DEVICE-TO-DEVICE COMMUNICATIONS WITH SOCIAL AWARENESS

TAG-ASSISTED SOCIAL-AWARE OPPORTUNISTIC DEVICE-TO-DEVICE SHARING FOR TRAFFIC OFFLOADING IN MOBILE SOCIAL NETWORKS

XIAOFEI WANG, ZHENGGUO SHENG, SHUSEN YANG, AND VICTOR C. M. LEUNG

ABSTRACT

In recent years, the service demand for rich multimedia over mobile networks has continually been soaring at a tremendous pace. To solve the critical problem of mobile traffic explosion, substantial efforts have been made by researchers to try to offload mobile traffic from infrastructured cellular links to direct short-range communications locally among nearby users. In this article, we discuss the potential of combining users' online and offline social impacts to exploit D2D opportunistic sharing for offloading mobile traffic. We propose a tag-assisted social-aware D2D sharing framework, TASA, with corresponding optimization models, architecture design, and communication protocols. Through extensive simulations based on real data traces, we demonstrate that TASA can offload up to 78.9 percent of mobile traffic effectively.

INTRODUCTION

Recently, mobile networks have been supporting a growing number of multimedia services, and mobile users are always downloading huge amounts of multimedia content files onto their mobile devices (e.g., smartphones). This results in explosively growing traffic load [1], which becomes one of the most severe concerns of mobile network operators (MNOs). To effectively support the growing traffic load, MNOs and network equipment vendors have been trying to adopt more and more sophisticated communication techniques, but the efficiency of utilizing authorized radio spectrum resources is nearly reaching the theoretical capacity limits. One interesting observation that has gained the interest of the mobile industry is that, as indicated in [2], most of the traffic load on the Internet is generated by duplicate downloads of the same popular contents. For instance, the top 10 percent of popular videos on YouTube may account for nearly 80 percent of all view requests [2]. Thus, the MNOs have to deliver the same video streams from remote servers to mobile users multiple times, which causes a huge waste of resource in MNOs' networks. In addition to cellular communications, a mobile user can also obtain content files from other users in proximity through short-range communications such as Bluetooth and WiFi Direct. Therefore, due to the fact that people are always clustered as crowds in certain areas (e.g., offices, apartments, subways, cafes), the *offloading* solution, for offloading mobile traffic for sharing among people, has great potential to significantly reduce the cellular traffic load caused by duplicate contents, and has been attracting increasing research interest from both academia and industry. Until now, a number of techniques have exploited device-to-device (D2D) sharing opportunities during intermittent encounters of mobile users for offloading traffic into mobile social networks (MSNs) (a type of delay-tolerant network with social relations of users) [3, 4]. In MSNs, users are able to securely discover adjacent users [5] for establishing temporary local connectivities and thus sharing already downloaded contents with each other. It is worth mentioning that recently, the Third Generation Partnership Project (3GPP) has designed a D2D technique as a novel and attractive underlay to Long Term Evolution-Advanced (LTE-A) networks [6], by which mobile users can use the authorized spectrum of MNOs for communicating directly without infrastructure support. For brevity, we use the term *D2D sharing* throughout this article for any kind of local direct short-range communication technique.

Toward such a D2D-based traffic offloading concept, one essential issue is how to initialize the content dissemination by appropriate *users*, called *seeds*, which should have high potential to download the content files and to move around for sharing with others by D2D communications [3, 4]. In this regard, two challenges should be further elaborated:

1 How do we model and predict the content sharing pattern of each user over interesting contents regarding individual social activities and life patterns [7]?

Xiaofei Wang is with Tianjin Key Laboratory of Advanced Networking, School of Computer Science and Technology, Tianjin University.

Zhengguo Sheng is with the University of Sussex.

Shusen Yang (corresponding author) is with the School of Mathematics and Statistics at Xi'an Jiaotong University.

Victor C.M. Leung is with the University of British Columbia.

1536-1284/16/$25.00 © 2016 IEEE

2 How do we design an effective content dissemination strategy by exploiting realistic social relationships among users?

Although there have been some related studies for social D2D sharing in MSNs, such as the studies in [4, 8], which both discuss the optimality of social-aware D2D-based content dissemination by selecting initial seed users, they all define the social relations based on mobility statistics, which cannot indicate users' real social relationships. Therefore, for exploring the natural social aspect of MSNs, we observe the dramatic rise in the number of people who frequently participate in activities in online social network services (SNSs), such as Facebook, Twitter, and Sina Weibo, where more and more popular contents are recommended and propagated widely and rapidly [9, 10]. Social relationships among SNS users have been researched intensively in the literature, and it has been determined that the "opinion leaders" with strong social impact always perform essential roles in spreading interesting and popular contents due to the "word-of-mouth" effect [11]. We are thus motivated to combine users' social properties from SNSs and real mobility patterns from MSNs for social-aware D2D sharing and traffic offloading.

Consequently, we utilize several important properties and observations of online SNSs and offline MSNs as follows:

• For quantifying and predicting the social-aware D2D sharing potential, the *spreading impact* can be obtained from SNS behavior histories, based on either probabilistic modeling [9, 12] or tag systems [8], and the *mobility impact* among users can be analyzed based on mobility traces [3, 4, 8].

• Studies in [7, 9, 10] have pointed out that there are often certain delays in sharing and re-sharing behaviors. The users' **access delays** are mostly user-dependent due to their differences of lifestyles, and hence can be measured and modeled for guaranteeing the quality of service (QoS) via delay-tolerant D2D content sharing [8, 9].

