TOSS: Traffic Offloading by Social Network Service-Based Opportunistic Sharing in Mobile Social Networks

Xiaofei Wang†°, Min Chen§, Zhu Han*, Dapeng Oliver Wu♯, Ted "Taekyoung" Kwon†
†School of Computer Science and Engineering, Seoul National University, Seoul, Korea
°Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, Canada
§School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, China
*School of Electrical and Computer Engineering, University of Houston, TX, USA
♯Department of Electrical and Computer Engineering, University of Florida, FL, USA

***Abstract*—The ever increasing traffic demand becomes a serious concern of mobile network operators. To solve this traffic explosion problem, there have been many efforts to offload the traffic from cellular links to direct communications among users. In this paper, we propose the framework of *T*raffic *O*ffloading assisted by *S*ocial network services (SNS) via opportunistic *S*haring in mobile social networks, *TOSS*, to offload SNS-based cellular traffic by user-to-user sharing. First we select a subset of users who are to receive the same content as initial seeds depending on their content spreading impacts in online 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, Device-to-device in LTE) with each other. The observation of SNS user activities reveals that individual users have distinct access patterns, which allows TOSS to exploit the user-dependent access delay between the content generation time and each user's access time for traffic offloading purposes. We model and analyze the traffic offloading and content spreading among users by taking into account various options in linking SNS and MSN trace data. The trace-driven evaluation demonstrates that TOSS can reduce up to 86.5% of the cellular traffic while satisfying the access delay requirements of all users.**

***Index Terms*—Traffic Offloading, Mobile Social Networks, Social Network Service, Opportunistic Networks**

I. Introduction

Due to the fast development of mobile communication technologies, increasingly 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), but studies [1] [2] point out that much of the traffic load is due to the downloads of the same popular files. For instance, top 10% of videos in YouTube account for nearly 80% of all the views [2]. Therefore, how to effectively reduce the duplicated download via cellular links by *offloading* the traffic via other local short-range communications so that users may re-use the content that is already downloaded by other users becomes a hot research topic.

Recently there have been many studies to exploit the people-to-people (user-to-user, or device-to-device) 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 users [3] [4] [5] [6]. Note that a MSN/DTN can be considered as the opportunistic network [7] as well.

In MSNs, users are able to discover the adjacent neighbors [8] and to set up temporary local connectivities, e.g., Bluetooth, Wi-Fi Direct, small cell network, Near-Field-Communication (NFC) [9], and Device-to-Device (D2D) [10], for sharing content with each other. Due to the user-to-user opportunistic sharing, it is advocated that by selecting an appropriate initial set of *seeds* the traffic load can be reduced significantly as studied in [5]. However, there are still several important issues in related research which are not fully elaborated:

- *How to know, or how to predict the dissemination delay of each user for each content?* Recent studies [5] [6] [11], assume the same dissemination deadline of the same content for all users; however, users indeed have various delay requirements [12].
- *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 [4] [13], but most of them focus on user mobility ignoring the practical social relationships among users.
- *Why mobile users share content with others?* Studies in [4] [5] [6] assume people always exchange content gratuitously. However, in reality, people mostly share information by "word-of-mouth" propagation [14], while they impact others due to their social impacts, and thus we need to exploit the realistic social relationship.

Regarding above issues, while we seek to exploit the "social" aspect of MSNs, we discover that there is a dramatic rise in the number of mobile users who participate in the online Social Network Services (SNSs), e.g., Facebook, Twitter, Sina Weibo, and so on, where more and more content is recommended and spread rapidly and widely [15] [16]. By investigating related measurements and modeling studies of the MSNs and SNSs, we have discovered the following key points, which can be utilized for content dissemination:

- In online SNSs, the access pattern of each user can be measured, and statistically modeled. That is, we can analyze the **access delay** between the content generation time and the user access time [17], which is per-user dependent mainly due to people's different life styles [12] [15]. We can disseminate the content of interest to users considering their different delay sensitivities.
- In online SNSs, a user's influence, or **spreading impact** to other users, can be modeled based on the analysis of social behavior histories, i.e., the forwarding probability.
- In offline MSNs, the mobility patterns of users can be measured and modeled as well [5] [11] [18], and hence a different offline **mobility 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 Sec. II), which is similar to those of offline MSNs [14][19]. Users are mostly both clustered by regions and interests, which can be exploited for content sharing and thus traffic offloading.

Therefore, we are 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.

