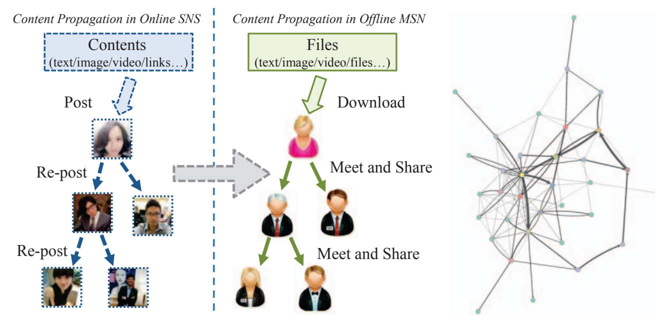
Measurement and Analytics on Social Groups of Device-to-Device Sharing in Mobile Social Networks

Hui Wang[1], Shanjia Wang[1], Yuhua Zhang[1], Xiaofei Wang[1], Keqiu Li[1], Tianpeng Jiang[2]
[1]Tianjin Key Laboratory of Advanced Networking, Tianjin University, Tianjin, China.
[2]Beijing Anqi Zhilian Technology Co. Ltd., Beijing, China.

***Abstract*—Recently many studies demonstrate that exploiting the Device-to-Device (D2D) content sharing in offline Mobile Social Networks (MSNs) is a promising solution to offload cellular data to local connectivities in proximity to reduce the duplicated cellular transmissions via the backbone network demanded by nearby users and hence to improve users' quality of service (QoS). However, related D2D-based social sharing and offloading proposals are based on either assumptions and theoretical models, or limited data-driven analysis caused by small scale of data sets (e.g. hundreds of MSN users) or single-dimensional feature (e.g. human mobility only), which severely restricts applications in practice. In this paper, we perform the *world-first* large-scale measurement and analytics on D2D-based content sharing groups from the perspective of social networks via the platform of *Xender*, a leading global D2D sharing platform in Asia. We analyze the behaviours of about 30 million users with 443 million D2D transmissions of 17 million files in 884 thousand social groups, and unveil the details of social structure properties, network motifs, cascade trees of friendship and propagation, which are helpful for improving the service of social D2D sharing. Finally we discuss challenges and opportunities for improving social D2D sharing services.**



(a) Content spreading in online SNSs *v.s.* D2D sharing in *Xender* MSNs

(b) A typical social graph in *Xender*'s MSNs

Fig. 1: *Xender*'s D2D sharing-based mobile social networks

I. Introduction

Mobile networks are being faced with supporting a growing number of multimedia services, which enables mobile users to frequently download content files onto mobile devices. This challenge has resulted in explosive growth of mobile traffic, which poses great pressure on the front-haul and the back-haul network of mobile network operators (MNOs) [1]. However, the wireless link capacity, MNOs' infrastructural network facilities suffer from difficulties to cope with the skyrocketing traffic load effectively. The *traffic explosion problem* becomes a severe concern, especially in the areas with high user densities but limited cellular network capacities.

One essential found is that most of the traffic load over the Internet is generated by downloading of the same popular content files [1] [2]. For example, the top 10% videos in YouTube account for around 80% of all the views [2]. The mobile networks have to always transmit the same files to users for many times, inducing network resource wastes.

Therefore, an effective way to reduce such duplicate downloads is to cache and share the multimedia content files among geographically proximal mobile devices through device-to-device (D2D) communications (e.g. WiFi). In doing so, each user is likely to obtain popular content files from nearby devices, and only needs to use the expensive cellular to download the files that are unavailable in proximity, resulting in cellular traffic offloading. Note that in this paper we consider all types of device-to-device (D2D) direct communication techniques between devices, e.g., Bluetooth, Wi-Fi Direct etc. Also the recent D2D communication underlaying the 3GPP LTE cellular network in the operator authorized spectrum becomes emerging, and is thus included as well [3].

It has been shown that exploiting the social behaviors of mobile users in Mobile Social Networks (MSNs) can significantly improve the efficiency of D2D sharing and cellular offloading [4], where socially-close users can exchange content files during frequent encounters, without relying on MNOs' communication infrastructures [5] [6] [7]. For instance, [5] *et al* reports that up to 86.5% cellular offloading can be achieved in small scale tests, by exploiting the social behaviors of mobile users. However, most current approaches [8] [9] are based on either unrealistic assumptions, or limited data analytics caused by small data size (e.g. hundreds of MSN users) or single-dimensional feature (e.g. large-scale human mobility measurement [10]), which severely restricts their applications in real-world D2D sharing scenarios.

