Optimal Data Partitioning and Forwarding in Opportunistic Mobile Networks

Ning Wang and Jie Wu

Center for Networked Computing, Temple University, USA

Email: {ning.wang, jiewu}@temple.edu

Abstract—In opportunistic mobile networks, existing schemes rely on the assumption that data can be entirely transmitted at each contact. However, in an opportunistic mobile network, the transmission probability exponentially decreases as the data size increases. That is, the contact duration in each contact might be insufficient to deliver large data. Therefore, it is reasonable to partition original data into small data chunks and each chunk is forwarded through an opportunistic path. The objective of this paper is to find an optimal data partition strategy where the data delivery ratio is maximized under a given deadline. There is a trade-off in data partitioning. Each small chunk in a path has a higher delivery probability than the original data, and consequently, a shorter delivery latency under the persistent transmission model with re-transmission. However, the destination needs to receive all chunks in multiple paths (a path is a sequence of contacts) to retrieve the data. A delay in any path will lead to a longer delivery latency. We formulate the data partitioning problem and propose solutions in blind flooding. In the blind flooding scenario, we find the optimal data partitioning size. Network coding technique is used to the proposed method to further improve the performance. Extensive experiments on realistic traces show that our scheme achieves a much better performance than those without partitioning.

Index Terms—Data offloading, data forwarding, delay tolerant networks, opportunistic mobile networks.

I. INTRODUCTION

The widespread availability of personal mobile devices has generated an opportunistic mobile network in which mobile users walk around and communicate with each other via Bluetooth or Wi-Fi in their carried short-distance wireless communication devices. The Cisco 2014-2019 White Paper [1] points out that as of 2014 the number of mobile-connected devices has exceeded the world's population and has led to many contact opportunities. New technologies, Wi-Fi Aware [2, 3], extend Wi-Fi's capabilities in its discovery mechanism and provides better user experiences. As the result, the opportunistic mobile network has become a promising traffic offloading approach to relieve the overloaded 3G/4G cellular communication system currently in place. Most previous data forwarding schemes [4–7] assume that content can always be entirely transmitted over a single opportunistic contact. With rapidly increasing size of multimedia contents in recent years, such an assumption may no longer hold. For example, the average size of the top ten most popular videos with a 1080-pixel resolution on YouTube in 2016 is nearly 100MB. As a result, it may need a certain contact duration to receive the data. However, in reality the contact duration is usually short. We verified the INFOCOM06 and SIGCOMM09 traces
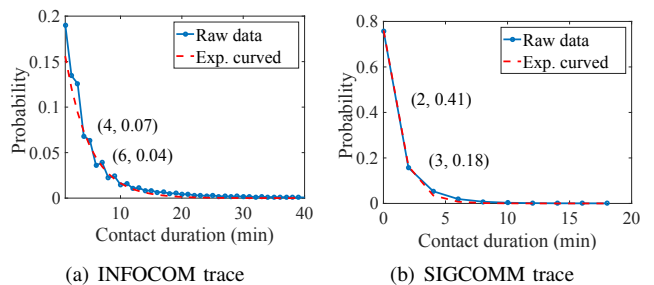

Fig. 1. Probability density function of contact duration in INFOCOM and SIGCOMM datasets.
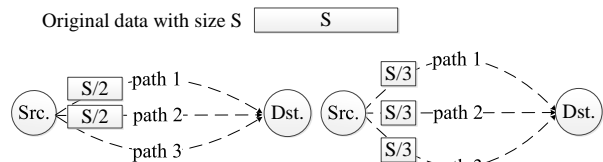

Fig. 2. An illustration of data partitioning.

[8, 9], two human mobility traces in international conferences, and got the following observation as shown in Fig. 1: *longer contacts are just a few while short contacts are many*.

Based on the aforementioned short contact duration observation and the common assumption that the limited contact bandwidth in opportunistic mobile networks, there is a small forwarding probability for large data in an opportunistic contact. Therefore, we try to determine *which* data and *how much* data should be forwarded to *whom* through data partitioning to fully utilizes the abundance of contact opportunities that are relatively short. This paper proposes an optimal data partitioning scheme to maximize the data delivery ratio within a given deadline. We assume that the transmission model used at each contact is persistent; that is, re-transmission at the next contact is applied until the transmission is successful.