The rest of the paper is organized as follows. After reviewing the related work in Sec. II, we detail the TOSS framework in Sec. III. Optimization issues are discussed in Sec. IV. The trace-driven evaluation and analysis are shown in Sec. V and Sec. VI, respectively, followed by conclusion in Sec. VII.

II. Related Work

A. Opportunistic Sharing and Traffic Offloading in MSNs

The epidemic content delivery in DTNs/MSNs has been extensively studied in recent years as it can be effectively utilized for traffic offloading. Zhang et al. [11] have developed a differentiation-based model to study the delay of epidemic content delivery. By similar modeling, Li et al. [6] also have designed an energy-efficient opportunistic content delivery framework in DTNs. The optimality of content dissemination by exploiting user-to-user contacts has been modeled as a social welfare maximization problem in [4]. Regarding the slow start and long completion time of the epidemic delivery, strategic pushing is studied to expedite the dissemination as studied in [5], while it extends the opportunistic sharing model into multi-cell mobile networks.

Accelerating the content dissemination by leveraging users' social relationships becomes a popular research topic recently. BUBBLE Rap [20] utilizes social grouping characteristics for content dissemination in DTNs. And recent study in [13] analyzes the social participation for content dissemination in MSNs based on selection of the optimal initial seed users. Also the study in [21] proposes to assign interest tags to the users and content objects to identify their preferences of content, and then to utilize users local centrality for efficient content sharing in DTNs. Bao et al. carried real tests in Manhattan and identified the sharing-based offloading can reduce 30% to 70% mobile traffic [22].

The sharing in MSNs mainly relies on infrastructureless user-to-user local communication techniques, e.g., Bluetooth, Wi-Fi Direct, NFC [9], and so on. For example, Apple's Airdrop [23] provides convenient local content sharing functionality via Wi-Fi Direct. Recently the Device-to-Device (D2D) communication underlaying the 3GPP LTE cellular network in the operator authorized spectrum, which is an efficient enabler of local services with high throughput and efficient content delivery but limited interference impact on the primary cellular network, and thus becomes quite popular [10] as well. However, the sharing-based offloading needs effective incentive-based business model, to encourage users to share content with others locally, such as the pricing study Win-Coupon in [24].

B. Information/Content Spreading in SNSs

The study in [25] indicates that in the real world, the "opinion leaders" who have strong social impact may perform the key roles for spreading information to most people, due to the effect of "word-of-mouth" [14]. Similarly in online SNSs, only a small amount of users significantly impact the information spreading to the most of the other users [15] [26]. There has been some studies using probabilistic modeling to analyze the information/content commenting or re-sharing activities, in order to quantify the spreading impact among users [17] [26] [27]. The measurement studies in [12] [15] [17] point 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 mainly depends on people's different life styles. Researchers can analyze, and predict the sharing activities and the access delays of SNS users based on measurement traces [15] [16] [17].

Studies in [14] [19] report that user relationships and interests in SNSs have significant **homophily** and **locality** characteristics as similar to those in MSNs. Homophily here means the online and offline users are both clustered by regions and interests, which is also called "birds-of-a-feather" effect [28]. People with similar interests like to share and transfer the interesting information with each other. The locality here means that people who are geographically close may have similar trends of accessing the content and sharing with each other. Even in online SNSs, users may significantly interact with and thus impact others in proximity [14] [19]. Thus, the online and offline locality nature of users has been utilized to facilitate the content delivery in [1] [19].

III. The TOSS Framework

A. Preliminaries

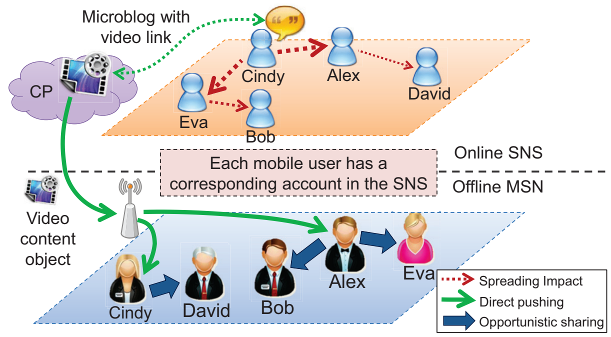


Fig. 1. Illustration of TOSS with two layers: online SNS and offline MSN

As shown in Fig. 1, the TOSS framework entails both an online SNS and an offline MSN. Note that we declare that "online" is for the accounts in the Internet-based virtual world, and "offline" is for the real people in the physical world. 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, e.g., Twitter, Sina Weibo. TOSS can also work with SNSs based on the bidirectional graph (e.g., Facebook), as it is a subset of the directional graph. The online SNS can thus be represented by, $G(V, E)$, where V is the set of users, 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 from u_i to u_j, denoted by v_{ij}, indicates that u_i has a direct impact to u_j for content spreading.