• The **homophily**[1] and **locality**[2] properties of user relations and interests in SNSs and MSNs indicate that users are clustered mostly by regions and interests both online and offline, and thus prove the feasibility of combining online and offline user data for facilitating social D2D sharing-based traffic offloading [3, 11].

In this article, we propose a novel framework to offload cellular traffic via tag-assisted social-aware (TASA) opportunistic sharing by exploiting D2D communications in MSNs. TASA is based on consideration of the *tags* of users and contents, and hence the modeling of social and mobility impacts; the tags indicating the social ties among users can accelerate the choosing of initial seeds and the spreading procedure of contents. In addition, systems architecture and protocols of the TASA framework are designed by considering the aforementioned issues. We also evaluate TASA with real traces to prove its effectiveness.

TASA Framework

We illustrate the TASA framework by one typical scenario, shown in Fig. 1, which entails both a layer of online SNS and a layer of offline MSN.

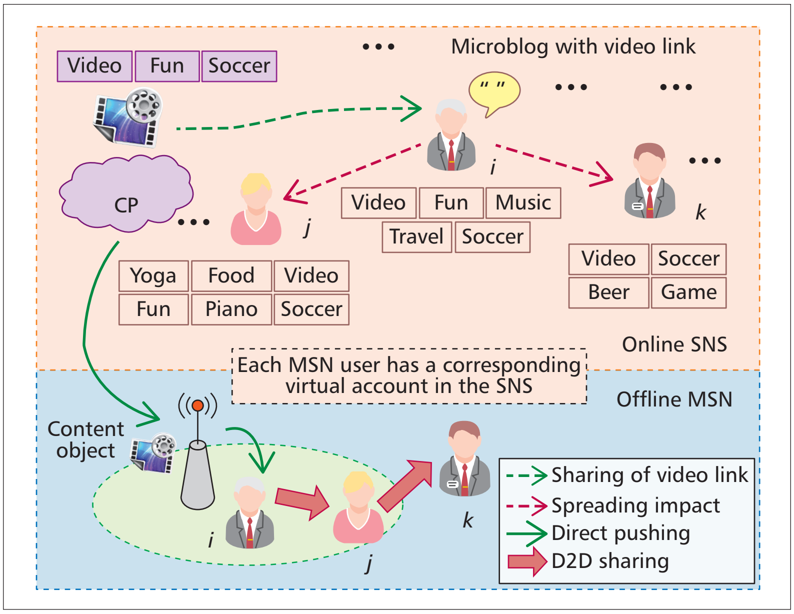

Figure 1. Illustration of TASA framework.

Here we again declare that the term *online* means the virtual accounts in the SNSs on the Internet, and the term *offline* means real human beings in the physical world (i.e., MSNs), corresponding with their online accounts. We assume that in this typical scenario, there are total N mobile users, u_i, $i = 1, \ldots, N$, who are real people, and each has a corresponding SNS account. The dashed arrow in the online SNS layer in Fig. 1 from u_i to u_j means u_i has a direct social impact on u_j for content spreading. In the offline MSN layer, we use a solid red arrow to indicate practical D2D-based sharing of a content object if any two users are within D2D communication range.

Referring to the fact that many popular SNSs have tagging systems (e.g., Twitter and Sinca Weibo), in TASA, contents can be assigned (either automatically or manually) by several tags, including *video*, *fun*, *soccer*, and so on, while users can also be characterized with a certain amount of tags based on the contents that they have published or re-posted, such as *video*, *soccer*, *beer*, and *game*. Furthermore, contents' tags can be interactively updated based on the tags of users that have accessed the content. The modeling of social relationships based on the tags will be detailed later.

For user u_i, TASA defines two online SNS factors, the *spreading impact* factor, I_i^S, and the *social similarity* factor between two users based on tags, τ_{ij}, indicating a user's importance in propagating the posts, and one offline MSN factor, the *mobility impact*, I_i^M, indicating a user's importance in sharing content objects (files) via encounters.

Based on the above factors, what TASA tries to achieve is that while the posts with the contents (i.e., posts) are being spread through the online SNS, by selecting a proper subset of users from the potential receiver group as seeds for directly pushing the real content object via cellular links, the object will then be delivered among users in the offline MSN by exploiting the social D2D sharing. We define a vector $\vec{p}$ to indicate

[1] http://en.wikipedia.org/wiki/Homophily

[2] http://en.wikipedia.org/wiki/Locality_of_reference

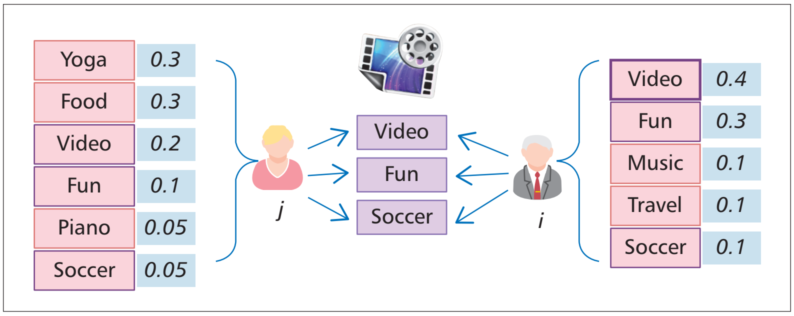

Figure 2. Illustration of social tags.

whether TASA will push the object to any user by using the cellular link or not. For example, $p_i = 1$ means to download the content object directly to user u_i via the cellular link. Note that the TASA framework is not strictly confined to the dissemination of one popular content object to all users, but can extend to apply to normal deliveries of any content object(s) to any group of potential recipients. The sharing in TASA can be considered as "prefetching" in advance before users' practically accessing activities.