Data collection with analysis of *large-scale* real-world D2D sharing activities is essential to deeply understand the user behaviors in social D2D sharing, and to provide meaningful

978-1-4673-8999-0/17/$31.00 ©2017 IEEE

implications and guidelines. Therefore, to bridge this gap, our paper presents comprehensive measurement and analytics based on *Xender* [1], one of the leading mobile applications (APP) for D2D content sharing in the world (the largest one in India). As shown in Fig. 1(a), D2D sharing activities among *Xender* users form offline MSNs, which have similar content delivery pattern and social properties as online Social Network Services (SNSs) [11]. Particularly we regard the sharing groups as social networks and thus methodologies from complex networks can be applied for insightful analytics as shown in Fig. 1(b). Hereby, we define the ***social group*** (the same as ***sharing group***) as the set of users who have shared content at least once with another user within a certain time window (e.g., 1 day or the whole data set), and obviously every user only belongs to one group. For instance, if user A shares files with user B, and user B shares files with user C, we regard user A, B, C belonging to the same social group.

The contributions of this paper are two folds: 1) we measure and analyze 4-week data from *Xender*, which covers over 30 million active users, 443 million D2D transmissions, 17 million content files, and 884 thousand social groups. To the best of our knowledge, *this is the first work on large-scale data analytics for real-world D2D sharing from the social network aspects*; 2) we further discuss motifs of social groups as well as cascade trees of friendship and propagation, and hence demonstrate several meaningful implications for improving the Quality of Service (QoS) of D2D sharing services.

The remainder of this paper is organized as following: related studies are introduced in Sec. II, and we describe the dataset and methodology in Sec. III. Structural properties of D2D social sharing groups will be studied in Sec. IV and Sec. V, respectively. We study the cascade trees of propagation and friendship in Sec. VI, and conclude the paper in Sec. VII.

II. Related Work

Exploiting D2D communications for MSN content sharing not only attracts an increasing research interest from academia [8] [11] [12] [13], but also promotes a number of mobile applications (APPs) such as Apple's Airdrop and APPs like *Xender*. Many studies focus on the D2D epidemic content dissemination in MSNs over recent years for traffic offloading purpose. For example, Zhang et al. [14] and Li et al. [15] have utilized a differentiation-based model to analyze the performance of popular content sharing in MSNs. However, none of them tests over a large scale user base but a small group of people.

There exist several measurement-based studies for online SNSs and offline MSNs. The studies in [13] [16] [17] shows the homophily and locality characteristics of users are observed in both MSNs and SNSs, which can be utilized to facilitate the D2D content dissemination [17]. In addition, Kwak *et al* [18] points out that there are obvious delays of re-sharing behaviors, while the spreading impact of each user is accumulated hop by hop. Such observation results enable the analysis, modeling, and prediction of the sharing activities and the content spreading of SNS users based on measurement traces [19] [20] [22]. There also exist large-scale measurements for human mobility [10] that are related to our work. In reality, D2D sharing in offline MSNs is much more complex than information sharing in online SNS, due to temporal and spatial constraints [16] [20]. However none of aforementioned studies focuses on multi-dimensional feature analytics on large-scale D2D sharing data.

III. Details of Datasets and Platform

A. Details of Datasets and Platform

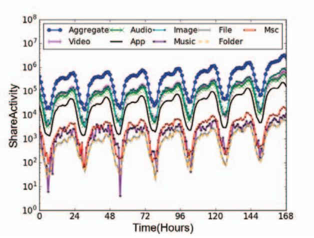
Xender is a world-wide popular mobile application (APP) for D2D content deliveries, which provides users with the convenience of sharing various types of content files without using 3G/4G cellular network across a large diversity of system platforms. The D2D connection in *Xender* is based on Wi-Fi tethering at a speed around several MBps, and transmissions are free since they consume no mobile data. Currently *Xender* has around 10 million daily and 100 million monthly active users, as well as about 110 million daily content deliveries. We capture *Xender*'s trace for four weeks (from 01/02/2016 to 28/02/2016). Note that around 70% users are from India, which motivates us to concentrate on analyzing Indian users in this paper. After cleaning entries without file's MD5, with zero file's size, or with invalid timestamps, we will work on the target data set, $\mathbb{X}$, including 30,485,335 users, 443,440,043 transmission timestamps conveying 16,785,175 content files, with data entry format: $\langle sender, receiver, content\langle MD5, size\rangle, timestamp\rangle$. To process our large-scale data (with total size of 843 GB) efficiently and reliably, we use Python 2.7 along with Anaconda scientific package and *graph-tool* [21] library (v2.18), on four Dell PowerEdge R730 servers, each of which has two 2630v3 CPUs, 32 cores, 64G RAM and 16T SAS hard disk.

