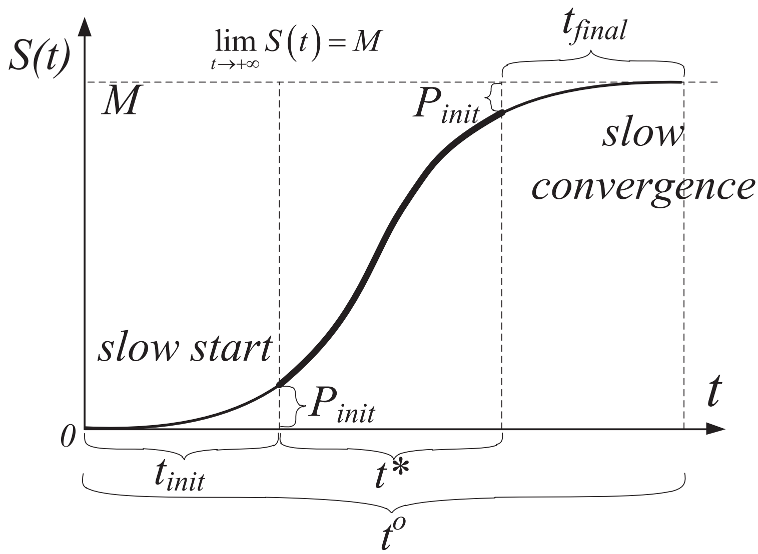
$S(t)$ from Eq. (4.5) in real domain cannot reach M in a finite time, which means $\lim_{t\to+\infty} S(t) = M$, and thus the dissemination completion time with only sharing, denoted by t^O, would be $t^O = S^{-1}(M) = +\infty$. However $S(t)$ actually takes integer values, so we define that the dissemination will be completed when $S(t) = M - \eta$, where η, $0 < \eta \ll 1$, takes a sufficiently small value, (e.g., $\eta = 1$). Then,

$$t^O = S^{-1}(M) \approx S^{-1}(M-\eta) = \frac{\log\left(\frac{(M-\eta)(M-1)}{\eta}\right)}{M\lambda}. \tag{4.8}$$

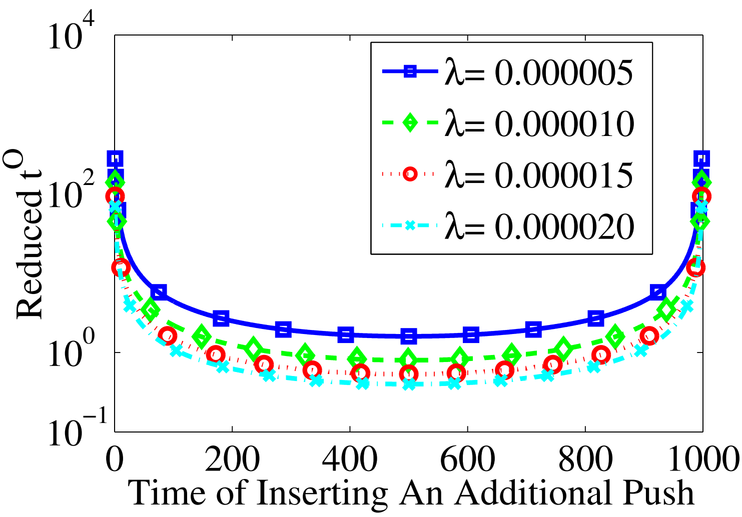
4.3.2 Content Dissemination with Initial Push and Final Push

We illustrate $S(t)$ from Eq. (4.5) in Fig. 4.2(a), and we observe that the content dissemination only by sharing (starting with only a single seed) suffers from both a slow start and a slow convergence, due to the limitation of the opportunistic sharing. Therefore, we propose to increase the number of MSs who receive the content from the BS to reduce the delay..

In order to investigate "when" the BS should push the content for efficient dissemination, we evaluate how much dissemination completion time is reduced by pushing the content to one more MS at an arbitrary time (X-axis), as shown in Fig. 4.2(b). We observe that the additional pushing at the



(a) Content dissemination curve by sharing with a single seed



(b) How much can one more push reduce delay? ($M = 1000$)

Fig. 4.2. How much can one more push accelerate the content dissemination?

beginning and at the end can reduce the dissemination completion time most compared.

Therefore, we propose to disseminate content by three steps: (1) the BS pushes the file to a certain number of MSs at the beginning, denoted by P_{init}, which is actually $S(0)$ in Eq. (4.5); (2) the MSs will share the content file via opportunistic meeting; (3) when most of the MSs have received the content and there are still P_{final} not-yet-updated MSs, the BS finally pushes the file to them.

How to choose which MSs appropriately for initial pushing is out of the scope (see related work in [10] and [23]). Here we use a random strategy as follows: for each BS, the DS will calculate the optimal number of initial pushing P_{init} based on the environments in each cell (refer to the optimization framework in Sec. 4.5), and send the file to the BS, along with the ratio of $\frac{P_{init}}{M}$, the interest identifier and the dissemination deadline. Then each BS broadcasts a short message containing the information, and each MS who is interested in the content will reply to the BS with the probability of $\frac{P_{init}}{M}$ to confirm the initial pushing. In this way, the BS can push the file to P_{init} MSs probabilistically. At the deadline, the MSs who have not obtained the content will ask the BS to push the content to them finally. Each BS does not need to track the status of each MS and the dissemination progress.

Therefore, given the estimated P_{init} and P_{final}, the time to push the content to all the P_{final} MSs who have not received the content, denoted by t^*, is when the number of updated MSs $S(t)$ becomes $M - P_{final}$. Thus t^* becomes the practically dissemination completion time with both pushing and sharing. Based on Eq. (4.7), we have

$$t^* = S^{-1}\left(M - P_{final}\right) = \frac{\log\left(\frac{\left(M - P_{final}\right)\left(M - P_{init}\right)}{P_{final} P_{init}}\right)}{M\lambda}. \tag{4.9}$$

Finally the content dissemination function $S(t)$ becomes a piece-wise function as follows,

$$S(t) = \begin{cases} P_{init} & t = 0, \\ \frac{P_{init} M e^{M\lambda t}}{M - P_{init} + P_{init} e^{M\lambda t}} & 0 < t < t^*, \\ M & t^* \leq t. \end{cases} \tag{4.10}$$

4.3.3 Content Dissemination Energy Cost

The energy consumption is a critical issue for mobile networks because of the limited power supply of mobile devices. We mainly discuss the energy consumed at MSs for the content dissemination, which consists of:

- **Probing**: MSs periodically wake up with a sufficient duty cycle to detect whether there are nearby MSs with the content. We use ρ to denote the energy cost per time unit for probing, which is much smaller than those of receiving the content via a cellular link and sharing the content via a Wi-Fi link.

- **Pushing via cellular link**: We use Φ to denote the energy cost for receiving a file by BS's unicast via a cellular link.

- **Sharing via Wi-Fi link**: We use ϕ to denote the energy cost for transmitting and receiving a file from one MS to another via a Wi-Fi link.

In practical, transmitting and receiving may consume different energy cost, but as they are will be just constants in our model, we hence assume the same value of them for simplicity, which will not affect our modeling. Thus the sharing of a file by Wi-Fi will cost 2ϕ. From the measurements in [115] and [116], Φ is greater than ϕ, and both are greater than ρ.

Therefore the accumulative energy cost for all MSs until time t can be then derived from Eq. (4.10) as follows:

$$C(t) = \begin{cases} \Phi P_{init} & t = 0, \\ \Phi P_{init} + 2\phi\left(S(t) - P_{init}\right) + M\rho t & 0 < t < t^*, \\ \Phi\left(P_{init} + P_{final}\right) + 2\phi\left(M - P_{init} - P_{final}\right) + M\rho t^* & t^* \leq t. \end{cases} \tag{4.11}$$

And after t^*, the energy cost for dissemination completion, denoted by C^*, can be calculated as

$$C^* = \Phi\left(P_{init} + P_{final}\right) + 2\phi\left(M - P_{init} - P_{final}\right) + M\rho t^*. \tag{4.12}$$

4.4 Content Dissemination in Multiple Cells

If we consider a number of BSs covering a large area, we should model the handovers among the cells, which strongly affect the content dissemination collectively. For instance, a BS covering a subway station will have many incoming and outgoing handover MSs, which either have the content or not.

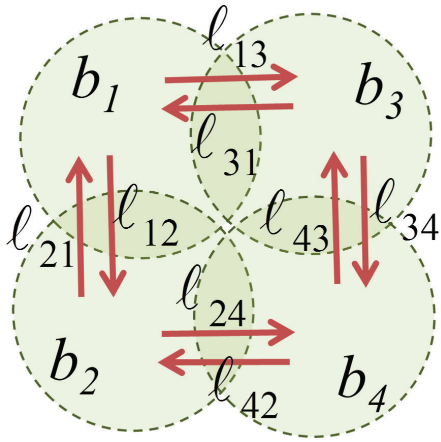
Thus we propose to adopt the multi-compartment model [117] to describe the content dissemination in a multi-cell scenario with handovers, based on the assumption that MSs' handovers follow a certain random process [118] [119].

The multi-compartment model is commonly used in the biology fields (e.g., pharmacokinetics and biomedicine) to investigate the density of materials (e.g., drugs) in blood among different cells or parts of the organism, called compartments, and to track how the blood with the materials is circulating among compartments with some transition rates [117] [120]. These transitions from one compartment to another are similar to the handovers of MSs.

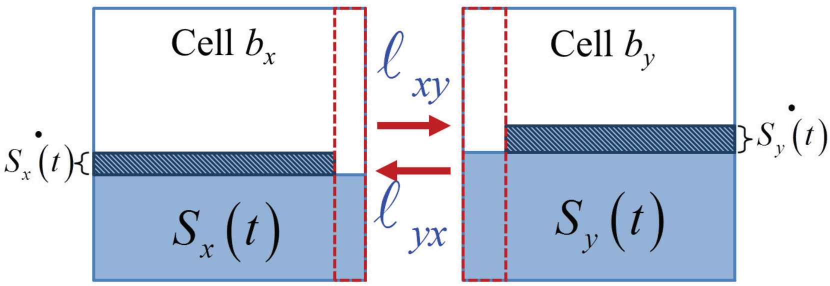
There have been some related studies for modeling handovers as a random process in [118] [119], [121] [122] and [123]. According to these studies, the cell dwell time of an MS statistically follows a certain probability distribution (e.g., exponential distribution). We use the average rate of the random process model to represent the handover rate.

In an example scenario in Fig. 4.3(a), there are four cells b_1, b_2, b_3 and b_4, and between two neighbor BSs, the MSs are performing handovers in or out with a certain rate, denoted by ℓ_{xy}, which is defined as the probability that an MS moves from BS b_x to another BS b_y during a time unit. Note that handover rates can be obtained or estimated based on practical measurements by BSs and the LME in the MNO.