From TASA's illustration in Fig. 1, u_i shares a video (link) in the online SNS with the tags of *video*, *fun* and *soccer* to u_j and u_k, while the video file is first downloaded by u_i via a cellular link and stored in u_i's phone. In the offline MSN, u_k is distant from u_i geographically, but u_j is in proximity to u_i. TASA analyzes that the I^S and I^M impact factors and τ similarity between u_i and u_j are very strong as well, so it directly lets u_i's phone share the video with u_j by D2D connectivity. TASA further detects that the impact factors and similarity between u_j and u_k via u_i are strong, so although u_j and u_k are not direct friends, TASA will trigger u_j to deliver the content to u_k by D2D sharing. Apparently, TASA reduces 2/3 of the cellular traffic in this simple scenario.

Social-Awareness for D2D Sharing: From Online to Offline

Online Spreading Impact

Tag-based impact is quantified based on related probabilistic models [12], and tags are estimated as a comprehensive probability from user interests. TASA generally allows a user to have various interests in different content files. As illustrated in Fig. 2, tags have ample information to compute the similarity of two users; the weights of tags will be especially important to indicate users' preferences in contents.

We hence model the tags within a fixed keyword space with size N_T to describe how the probability of user interests can be calculated. The tag profile of a user u_i is an N_T probability vector, $G_i = [g_{i1}, g_{i2}, \ldots, g_{iN_T}]$, where g_{ik} indicates the user weight (probability) to be interested in the kth tag. In practice, g_{ik} is used to compare the user's interests in different tags. Hence,

$$\sum_{k=1}^{N_T} g_{ik} = 1$$

for $\forall i$,

$$\tau_{ij} = \sum_{k=1}^{N_T} \left(g_{ik} \cdot g_{jk} \right)$$

and τ is calculated:

$$I_i^S = \sum_{j=1, j\neq i}^{N} \tau_{ij}.$$

For example, from Fig. 2, user u_i has a *video* tag with weight 0.2, a *fun* tag with weight 0.1, and a *soccer* tag with weight 0.05. Then the social similarity τ between the two person will be calculated by $0.2 \times 0.4 + 0.1 \times 0.3 + 0.05 \times 0.1 = 0.115$.

For obtaining realistic tag data from SNSs, we chose one of the most popular Chinese SNSs, Sina Weibo, and we kept tracking of 2.2 million users' activities for about one month during July, 2012, finally collecting 37 million posts along with their tag profiles. In Sina Weibo, each user can specify up to 10 tags for indicating his/her interests and characteristics. The tags in Weibo are assigned with tag IDs, and each tag has a value of weight. Also, each content has certain values of tags as well. However, the weights of tags are not float values between 0 and 1 as expected, but a much larger number calculated by Sina, depending on each user's activities and characteristics, which we define as w_{it} from u_i to a tag t, and thus we calculate the weight probability of a tag as its weight to the sum of the total weights of all tags:

$$w_{it} \Big/ \sum_{t=1}^{N_T} w_{it}, \forall g_i, i = 1 \ldots N_T.$$

Offline Mobility Impact

In order to model the various moving and meeting patterns of mobile users in offline MSNs, TASA adopts the inter-contact time (ICT)-based methodology from previous studies [3, 4, 8]. The mobility impact, I^M, can be defined to quantify the capability of a user to share a content object with other users by opportunistic D2D sharing, while moving in the MSN. The temporary connectivities with other users in proximity mostly rely on certain discovery mechanisms in a neighborhood, and we assume all users are synchronized by a low duty cycle for probing as proposed by eDiscovery [5]. Also, we assume that the inter-contact intervals of any two mobile users follow an exponential distribution. We use λ_{ij} to denote the opportunistic contact rate of user u_i with user u_j. TASA utilizes similar epidemic modeling from [3]; we skip calculation details due to limited space. Finally, we can get the mobility impact factor I_i^M for u_i to the whole user base as I

$$I_i^M = \sum_{j=1}^{N} \lambda_{ij}.$$

Therefore based on the evaluation of the online spreading impact and offline mobility impact of users, given an initial pushing vector $\vec{p}$, we can finally estimate how long it may take for any user to obtain the content object via meetings, which is defined as the *content obtaining delay* of u_i, t_i^*, presented as

t_i^* = ObtainingDelay$_i$ (OnlineSpreadingImpact$_i$, OfflineMobilityImpact$_i$]

TASA does not have to shorten the content obtaining delays for all users, but can seek the optimal pushing vector $\vec{p}$ to induce proper content obtaining delays to match the delay sensitivity of each user (i.e., satisfactory function), which is detailed in the next subsection.