We define the home-site, where a user creates 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**, e.g., a tweet in Twitter or a post in Facebook. And then 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 the publishing 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, and thus we only consider the re-share activities. Afterwards, what TOSS tries 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. Also because of the emerging trend of integration of SNSs with Content Providers (CPs), more traffic will be inflowed through SNSs inherently. And hence TOSS will benefit, and thus becomes more effective for offloading traffic.

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 encounters. We will discuss their calculation in Sec. III-B and Sec. III-D.

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. The sharing in TOSS can be considered as "prefetching" [29] before the user access. 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. 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 always meets 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. Therefore, TOSS reduces 3/5 cellular traffic in this scenario, because the cellular network only transfers two copies instead of five copies, to the five users.

B. Spreading Impact in the Online SNS

We extend the previous probabilistic models [17] [26] [27] 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}, $0 \leq \gamma_{ij} \leq 1$, as 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. And thus, γ_{ij} is the probability that u_j will re-share the microposts from u_i.

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 himself/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:

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

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

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

$$\gamma_{ij}^h = 1 - \prod_{k \in D_i^{h-1} \cap U_j^1} (1 - \gamma_{ik}^{h-1} * \gamma_{kj}^1). \tag{4}$$

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:

$$\gamma_{ij}^* = 1 - \prod_{n=1}^{H} (1 - \gamma_{ij}^n), \tag{5}$$

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

$$I_i^{S\rightarrow} = \sum_{j=1}^{N} \gamma_{ij}^*, \quad I_i^{S\leftarrow} = \sum_{j=1}^{N} \gamma_{ji}^*. \tag{6}$$

Note that it is reported in [15] [28] that the average path length in SNSs is normally 4.12 and the spreading impact after 3 hops becomes negligible.

C. Access Delays of Users in the SNS

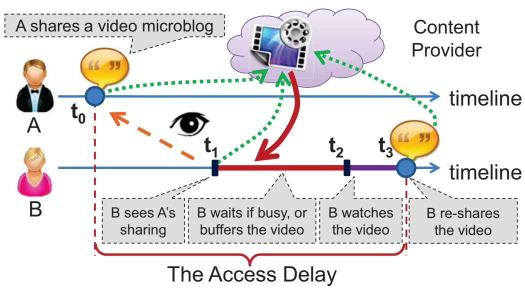
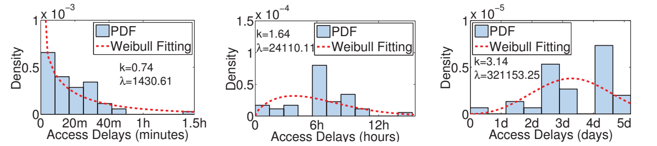


Fig. 2. Illustration of the content access delay

In SNSs, some users 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 [12] [15]. As illustrated in Fig. 2, user A creates a microblog for an interesting video in the SNS at t_0. One of A's followers, B, happens to see A's microblog 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. So we consider B's access delay as $t_3 - t_0$, which can be captured from the SNS measurement trace.

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 Sec. V). The access delay is gathered as the time difference between the generation time of the original microblog and the time of re-sharing by a follower. We pick up three real users from the trace, and plot their access delays by probability distribution function (PDF) in Fig. 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.



(a) User u_1 (b) User u_2 (c) User u_3

Fig. 3. The access delay distributions of three real users with Weibull fitting

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 [4], $A_i(t)$ can be considered as the access utility function, in order to calculate the user satisfaction performance. If the content object is already obtained locally in the user's device when he/she has the highest probability to access the content, he/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, which is commonly used for profiling user behaviors in SNSs [30]:

$$A_i(t, \beta_i, k_i) = \frac{k_i}{\beta_i}\left(\frac{t}{\beta_i}\right)^{k_i - 1} e^{-\left(\frac{t}{\beta_i}\right)^{k_i}}, \quad t \geq 0, \tag{7}$$

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).