There is a trade-off in the optimal data partitioning. A small chunk size has a higher forwarding probability, and as a result, a shorter deliver latency. However, the destination needs to collect all the data chunks from multiple paths to retrieve the data. A long delay in any of the paths will lead to a longer delivery latency. A motivation example is shown in Fig. 2, where the data size (i.e., proportional to transmission time), S, is 12 and 6, respectively. In this toy example, we consider direct contact with the destination through three independent and identical paths. There are two data partitioning strategies: partitioning

data into two chunks or three chunks. The corresponding delivery probabilities are $\{0.04, 0.07\}$, $\{0.18, 0.41\}$ (derived from Fig. 1(a) and 1(b), respectively for $S/2$, and $S/3$. Then, the final delivery probabilities are $\{0.0016, 0.0003\}$ and $\{0.032, 0.069\}$, respectively. In Fig. 1(a), partitioning data into two chunks is better; however, in Fig 1(b), partitioning data into three chunks is better.

This paper discusses the optimal data partitioning scheme in opportunistic network. A blind flooding algorithm is proposed. The optimal homogeneous data chunk size is analyzed in terms of the maximum data delivery ratio under a given deadline. In addition, to avoid receiving the same data chunk multiple times (i.e., the coupon collection problem [10]), we further use the network coding technique [11, 12] to improve the performance with a little extra computation cost.

The contributions of this paper are summarized as follows:

- To our best knowledge, we are the first to discuss optimal homogeneous data partitioning strategies in opportunistic mobile networks.
- We adopt the homogeneous data partitioning in the blind flooding and analyze the optimal data chunk size.
- We verify the optimality of the proposed scheme using extensive experiments in real traces.

The remainder of the paper is organized as follows. The related works are summarized in Section II. The problem statement is introduced in Section III. Then, the blind flooding algorithm is proposed in Section IV. The experiment results are shown in Section V. We conclude the paper in Section VI.

II. RELATED WORKS

In this section, we capture some important data routing issues in opportunistic mobile networks [4, 5, 13–17]

A. Replication-based routing

The earliest routing scheme in opportunistic mobile networks is blind flooding routing [4], which does not control data copy replication. To control the number of data copies, researchers proposed data replication routing based on a quality metric. In delegation forwarding [5], relay replicates data when an encountered node has the best quality metric that it has seen. Basically, we can categorize these routing schemes into two types: contact-capability-based schemes including [18, 19] and social-concept-based schemes including [6, 7]. They used physical distance or social distance with the destination to guide data replication. Many works use different quality metrics and copy control methods [20–22]. However, the network resource consumption of the aforementioned schemes is still high. Our approach differs in that we partition data rather than replicating it, which consumes fewer network resources (e.g., buffer and bandwidth).

B. Contact-duration-aware routing

In [13], the authors optimized the data routing based on the contact duration between nodes. However, the assumption that there is a time-ordered contact graph may not be practical, since the mobile social network is an autonomous network

data size	delivery probability	
	single path	overall
S	0.22	0.220
$S/2$	0.67	$0.67^2 = 0.449$
$S/3$	0.74	$0.74^3 = 0.405$

with much uncertainty. In [14], the authors proposed data partitioning on all better routes except the direct route to the destination. However, selecting too many forwarding paths may lead to poor performance as this paper will later explain. In addition, they failed to discuss the route dependency problem. In this paper, we discuss the optimal data partitioning strategy and apply the concept of hypercube routing [6]. Multiple paths are guaranteed to be edge-disjoint in the hypercube routing. Note that such a property is absent in existing mobile opportunistic network multi-path routing protocols. In addition, [6] is replication-based routing.

III. PROBLEM STATEMENT

In this section, we introduce the contact-aware network model and problem. We use a motivational example to illustrate it. We conclude this section with the problem formulation.