Content Access Delays

In an SNS, while some users may access the SNS frequently, others access the SNS at relatively longer intervals. As studied in [10], the retweeting delays in Twitter can be mostly within the range of hundreds of seconds to thousands of seconds, or even up to several days. Therefore, users may have their own patterns of accessing contents via SNSs, and thus may have different sensitivity requirements on the contents in which they are interested [6, 8]. We consider such a period the *access delay*, which can be measured from online SNS datasets. From previous measurements in [3] on 2 million Sina Weibo users, nearly half of the user base may have access delay longer than 6.5 h. 20.38 percent users have delays shorter than 1 h, and 26.79 percent users access SNSs with delays larger than even 1 day. Therefore, TASA is able to send content to users who have strong spreading and mobility impact factors but short access delays via cellular links, and then can disseminate content by D2D sharing to a certain portion of users with sufficiently large delays for accessing the content.

TASA will be able to send content to users who have strong spreading and mobility impact factors but short access delays via cellular links, and then can disseminate content by D2D sharing to a certain portion of users with sufficiently large delays for accessing the content.

In order to quantify the access delay parameters in the TASA framework, we model user delays in terms of the probability to intend to access the content at t, which can be considered as access satisfactory function for each user, *Satisfactory_i()* [3, 4]. If the content object can immediately be obtained in a user's mobile device locally as soon as he/she hopes to access (with the highest probability), he/she will be mostly satisfied. We use a fitting method based on Weibull distribution over the probability distribution function (PDF) of the access delays, which is popularly used to profile user behaviors in SNSs:

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

where fitting parameters β_i and k_i can identify u_i's access pattern (the PDF curve shape), and thus are considered as *access delay properties*. Therefore, the satisfactory function can be presented as *Satisfactory$_i$*(*ObtainingDelays*, *AccessDelayProperties*), where the calculated obtaining delays will be the input of the satisfactory functions.

Optimization on Social-Aware D2D Sharing for Traffic Offloading

TASA's objective is to choose a proper initial pushing set of seeds, $\vec{p}$, by evaluating I^S and I^M of all users to estimate their content obtaining delays, t^*, so that the satisfactory functions of all users can be maximized based on the matching between content obtaining delays with the content access delay properties (Box 1) where the amount of initial pushing seeds must be constrained by C. Analytically solving the above problem is hard, but we can get near-optimal results by general power series numerical methods. Also, we can reversely tune and seek the needed C given a targeted total satisfactory level. We further design a heuristic algorithm referring to [3] for finding the solution numerically, based on the well-known hill-climbing algorithm. Initially we choose an easy-starting strategy, such as selecting the top C users with high I^M values, and go through the user base iteratively to exchange the selecting decisions (0 or 1) of any two users if the exchange can result in a higher value of totally satisfactory. We terminate the iteration when the results converge. However, we have to skip the algorithm details in this article due to the space limit.

TASA System Infrastructure

System Architecture

Practical deployment of TASA is challenging, as it requires collection and processing huge amounts of data from online SNSs and offline MSNs, and needs tight consolidation of SNS platforms, MNOs' management of user mobility, and users' D2D transmissions. Figure 3 shows a top-down approach to the system architecture of TASA's social D2D sharing including three key layers: D2D management platform, cloud gateway layer, and wireless infrastructure.

D2D Management and Operation Platform: We design the TASA-based data analysis engine as a key service in the basic service management layer. Considering the increasing demand for D2D-based services, contextual data with high diversity need to be collected and processed at the D2D cloud platform for further online processing and optimization. Further data mining (e.g., social similarity matching, mobility, and spreading impacts) via the TASA analysis engine on the cloud platform have to be addressed before an effective content dissemination strategy can be delivered. Thus, to make the contextual data from different users context-aware, a feasible way is to require SNSs and content providers (CPs) to pre-specify the definition of context for their D2D services and register them to the cloud. Furthermore, the TASA framework needs to be jointly implemented and synchronized on both user devices and the cloud platform to maintain the online and offline social

$$\underset{\text{InitialPushingVector}:\vec{p}}{\text{Maximize}:}\ \sum_{i=1}^{\text{All Users}} Satisfactory\left(\begin{matrix}ObtainingDelay_i\left(\text{OnlineSpreadingImpact}_i,\text{OfflineMobilityImpact}_i\right)\\ \text{AccessDelayProperties}_i\end{matrix}\right)$$

Subject to : InitialPushingVector $\vec{p}$ should be constrained by C, The wireless infrastructure needs to further consider realistic social relationships and individual user behaviors, such as selfishness and hostility, to create more effective incentive and pricing strategies and thus achieve optimal content delivery services in TASA's sharing-based D2D networks.