B. Preliminary Analysis on Time Series

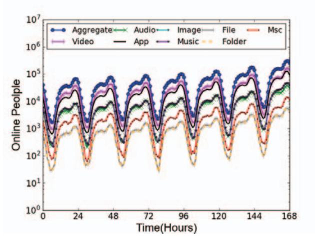
We analyze the time-varying performance basics of *Xender* data, and for clear visualization, we only plot the data of the 1st week as the representative of the four weeks. Fig. 2(a) shows the number of sharing activities over time with regard to various types and aggregated value. Both the weekly and daily observations show typical life cycles with temporal regularity, and further verifies the fact that families and friends always get together in the weekend in India. This implies that there exists large room of optimizing the time-varying system resource and proposing effective family-focused marketing strategies.

Fig. 2(b) represents the time series of the number of online individuals, where a user will be only counted for once if this user shares at least one content file in this time window. Notably, users who share videos and pictures take the highest portion. This is mainly because of the inconvenient centralized mobile APP market services via the MNOs' networks in India, but instead, people often rely on exchanging interesting APPs via D2D sharing activities. The time series of ongoing traffic load is plotted in Fig. 2(c). The traffic load on Sunday raises to

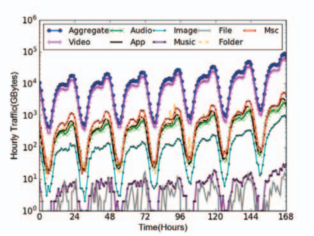
[1] http://www.xender.com



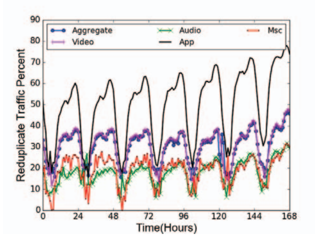
(a) Number of sharing activities



(b) Number of online individuals



(c) Disseminated traffic load



(d) Reduplicate traffic percentage

Fig. 2: Time series of *Xender*'s dataset for the 1st week

5 to 10 times than that in working days. Videos still count for the majority, and obviously *Xender* helps the mobile networks to offload a large number of large-size videos into D2D sharing activities because it is free and fast.

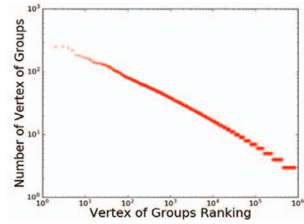
From the plots of portion of duplicate traffic (selected content types) in Fig. 2(d), which is computed as one minus the ratio of total size of involved individual content files to the total traffic load, approximately up to 40% aggregated traffic in *Xender* is proved to be redundant. This implies that huge amount of mobile users request the same popular content files via *Xender*, especially APPs and videos. Particularly APPs have the largest the duplicate traffic ratio at about 60%, revealing a substantial number of users disseminate limited number of popular APPs. Therefore, *Xender*-like services become a more effective method for video dissemination and APP marketing, and strategies encouraging sharing the trending videos and APPs will be beneficial [5] [23].

IV. Structural Properties of Sharing Groups

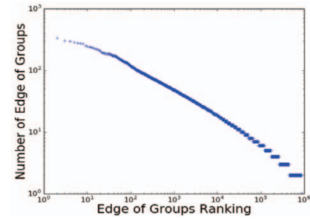
D2D content sharing activities form opportunistic networks [18] [24], where the social links among users are represented by their accumulated sharing activities. Therefore, we carry out the analysis on the social characteristics of users in *Xender*'s sharing groups [25]. By considering the sharing directions, we thus define a directed social graph of a D2D social group (sharing group) by $\overrightarrow{G} = \{V, \overrightarrow{E}, \overrightarrow{W}\}$, where a user is reflected as a vertex, $v_i \in V$, transmissions are reflected as an edge, $\overrightarrow{e_{ij}} \in \overrightarrow{E}$, and $\overrightarrow{w_{ij}} \in \overrightarrow{W}$ is the edge weight (i.e., sharing frequency). We further define the undirected version of graphs, $G = \{V, E, W\}$, where e_{ij} and e_{ji} are the same undirected edge, but $w_{ij} = w_{ji} = \overrightarrow{w_{ij}} + \overrightarrow{w_{ji}}$, $w_{ij} \in W$.