D. Mobility Impact in the Offline MSN

It has been studied that mobile users in the offline MSNs (or DTNs), have different mobility patterns [5] [11] [18], 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 [8].

Referring to [4] [5] [6] [11] [18] [21], 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. The contact duration is ignorable in TOSS, because we assume the content delivery is always successfully finished during the contact due to the high bandwidth of local communications (e.g., Wi-Fi) [4] [5] [11] [21].

We adopt the epidemic modeling from [5] [6] [11] with the continuous time Markov chain to model the opportunistic sharing in TOSS. We assume that file objects can be shared instantly via high-capable device-to-device link when two users contact, and hence the contact duration time is thus not considered. We let $S_i(t)$ be the probability that user u_i may have the content at t, $0 \le S_i(t) \le 1$, while $1 - S_i(t)$ is the probability that user u_i has not received the content until t. The increment of $S(t)$ within a period Δt, that is $S_i(t+\Delta t) - S_i(t)$, will be calculated as following.

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:

$$\varepsilon_{ij} = \left(1 - e^{-\lambda_{ij}\Delta t}\right) \cdot \gamma_{ji}^{*} \cdot S_j(t), \tag{8}$$

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 microblogs from u_i and *(ii)* the sharing probability that u_i can obtain the content object from u_j.

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

$$1 - \prod_{j=1, j\neq i}^{N} (1 - \varepsilon_{ij}). \tag{9}$$

The increment of the probability that u_i has the content is,

$$S_i(t+\Delta t) - S_i(t) = (1 - S_i(t)) \cdot \left(1 - \prod_{j=1, j\neq i}^{N} (1 - \varepsilon_{ij})\right). \tag{10}$$

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

$$\begin{aligned} \overset{\bullet}{S_i}(t) &= \lim_{\Delta t \to 0} \frac{S_i(t+\Delta t) - S_i(t)}{\Delta t} \\ &= (1 - S_i(t)) \cdot \sum_{j=1, j\neq i}^{N} \lambda_{ij} \cdot \gamma_{ji}^{*} \cdot S_j(t), \end{aligned} \tag{11}$$

where initially $S_i(0) = p_i$ 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 [31]. Due to the limited space, we skip the details of the procedure for getting numerical solutions.

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^*:

$$t_i^* = S_i^{-1}\left(\left\{\gamma_{ji}^{*}\right\}, \left\{\lambda_{ij}\right\}, \overrightarrow{p}\right) \quad , \quad j = 1, ..., N, j \neq i, \tag{12}$$

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

$$I_i^{M\rightarrow} = I_i^{M\leftarrow} = \lambda_i^* = \sum_{j=1}^{N} \lambda_{ij}. \tag{13}$$

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 $S_i(t)$.

IV. System Optimization and Heuristic Algorithm

TOSS's objective is to choose proper set of seeds for initial pushing, $\overrightarrow{p}$, by evaluating I^S (both incoming and outgoing) and I^M values of all users, 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

$$\begin{aligned} \underset{\overrightarrow{p}}{\text{Maximize}} : & \sum_{i=1}^{N} A_i\left(t_i^*, \beta_i, k_i\right) \\ & = \sum_{i=1}^{N} A_i\left(S_i^{-1}\left(\left\{\gamma_{ji}^{*}\right\}, \left\{\lambda_{ij}\right\}, \overrightarrow{p}\right), \beta_i, k_i\right) \\ & \quad (j = 1, ..., N, j \neq i) \\ \text{Subject to} : & |\overrightarrow{p}| \le C \quad , \end{aligned} \tag{14}$$

where the number of initial pushing seeds is constrained by C, 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 [4]. 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.

We design a heuristic algorithm to find the near-optimal solution $\overrightarrow{p}$ for maximizing $\sum A_i(t)$ numerically, based on the popular hill-climbing method, but due to the space limit, we will skip to list it in this paper. 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.