The multi-compartment model is based not only on the handover rates but also on the number of the MSs, M_i, at each BS, b_i. We consider two kinds of scenarios for handovers to calculate M_i: (a) **non-steady-state scenario**,



(a) A multi-cell scenario



(b) How does the handover affect the content dissemination?

Fig. 4.3. Modeling the handovers in the content dissemination

where the number of the MSs at each cell dynamically changes, for instance, a BS in a residential area during commuting time; (b) **steady-state scenario**, where the number of MSs at each cell can be assumed to be unchanged if the incoming handovers and outgoing handovers balance.

We also define the neighborhood set, Ω_i, which includes all neighbor BSs of b_i, for instance, $\Omega_1 = \{b_2, b_3\}$ in Fig. 4.3(a).

4.4.1 Non-steady-state Modeling of MSs in Multiple Cells

In non-steady-state scenarios, M_i of each BS b_i is dynamically changing; thus we use function $M_i(t)$ since M_i is changing over time t. Therefore, in a short period, its derivative, $\overset{\bullet}{M}_i(t)$, can be calculated based on the difference between incoming MSs and outgoing MSs as follows,

$$\overset{\bullet}{M}_i(t) = -M_i(t) \sum_{b_k \in \Omega_i} \ell_{ik} + \sum_{b_k \in \Omega_i} (\ell_{ki} M_k(t)). \tag{4.13}$$

Thus, for the n BSs, there will be n equations, which formulate a 1st-order linear homogeneous ODE system. Referring to [124] and [125], the general solution is given by

$$M_i(t) = \sum_{z=1}^{n^2} A_z e^{B_z(t-C_z)}, \tag{4.14}$$

where the coefficients A_z, B_z and C_z are coefficient constants that can be calculated straightforward, but we will skip the details due to space limit (see related work in [124] and [125]).

4.4.2 Steady-State Modeling of MSs in Multiple Cells

When the BSs are in a steady-state, the incoming and outgoing MSs practically make no change to the number of MSs at each BS. Then $M_i(t)$ of any

BS b_i will be fixed to a static number M_i, which simplifies Eq. (4.13) to,

$$-M_i \sum_{b_k \in \Omega_i} \ell_{ik} + \sum_{b_k \in \Omega_i} (\ell_{ki} M_k) = 0. \tag{4.15}$$

Therefore, n BSs will generate a linear system with n equations, which can be easily solved to get M_i of each BS in the steady-state scenario.

4.4.3 How Handovers Affect the Content Dissemination

From the previous two subsections, we obtain the number of MSs at each cell in either non-steady-state or steady-state scenario. Thus, along with the known handover rates, we analyze how the handovers affect the content dissemination among cells based on the multi-compartment model. Note that we change $S(t)$ to $\mathbb{S}(t)$ to describe the dissemination function with handovers in multi-cell scenarios.

As illustrated in Fig. 4.3(b), at an arbitrary time t, there are $\mathbb{S}_x(t)$ updated MSs in cell b_x and $\mathbb{S}_y(t)$ updated MSs in cell b_y, which are represented by the light blue solid rectangles. Then during a short period, there will be two types of MSs in the cell: (a) MSs who are performing handovers; (b) MSs who are sharing the content. Note that we assume that during the period, (a)-type MSs will not share the file, and (b)-type MSs will not perform handovers. Then the red dashed rectangles represent the (a)-type MSs, who move from one cell to another, and the dark blue shadowed rectangles represent the newly updated MSs during the period shared by (b)-type MSs.

In the non-steady-state scenario, considering those two types of MSs, the derivative function of $\mathbb{S}_i(t)$ of BS b_i can be extended based on Eq. (4.4) as follows,

$$\begin{aligned}
&\overset{\bullet}{\mathbb{S}}_i(t) = \\
&(M_i(t) - \mathbb{S}_i(t)) \left(1 - \sum_{b_k \in \Omega_i} \ell_{ik}\right) \left(\lambda \left(1 - \sum_{b_k \in \Omega_i} \ell_{ik}\right) \mathbb{S}_i(t)\right) \\
&- \sum_{b_k \in \Omega_i} \ell_{ik} \mathbb{S}_i(t) + \sum_{b_k \in \Omega_i} (l_{ki} \mathbb{S}_k(t)).
\end{aligned} \tag{4.16}$$

And for the steady-state scenario, the $M_i(t)$ becomes M_i.

Finally, there will be a complicated ODE system with n differential equations for modeling the content dissemination with both pushing and sharing in multi-cell scenario.

In the steady-state scenario, the number of MSs at each BS is constant; thus, the above ODE system is a 1st-order quadratic homogeneous ODE system with constant coefficient, which is a Riccati type matrix differential equation system. Jodar et al. [126] discussed its closed analytical approximation solution. Also Darling [127] proposed to convert the Riccati matrix different equations to 2nd-order linear ODE system to obtain explicit solutions. In non-steady-state scenario, the ODE system becomes a 1st-order quadratic homogeneous ODE system with variable coefficients, which is difficult to obtain its exact analytical solution, but can be approximated by the power series methodology (see [85]). Furthermore, the homotopy perturbation method can be also applied to obtain the approximation of $\mathbb{S}(t)$ (see [84]). Due to the limited space, we skip the details of the solving procedure.

Regarding the energy cost for content dissemination in multiple cells with handovers, based on the above $\mathbb{S}_i(t)$ in Eq.(4.16), we can also easily extend $C_i(t)$ in Eq. (4.11) and C_i^* in Eq. (4.12), which are denoted by $\mathbb{C}_i(t)$ and $\mathbb{C}_i^*$, respectively.

4.5 Optimization Framework

From previous modeling of the content dissemination in a single cell and multiple cells with handovers, we discuss the optimization framework for the DS in the MNO to allocate the P_{init} and P_{final} to all BSs, in order to achieve the minimum dissemination completion time and energy cost.

4.5.1 Minimum Dissemination Completion Delay

From Fig. 4.2(a), the effective allocation of the number of initial pushing and final pushing becomes critical for accelerating the content dissemination procedure for a shorter completion time. Then the problem becomes that, at any BS, by given a specific upper bound of the number of MSs that are going to be pushed, P_{total}, how to find the optimal values of P_{init} and P_{final} to achieve the minimum dissemination completion time t^* referring to Eq. (4.9):

$$\begin{aligned} &\min_{P_{init},P_{final}} \{t^*\} \\ &\text{Subject to :} \quad P_{init} + P_{final} = P_{total}. \end{aligned} \tag{4.17}$$

We replace P_{final} by $P_{final} = P_{total} - P_{init}$, and find the minimum value by letting $\frac{\partial t^*}{\partial P_{init}} = 0$, so that the optimal value of P_{init} is found as $P_{init} = \frac{P_{total}}{2}$, which means that the BS should always equally allocate the number of initial pushing and that of final pushing so that the dissemination completion time t^* can be minimized, regarding a limited total number of pushing.

Therefore, in the rest, we will just focus on the number of initial pushing, P_{init}, and consider $P_{final} = P_{init}$ by default. Note that the values of P_{init} and P_{final} should be less than the $\frac{M}{2}$.

4.5.2 Minimum Dissemination Completion Cost

Referring to the measurements in [115] and [116], Φ is several times larger than ϕ for one content delivery. If a BS pushes the content to more MSs via the cellular link in order to get a smaller t^*, it may consume more energy; otherwise if a BS pushes to less MSs inducing a larger t^*, it may also consume a large amount of probing energy over time. Thus we have the problem on how to find the optimal value of P_{init} to minimize the energy cost for completing the dissemination as follows,

$$\min_{P_{init}} \{C^*\}. \tag{4.18}$$

Based on Eq. (4.12), we use the similar method in the previous subsection to solve $\frac{\partial C^*}{\partial P_{init}} = 0$, and find the optimal P_{init} for the minimum C^* as,

$$P_{init} = \frac{M}{2} - \frac{\sqrt{M\lambda\left(M\lambda\left(\Phi-2\phi\right)-8\rho\right)\left(\Phi-2\phi\right)}}{2\lambda\left(\Phi-2\phi\right)}, \tag{4.19}$$

under the condition of,

$$\left(M\lambda\left(\Phi-2\phi\right)-8\rho\right)\left(\Phi-2\phi\right) \geq 0. \tag{4.20}$$

Then we can obtain the minimum C^* referring to Eq. (4.12). When the condition in Eq. (4.20) equals to or less than 0, the optimal P_{init} with $P_{init} = \frac{M}{2}$

will lead to the minimum C^*.

In the multi-cell scenario, each BSs, b_i, can locally calculate the optimal P_{init_i} to minimize the energy cost $\mathbb{C}_i^*$, unless there is a limitation on the total number of the MSs being pushed among all BSs, P_{budget}, which is smaller than the sum of the local optimal values of P_{init_i}, that is

$$\sum_{\forall b_i} P_{init_i} < P_{budget} < \sum_{\forall b_i} \left(\arg\min_{P_{init_i}} \mathbb{C}_i^* \right). \tag{4.21}$$

With this constraint, the local optimization for each cell will not guarantee the minimum energy cost among all BSs. So the problem extends as

$$\begin{aligned} &\min_{\vec{P}_{init}} \left\{ \sum_{\forall b_i} \mathbb{C}_i^* \right\} \\ &\text{subject to}: \sum_{\forall b_i} P_{init_i} < P_{budget}. \end{aligned} \tag{4.22}$$

It is hard to verify the convexity of $\mathbb{C}^*$ to P_{init}. So we will firstly approximate the above objective function based on the power series methodology (see [85]), and then carry out numerical analysis.