A. Contact-Aware Network Model

In this paper, we consider an opportunistic mobile network with N mobile nodes, where mobile nodes opportunistically contact each other through short range interfaces (e.g., Bluetooth or Wi-Fi). In each contact, each node makes a data-forwarding decision, i.e., which data to send and how much data to send. Note that two encountered nodes do not know the contact duration. If the data transmission is not finished when the encountered node leaves, the data forwarding fails. The current node will transmit this data again during the next contact opportunity between them, it is called persistent transmission since the node will conduct repeated transmissions in next contacts until the first successful transmission. Note that re-transmission will add extra stress to the other existing data chunks to compete the scarce resource: transmission opportunities. If the data transmission has finished when the encountered node leaves, it is a successful data forwarding. Due to the low probability to successfully transmit another data chunk after one data chunk, (which causes high transmission power waste [23]), we use a conservative transmission model. That is, there is at most one transmission in one contact. A path is a sequence of successful contacts from the source to destination.

Based on the observation [8, 9] shown in Fig. 1, we assume that the number of the contacts decreases exponentially as the contact duration increases. In addition, we assume that the contact bandwidth between nodes is a constant value. Then, the data delivery probability exponentially increases as the data size decreases over the opportunistic path/route. When the data size is small enough, it can utilize all contact opportunities. Let us denote this delivery probability as the *base delivery*

probability of the path i as p_i during one contact opportunity. Then, as the data size increases, the delivery probability of a path decreases exponentially, called *decay function*, $\beta(s)$, where s is the data size. Therefore, the delivery probability of a path for data of size s is the product of the base delivery probability and the decay function, as shown in Eq. 1.

$$p_i(s) = p_i \cdot \beta(s). \tag{1}$$

Therefore, the overall delivery probability, P, of a data with m paths is $P = \prod_{i=1}^{m} p_i(s_i)$.

B. Basic Idea and Challenges

Aware of the short contact duration and the large size of content, the source node can either forward data directly or partition data into small chunks that will be sent through multiple paths. In the latter case, each path has a larger contact opportunity than sending data directly, but the destination node needs to receive all the chunks from the multiple paths to recover the original data.

An illustration of the basic idea and challenge of the problem is shown in Fig. 2. In Fig. 2, the source node has a data item with size S to the destination. There are three disjointed opportunistic paths labeled in Fig. 2. To simplify this explanation, we assume all three paths have identical forwarding abilities. Specifically, as Table I shown, each path has delivery probabilities of 0.22, 0.67, and 0.74 to deliver data with sizes of S, $S/2$, and $S/3$, respectively. Without data partitioning, the forwarding probability of a single copy is only 0.22. If we partition the data into two chunks of the same size and use any two paths to conduct forwarding, the delivery probability is $0.67^2 = 0.449$, double that without data partitioning. We can further partition the data into three chunks and use all three paths, the delivery probability becomes $0.74^3 = 0.405$. Its performance decreases, but it is still much better than that without partitioning. There is a trade-off: if we partition the data into smaller chunks, each small chunk has a higher delivery probability to the destination. However, the destination needs to receive chunks from more routes to recover the original data.

C. Problem Statement

The proposed optimal data-partitioning problem is defined as follows, data of size S can be partitioned/fragmented into small data chunks, and each data chunk is forwarded along one forwarding path. That is, for each selected path i, we will assign a chunk with a certain size, s, to it. Then, the offloading delivery probability regarding maximizing the delivery ratio given a deadline T, can be mathematically formulated as

$$\max \quad \prod_{i=1}^{m} \prod_{j=1}^{n} p_j(s)^{x_{i,j}}$$

$$\text{s.t.} \quad \sum_{i=1}^{m} s = S, \tag{2a}$$

$$\sum_{j=1}^{n} x_{i,j} = 1, \quad \sum_{i=1}^{m} x_{i,j} \in \{0,1\}, \quad x_{i,j} \in \{0,1\}, \tag{2b}$$

where n is the number of disjoint paths between the source and destination, m is the number of paths that we selected in the routing, $m \leq n$. $p_j(s)$ is the probability that a chunk with size s is transmitted through path j with the given deadline. $x_{i,j} = 1$, if a chunk is assigned to path j, otherwise $x_{i,j} = 0$. Eq. (2a) means that the destination must receive all the chunks to recover the original data. Eq. (2b) means that each chunk is assigned once, each path is assigned only once, respectively. Thus, the main problem of data offloading is to find an optimal strategy so that the delivery ratio of data with partitioning can be maximized in opportunistic mobile networks.