V. Measurement Results

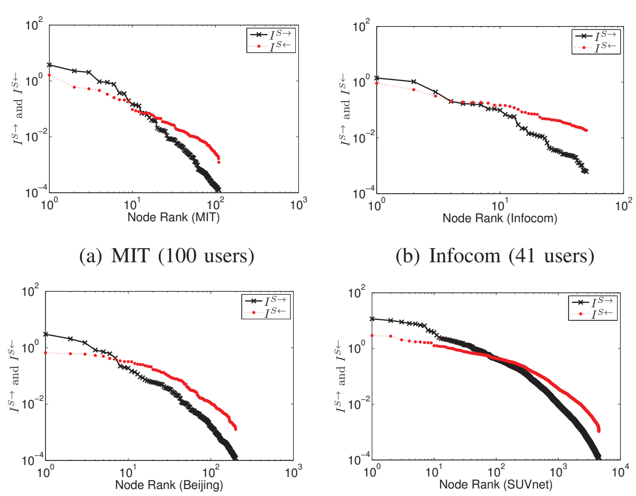
To evaluate the effectiveness of TOSS framework, we need SNS trace data to quantify the spreading impact factors and access delays, as well as MSN trace data to analyze the mobility impact. However, there is no publicly available trace data that contains both the SNS and the MSN activities. Thus, we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

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. 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 robot accounts in Sina Weibo, which are with artificial intelligence, and 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 robot accounts is out of the scope of this paper. In summary, we believe that the 2.2 million user base can reflect the ground-truth of the I^S impact factors and the access delay statistics.

A. Online SNS - Spreading Impact, γ_{ij} and I^S

Recall that I^S is the overall spreading impact of the user to all users in the SNS, calculated by Eq. (6). However, calculating I^S for the whole user base takes substantially long time. Thus, we analyze the sub-graphs of corresponding number of users from the whole social graph by random walking according to the scale of the mobility traces (to be detailed in Sec. V-C). We consider up to 4-hop paths ($H = 4$) among the users in the graphs as suggested in [15].

There have been some related measurement studies in [15] [32] and [33] pointing out that: the SNS is a **scale-free** complex network, in which node strength distribution follows the power law, at least asymptotically. That is 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 strength of a node to the network [32] and [33]. So due to the nature characteristics of scale-free complex networks, sub-graphs (with not too small size) from the whole graph by random walking can still obtain similar power law characteristics.



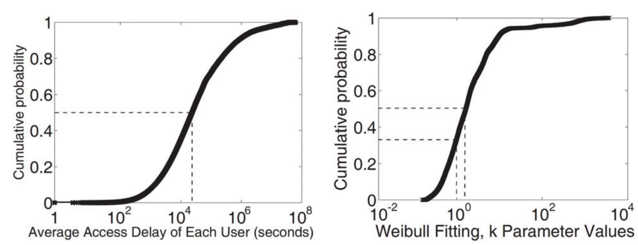
(a) MIT (100 users) (b) Infocom (41 users)

(c) Beijing (182 users) (d) SUVnet (4,311 users)

Fig. 4. Measurement values of I^S for sub-graphs sampled from the SNS graph with different sizes corresponding to the mobility traces

We then check the sub-graphs that we abstract from the online SNS graph with the sizes corresponding to the mobility traces (to be detailed in Sec. V-C), and for each trace we abstract sub-graphs for five times, and then make average value. We draw the log-log plots for $I^{S\rightarrow}$ and $I^{S\leftarrow}$ of the nodes from the sampled sub-graphs as shown in Fig. 4. We can see that a smaller number of people have significant outgoing impact ($I^{S\rightarrow}$) to the whole SNS, while many users have very small impact. Also we see that many users are more likely to be impacted rather than impacting others ($I^{S\rightarrow} < I^{S\leftarrow}$). All of the figures are able to reflect the asymptotical power law trend. So conclusively, all of the sub-graphs with different sizes can still represent the SNS characteristics, 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.

B. Online SNS - Access Delay of u_i, $A_i(t)$



(a) Average access delays (b) Weibull parameter, k

Fig. 5. Access delays and fitting parameters

Measurement results of the access delays on the whole user base are shown in Fig. 5. From the cumulative distribution function (CDF) of the average of all the access delays of each user in Fig. 5(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 small 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. 5(b). Therefore, we verify that a large portion of users access the SNSs with sufficiently large delays, which TOSS can utilize to disseminate the content object by offline opportunistic sharing.

C. Offline MSNs - λ_{ij} and I^M

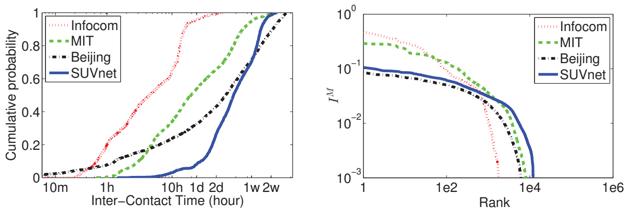
We choose four mobility traces, MIT [34], Infocom [35], Beijing [36], and SUVnet [37], in order to evaluate the performance of TOSS. These traces record either direct contacts among users carrying mobile devices or GPS-coordinates of each user's mobile route, and traces details are shown in Table I. 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. Beijing and the SUVnet traces have no contact records, but only GPS coordinates by time; we assume a contact once two users are within 20 meters during 20s.