4.5.3 Conjunctive Minimization of Delay and Cost

Because CPs, MNOs and MSs all desire for both minimum delay t^* and cost C^*, we try to carry out overall optimization on both of them. Due to the different unit of time and energy, we bring a weight factor, w, to combine t^* and C^* conjunctively, which is also considered as the Pareto-optimality , and w indicates the emphasis on either the cost or delay. Thus, for a single cell, we have the following minimization problem:

$$\min_{P_{init}} \{t^* + w \cdot C^*\}. \tag{4.23}$$

We solve it by letting $\frac{\partial (t^* + w \cdot C^*)}{\partial P_{init}} = 0$, and then obtain the solution as:

$$P_{init} = \frac{M}{2} - \frac{\sqrt{\lambda (M^2 \lambda (\Phi - 2\phi) - 8M\rho - 4w)(\Phi - 2\phi)}}{2\lambda (\Phi - 2\phi)}, \tag{4.24}$$

with a condition that

$$\left(M^2 \lambda (\Phi - 2\phi) - 8M\rho - 4w\right)(\Phi - 2\phi) \geq 0. \tag{4.25}$$

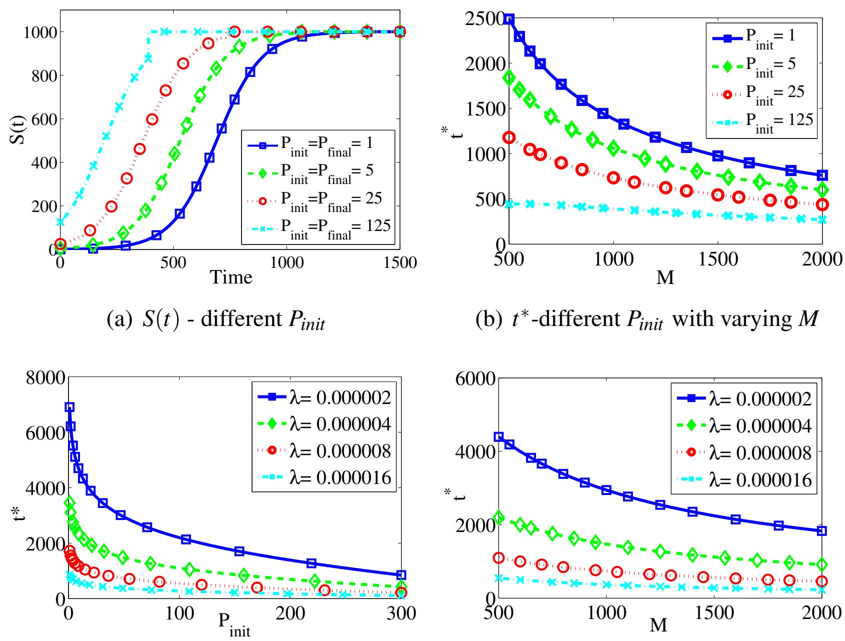
For the multi-cell scenario, if there is a constraint on the total amount of pushing, P_{budget}, the same to Eq. (4.21), each BS cannot push as it wants for local optimality between the delay and the cost; instead, all BSs must endeavor for the global optimization for following problem,

$$\begin{aligned} &\min_{\vec{P}_{init}} \left\{ \sum_{b_i} (t_i^* + w \cdot \mathbb{C}_i^*) \right\} \\ &\text{subject to} : \sum_{\forall b_i} P_{init_i} < P_{budget}, \end{aligned} \tag{4.26}$$

which is hard to find the minimum value by open-form solutions. Similar to previous subsection, we carry out approximation on the objective function for just numerical results.

4.6 Evaluation Results

We simulate the continuous-time Markov system of our proposed model in Mathematica 8 [102] along with the support of MATLAB 2010 [82] and Maple 14 [128]. For the purpose of evaluating the model realistically, we set the parameters with reasonable values based on previous mobility work in [114]: the meeting rate λ_i among MSs is from 0.0000001 to 0.0001 per second, and the number of MSs under one BS, M_i, is within the range from 300 to 3000. Also referring to [115] and [116], we set $\phi = 1$, $\Phi = 4$, and $\rho =$ 0.001 per second by default. Note that the default values of the parameters are shown in Table 1.



(a) $S(t)$ - different P_{init}

(b) t^*-different P_{init} with varying M

(c) t^*-different λ with varying P_{init}

(d) t^*-different λ with varying M

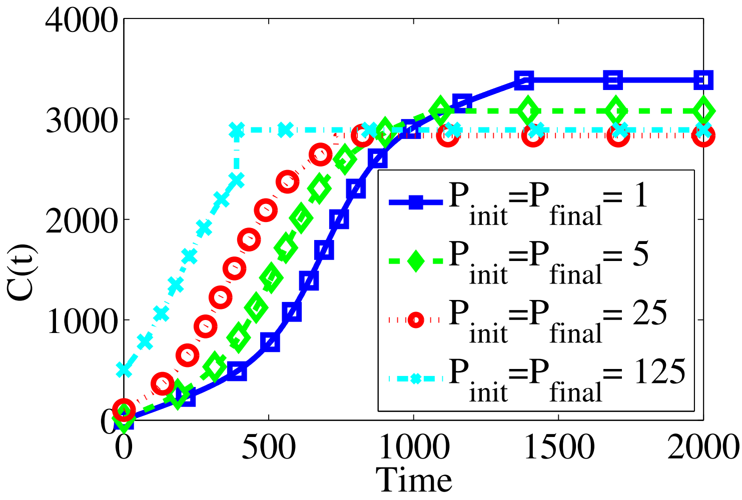
Fig. 4.4. Evaluation on $S(t)$ and t^*

4.6.1 Content Dissemination within One Single Cell

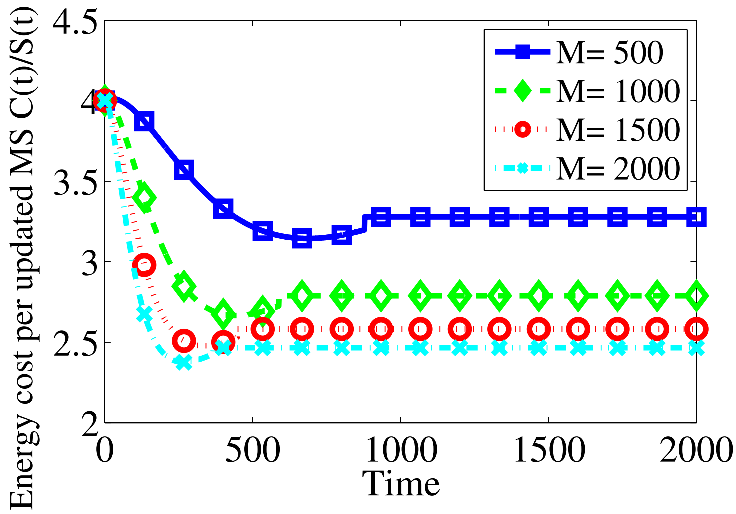
The evaluation of the dissemination function $S(t)$ in Eq. (4.10) and the completion time t^* in Eq. (4.9) are shown in Fig. 4.4. From Fig. 4.4(a), when there is only one push ($P_{init} = 1$) at the beginning, the number of updated MSs starts to grow slowly, and converges to the dissemination completion slowly as well. When we increase the value of initial pushing, P_{init}, the dissemination procedure can be greatly shortened.

Regarding the completion time t^*, we observe that a cell with a small number of MSs will suffer from a large t^*, but a larger value for initial pushing P_{init} can reduce t^* dramatically as shown in Fig. 4.4(b). This indicates that when adjusting the values of initial pushing for the BSs, it is more beneficial to push more copies to small cells from the perspective of dissemination completion time. Fig. 4.4(c) and 4.4(d) both show that larger values of λ and M can significantly accelerate the dissemination and thus shorten t^*, because larger λ and M mean the higher probability that the MS can meet other MSs and thus be able to get the content by sharing. However, the benefit of increasing P_{init} is not significant when the meeting rate is high, as shown in Fig. 4.4(c).

The accumulative energy cost function of $C(t)$ in Eq. (4.11) is evaluated as shown in Fig. 4.5. From Fig. 4.5(a), we observe that the value of P_{init} has two-side impact on the $C(t)$: a small P_{init} ($P_{init} = 1$) will induce a long completion time, but the probing will consume a lot and thus $C(t)$ becomes quite large; however a large value of P_{init} ($P_{init} = 125$) can reduce t^* dramatically, but because of the more expensive energy cost for cellular



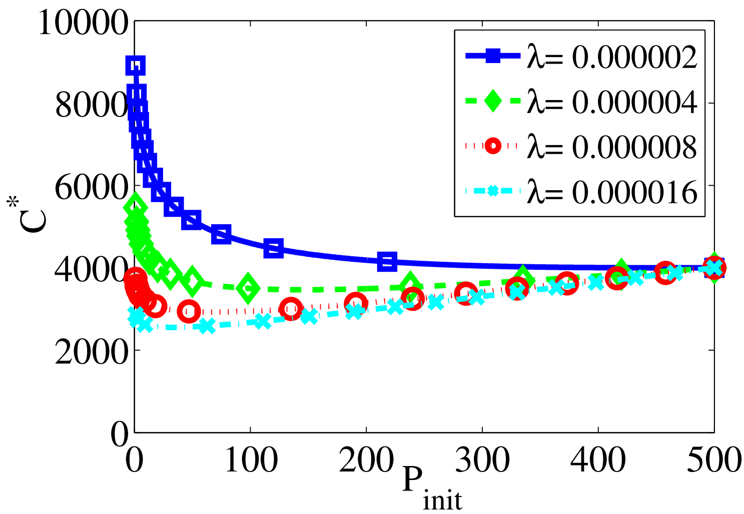
(a) Different P_{init}



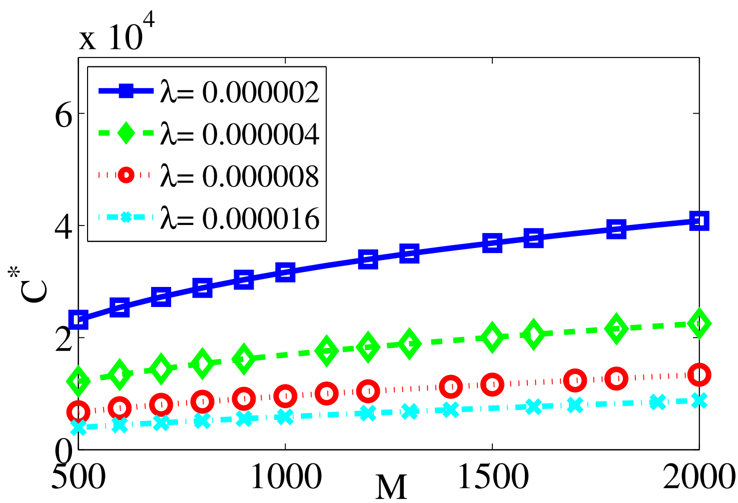
(b) Different M regarding $C(t)/S(t)$

Fig. 4.5. Evaluation of $C(t)$ - accumulative energy cost by time

links, it still consumes more C^* than that when $P_{init} = 25$. This falls into the optimization framework on C^* in Sec. 4.5.2, which we will discuss in later paragraphs. Furthermore, we calculate the energy cost per updated MS over time as shown in Fig. 4.5(b), and we discover that a large group will actually reduce the energy cost for each individual MS due to sharing.