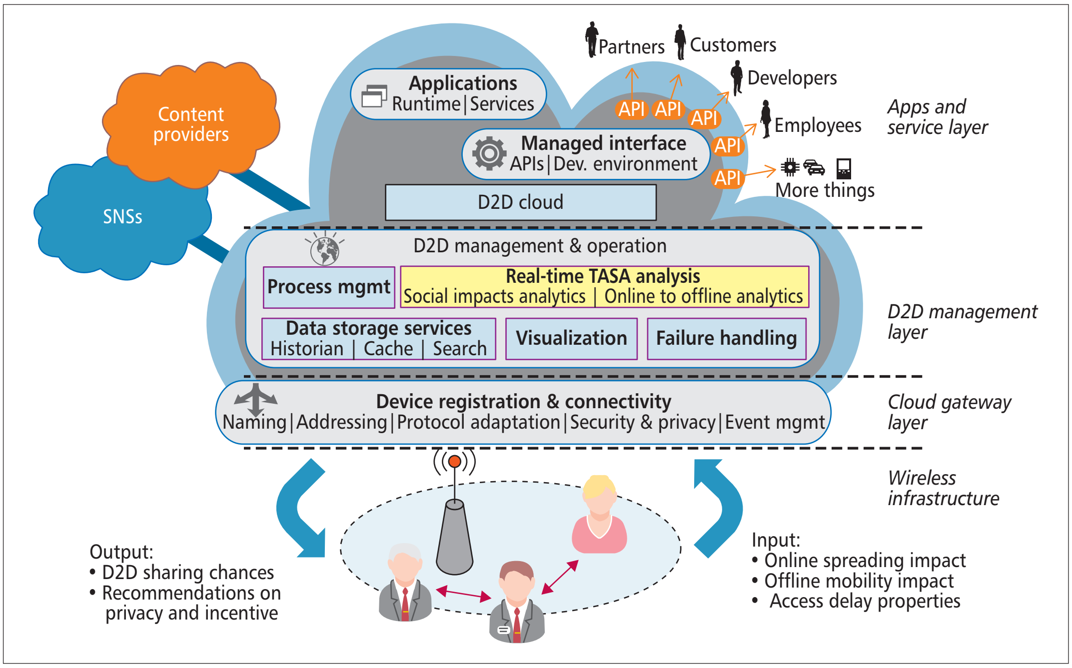

Figure 3. User-infrastructure-cloud interactions in TASA architecture.

data integrity. The engine of data analysis can thus understand the context of data, and model for predicting the content sharing pattern regarding individual social activities and life patterns. Note that there may be multiple roles in such a D2D cloud, as shown in Fig. 3; for example, partners are cooperating entities with MNOs (e.g., other MNOs, SNSs, and CPs), customers are mobile users, developers are from third parties to improve their programs by interacting with the D2D cloud, and employees are those who manage the systems internally in MNOs, while more things form the concept that may allow social D2D to collaborate with the Internet of Things.

Cloud Gateway Layer: This layer can perform as a bridge between wireless infrastructure and D2D cloud platform for forming a seamless online SNS to offline MSN transformation. Additionally, in order to ensure satisfactory quality of D2D service, such as maintaining an acceptable content access delay, well defined application interfaces along with efficient management protocols should be designed and agreed between both CPs and MNOs. Although current service interactions in the cloud are mostly based on Simple Object Access Protocol (SOAP), the lightweight Representation State Transfer (REST) style web service [13] is quite suitable for user mobile devices to realize D2D-based services that are more sharable, reusable, and loosely coupled. Therefore, we achieve SOAP-REST transformation by additional adapters on the cloud gateway layer.

Wireless Infrastructure: The interconnections between cellular networks and user devices form a heterogeneous wireless infrastructure to offload content objects, in which the cellular networks exert a light touch by managing and controlling the allocation of secure D2D resource with alternative short-range radio interfaces. TASA will maintain the radio information of each user based on their reports and thus will manage the radio techniques for any potential connectivities. Specifically, the wireless infrastructure can provide a decentralized approach to proximity discovery and D2D communication, which is efficient, flexible, dynamic, and secure, as well as privacy-enhanced, to enable proximity-based services to flourish. The wireless infrastructure needs to further consider realistic social relationships and individual user behaviors, such as selfishness and hostility, to create more effective incentive and pricing strategies and thus achieve optimal content delivery services in TASA's sharing-based D2D networks.

Components and Mechanisms for TASA

Furthermore, to support TASA in practice, new components and mechanisms should be developed based on existing network infrastructure and D2D communication techniques. We highlight three key designs for TASA.

Social Neighbor and Temporary D2D Neighbor Tables: Both of the tables are key data structures to represent the information of online SNSs at the cloud and physical mobile D2D networks in offline MSNs. Examples of the two tables for user i are illustrated in Fig. 4. Here, the social neighbor table can be straightforwardly obtained according to the users' online activities, while D2D neighbor table establishment requires discovery of opportunistically encountered devices over all possible D2D interfaces. There are many practical methods to track D2D opportunities, such as centralized localization by the mobility management entity (MME) in the MNO, and the eDiscovery scheme in [5].

Dynamic and Distributed Networking: From the networking point of view, data sharing decisions made by TASA will result in corresponding data traffic being injected into the physical mobile D2D networks in offline MSNs. To optimally balance the trade-off between user experience and cellular offloading efficiency, distributed and adaptive decision mechanisms should be developed to jointly optimize the traffic admission control, the selection of D2D peer devices, and the scheduling of multiple available D2D radios as shown in Fig. 4.

Incentivization and Privacy: D2D transmissions cost computing resources (CPU, memory, and I/O) and battery energy, and lead to a risk to privacy disclosure. Therefore, effective incentivization is vital to encourage users to fulfill the D2D tasks allocated by TASA. Possible incentivization methods include reputation-based trust management, a game-theoretical auction approach, market-centric pricing schemes for data content trading, and so on, which are quite interesting and challenging research directions in the future.

Evaluation Results

For practically evaluating TASA, along with the Sina Weibo dataset, we choose three mobility traces: Infocom,[3] MIT,[4] and SUVnet.[5] The three traces have different durations, scales, and patterns. The Infocom and MIT traces are collected by persons in conference and campus spots, who have high contact rates, but the SUVnet trace has low contact rates as it is collected by vehicles in a big city. However, since there is no any trace that may contain the activities of the same users in both SNSs and MSNs, we have to consider three possible schemes for mapping the SNS users with MSN users regarding the three mobility traces:

1. **Random**: SNS users are randomly mapped with MSN users.
2. **h-h**: Both MSN and SNS users are sorted in descending order of I^S and I^M, respectively, and are mapped.
3. **h-l**: Both SNS and MSN users are sorted in the same way as **h-h**, but SNS users with higher I^S are mapped to MSN users with lower I^M.