We analyze the traces and obtain the inter-contact intervals of all user pairs, as shown in Fig. 6(a). The Infocom trace

TABLE I
MOBILITY TRACES (BL. IS FOR BLUETOOTH)

Trace	Link	Users	Days	Contacts	Avg.λ
MIT[34]	Bl.	100	246	54,667	0.01532
Infocom[35]	Bl.	41	4	22,459	0.14167
Beijing[36]	/	182	150	8,894	0.00023
SUVnet[37]	/	4,311	30	169,762	0.00131

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 rates 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. I^M values of all users of the traces (values smaller than 0.001 are ignored) are plotted in Fig. 6(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. 6. λ_{ij} and I^M values

VI. TOSS EVALUATION RESULTS

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 the TOSS framework.

A. How $\overrightarrow{p}$ Impacts the Total Access Utility, $\sum A_i(t)$

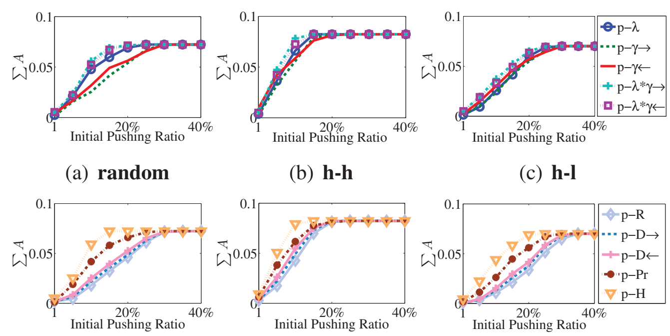
Due to the lack of a trace that contains the activities of the same users in both online SNSs and offline MSNs, we consider three schemes for mapping SNS users to MSN users in each of the four mobility traces: (1) **random**: SNS users are randomly mapped to MSN users; (2) **h-h**: both SNS and MSN users 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. As discussed in Sec. V-A, since the number of SNS users is much larger than that of MSN users in each trace, we pick accounts from the sub-graphs of the SNS by the random walking method to match the number of MSN users in each trace.

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 a descending order and choose the top C ones (similar to [4]);
- **p-**$\gamma^{\rightarrow}$: we sort users by $I^{S\rightarrow}$ ($\sum \gamma_{ij}^*$) in a descending order and choose the top C ones (similar to [20]);
- **p-**$\gamma^{\leftarrow}$: we sort users by $I^{S\leftarrow}$ ($\sum \gamma_{ji}^*$) in a descending order and choose the top C ones;
- **p-**$\lambda * \gamma^{\rightarrow}$: we sort users by $I^M * I^{S\rightarrow}$ conjunctively in a descending order and choose the top C ones;
- **p-**$\lambda * \gamma^{\leftarrow}$: we sort users by $I^M * I^{S\leftarrow}$ conjunctively in a 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 [38] 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 SNS sub-graphs and obtain the PageRank scores. We also consider the random pushing and the heuristic algorithm, and hence we have five more pushing strategies:

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



(a) **random** (b) **h-h** (c) **h-l**

(d) **random** (e) **h-h** (f) **h-l**

Fig. 7. As C increases, $\sum A_i(t)$ converges - 3 mapping schemes; 5 pushing strategies from γ and λ; 5 more from the graph; MIT trace as an example

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. 7, and we skip to show the results of other traces since they show very similar trends. The percentage in the figures is the pushing ratio of C to 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 a 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 MSN.

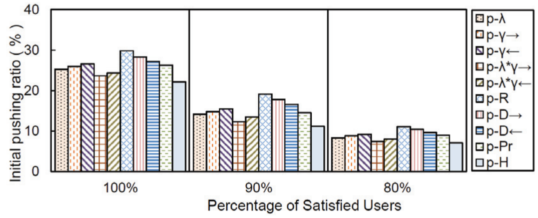
B. Satisfying 100%, 90%, and 80% of Users

Recall that the access utility function of u_i is $A_i(t)$. A user is **satisfied**, if he/she can obtain the content when his/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). 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.