IV. THE BLIND FLOODING ROUTING

In the classic blind flooding model, the contact delivery probability of two nodes is denoted as a constant value to data of any size. In this paper, we add the decay function $\beta(s)$, as explained in Section III, to model the observation in mobile opportunistic networks. Here, the original data is partitioned into homogeneous data chunks to increase the data delivery probability for each chunk. Once two nodes meet, each of them will forward one chunk from their buffer to the other node in a round-robin fashion.

The optimal chunk size s for each path in the random-waypoint model [24] is analyzed. In the random-waypoint model, every node has the same delivery probability for all the other nodes in the network. Therefore, if data is partitioned into several chunks and sent to the destination through multiple paths, each forwarding path has the same forwarding ability and should have the same chunk size. Then, the problem is to find the optimal homogeneous chunk size.

Theorem 1. *To maximize data delivery probability if nodes' mobilities follow the random-waypoint model and $\beta(s)$ is a decreasing function, the optimal data-partitioning strategy within deadline T in the flooding routing is:*

$$s = -p\frac{d\beta(s)}{ds}T,$$

Proof. Let us denote that the percentage of nodes with a certain chunk at time t in the network as $P(t,s)$. Then, the percentage of node without a certain chunk at time t in the network is $1 - P(t,s)$. Based on the random-waypoint model [24] with an encountering probability, p for every two nodes, the probability increase of a node buffered a certain chunk can be calculated as follows:

$$\frac{d(P(t,s))}{dt} = p\beta(s)P(t,s)(1 - P(t,s)), \tag{3}$$

which means the delivery probability is proportional to the number of nodes with a certain chunk, the number of remaining nodes and the delivery probability. Then, through solving the differential equation, we get the following result for a certain data chunk.

$$P(t,s) = \frac{x_0 e^{p\beta(s)t}}{1 - x_0 + x_0 e^{p\beta(s)t}} = \frac{x_0}{(1 - x_0)e^{-p\beta(s)t} + x_0}, \tag{4}$$

where the x_0 equals to $\frac{1}{N}$. When t equals T, Eq. 5 is the delivery ratio for a certain chunk. Then, the delivery ratio for a data size S, which is partitioned into chunks with size s is

$$P(T,s) = [\frac{x_0}{(1-x_0)e^{-p\beta(s)T}+x_0}]^{S/s}. \quad (5)$$

To get the optimal m in such a situation, we need to get the differential for the Eq. 5 for variable s. Since the function $P(T,s)$ is a concave function in terms of s, when the first differential reaches 0, the function is maximal,

$$\frac{\mathrm{d}P(T,s)}{ds} = \frac{[\frac{x_0}{(1-x_0)e^{-p\beta(s)T}+x_0}]^{\frac{S}{s}}}{ds}. \quad (6)$$

Let us denote $\frac{x_0}{(1-x_0)e^{-p\beta(s)T}+x_0}$ as $I(s)$ for readability. Then, Eq. 6 is changed into the following format,

$$\begin{aligned}
\frac{\mathrm{d}P(T,s)}{ds} &= \frac{\mathrm{d}I(s)^{\frac{S}{s}}}{ds} = \frac{\mathrm{d}e^{\frac{S}{s}\ln I(s)}}{ds} \\
&= e^{\frac{S}{s}lnI(s)}[-\frac{S}{s^2}\ln(I(s)) + \frac{S}{s}\frac{1}{I(s)}\frac{\mathrm{d}(I(s))}{ds}].
\end{aligned} \quad (7)$$