Since there are many more SNS users than MSN users in each trace, we pick a number of accounts from the sub-graphs of the SNS graph by using the random walking method, according to the number of MSN users in each trace. This is proven to be an effective way to get a representative sample sub-graph from a large user base [14]. Note that the corresponding social tags and access delay patterns are assigned as well. We also take the average values over the evaluation results across the three mapping schemes to reflect the general trends in performance.

We consider a number of varied initial pushing strategies by evaluating users' impact factors, as well as several general graph-based viral marketing strategies, as shown in Table 1. We will investigate what percentage of users should be selected as initial seeds for satisfying the access delay requirements from 100, 90, and 80 percent of users.

Figure 5 shows that **p-H** can find the best initial pushing vector (i.e., the smallest number of seeds) to satisfy 100 percent of the user base. However, **p-R**, **p-**$D^{\rightarrow}$, and **p-**$D^{\leftarrow}$ perform poorly, demonstrating the ineffectiveness of simply pushing by using the graph-based strategies. In most cases, **p-**I^M*I^S performs well, which implies that conjunctively considering the I^S and I^M factor by multiplication can achieve near-optimal performance in practice. **p-**Pr does not perform better

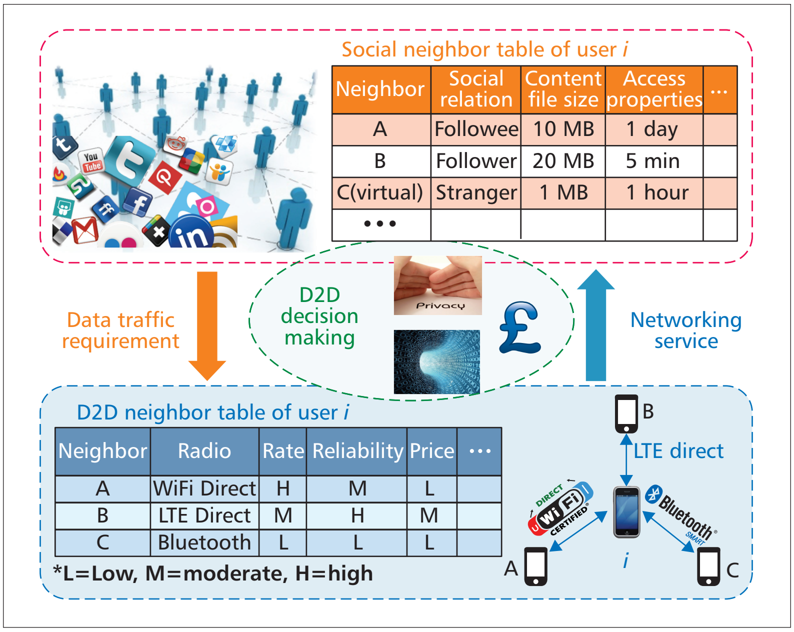

Figure 4. Network protocol support for TASA.

p-I^M	Top C users by sorting all users by I^M descendingly (similar to [4])
p-I^S	Top C users by sorting all users by I^S descendingly)
p-I^M * I^S	Top C users by sorting all users by I^M * I^S descendingly (similar to [3, 7]
p-R	Randomly choose C users
p-$D^{\rightarrow}$	Top C users by sorting all users by outgoing node degree descendingly
p-$D^{\leftarrow}$	Top C users by sorting all users by incoming node degree descendingly
p-Pr	Top C users by sorting all users by PageRank score of Google descendingly
p-H	Selected C users from the hill-climbing heuristic algorithm

Table 1. Strategies for initial pushing.

compared to strategies based on impact factors. In the Infocom and MIT traces, I^M-based pushing strategies can perform better than those based on I^S, which indicates that the mobility factor determines more of the sharing process when nodes have high mobility. In the SUVnet traces, I^S-based ones perform better, which means the social factor controls more when nodes have low mobility.

The required initial pushing ratio is reduced significantly if we target to satisfy 90 percent of all users, as the worst case users are less considered and mostly given lower priority for satisfaction. For the MIT and Infocom traces, only about 21.0 and 17.7 percent of users need to be the initial seeds on average, respectively. More interestingly, the number of initial seeds is further reduced dramatically when satisfying 80 per-

[3] Infocom Trace, http://crawdad.org/cambridge/haggle/20090529/

[4] MIT Trace, http://realitycommons.media.mit.edu/realitymining.html

[5] SUVnet Trace, http://wirelesslab.sjtu.edu.cn/taxi_trace_data.html.