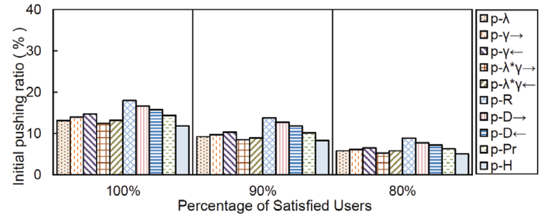
From Sec. IV and Fig. 7, $\sum A_i(t)$ is an increasing function of C (i.e., $|\overrightarrow{p}|$), and the number of satisfied user is also an increasing function of C. The C value that makes $\sum A_i(t)$ approach its maximum will be the standard number of initial pushing seeds for satisfying 100% 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. 8, 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 ignores the historical spreading impact, while our proposed factors (γ) make better sense. In MIT and Infocom traces, λ-based strategies perform 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**.

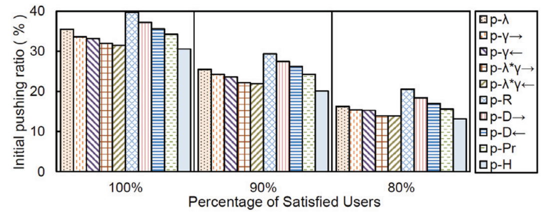
When we target to satisfy 90% of all users, the required initial pushing ratio is reduced significantly. With simple pushing strategies, for the MIT and the Infocom traces, only 15.4% and 10.5% of users need to be the initial seeds on average. The number of initial seeds is further dramatically reduced, 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.



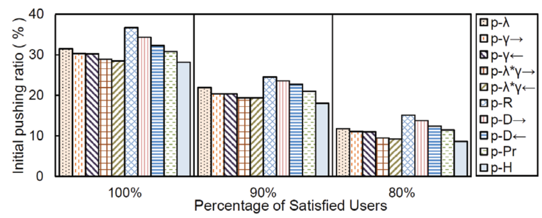
(a) MIT



(b) Infocom



(c) Beijing



(d) SUVnet

Fig. 8. Initial pushing ratios to satisfy 100%, 90%, and 80% of all users

TABLE II
PERCENTAGE (%) OF TRAFFIC REDUCTION WITH ON-DEMAND DELIVERY - AVERAGE OF 9 SIMPLE PUSHING STRATEGIES / HEURISTIC

Trace	100%	90%	80%
MIT[34]	73.6 / 76.3	**74.6 / 76.9**	70.9 / 72.2
Infocom[35]	**85.3 / 86.5**	79.5 / 80.4	73.4 / 74.1
Beijing[36]	65.3 / 68.4	**65.0 / 68.9**	63.8 / 65.2
SUVnet[37]	68.5 / 70.3	**68.7 / 71.0**	68.3 /70.7

C. On-Demand Delivery

If a user who has not obtained the content (by initial pushing or sharing) until he/she actually accesses it, we have to deliver it over a cellular link, which is called on-demand delivery. Then the cellular 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. Table II shows how much traffic is offloaded from cellular links for the three cases, where the offloaded traffic ratios of the nine pushing strategies are averaged, which are followed with that of **p-H** after "/". 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 the remaining 10% and 20% of users may induce the increment of the total cellular traffic instead. 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. It is the issue about how to deal with those users with low I^S and I^M, who will burden TOSS to select the optimal initial seeds, as they are hard to reach by even many hops of sharing. Instead, it will be better for TOSS to exclude them for a better initial pushing solution to satisfy most of users, and finally to carry out on-demand delivery for them.

VII. Conclusion

In this paper, 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. Also the user access delays are exploited and utilized for content sharing. Trace-driven evaluation reveals that TOSS can reduce up to 86.5% of the cellular traffic while guaranteeing the access delay requirements of all users.

Acknowledgment

This research was partially supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2013R1A2A2A01016562), and partially supported by US NSF CNS-1265268, CNS-1117560, ECCS-1028782, and CNS-0953377. Prof. Min Chen's work was supported by Youth 1000 Talent Program, China National Natural Science Foundation (No. 61300224), China Ministry of Science and Technology (MOST), the International Science and Technology Collaboration Program (Project No.: S2014GAT014), and the Hubei Provincial Key Project (No. 2013CFA051).

References

[1] S. Scellato, C. Mascolo, M. Musolesi, and J. Crowcroft, "Track Globally, Deliver Locally: Improving Content Delivery Networks by Tracking Geographic Social Cascades," *WWW*, 2011.