A. Properties of Vertices and Edges

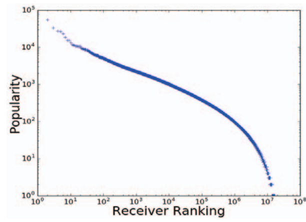
1) Group Size: First we denote the size of group $G = \{V, E, W\}$ by $|V|$, and we finally discover 883,772 groups in



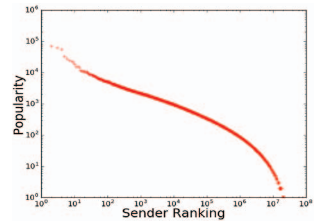
(a) Group size by vertex



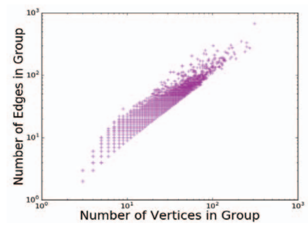
(b) Group size by edge



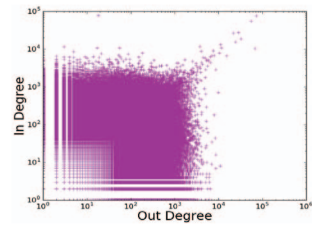
(c) In-degree



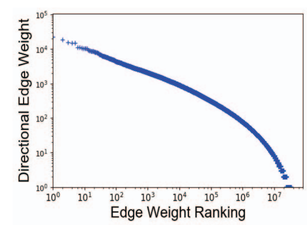
(d) Out-degree



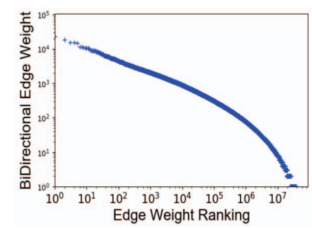
(e) Correlation of edge and vertex



(f) Correlation of in-degree and out-degree



(g) Weights of directed edges



(h) Weights of undirected edges

Fig. 3: Properties of groups, vertices and edges

$\mathbb{X}$, and note that we choose to skip extremely small groups with size of 2 as it will be meaningless to analyze them. We define the number of users within a group, $|V|$, is the size of the group,1 and illustrate the distribution of the sizes of groups in Fig. 3(a). We can define another size of the group by the number of edge, $|E|$, and the distribution is shown in Fig. 3(b). Also we plot their correlation in Fig. 3(e). Notably the sizes follow perfect power law property. As some social structural analytics will be meaningless over quite small graphs, for effective analysis on the graphs, we cluster all D2D sharing graphs according to their sizes of vertices and edges by k-mean method. We choose $K = 3$ and then get following groups: and in following analytics, we will investigate groups selectively:

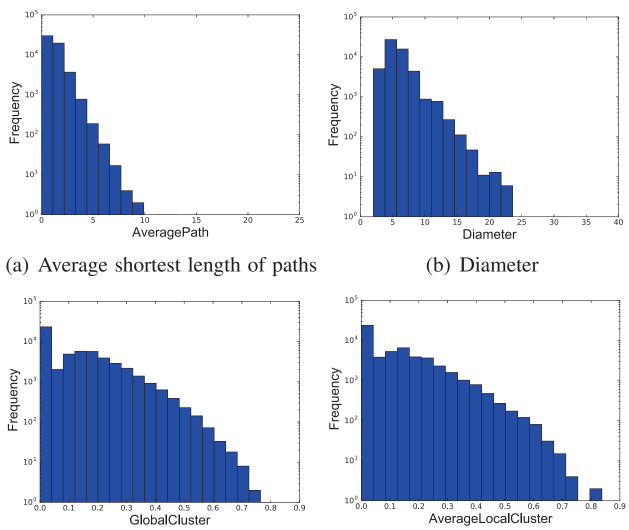
- **Small:** $\forall G, |V| \in [3, 8]$, in total 829,229 graphs,
- **Medium:** $\forall G, |V| \in [9, 27]$, in total 51,991 graphs,
- **Large:** $\forall G, |V| \in [28, 309]$, in total 2,549 graphs.

2) Vertex Degree: The in-degree of v is the number of incoming edges representing the records of being shared, while the out-degree of v is the number of outgoing edges representing the records of sharing to others. We draw the in-degree and out-degree of all vertices in all graphs in Fig. 3(c) and Fig. 3(d), as well as their correlation in Fig. 3(f) Power law effect is observed clearly meaning small amount of vertices taking large amount of incoming and outgoing edges, but most of the vertices have limited degrees. However, there is very little correlation between vertices' in-degree and out-degree; a small number of users have reciprocal in-degree and out-degree values, but most don't.