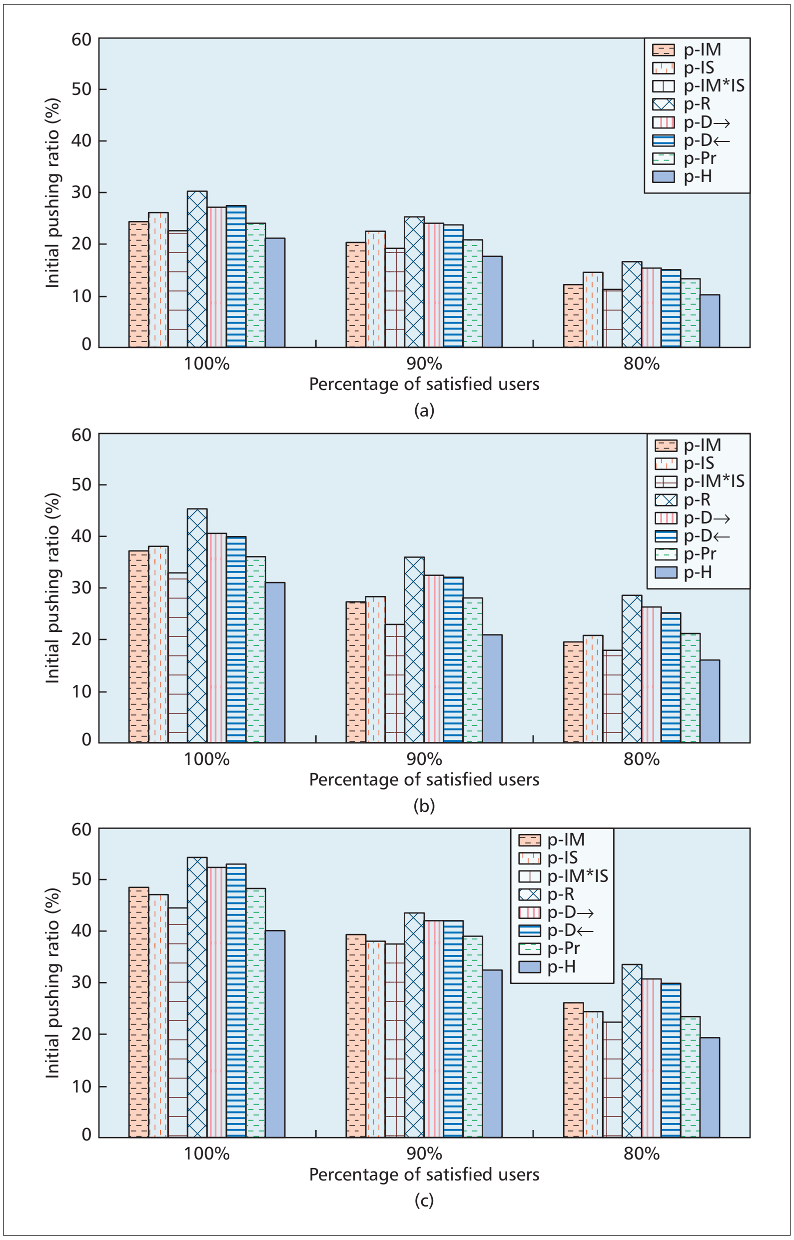

Figure 5. Initial pushing ratios to satisfy 10, 90, and 80 percent of users: a) Infocom; b) MIT; c) SUVnet.

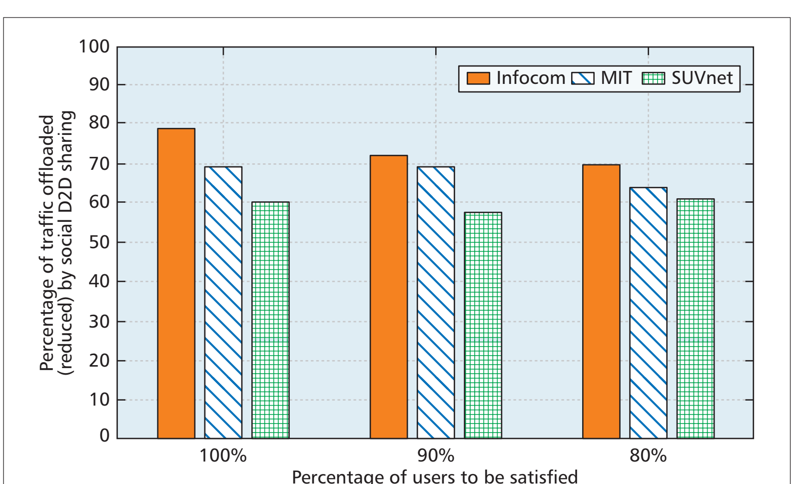

Figure 6. Percentage of traffic offloaded (reduced) by TASA.

cent of users. The SUVnet trace always needs a relatively large number of initial seeds due to the low contact rates of a large number of users.

We notice that some worst case users with both low I^S and I^M will make opportunistic D2D sharing ineffective. It is better to push the content to them in the beginning if they have hard access delay requirements; otherwise, it is better to let them carry out on-demand fetching when they get to the peaks of their access delay PDF. Generally, **p-H** performs best but with high complexity, and **p-**$I^M * I^S$ will be a proper solution in practice with acceptable complexity and deployment efficiency.

Therefore, if a user has not obtained the content (by initial pushing and thus D2D sharing) until he/she hopes to access it, TASA has to deliver it over a cellular link in the on-demand manner, resulting in less cellular data offloading. We now compare the performance of 100, 90, and 80 percent of satisfied users in terms of total offloaded traffic. For example, targeting the satisfaction of 90 percent of users, the remaining 10 percent (i.e., those who have not obtained the content object) will get the content object via the cellular link by on-demand delivery when they access the content with the highest probability.