[2] 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," *ACM IMC*, 2007.

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

[4] S. Ioannidis, A. Chaintreau, and L. Massoulie, "Optimal and Scalable Distribution of Content Updates over An MSN," *IEEE INFOCOM*, 2009.

[5] X. Wang, M. Chen, Z. Han, T. Kwon, and Y. Choi, "Content Dissemination by Push and Share in Mobile Cellular Networks: An Analytical Study," *IEEE MASS*, 2012.

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

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

[8] B. Han and A. Srinivasan, "eDiscovery: Energy Efficient Device Discovery for Mobile Opportunistic Communications," *IEEE ICNP*, 2012.

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

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

[11] X. Zhang, G. Neglia, J. Kurose, and D. Towsley, "Performance Modeling of Epidemic Routing," *Computer Networks*, no.10, pp.2867–2891, 2005.

[12] J. Yang and J. Leskovec, "Patterns of Temporal Variation in Online Media," *ACM WSDM*, 2011.

[13] 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," *IEEE TMC*, vol.11, no.5, pp.821–834, 2011.

[14] T. Rodrigues, F. Benvenuto, M. Cha, K. Gummadi, and V. Almeida, "On Word-of-Mouth Based Discovery of the Web," *ACM IMC*, 2011.

[15] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a Social Network or a News Media?" *WWW*, 2010.

[16] G. V. Steeg and A. Galstyan, "Information Transfer in Social Media," *WWW*, 2012.

[17] G. Comarela, M. Crovella, V. Almeida, and F. benevenuto, "Understanding Factors that Affect Response Rates in Twitter," *ACM HT*, 2012.

[18] T. Karagiannis, J. Y. Le Boudec, and M. Vojnovic, "Power Law and Exponential Decay of Inter Contact Times between Mobile Devices," *ACM MobiCom*, 2007.

[19] M. P. Wittie, V. Pejovic, L. Deek, K. C. Almeroth, and B. Y. Zhao, "Exploiting Locality of Interest in Online SNS," *ACM CoNEXT*, 2010.

[20] P. Hui, J. Crowcroft, and E. Yoneki "BUBBLE Rap: Social-Based Forwarding in Delay-Tolerant Networks," *ACM MobiHoc*, 2008.

[21] W. Gao and G. Cao, "User-Centric Data Dissemination in Disruption Tolerant Networks," *IEEE INFOCOM*, 2011.

[22] X. Bao, Y. Lin, U. Lee, I. Rimac, and R. Choudhury, "Exploiting Naturally Clustered Mobile Devices to Offload," *IEEE INFOCOM*, 2013.

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

[24] X. Zhuo, W. Gao, G. Cao, and Y. Dai, "Win-Coupon: An incentive framework for 3G traffic offloading," *IEEE ICNP*, 2011.

[25] D. J. Watts and P. S. Dodds, "Influentials, Networks, and Public Opinion Formation," *Journal of Consumer Research*, vol.34, pp.441–458, 2007.

[26] A. Goyal, F. Bonchi, and L. Lakshmanan, "Learning Influence Probabilities in Social Networks," *ACM WSDM*, 2010.

[27] R. Xiang, J. Neville, and M. Rogati, "Modeling Relationship Strengthin Online Social Networks," *WWW*, 2010.

[28] 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?" *CoRR*, 2010.

[29] X. Wang, T. T. Kwon, Y. Choi, H. Wang, J. Liu, "Cloud-Assisted Adaptive Video Streaming and Social-Aware Video Prefetching for Mobile Users," *IEEE Wireless Communication*, vol.20, pp.72–79, 2013.

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

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

[32] R. P. Satorras and A. Vespignani, "Epidemic Spreading in Scale-Free Networks," *Phys. Rev. Letter*, vol.86, issue.14, 2001.

[33] D. Towsley, "A Walk in the Dark: Random Walks and Network Discovery," *IFIP Networking, Keynote*, 2013.

[34] N. Eagle and A. Pentland, "Reality mining: Sensing complex social systems," *Pers. Ubiquitous Comput.*, vol.10, no.4, pp.255–268, 2006.

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

[36] Y. Zheng, L. Zhang, X. Xie, and W. Ma, "Mining Interesting Locations and Travel Sequences from GPS Trajectories," *WWW*, 2009.

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

[38] 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.