3) Edge Weight: We investigate the weight of both directed and undirected edges as illustrated by ranking plots in Fig. 3(g) and Fig. 3(h). They show the power-law-like curves with drop tails, which indicates there exists heavy user pairs sharing lots of files, but a great portion of users don't share much frequently. This implies the differentiation of heavy and light user pairs and hence customized recommendation strategies are needed [26].

B. Properties of Social Graphs

From the perspective of social graph, we apply complex network theory onto the groups focusing on path and structure over the large and medium groups.



(a) Average shortest length of paths

(b) Diameter

(c) Global cluster

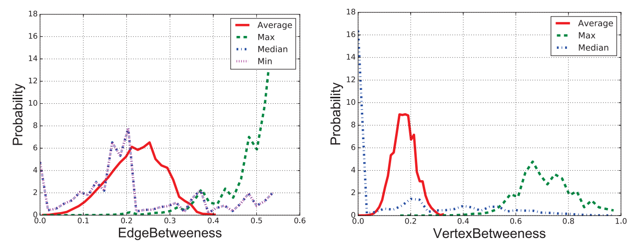
(d) Average local cluster

Fig. 4: Properties from the aspect of social graphs

1) Average Shortest Length of Paths (ASLP): ASLPs is defined as the average number of steps along the shortest paths for all possible pairs of network nodes. It is a measure of the efficiency of information or mass transport on a network. A network with low ASLPs provides efficient data transmission. From Fig. 4(a), ASLPs in *Xender*'s social graphs are mostly as short as up to 6 hops, which indicates they are small-world-networks, and follow the "six degrees of separation" [25].

2) Diameter: The diameter of the graph is the largest of the shortest paths of all paths in the graph, which briefly shows the scale of the graphs. From Fig. 4(b), diameters mostly range from 2 to 20, with an average of 5.49, which is correlated to the ASLP values and the "six degrees of separation" significantly.

3) Clustering Coefficient: We investigate two clustering coefficients, global one of a graph and the average local one of vertices of a graph. The global clustering coefficient is the number of closed triplets over the total number of triplets (both open and closed), illustrated in Fig. 4(c) indicating the graphs in *Xender* is not well clustered with low coefficient values. The average local clustering coefficient indicates whether the vertices in each graph are having close neighbours to being a clique (complete graph). As shown in Fig. 4(d), the coefficient values are not high as well. Therefore, *Xender*'s users and groups are not closely clustered but user-to-user pair-concentrated activities may be more important, probably due to the security and privacy issues during D2D sharing.



(a) Betweenness by edge (max., min., medium, avg.)

(b) Betweenness by vertex (max., min., medium, avg.)

Fig. 5: Betweenness properties from the aspect of social graphs

4) Betweenness: Betweenness centrality is an indicator of a vertex's centrality in a network, which equals to the number of shortest paths from all vertices to all others that pass through that vertex. A vertex with high betweenness centrality has a large influence on the transfer of items through the network, We investigate both the edge-betweenness and vertex-betweenness of all edges and vertices in all graphs. And we illustrate the probability density functions of max., min., avg., and medium values of betweenness of vertices for all graphs in Fig. 5(a), those of edges in Fig. 5(b). The avg. plots in red show that the average value of avg. edge betweenness of all graphs is about 0.24 and the vertex one is about 0.18.

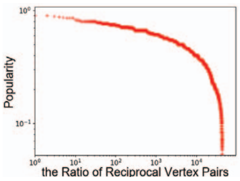
V. Motif Analytics of Social Groups

A. Sharing Reciprocity

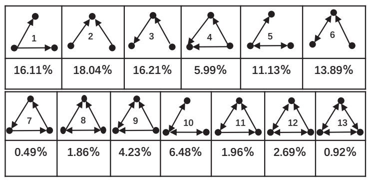
In the directed graph $\overrightarrow{G}$, a vertex pair (u_i,u_j) is called **reciprocal** if there exist edges pointing in both directions. The reciprocity of a network is the fraction of reciprocal vertex pairs among all connected vertex pairs. From the illustration of the reciprocity values of all graphs in Fig. 6(a), a few graphs have high reciprocity values, but most of the graphs have low reciprocity values, i.e., nearly 91% graphs have reciprocity values smaller than 0.5, which means a large number of user pairs have only one sharing direction. Thus there is great potential to apply effective method to encourage "reciprocal-sharing" to bring users tighter.