(a) Different λ with varying P_{init}



(b) Different λ with varying M

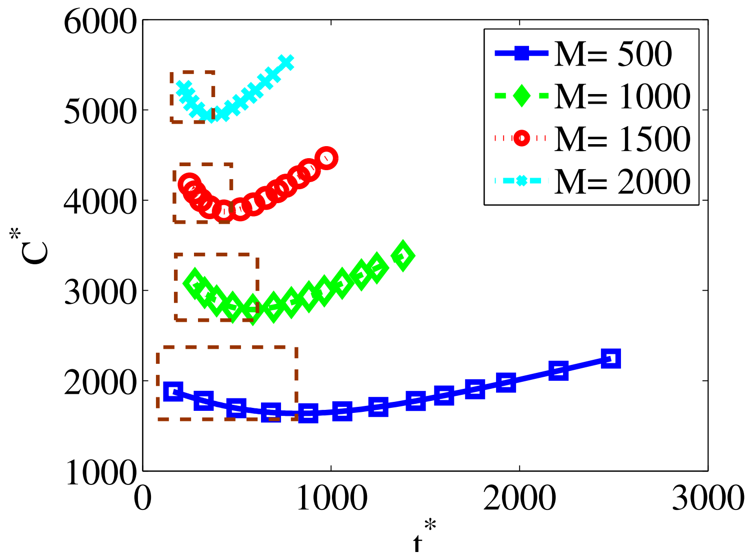
Fig. 4.6. Evaluation of C^* - the energy cost for dissemination completion

The evaluation on the energy cost for dissemination completion C^* in Eq. (4.6) is shown in Fig. 4.6. The relationship between P_{init} and C^* in Fig. 4.6(a) reflects our optimization framework in Sec. 4.5.2; P_{init} can be adjusted for a minimized C^* under the condition in Eq. (4.20). In the case that the condition is not satisfied ($\lambda = 0.000002$ in Fig. 4.6(b)), the optimal P_{init} for minimizing C^* will be $\frac{M}{2}$. Also when λ is larger, the optimal P_{init} for minimum C^* is smaller. And also from Fig. 4.6(b), a higher meeting rate λ means more frequent social sharing via Wi-Fi, and it can significantly reduce the C^*, due to the lower energy cost of Wi-Fi links.

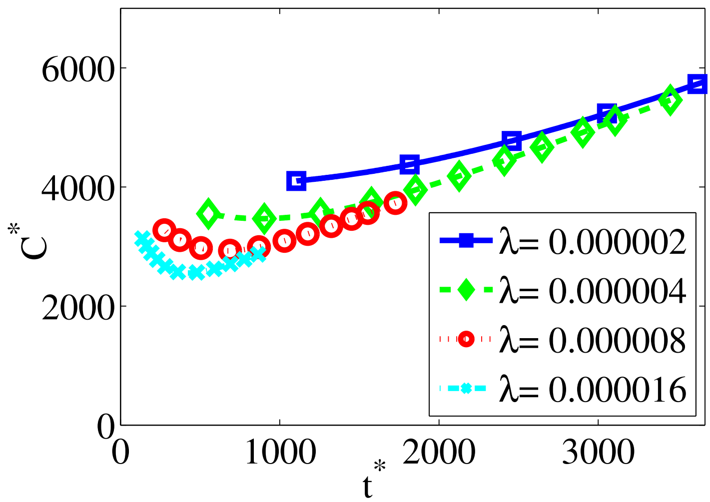
The trade-off between C^* and t^* is explored in Fig. 4.7(a) and 4.7(b), when we adjusting P_{init} with different numbers of MSs and meeting rates. there is always a valley in the C^*-t^* curve, where C^* gets minimized (referring to Eq. (4.18)). The part of the curve on the left of the valley, where C^* and t^* are in an inverse relation, defines the boundary of the achievable delay-energy region (emphasized within the dashed rectangles) when P_{init} is higher than $\arg\min_{P_{init}}\{C^*\}$. This reflects the Pareto-optimal between C^* and t^* discussed in Sec. 4.5.3, and depending on the weight factor w, it is easy to find an optimal balance between C^* and t^* within the rectangle areas. On the right part of the curve, when P_{init} is not sufficiently large, the system will suffer from both high energy cost and long dissemination completion time.

4.6.2 Content Dissemination within Multiple Cells

For investigating the content dissemination in multi-cell scenario with handovers, we evaluate the MNO network in Fig. 4.3(a) as a typical example. At the beginning, for M_i and λ_i for each cell, we randomly assign practi-



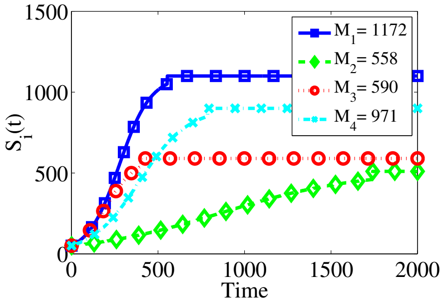
(a) Balance between C^* and t^* with different M



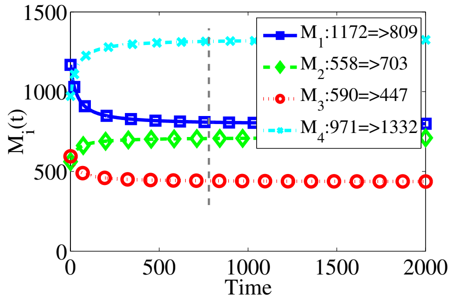
(b) Balance between C^* and t^* with different λ

Fig. 4.7. Trade-off between C^* and t^* for completing the dissemination

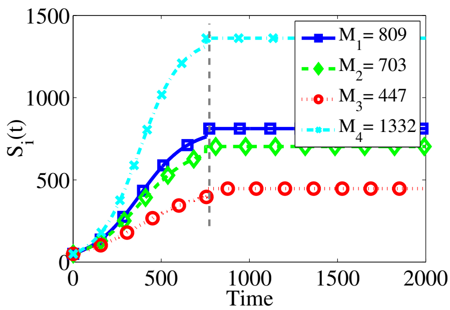
cal values as introduced previously. Also the handover rates are set between 0.01 to 0.2 randomly, because the handover rates are not too high in real



(a) $\mathbb{S}_i(t)$ without Handovers



(b) How handovers impact $M_i(t)$



(c) $\mathbb{S}_i(t)$ with handovers

Fig. 4.8. Content dissemination in multi-cell scenario with handovers

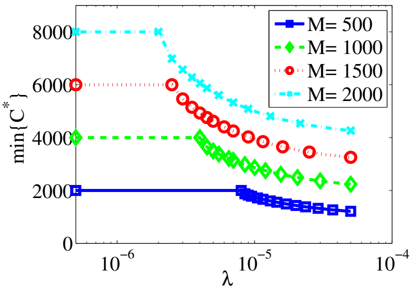
measurements [118] [121] [122] and [123].

We firstly plot $\mathbb{S}_i(t)$ of each cell without applying the handover rates as shown in Fig. 4.8(a). We can see that each BS completes the content dissemination separately, regardless of either the very slow dissemination of b_2 in green color (the diamond dashed curve) with $M_2 = 558$, or the very fast one of b_3 in red color (the circle dotted curve) with $M_3 = 590$.

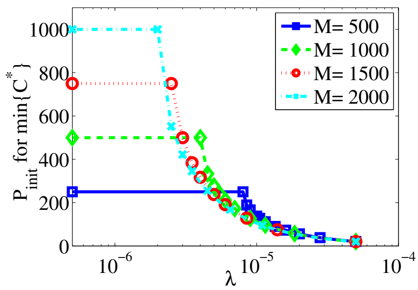
Then we apply the handover rates to the model and examine $M_i(t)$ as shown in Fig. 4.8(b). Each BS changes the number of MSs due to the handovers of the MSs, and finally $M_i(t)$ converges to a steady-state around 520 seconds. Note that we approximately assume the steady-state when the change of $M_i(t)$ per second is small than 1. The corresponding plot of $\mathbb{S}_i(t)$ is shown in Fig. 4.8(c), and we can see the BSs complete the dissemination at the same time around 783 seconds. This is mainly because when MSs are performing handovers, some of them carry the content but the other do not; each cell will then exchange its both not-yet-updated MSs and updated MSs with its neighbor cells. The cells, which originally disseminate content fast, will "help" those who suffer from slow dissemination. Therefore, $\mathbb{S}_i(t)$ of BSs together grow and finally complete with same t^* in a harmonized manner.

4.6.3 Optimization Framework

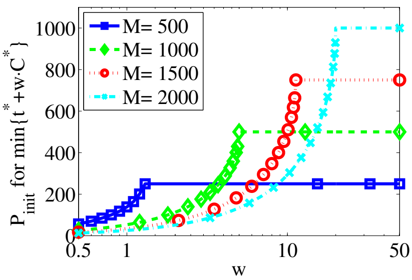
The minimization of C^* in a single cell is shown in Fig. 4.9(a) , 4.9(b) and 4.9(c). Note that the X-axis is in log scale. We see that with a larger meeting rate λ, the BS can adjust P_{init} to a smaller value for getting the minimum C^*. But when λ goes smaller below a boundary (referring to the condition in Eq. (4.20)), the BS will only have to set P_{init} to $\frac{M}{2}$ for the minimum C^*, which means to push the content to all of its MSs. The Pareto-optimality between the delay and energy cost is evaluated in Fig. 4.9(c), which indicates if the



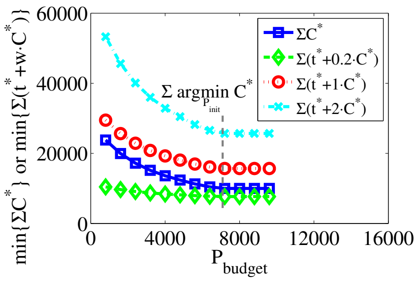
(a) $\min\{C^*\}$ - different M with varying λ



(b) P_{init} for $\min\{C^*\}$ - different M with varying λ



(c) P_{init} for $\min\{t^* + w \cdot C^*\}$ - different M with varying w



(d) $\min\{\Sigma\mathbb{C}^*\}$ and $\min\{\Sigma(t^* + w \cdot \mathbb{C}^*)\}$ - different w with varying P_{budget}

Fig. 4.9. Optimization of C^*, $t^* + w \cdot C^*$, and $\Sigma(t^* + w \cdot \mathbb{C}^*)$

MNO system emphasizes more on the energy cost (a higher value of w), P_{init} should be set to a higher value until $\frac{M}{2}$. Fig. 4.9(d) shows the evaluation on the P_{budget}-constrained optimization in the multi-cell scenario (20 cells with reasonable parameters). Depending on the boundary condition in Eq. (4.21), when P_{budget} is sufficient, BSs can freely adjust P_{init_i}'s values individually for both local and global minimum cost; when P_{budget} is not enough, the minimum energy cost increases. Also $\Sigma(t_i^* + w \cdot \mathbb{C}_i{}^*)$ follows the same trend. Note that when P_{budget} is quite small, BSs will have small P_{init_i}, so BSs will mostly rely on the ICT-based sharing, and thus suffer from high energy cost and large delay.