Figure 6 shows how much traffic can be offloaded maximally on average from cellular links to D2D communications in the three cases, where the heuristic **p-H** strategy is used. When setting the percentage of satisfied users to 100, 90, and 80 percent, although we reduce the initial pushing ratios, the remaining 10 and 20 percent of users still take on-demand deliveries and may instead increase the cellular traffic usage further. Generally, TASA manages to offload 57.6–78.9 percent of traffic onto social D2D sharing.

Conclusions and Discussions

In this article, we propose a framework, tag-assisted social-aware (TASA) opportunistic D2D sharing, for effective mobile data offloading in mobile social networks. By analyzing users' social tags measured from online SNSs, TASA selects an optimal subset of users according to their online social spreading impact and offline social mobility patterns. Furthermore, user-dependent access delays are utilized for matching the delay-tolerant content deliveries. Through extensive trace-driven simulations, we demonstrate that TASA can reduce up to 78.9 percent of mobile traffic load, while all users' access delay requirements can be satisfied.

In the near future, we will focus on large-scale practical implementations and experiments, while investigation will also be put to the security and privacy issues. We expect to extend TASA by exploring other promising techniques as well, such as cooperative integration of D2D sharing with a base station caching framework [15], autonomous, transparent, and trustworthy D2D connection management, and incorporation between D2D-based mobile crowdsensing and IoT-enabled mobile pervasive computing techniques.

Acknowledgment

This work was supported by the Canadian Natural Sciences and Engineering Research Council under grants RGPIN-2014-06119 and RGPAS-462031-2014.

REFERENCES

[1] "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2014–2019," tech. rep., 2014.

[2] M. Cha *et al.*, "I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System," *ACM IMC*, 2007.

[3] X. Wang *et al.*, "TOSS: Traffic Offloading by Social Network Service-Based Opportunistic Sharing in Mobile Social Networks," *IEEE INFOCOM*, 2014.

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

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

[6] C. Yu *et al.*, "Resource Sharing Optimization for Device-to-Device Communication Underlaying Cellular Networks," *IEEE TMC*, vol. 10, no. 8, 2011, pp. 2752–63.

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

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

[9] H. Kwak *et al.*, "What Is Twitter, a Social Network or a News Media?" *ACM WWW*, 2010.

[10] G. Comarela *et al.*, "Understanding Factors that Affect Response Rates in Twitter," *ACM HT*, 2012.

[11] T. Rodrigues *et al.*, "On Word-of-Mouth Based Discovery of the Web," *ACM IMC*, 2011.

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

[13] Z. Sheng *et al.*, "Lightweight Management of Resource-Constrained Sensor Devices in Internet of Things," *IEEE Internet of Things J.*, 2015.

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

[15] X. Wang *et al.*, "Cache in the Air: Exploiting Content Caching and Delivery Techniques for 5G Systems," *IEEE Commun. Mag.*, vol. 52, no. 2, 2014, pp. 131–39.

BIOGRAPHIES

XIAOFEI WANG [M'13] (xiaofeiwang@ieee.org) received M.S. and Ph.D. degrees from the School of Computer Science and Engineering, Seoul National University in 2008 and 2013, respectively. He received his B.S. degree from the Department of Computer Science and Technology, Huazhong University of Science and Technology in 2005. He is currently a professor in Tianjin Key Laboratory of Advanced Networking, School of Computer Science and Technology, Tianjin University, China. His current research interests are social-aware multimedia service in cloud computing, cooperative backhaul caching, and traffic offloading in mobile content-centric networks.

ZHENGGUO SHENG [M'14] (z.sheng@sussex.ac.uk) received his B.Sc. degree from the University of Electronic Science and Technology of China, Chengdu, in 2006 and his M.S. and Ph.D. degrees (with distinction) from Imperial College London, United Kingdom, in 2007 and 2011, respectively. He is currently a lecturer with the Department of Engineering and Design, University of Sussex, Brighton, United Kingdom. He is also a visiting faculty member at the University of British Columbia (UBC), Vancouver, Canada. Previously, he was with UBC as a research associate and with France Telecom Orange Labs as a senior researcher and project manager for machine-to-machine (M2M)/Internet of Things. His current research interests cover IoT/M2M, cloud/edge computing, vehicular communications, and power line communication.

SHUSEN YANG [M'13] (cuiangmu@gmail.com) received his Ph.D. in computing from Imperial College London in 2014. He is currently a professor in the Institute of Information and System Science at Xi'an Jiaotong University (XJTU). Before joining XJTU, he worked as a lecturer (assistant professor) at the University of Liverpool from 2015 to 2016, and a research associate at Intel Collaborative Research Institute for sustainable connected cities from 2013 to 2014. His research interests include mobile and wireless networks, networks with human in the loop, and data-driven networked systems. He is a recipient of China's "1000 Young Talents Program and an honorary research fellow at Imperial College London. He is a member of ACM and IEEE Communications Society.

VICTOR C. M. LEUNG [S'75, M'89, SM'97, F'03] (vleung@ece.ubc.ca) is a professor of electrical and computer engineering and holder of the TELUS Mobility Research Chair at UBC. His has co-authored more than 900 journal and conference papers in the areas of wireless networks and mobile systems. He is a Fellow of the Royal Society of Canada, the Canadian Academy of Engineering, and the Engineering Institute of Canada. He is an editor of *IEEE JSAC-SGCN*, *IEEE Wireless Communications Letters*, and several other journals.