B. Sharing Motifs

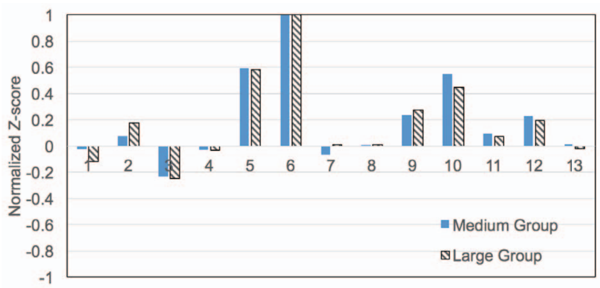
Triad census of the thirteen distinct motifs are carried out to analyze the interaction dynamics among users in the groups referring to popular online SNSs studies [27]. The benchmark for the relative abundance or scarcity of a motif is, naturally, the random graph. Specifically, the relative frequency of each motif against their expected number in the random model will be quantified via the normalized version of Z-score [25].



(a) Sharing reciprocity



(b) Motif types and statistics



(c) Normalized z-scores

Fig. 6: Motif analytics

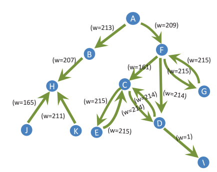
By checking the statistics of existences of 13 motif types as shown in Fig. 6, motif types in forms of open triads, i.e., type 1, 2, 3, 5, and 6, take very high portions, and those in forms of closed triads are with less portions. Friends of friends are not establishing enough interactions and thus are not close enough, which unveils a great room of deploying friend recommendation schemes for bringing users closer. From Fig. 6(c), medium and large groups don't have distinct differences, but motif type 5, 6, and 10 have quite high Z-scores, indicating in *Xender*, there are very huge amount of relationships in forms of nearly closing triplets. A high probability of establishing new triangle friendship can be expected but not yet. This is probably due to some blocking issues in the designs of user interface and experience, which motivates us to exploit some more effective friend discovery algorithms to "enclose" those nearly-closed-triplets.

VI. Cascade Trees of Friendship and Propagation

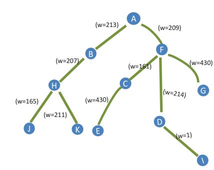
To investigate *Xender*'s content propagation and friendship expansion, we define two types of trees according to Fig. 7:

- **Friendship Extension Tree (FET)**, which is undirected, and will be obtained by calculating the *Maximum Spanning Tree* of the original undirected graph G with classical minimum spanning tree algorithms (e.g., Prim's and Kruskal's algorithm). FET mostly indicates the expansion of social friendship because the establishment of friendship doesn't rely on the direction of sharing. And hence FET can be used for estimating influencing potentials by exploring user friendships.
- **Content Propagation Tree (CPT)**, which is actually directed, and will be obtained by calculating the *Directed Maximum Spanning Tree* of the original directed graph $\overrightarrow{G}$, with a modified version of Edmonds's algorithm [28] [29], as shown in Algorithm 1. Note that the CPT is also mentioned as *Minimum Propagation Arborescence* in [28] [29]. CPT mostly indicates the directions of content propagation and can be used for estimating the propagation potentials for spreading files.

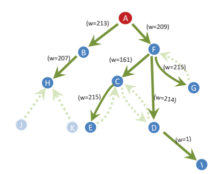
Cascade tree analysis over small graphs provides limited implications, so we analyze large and medium graphs. And note that FETs have exactly the same amount of vertices as the original graphs have, but CPTs may not since some vertices with less frequency and inappropriate directions may be deleted from the algorithms to maintain the main cascade flow. Also although in notations and algorithms we mention "minimum", but actually we assign a new weight $w'_{ij} = max\{W\}+1-w_{ij}$ to each edge, since our realistic purpose is to maximize the friendship expansion and content spreading.



(a) Original social graph



(b) Friendship Expansion Tree



(c) Content Propagation Tree

Fig. 7: A typical graph from trace transferring to FET & CPT

Algorithm 1 Algorithm for Getting CPTs

1. find the vertex with the largest value of out-degree, as the root;
2. put all vertices that root reaches by DFS into a subgraph;
3. find another subgraph in the remaining vertices set;
4. repeat 1-3, until there is no remaining vertex;
4. choose the subgraph with maximal size for later algorithm;
5. get the precursor with minimum weighted edge for each vertex, put precursor into V', put the edge into E';
6. if $< V', E' >$ is a tree, done.
7. else: // there exists loop!
   8. transfer the loop into a super vertex, v_s;
   9. find edges connected with real vertices inside v_s;
   10. mark edges in v_s making vertices with 2 or more precursors;
   11. repeat 5-10, until there is no loop;
12. expand the super vertices, delete all marked edges.