4.7 Conclusion

In this chapter, I proposed **Push-Share** framework to reduce the traffic load on cellular links by coordinating pushing and sharing for disseminating delay-tolerant content. Content dissemination can be adaptively accelerated or decelerated to satisfy performance requirements by adjusting the initial and final pushing rates. The multi-compartment model can be adopted for modeling the content dissemination among multiple cells with handovers in cellular networks. The proposed optimization framework can be used by MNOs to control the pushing strategy for the objectives such as the minimum delay or minimum cost.

The lessons from the analytic studies are summarized as follows: pushing more copies to cells with the fewer MSs can be more beneficial for reducing the completion delay (Fig. 4.4(b)); the more users participate in sharing, the more energy saving can be achieved due to the sharing (Fig. 4.5(b)); the completion delay and energy cost exhibit an inverse relation, which reflects the Pareto-optimality when the required completion delay is small (the dashed boxes in Fig. 4.7(a)); if the requirement of completion delay is long, the energy cost of neighborhood monitoring will be overwhelming as shown in Figs. 4.7(a) and 4.7(b); the handovers among cells mix the MSs with or without the content, which implies a balance of overall completion delays among cells, and hence BSs can finish the content dissemination to their MSs with almost similar delays (Fig. 4.8(c)). In the future, we will extend the model for more practical scenarios such as heterogeneous mobility and transmission failure probability.

Chapter 5

Summary and Future Work

In this thesis, I mainly focused on the user-to-user opportunistic sharing and tried to elaborate its effectiveness and efficiency for mobile traffic offloading, in order to solve the traffic explosion problem.

In the first work, I proposed the ***T***raffic ***O***ffloading assisted by ***S***ocial network services via opportunistic ***S***haring in mobile social networks, **TOSS** framework, to select optimal seed users for initial content pushing, depending on their content spreading impact in online social network services (SNSs) and their mobility patterns in offline mobile social networks (MSNs). Then users share the content via opportunistic local connectivity (like Wi-Fi Direct and D2D) with each other. Also TOSS exploited the user-dependent access delay between the content generation time and each user's access time for traffic offloading purposes. We modeled and analyzed the traffic offloading and content spreading among users by taking into account various options in linking SNS and MSN trace data. And the trace-driven evaluation showed that TOSS can reduce up to 86.5% of the cellular traffic while satisfying the access delay requirements of all users.

Furthermore, I focused on the analytical research on **Push-Share** content disseminating in the second work, which is highly correlated with the first study from the theoretical study perspective. In Push-Share, a content is firstly pushed the to a subset of subscribers via cellular links, and mobile subscribers share the content via opportunistic local connectivity. We theoretically modeled and analyzed how the content can be disseminated across multiple cells, where handovers are modeled based on the multi-compartment model. We also formulated mathematical framework to optimize the system, by which the trade-off between the dissemination delay and the energy cost is explored.

From the measurement study, trace-driven analysis, theoretical modeling and system optimization in above studies, the traffic offloading by user-to-user opportunistic sharing in mobile social networks is proved to be effective and efficient.

5.1 A Comparison with Traffic Offloading based on Wi-Fi APs

As already discussed in Sec. 2.2, there actually have been many research studies and realistic deployment cases for the mobile traffic offloading based on Wi-Fi Access Points (APs), such as [55] [59]. and [58] in Korea, and [56] [57] in USA. Also economics of traffic offloading by Wi-Fi APs have been studied in [60] and in [61] in detail.

Regarding the realistic deployment of the proposed offloading by user-to-user sharing, how to promote and encourage people to share content during moving becomes one important issue, which is not only an issue of technology. As the sharing-based offloading will help the MNO to reduce their traffic load significantly, a popular research trend is to design incentive-based business model for mobile operators and mobile users. It is advocated that the financial benefit will drive the motivation for users to cache content and share with nearby user either to reduce their cellular data plan, or even to earn some money, such as research work in [39] [62] to utilize the benefit-based data caching and forwarding in mobile networks. Also there are new incentive-based designs for sharing-based offloading with further consideration of user privacy, such as [63], IPAD [64] and iDEAL [65].

In order to comprehansively study the advantage and disadvantage of the traffic offloading by opportunistic sharing and that by Wi-Fi APs, we hereby compare the major pros and cons between them , as shown in Table. 5.1. It is clear that the offloading by Wi-Fi APs still consumes 100% backbone traffic, which is not completely solving the "traffic explosion problem" but is keeping the high load to provider's backhaul network. Furthermore, it needs large-scale deployment, serving people indoor in most cases. But the offloading by opportunistic sharing consumes little backbone traffic, and it offloads traffic from cellular link to other local short-range links without any infrastructure deployment. Sharing-based offloading is not replacing "Wi-Fi APs", but they will work together to solve the "traffic explosion problem".

5.2 Practical Deployment and Application

A very easy beginning for deploying the sharing-based offloading framework can be a mobile SNS application, with extra functions for discovering nearby SNS friends, friends of friends, and even strangers, for exploring and transmitting files with them by both active "request-to-share" and proactive "background-share" mechanisms. Note that the MNO should track user

Table. 5.1. Comparison between traffic offloading via user-to-user sharing and that via Wi-Fi APs

	via user-to-user Sharing	via Wi-Fi AP
Cost to operators	near zero	Large scale deployment
Cost to users	near zero (even with incentive)	zero (most cases)
Backhual traffic	13.5 to 36.2% (initial push + final on-demand)	100%
Cellular traffic	13.5 to 36.2% (initial push + final on-demand)	zero
Energy consumption	Probing + sending + receiving	probing + receiving
Availability	Whenever there are users in proximity	9 to 18% in US, 24 to 64% in KR
Business model	Free or incentive-based	Free or complementary to data plan

sharing activities via the application and count their incentives.e

When a user moves with the mobile device, most of content objects that 1) accessed and shared by good friends in the SNS and MSN, and 2) published by interesting or famous publishers that the user has subscribed will be collected in the background already. Because it is expected to be able to obtain the content object before users may access it, the sharing can be carried out in the background, which can be considered as **prefetching**.

The user can even actively further explore more content resource via the application, click interesting content objects for “accessing it later” depending on the user’s life style and activity pattern, and set up a deadline for obtaining it by opportunistic sharing. While the user moves around and meets people, the pending content objects will be collected, i.e., prefetched, opportunistically. When the user wants to directly access the object, which is not prefetched yet, on-demand delivery will be carried out then.

5.3 Future Work and Vision

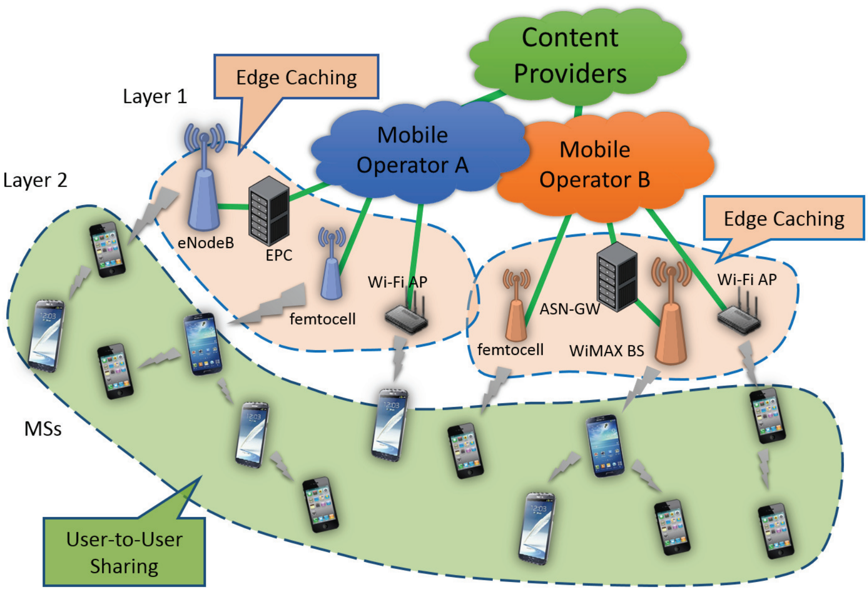


Fig. 5.1. Illustration of the 2-tier structure for future mobile network

This thesis mainly focuses on the user-to-user opportunsitic sharing for mobile traffic offloading, which is actually a part of my vision for the "two-layered caching and sharing infrastructure in the future mobile network", as illustrated in Fig. 5.1. The layer 1 is actually the in-network caching of mobile backhaul networks, which is the edge of the mobile networks consisting a lot of Wi-Fi APs, macrocells and femtocells, the base stations of which forms a "cooperative buffer" for mobile users. The topic is related to the extension of the concept of Content-Centric Networking (Named Data

Networking [129] [130]) into future mobile networks. And my focus in this thesis is then the layer 2, that is the opportunistic user-to-user sharing under the cells, based on the user mobility impact and social spreading impact. Due to the cooperative caching and thus the resource re-utilization at the cells, the mobile network can significantly reduce the traffic to the Internet and to other providers, also the backbaul traffic can be further reduced since mobile users frequently share popular and interesting content with each other by local short-range communication.

Furthermore, since the D2D technique is being hotly discussed in 3GPP standards for 4G LTE (LTE-advanced) networks [19], by which users use operator authorized spectrum for direct communication without the support of infrastructure. The transmission range of D2D communication can be much larger than other local range communications (such as Wi-Fi Direct). Therefore by optimal resource allocation and interference management, the new D2D communication can increase the total throughput (resource utilization) in the cell area as studied in [19] [53] and [54]. Therefore, based on the analysis in this thesis, along with the trend of the significant growth of the number of mobile devices, the D2D technique will significantly facilitate the sharing-based offloading in mobile networks in the very near future.

Bibliography

[1] CISCO, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2010-2015," CISCO, Tech. Rep., 2011.

[2] S. Scellato, C. Mascolo, M. Musolesi, and J. Crowcroft, "Track Globally, Deliver Locally: Improving Content Delivery Networks by Tracking Geographic Social Cascades," in Proc. of *The 20th International World Wide Web Conference (WWW)*, Bangalore, India, 2011.

[3] Y. Chen, L. Qiu, W. Chen, L. Nguyen and R. Katz, "Clustering Web Content for Efficient Replication," in Proc. of *The 10th IEEE International Conference on Network Protocols (ICNP)*, Paris, France, 2002.

[4] M. Cha, H. Kwak, P. Rodriguez, Y. Y. Ahn, and S. Moon, "I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System," in Proc. of *The 7th ACM Internet Measurement Conference (IMC)*, San Diego, USA, 2007.

[5] X. Wang, S. Kim, T. Kwon, H. Kim, Y. Choi, "Unveiling the BitTorrent Performance in Mobile WiMAX Networks," in Proc. of *The 12th Passive and Active Measurement Conference*, Atlanta, USA, 2011.

[6] K. Fall, "A Delay-Tolerant Network Architecture for Challenged Internets," in Proc. of *ACM Special Interest Group on Data Communication (SIGCOMM)*, Karlsruhe, Germany, 2003.

[7] L. Psaras, L.Wood and R. Tafazolli, "Delay-/Disruption-Tolerant Networking State of the Art and Future Challenges," in *Tech. Rep.*, University of Surrey, UK, 2010.

[8] A. Voyiatzis, "A Survey of Delay- and Disruption-Tolerant Networking Applications," in *Journal of Internet Engineering*, vol.5, no.1, pp.331–344, 2012.

[9] J. Whitbeck, M. D. Amorim, Y. Lopez, J. Leguay, V. Conan, "Relieving the Wireless Infrastructure: When Opportunistic Networks Meet Guaranteed Delays," in Proc. of *The 12th IEEE International Symposium on World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, Lucca, Italy, 2011.

[10] S. Ioannidis, A. Chaintreau, and L. Massoulie, "Optimal and Scalable Distribution of Content Updates over A Mobile Social Network," in Proc. of *The 28th IEEE International Conference on Computer Communications (INFOCOM)*, Rio de Janeiro, Brazil, 2009.

[11] F. Malandrino, M. Kurant, A. Markopoulou, C. Westphal, and U. C. Kozat, "Proactive Seeding for Information Cascades in Cellular Networks," in Proc. of *The 31st IEEE International Conference on Computer Communications (INFOCOM)*, Orlando, USA, 2012.

[12] X. Wang, M. Chen, Z. Han, T. Kwon, and Y. Choi, "Content Dissemination by Pushing and Sharing in Mobile Cellular Networks: An Analytical Study," in Proc. of *The 9th IEEE International Conference on Mobile Ad hoc and Sensor Systems (MASS)*, Las Vegas, USA, 2012.

[13] Y. Li, Y. Jiang, D. Jin, L. Su, L. Zeng, and D. Wu, "Energy-efficient Optimal Opportunistic Forwarding for Delay-Tolerant Networks," in *IEEE Transactions on Vehicular Technology (TVT)*, vol.58, no.9, pp.4500–4512, 2010.

[14] L. Pelusi, A. Passarella and M. Conti, "Opportunistic Networking: Data Forwarding in Disconnected Mobile Ad-hoc Networks," in *IEEE Communications Magazine*, vol.44, issue.11, pp.134-141, 2006

[15] C. Huang, K. Lan and C. Tsai, "A Survey of Opportunistic Networks," in Proc. of *The 22nd International Conference on Advanced Information Networking and Applications - Workshops*, Okinawa, Japan, 2008.

[16] B. Han and A. Srinivasan, "eDiscovery: Energy Efficient Device Discovery for Mobile Opportunistic Communications," in Proc. of *The*

20th IEEE International Conference on Network Protocols (ICNP), Austin, USA, 2012.

[17] F. Michahelles, F. Thiesse, A. Schmidt and J. R. Williams, "Pervasive RFID and Near Field Communication Technology," in *IEEE Pervasive Computing*, vol.6, issue.3, pp.94–96, 2007.

[18] C. Yu, K. Doppler, C. Ribeiro and O. Tirkkonen, "Resource Sharing Optimization for Device-to-Device Communication Underlaying Cellular Networks," in *IEEE Transactions on Mobile Computing (TMC)*, vol.10, no.8, pp.2752–2763, 2011.

[19] K. Doppler, M. Rinne, C. Wijting, C. Ribeiro and K. Hugl, "Device-to-Device Communication as An Underlay to LTE-Advanced Networks," in *IEEE Communications Magazine*, vol.47, issue.12, pp.42–49, 2009.

[20] X. Zhang, G. Neglia, J. Kurose, and D. Towsley, "Performance Modeling of Epidemic Routing," in *Elsevier Computer Networks*, no.10, issue.10, pp.2867–2891, 2007.

[21] Y. Li, G. Su, P. Hui, D. Jin, L. Su and L. Zeng, "Multiple Mobile Data Offloading Through Delay Tolerant Networks," in Proc. of *The 13th ACM International Conference on Mobile Computing and Networking (MobiCom), Workshop on Challenged Networks (CHANT)*, 2011.

[22] J. Yang and J. Leskovec, "Patterns of Temporal Variation in Online Media," in Proc. of *The 4th ACM International Conference on Web Search and Data Mining*, Hongkong, 2011.

[23] B. Han, P. Hui, V. S. A. Kumar, M. V. Marathe, J. Shao, and A. Srinivasan, "Mobile Data Offloading through Opportunistic Communications and Social Participation," in *IEEE Transaction on Mobile Computing (TMC)*, vol.11, no.5, pp.821–834, 2012.

[24] T. Rodrigues, F. Benvenuto, M. Cha, K. Gummadi, and V. Almeida, "On Word-of-Mouth Based Discovery of the Web," in Proc. of *The*

11th ACM Internet Measurement Conference (IMC), Berlin, Germany, 2011.

[25] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a Social Network or a News Media?" in Proc. of *The 19th International World Wide Web Conference (WWW)*, Raleigh, USA, 2010.

[26] Facebook, `http://www.facebook.com`

[27] Twitter, `http://www.twitter.com`

[28] Tumblr, `http://www.turmblr.com`

[29] Sina Weibo, `http://weibo.com`

[30] G. V. Steeg and A. Galstyan, "Information Transfer in Social Media," in Proc. of *The 21st International World Wide Web Conference (WWW)*, Lyon, France, 2012.

[31] G. Comarela, M. Crovella, V. Almeida, F. benevenuto, "Understanding Factors that Affect Response Rates in Twitter," in Proc. of *The 23rd ACM Conference on Hypertext and Social Media (HT)*, Milwaukee, USA, 2012.

[32] S. Myers, C. Zhu, and J. Leskovec, "Information Diffusion and External Influence in Networks," in Proc. of *The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, Beijing, China, 2012.

[33] T. Karagiannis, J. Y. Le Boudec, and M. Vojnovic, "Power Law and Exponential Decay of Inter Contact Times between Mobile Devices," in Proc. of *The 13th ACM International Conference on Mobile Computing and Networking (MobiCom)*, Montreal, Canada, 2007.

[34] W.J. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy, "Modeling Time-Variant User Mobility in Wireless Mobile Networks," in Proc. of *The 26th IEEE International Conference on Computer Communications (INFOCOM)*, Anchorage, USA, 2007.

[35] W. Gao and G. Cao, "On Exploiting Transient Contact Patterns for Data Forwarding in Delay Tolerant Networks," in Proc. of *The 18th IEEE International Conference on Network Protocols (ICNP)*, Kyoto, Japan, 2010.

[36] E. Jaho and I. Stavrakakis, "Joint Interest- and Locality-Aware Content Dissemination in Social Networks," in Proc. of *The 6th International Conference on Wireless On-Demand Network Systems and Services (WONS)*, Snowbird, USA, 2009.

[37] M. P. Wittie, V. Pejovic, L. Deek, K. C. Almeroth, and B. Y. Zhao, "Exploiting Locality of Interest in Online Social Networks," in Proc. of *The 6th International Conference on emerging Networking EXperiments and Technologies (CoNEXT)*, Philadelphia, USA, 2010.

[38] F. Chen, M. Johnson, A. Bar-Noy, I. Fermin and T. La Porta, "Proactive Data Dissemination to Mission Sites," in Proc. of *The 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, Rome, Italy, 2009.

[39] X. Zhuo, W. Gao, G. Cao, and Y. Dai, "Win-Coupon: An incentive framework for 3G traffic offloading," in Proc. of *The 19th IEEE International Conference on Network Protocols (ICNP)*, Vancouver, Canada, 2011.

[40] J. Ahn, B. Krishnamachari, F. Bai and L. Zhang, "Optimizing Content Dissemination in Heterogeneous Vehicular Networks," *USC CENG Technical Report CENG-2010-2*, 2010.

[41] P. Hui, J. Crowcroft, and E. Yoneki, "BUBBLE Rap: Social-Based Forwarding in Delay-Tolerant Networks," in *IEEE Transaction on Mobile Computing (TMC)*, vol.10, pp.1576-1589, 2011.

[42] W. Gao and G. Cao, "User-Centric Data Dissemination in Disruption Tolerant Networks," in Proc. of *The 30th IEEE International Conference on Computer Communications (INFOCOM)*, Shanghai, China, 2011.

[43] C. Boldrini, M. Conti, and A. Passarella, "ContentPlace: Social-aware Data Dissemination in Opportunistic Networks," in Proc. of *The 11th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, Vancouver, Canada, 2008.

[44] E. Daly and M. Haahr, "Social Network Analysis for Routing in Disconnected Delay-Tolerant MANETs," in Proc. of *The 8th ACM International Symposion on Mobile Ad Hoc Networking and Computing (MobiHoc)*, Montreal, Canada, 2007.

[45] E. Bulut and B. W. Szymanski, "Friendship Based Routing in Delay Tolerant Mobile Social Networks," in Proc. of *IEEE Global Communications Conference*, Miami, USA, 2010.

[46] D. Rothfus, C. Dunning and X. Chen, "Social-Similarity-based Routing Algorithm in Delay Tolerant Networks," in Proc. of *IEEE International Conference on Communications (ICC)*, Budapest, Hungary, 2013.

[47] X. Bao, Y. Lin, U. Lee, I. Rimac and R. R. Choudhury, "DataSpotting: Exploiting Naturally Clustered Mobile Devices to Offload Cellular Traffic," in Proc. of *The 32nd IEEE International Conference on Computer Communications (INFOCOM)*, Turin, Italy, 2013.

[48] A. Kate, G. M. Zaverucha, U. Hengartner, "Anonymity and Security in Delay Tolerant Networks," in Proc. of *The 3rd International Conference on Security and Privacy in Communication Networks (SecureCom)*, Nice, France, 2007.

[49] R. Lu, X. Lin, T. H. Luan, X. Liang, X. Li, L. Chen and X. Shen, "PReFilter: An Efficient Privacy-preserving Relay Filtering Scheme for Delay Tolerant Networks," in Proc. of *The 31st IEEE International Conference on Computer Communications (INFOCOM)*, Orlando, USA, 2012.

[50] A. Shikfa, M. Onen and R. Molva, "Privacy in Content-Based Opportunistic Networks," in Proc. of *International Conference on Advanced Information Networking and Applications Workshops*,2009.