The illustrations in Fig. 8 show the structural information of both FETs and CPTs considering the maximal depths and width. And furthermore we classify FETs and CPTs into two types: a) tall trees with large depths, and b) fat trees with large width. Tall trees indicate strong capability of long influencing and propagating paths with many hops, and instead fat trees indicate strong capability of fast influence and propagation of vertices with quite large centrality.

We evaluate the maximal widths and depths of FETs and CPTs in Fig. 9, and FETs are observed to have width ranging

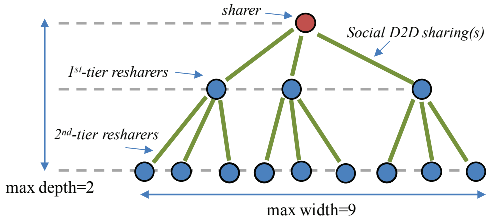
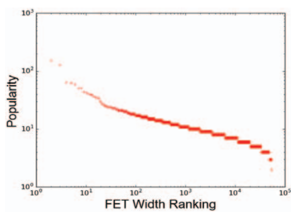
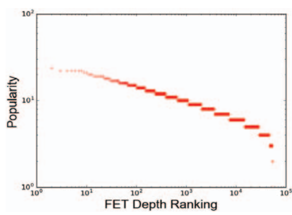


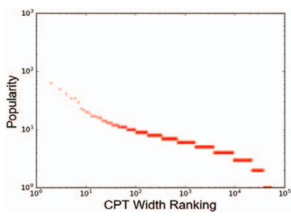
Fig. 8: Example tree structure



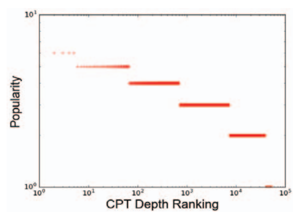
(a) Width values of FETs



(b) Depth values of FETs



(c) Width values of CPTs



(d) Depth values of CPTs

Fig. 9: Width and depth values of trees

from 2 to about 100, and depths raging from 2 to about 20. Because CPTs are formed with directions, they are obtained with much smaller values of width (ranging from 1 to about 60) and depth (ranging from 1 to 6), which corresponds to the previous analysis over ASLPs and diameters. As we calculated about 54.3% graphs have large values of depth than that of width, and thus we identify that regarding the friendship establishment and content propagation, tall trees take slightly larger portion. That means *Xender*'s users prefer to expand friendship and to propagate content farther rather than more users, which probably is due to the geographical constraint of D2D sharing, and thus people cannot often share to many others. Another aforementioned issue is *Xender*'s users prefer to concentrate on activities with not too many neighbors, probably due to the security and privacy issues during D2D sharing. And hence reliable and convenient D2D sharing schemes are highly demanded.

VII. Conclusion

In this paper, we present a large-scale measurement study for the offline device-to-device (D2D) content sharing in Mobile Social Networks (MSNs) with data collected from *Xender*. We carry out comprehensive analytics from the perspective of social networks including structural properties of sharing groups, characteristics of social graphs, motif dynamics, and cascade trees. This results in a set of meaningful implications and guidelines, which could be very useful to improve the quality of *Xender*-like D2D sharing services. Our future work will focus on learning-based friend recommendation.

References

[1] CISCO. Cisco visual networking index: Global mobile data traffic forecast update. Technical report, CISCO, Tech. Rep, 2014.
[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. In *ACM WWW*, 2007.
[3] Device-to-Device Communications in 3GPP LTE standard, release 12, http://www.3gpp.org/specifications/releases/68-release-12
[4] S. Ioannidis, A. Chaintreau, and L. Massoulie, "Optimal and Scalable Distribution of Content Updates over A Mobile Social Network," in *IEEE INFOCOM*, pp.1422–1430, 2009.
[5] X. Wang, M. Chen, Z. Han, D. Wu, T. Kwon, "TOSS: Traffic Offloading by Social Network Service-based Opportunistic Sharing in Mobile Social Networks," in *IEEE Infocom*, pp.2346–2354, 2014.
[6] X. Chen, B. Proulx, X. Gong, and J. Zhang, "Exploiting Social Ties for Cooperative D2D Communications: A Mobile Social Networking Case," *IEEE/ACM Trans. Netw.*, vol.23, no.5, pp.1471–1484, 2015.
[7] Andreev, S., Pyattaev, A., Johnsson, K., Galinina, O., Koucheryavy, Y. "Cellular Traffic Offloading onto Network-Assisted Device-to-Device Connections." *IEEE Commun. Mag.*, vol.52, no.4, pp.20-31, 2014.
[8] 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 Trans. Mobi. Comput.*, vol.11, 2011.
[9] D. Hristova, M. J. Williams, M. Musolesi, P. Panzarasa, and C. Mascolo, "Measuring Uban Social Diversity Using Interconnected Geo-Social Networks," in *ACM WWW*, pp.21–30, 2016
[10] D. Zhang, J. Huang, Y. Li, F. Zhang, C. Xu, and T. He. "Exploring human mobility with multi-source data at extremely large metropolitan scales," in *ACM Mobicom*, pp. 201-212, 2014.
[11] G. Steeg and A. Galstyan,"Information Transfer in Social Media," in *ACM WWW*, pp. 509–518, 2012.
[12] R. Xiang, J. Neville, and M. Rogati,"Modeling Relationship Strengthin Online Social Networks," in *ACM WWW*, pp. 981–990, 2010.
[13] T. Rodrigues, F. Benvenuto, M. Cha, K. Gummadi, and V. Almeida, "On Word-of-Mouth Based Discovery of the Web," in *ACM IMC*, 2011.
[14] X. Zhang, G. Neglia, J. Kurose, and D. Towsley, "Performance Modeling of Epidemic Routing," *Computer Networks*, vol.51, pp.2867–2891, 2007.
[15] Y. Li, Y. Jiang, D. Jin, L. Su, L. Zeng, and D. Wu,"Energy-efficient Optimal Opportunistic Forwarding for Delay-Tolerant Networks," *IEEE Trans. Veh. Technol.*, vol.59, no.9, pp.4500-4512, 2010
[16] K. Zhang and K. Pelechrinis. "Understanding Spatial Homophily: The Case of Peer Influence and Social Selection." in *ACM WWW*, 2014.
[17] M. Wittie, V. Pejovic, L. Deek, K. Almeroth, and B. Zhao, "Exploiting Locality of Interest in Online SNS," in *ACM CoNEXT*, 2010.
[18] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a Social Network or a News Media?", pp.591–600, in *ACM WWW*, 2010.
[19] E. Cho, S. Myers, and J. Leskovec." Friendship and Mobility: User Movement in Location-based Social Networks." in *ACM KDD'11*, 2011.
[20] L. Backstrom, E. Sun, and C. Marlow. "Find Me if You Can: Improving Geographical Prediction with Social and Spatial Proximity." in *ACM WWW*, pp.61–70, 2010
[21] Graph-tool, version 2.18, https://graph-tool.skewed.de/
[22] G. Steeg and A. Galstyan,"Information Transfer in Social Media,"in *ACM WWW*, pp.509–518, 2012.
[23] C. Yu, K. Doppler, C. Ribeiro, and O. Tirkkonen,"Resource Sharing Optimization for Device-to-Device Communication Underlaying Cellular Networks," *IEEE Trans. Mobi. Comput.*,vol.10, pp.2752–2763, 2011.
[24] J. Qiu, Y. Li, J. Tang, B. Chen, Q. Yang "The Lifecycle and Cascade of WeChat Social Messaging Groups", in *ACM WWW*, pp. 311–320, 2016.
[25] M. Newman. The structure and function of complex networks. *SIAM Rev.*, 45(2), 2003.
[26] V. Chaoji, S. Ranu, R. Rastogi and R. Bhatt, "Recommendations to Boost Content Spread in Social Networks," in *ACM WWW*, 2012.
[27] J. Han, D. Choi, B. Chun, T. Kwon, H. Kim, Y. Choi, "Collecting, Organizing, and Sharing Pins in Pinterest:Interest-driven or Social-driven?" in *ACM SIGMETRICS*, Vol. 42, No. 1, pp.15-27, Jun. 2014.
[28] J. Edmonds. Optimum branchings. *J. Res. Nat. Bur. Standards*, 71B:233–240, 1967.
[29] Y. J. Chu and T. H. Liu. On the shortest arborescence of a directed graph. In *Science Sinica*, 14:1396–1400, 1965.