[51] AirDrop, Apple Inc., `http://en.wikipedia.org/wiki/AirDrop`

[52] A. Noori and D. Giustiniano, "HyCloud: A Hybrid Approach Toward Offloading Cellular Content Through Opportunistic Communication," in Proc. of *The 11th International Conference on Mobile Systems, Applications, and Services (MobiSys)*, Taipei, Taiwan, 2013.

[53] M. Meibergen, "Device-to-Device Communications Underlaying a Cellular Network," *Master Thesis*, Delft University of Technology, 2011.

[54] M. Zulhasinine, C. Huang and A. Srinivasan, "Efficient Resource Allocation for Device-to-Device Communication Underlaying LTE Network," in Proc. of *The 6th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Niagara Falls, Canada, 2010.

[55] Y. Cho, H. Ji, J. Park, H. Kim and J. Silvester, "A 3W Network Strategy for Mobile Data Traffic Offloading," in *IEEE Communications Magazine*, vol.49, issue.10, pp.118–123, 2011.

[56] A. Balasubramanian, R. Mahajan, A. Venkataramani, "Augmenting Mobile 3G Using Wi-Fi," in Proc. of *The 8th International Conference on Mobile Systems, Applications, and Services (MobiSys)*, San Francisco, USA, 2010.

[57] S. Dimatteo, P. Hui and B. Han, "Cellular Traffic Offloading through Wi-Fi Networks," in Proc. of *The 8th IEEE 8th International Conference on Mobile Adhoc and Sensor Systems (MASS)*, Valencia, Spain, 2011.

[58] K. Lee, J. Lee, Y. Yi, I. Rhee and S. Chong, "Mobile Data Offloading: How Much Can WiFi Deliver?" in *IEEE/ACM Transactions on Networking (ToN)*, vol.21, issue.2, pp.536–550, 2013.

[59] Y. Im, C. J. Wong, S. Ha, S. Sen, T. T. Kwon and M. Chiang, "AMUSE: Empowering Users for Cost-Aware Offloading with Throughput-Delay Tradeoffs," in Proc. of *The 32nd IEEE International Conference on Computer Communications (INFOCOM)*, Turin, Italy, 2013.

[60] L. Gao, G. Iosifidis, J. Huang and L. Tassiulas, "Economics of Mobile Data Offloading," in Proc. of *The 32nd IEEE International Conference on Computer Communications (INFOCOM), Smart Data Pricing (SDP) Workshop*, Turin, Italy, 2013.

[61] J. Lee, Y. Yi, S. Chong and Y. Jin, "Economics of WiFi Offloading: Trading Delay for Cellular Capacity," in *Computing Research Repository (CoRR)*, 2012.

[62] B. Tang, H. Gupta and S. R Das, "Benefit-Based Data Caching in Ad Hoc Networks," in *IEEE Transaction on Mobile Computing (TMC)*, vol.7, no.3, pp.289–304, 2008.

[63] T. Ning, Z. Yang, H. Wu and Z. Han, "Self-Interest-Drive Incentives for Ad Dissemination in Autonomous Mobile Social Networks," in Proc. of *The 32nd IEEE International Conference on Computer Communications (INFOCOM)*, Turin, Italy, 2013.

[64] R. Lu, X. Lin, Z. Shi, B. Cao, and X. Shen, "IPAD: An Incentive and Privacy-Aware Data Dissemination Scheme in Opportunistic Networks," in Proc. of *The 32nd IEEE International Conference on Computer Communications (INFOCOM)*, Turin, Italy, 2013.

[65] W. Dong, S. Rallapalli, R. Jana, L. Qiu, K. Ramakrishnan, L. Razoumov, Y. Zhang and T. Cho, "iDEAL: Incentivized Dynamic Cellular Offloading via Auctions," in Proc. of *The 32nd IEEE International Conference on Computer Communications (INFOCOM)*, Turin, Italy, 2013.

[66] E. Katz and P. F. Lazarsfeld, "Personal Influence," in *New York: The Free Press*, 1955.

[67] D. J. Watts and P. S. Dodds, "Influentials, Networks, and Public Opinion Formation," in *Journal of Consumer Research*, vol.34, pp.441-458.

[68] S. Wu, J. M. Hofman, W. A. Mason, and D. J. Watts, "Who Says What to Whom on Twitter," in Proc. of *The 20th International World Wide Web Conference (WWW)*, Bangalore, India, 2011.

[69] W. Chen, Y. Wang, and S. Yang, "Efficient Influence Maximization in Social Networks," in Proc. of *The 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, Paris, France, 2009.

[70] A. Goyal, F. Bonchi, and L. Lakshmanan, "Learning Influence Probabilities in Social Networks," in Proc. of *The 3rd ACM International Conference on Web Search and Data Mining*, New York City, USA, 2010.

[71] R. Xiang, J. Neville, and M. Rogati, "Modeling Relationship Strengthin Online Social Networks," in Proc. of *The 19th International World Wide Web Conference (WWW)*, Raleigh, USA, 2010.

[72] M. Gomez-Rodriguez, J. Leskovec, and A. Krause, "Inferring Networks of Diffusion and Influence," in Proc. of *The 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, Washington DC, USA, 2010.

[73] M. D. Choudhurry, W. A. Mason, J. M. Hofman, and D. J. Watts, "Inferring Relevant Social Networks from Interpersonal Communication," in Proc. of *The 19th International World Wide Web Conference (WWW)*, Raleigh, USA, 2010.

[74] V. Chaoji, S. Ranu, R. Rastogi and R. Bhatt, "Recommendations to Boost Content Spread in Social Networks," in Proc. of *The 21st International World Wide Web Conference (WWW)*, Lyon, France, 2012.

[75] Homophily, http://en.wikipedia.org/wiki/Homophily.

[76] M.D. Choudhury, H. Sundaram, A. John, D.D. Seligmann, and A. Kelliher, "Birds of a Feather: Does User Homophily Impact Information Diffusion in Social Media?" in *Computing Research Repository (CoRR)*, vol.1006, 2010.

[77] Locality of Reference (Principle of Locality), http://en.wikipedia.org/wiki/Locality_of_reference.

[78] L. Gyarmati and T. A. Trinh, "Measuring User Behavior in Online Social Networks," in *IEEE Network Magazine*, vol.24, pp.26–31, 2010.

[79] C. Liu, R. White and S. Dumais, "Understanding Web Browsing Behaviors Through Weibull Analysis of Dwell Time," in Proc. of *The 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, New York, USA, 2010.

[80] R. Jain, D. Lelescu and M. Balakrishnan, "Model T: An Empirical Model for User Registration Patterns in A Campus Wireless LAN," in Proc. of *The 11th ACM International Conference on Mobile Computing and Networking (MobiCom)*, Cologne, Germany, 2005.

[81] P. Svoboda, W. Karner and M. Rupp, "Traffic analysis and modeling for World of Warcraft," in Proc. of *IEEE International Conference on Communications (ICC)*, Glasgo, UK, 2007.

[82] MATLAB, The Mathworks Inc.

[83] 3GPP, 3GPP TS 23.012, *Location Management Procedures*, 2011.

[84] H. Aminikhah and J. Biazar, "A New Analytical Method for System of ODEs," in *Numerical Methods for Partial Differential Equations*, vol.26, no.5, pp.1115–1124, 2010.

[85] N. Guzel and M. Bayram, "Power Series Solution of Non-linear First Order Differential Equation Systems," in *Trakaya University Journal of Science*, vol.6, no.1, pp.107-111, 2005.

[86] Hill Climbing Algorithm, `https://en.wikipedia.org/wiki/Hill_climbing`

[87] W. Wei, F. Xu, C. Tan and Q. Li, "SybilDefender: Defend against Sybil Attacks in Large Social Networks," in Proc. of *The 31st IEEE International Conference on Computer Communications (INFOCOM)*, Orlando, USA, 2012.

[88] Z. Chu, S. Gianvecchio, H. Wang and S. Jajodia, "Who Is Tweeting on Twitter: Human, Bot, or Cyborg?" in Proc. of *The 26th Annual Computer Security Applications Conference (ACSAC)*, Austin, USA, 2010.

[89] G. Stringhini, C. Kruegel and G. Vigna, "Detecting Spammers on Social Networks," in Proc. of *The 26th Annual Computer Security Applications Conference (ACSAC)*, Austin, USA, 2010.

[90] Power Law, `http://en.wikipedia.org/wiki/Power_law`.

[91] R. P. Satorras and A. Vespignani, "Epidemic Spreading in Scale-Free Networks," in *Phys. Rev. Letter*, vol.86, issue.14, pp2001.

[92] A. Barabasi and E. Bonabeau, "Scale-free Networks," in *Scientific American*, pp.50–59, May, 2003.

[93] A. Hernando, D. Villuendas, C. Vesperinas, M. Abad and A. Plastino, "Unravelling the Size Distribution of Social Groups with Information Theory in Complex Networks," in *The European Physical Journal B*, vol.76, no.1, pp.87-97, 2010.

[94] A. Mislove, M. Marcon, K. Gummadi, P. Druschel and B. Bhattacharjee, "Measurement and Analysis of Online Social Networks," in Proc. of *The 7th ACM Internet Measurement Conference (IMC)*, San Diego, USA, 2007.

[95] A. Barabasi, H. Jeong, Z. Neda, T. Vicsek and A. Schuber, "On the Topology of the Scientific Collaboration Networks," in *Physica A*, vol.311, pp.590–612, 2002.

[96] M. Girvan and M. E. Newman, "Community Structure in Social and Biological Networks," in *Proceedings of the National Academy of Science of USA (PNAS)*, vol.99, no.12, pp.7821–7826, 2002.

[97] D. Towsley, "A Walk in the Dark: Random Walks and Network Discovery," in *The 12th IFIP International Conference on Networking, Keynote*, New York, USA, 2013.

[98] N. Eagle and A. Pentland, "Reality Mining: Sensing Complex Social Systems," in *Springer Journal of Personal and Ubiquitous Computing*, vol.10, no.4, pp.255–268, 2006.

[99] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of Human Mobility on Opportunistic Forwarding Algorithms," in *IEEE Transaction of Mobile Computing (TMC)*, vol.6, no.6, pp.606–620, 2007.

[100] Y. Zheng, L. Zhang, X. Xie, and W. Ma, "Mining Interesting Locations and Travel Sequences from GPS Trajectories," in Proc. of *The 18st International World Wide Web Conference (WWW)*, Madrid, Spain, 2009.

[101] H. Huang, P. Luo, M. Li, D. Li, X. Li, W. Shu, and M. Wu, "Performance Evaluation of SUVnet With Real-Time Traffic Data," in *IEEE Transactions on Vehicular Technology (TVT)*, vol.56, no.6, pp.3381–3396, Nov. 2007.

[102] Mathematica, Wolfram Research Inc.

[103] SPSS, SPSS Inc.

[104] A. Pietilainen, "Opportunistic Mobile Social Networks at Work," *PhD Thesis*, Pierre-and-Marie-Curie University, 2010.

[105] F. Bjurefors, "Measurements in Opportunistic Networks," *PhD Thesis*, Uppsala University, 2012.

[106] V. Le, H. Scholten and P. Havinga, "Towards Opportunistic Data Dissemination in Mobile Phone Sensor Networks," in Proc. of *The 7th International Conference on Networks (ICN)*, Saint Gilles, Belgium, 2012.

[107] T. Karagiannis, J. Boundec, M. Vojnovic, "Power Law and Exponential Decay of Inter Contact Times between Mobile Devices," in Proc. of *The 13th ACM International Conference on Mobile Computing and Networking (MobiCom)*, Montreal, Canada, 2007.

[108] O. R. Helgason, "Opportunistic Content Distribution," *PhD Thesis*, Royal Institute of Technology, 2009.

[109] E. Cho, S. Myers and J. Leskovec, "Friendship and Mobility: User Movement In Location-Based Social Networks," in Proc. of *The 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, San Diego, USA, 2011.

[110] A. Noulas, S. Scellato, C. Mascolo and M. Pontil, "An Empirical Study of Geographic User Activity Patterns in Foursquare," in Proc. of *The 5th International AAAI Conference on Weblogs and Social Media (ICWSM)*, Barcelona, Spain, 2011.

[111] L. Page, S. Brin, R. Motwani and T. Winograd ,"The PageRank Citation Ranking: Bringing Order to the Web," Tech. rep., Stanford University, 1998.

[112] S. Brin and L. Page, "The Anatomy of A Large-scale Hypertextual Web Search Engine," *Computer Networks and ISDN Systems* , vol. 30, no.1-7, pp. 107–117, 1998.

[113] K. Suh, C. Diot, J. Kurose, L. Massoulié, C. Neumann, D. Towsley and M. Varvello, "Push-to-Peer Video-on-Demand system: design and

evaluation," in *IEEE Journal on Selected Areas in Communications (JSAC)*, no.25, issue.9, pp.1706–1716, 2007.

[114] R. Groenevelt, "The Message Delay in Mobile Ad-hoc Networks," in *Performance Evaluation*, no. 1-4, pp. 210–228, 2005.

[115] N. Balasubramanian, A. Balasubramanian and A. Venkataramani, "Energy Consumption in Mobile Phones: A Measurement Study and Implications for Network Applications," in Proc. of *The 9th ACM Internet Measurement Conference (IMC)*, Chicago, USA, 2009.

[116] S. Bohacek, "3G v.s WIFI Radio Energy with YouTube downloads," Tech. Report, 2009.

[117] A. Locker, "Multi-Compartment Models and Their Analysis," in *International Journal Of Clinical Pharmacology Therapy And Toxicology*, vol.6, no.1, pp.77–87, 1972.

[118] J. K. Kwon and D. K. Sung, "Soft Handoff Modeling in CDMA Cellular Systems," in Proc. of *The IEEE 47th Vehicular Technology Conference (VTC)*, pp. 1548–1551, 1997.

[119] P. V. Orlik, S. S. Rappaport, "On the Hand-off Arrival Process in Cellular Communications," in *Springer Wireless Networks (WINET)*, vol.7, issue.2, pp.147–157, 2001.

[120] A. Iggidr, "Multi-compartment Models," in *Discrete and Continuous Dyanmical Systems Supplement*, pp. 506–519, 2007.

[121] M. Zonoozi and P. Dassanayake, "User Mobility Modeling and Characterization of Mobility Patterns," in *IEEE Journal on Selected Areas in Communications (JSAC)*, vol.15, pp.1239-1252, 1997.

[122] N, Hegde and K. Sohraby, "On the Impact of Soft Hand-off in Cellular Systems," in Proc. of *ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS)*, pp.178–187, Santa Clara, USA, 2000.

[123] N. D. Tripathi, J. H. Reed and H. F. VanLandingham, “Handoff in Cellular Systems,”in *IEEE Personal Communications*, Dec, 1998.

[124] D. Keffer, “An Analytical Method for Solving Systems of Linear nth-order ODEs,” Lecture, 1999.

[125] S. Khorasani and A. Adibi, “Analytical Solution of Linear Ordinary Differential Equations by Differential Transfer Matrix Method,” in *Electronic Journal of Differential Equations*, vol.79, pp.1–18, 2003.

[126] L. Jodar and E. Navarro, “Closed Analytical Solution of Riccati Type Matrix Differential Equations,” in *Indian Journal of Pure Application Mathematics*, vol.23, no.3, pp.185–187, 1992.

[127] R. W. R. Darling, “Converting Matrix Riccati Equations to Second-Order Linear ODE,” in *Society for Industrial and Applied Mathematics (SIAM) Review*, vol.39, no.3, pp.508–510, 1997.

[128] Maple, Maplesoft Inc.

[129] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs and R. L. Braynard, “Networking Named Content,” in Proc. of *The 6th International Conference on emerging Networking EXperiments and Technologies (CoNEXT)*, Rome, Italy, 2009.

[130] B. Han, X. Wang, T. T. Kwon, Y. Choi and N. Choi, “AMVS-NDN : Adaptive Mobile Video Streaming with Offloading and Sharing in Wireless Named Data Networking,” in Proc. of *IEEE Conference on Computer Communications (INFOCOM), The 2nd Workshop on Emerging Design Choices inName-Oriented Networking (NOMEN)*, Turin, Italy, April 2013.

초 록

최근 모바일 트래픽의 빠른 증가는 이동통신 사업자에게 심각한 문제가 되고 있다. 이러한 문제를 해결하기 위해 단거리 통신 기술 및 모바일 소셜 네트워크 등을 이용하여 사용자간 직접 데이터를 주고 받는 오프로딩 기법을 사용하는 것에 대한 연구가 이루어지고 있다. 본 논문에서는 사용자간 직접 통신을 통한 효율적인 콘텐츠 공유 및 오프로딩 기법을 제안하고자 한다.

첫번째 연구로, 모바일 소셜 네트워크에서 사용자간 직접 전송기회를 활용해 데이터를 공유하는 모바일 트래픽 오프로딩의 핵심 기술인 TOSS를 제안 하였다. TOSS에서는 셀룰러 네트워크에서 급속히 증가하고 있는 트래픽으로 인한 네트워크 과부하를 경감시키기 위해 온라인 소셜 네트워크에서 사용자의 연결성 및 오프라인 네트워크에서 사용자의 이동성을 고려하여 콘텐츠를 전달할 사용자를 결정하고 블루투스나 와이파이 다이렉트 등의 기술을 이용해 콘텐츠를 직접 전달 하였다. 또한 소셜네트워크 서비스 사용자의 서로 다른 콘텐츠 접근 패턴, 즉 각 사용자가 콘텐츠 생성으로부터 오프로딩을 위해 콘텐츠에 접근하기까지의 시간을 고려 하였다. 본 연구에서는 이러한 요건을 고려하여 트래픽 오프로딩과 콘텐츠 확산을 모델링하고 분석 하였다. 모바일 소셜 네트워크의 데이타 셋을 기반으로 분석 결과에서 TOSS는 모든 사용자의 딜레이 요구조건을 만족시키면서 최대 86.5%의 셀룰러 트래픽을 경감시키는 것을 보였다.

두번째의 연구에서는 모바일 네트워크에서 멀티셀을 고려하여 콘텐츠를 배포하는 프레임워크에 대한 연구를 진행 하였다. 해당 프레임

워크에서 콘텐츠는 셀룰러 링크와 모바일 사용자간 로컬 링크를 통해 푸시-공유 기반의 통신으로 전달 되였다. 이러한 기법을 바탕으로 multi-compartment 모델을 이용하여 셀 간 핸드오버 및 콘텐츠 전달을 모델링 및 분석하고, 콘텐츠 전달 딜레이와 에너지 소모 사이의 trade-off를 수학적인 최적화 기법을 사용하여 해결 하였다.

본 논문에서는 이와 같이 기존의 측정 연구에 기반한 trace-driven 분석, 모델링 및 시스템 최적화에 대한 연구를 통해 모바일 소셜 네트워크에서 사용자간 직접 전송을 통한 오프로딩 기법이 고효율적임을 보였다. 또한 본 논문은 제안된 기법의 상용화 전망 및 이를 위한 이슈들에 대한 논의도 포함 하였다.

주요어: 모바일 소셜 네트워크, 트래픽 오프로딩, 기회적인 공유
학번: 2008-30709

Acknowledgement

Completing my Ph.D. degree is probably the most challenging activity of my first 30 years of my life. First, I must strongly express my heartfelt gratitude to my supervisor Prof. Yanghee Choi and Prof. Taekyoung Kwon. With their substantial guidance and constructive advice, I can diligently and continuously focus on my research work for years, and finally can complete my Ph.D. thesis. During these years I got infinite help from them in lectures and projects, as well as daily life. And I will keep their words on how to do meaningful research work and how to contribute to the community and society in my heart deeply.

I must thank Prof. Chongkwon Kim, Prof. Sunghyun Choi and Dr. Byoungjoon Lee, for their support and help to my Ph.D. thesis during this semester. Also I need to say many thanks to Prof. Min Chen, Prof. Zhu Han, and Prof. Hyunchul Kim, who helped and taught me a lot during my Ph.D. study.

I would like to thank all the lab mates and friends around me in Seoul National University, in Korea. I will remember every precious moment with everyone. During these years, I faced to a lot of problems and difficulties. I really appreciate those friends who always helped me with their best. It is my luck to have them in my life.

Sincerely I am indebted to my parents for their continuous support and encouragement. I must thank my parents for my birth to this wonderful world, for bringing me up and supporting me with their infinite love. I am so sorry that I left home very early to study in undergraduate school far away from home, and to pursue Ph.D. degree abroad. I always recall my father's motto "no pain, no gain" and my mother's optimistic life style. With their unlimited support behind me, I can keep trying my best to live well and work hard, to make my life colorful, and to devote my best effort to make significant contribution for my favorite research topics. Father and mother, I am your proud son forever. I love you.

My deepest thanks would go to my beloved wife, Qing, for your loving considerations and great confidence in me all through these years. It is my fortune to have you, and because you always understand me, trust me, tolerate me, comfort me, encourage me, and support me, unwaveringly with your best, I can continue to dedicate to my research work. These years, we together have faced to many hard periods; owning you with me, they become the happiest memories instead. You are always cheering me up and standing by me through the good times and bad. Loving you and being loved by you must be the most grateful matters in my life. Darling, I love you forever.

Xiaofei Wang

2013.07