From the differential of $I(s)$, we get the following result

$$\begin{aligned}
\frac{\mathrm{d}I(s)}{ds} &= \frac{\mathrm{d}[\frac{x_0}{(1-x_0)e^{-p\beta(s)T}+x_0}]}{ds} \\
&= \frac{x_0(1-x_0)Te^{-p\beta(s)T}\frac{\mathrm{d}\beta(s)}{ds}}{[(1-x_0)e^{-p\beta(s)T}+x_0]^2} \\
&= I(s)^2\frac{(1-x_0)Te^{-p\beta(s)T}\frac{\mathrm{d}\beta(s)}{ds}}{x_0}.
\end{aligned} \quad (8)$$

By Substituting Eq. 8 to Eq. 7, we can get the following result,

$$\begin{aligned}
\frac{\mathrm{d}p(s)}{ds} &= e^{\frac{S}{s}\ln I(s)}[\frac{-S}{s^2}\ln(I(s)) + \frac{S}{s}\frac{1}{I(s)}\frac{\mathrm{d}(I(s))}{ds}] \\
&\approx e^{\frac{S}{s}\ln I(s)}\frac{S}{s}[\frac{(I(s)-1)}{-s} \\
&+ I(s)\frac{(1-x_0)Te^{-p\beta(s)T}\frac{\mathrm{d}\beta(s)}{ds}}{x_0}],
\end{aligned} \quad (9)$$

the above equation equals 0, when

$$s = -p\frac{\mathrm{d}\beta(s)}{ds}T. \quad (10)$$

Then, we prove the optimal data size s in flooding routing. $\square$

The insight is that if the deadline or the contact probability between two nodes is large, we do not need to partition the data into many chunks, since the current contact opportunity is sufficient to achieve a high delivery ratio. If data partitioning can increase the delivery ratio significantly, the optimal solution will only need to partition the data into few chunks. Note that based on the trace analysis from the traces [8, 9], $\beta(s)$ is an exponential distribution function. However, we only assume $\beta(s)$ is a decreasing function in the proof of Theorem 1. Therefore, Theorem 1 has a wide application scenario. The proposed blind flooding algorithm with data partitioning is shown in Algorithm 1. The average base contact probability

---

Algorithm 1 Blind flooding routing with data partitioning

Input: The average base contact probability between nodes, $\bar{p}$, decay function $\beta(s)$, the delivery deadline T.
1: Calculate the chunk size $s = -\bar{p}\frac{\mathrm{d}\beta(s)}{ds}T$.
2: **if** There exists some chunks that the encountered node does not have **then**
3:     Replicate data chunk in a round-robin fashion.

---

in the network is used to replace p in Theorem 1 to get the optimal data chunk size s in general case.

Note that one potential drawback of data partitioning is that the node might receive the same data chunk multiple times (i.e., the coupon collection problem [10]), especially in the scenarios where data chunk size is small or there are many different data from different sources to different destinations. To address such problem, we propose to use the random linear coding [25] in the proposed method to further improve the performance. The idea of the random linear network coding is that all data chunks in a node's buffer are coded. If the coefficients of the linearly coded packets are selected randomly and in a distributed fashion at each intermediate node, there is a high probability that the generated coded packets will be linearly independent. As a result, it is high possible that receiving any m data chunks can decode the original data.

V. Experiments

In this section, we conduct experiments in two real traces, and we compare the proposed algorithm with other algorithms in terms of the delivery ratio.

A. Trace Introduction

The INFOCOM06 dataset consists of two parts: contacts between the iMote devices carried by participants. There are 74,981 contacts between these participants over a period of 337,418 time slots in seconds during Infocom 2006 in Barcelona. The SIGCOMM09 dataset was collected during the SIGCOMM 2009 conference in Barcelona, Spain. Around 76 smartphones were distributed to a set of volunteers during the first two days of the conference. Participants were recruited on-site in conjunction with the conference registration. In this trace, each device detects its surroundings every 120 ± 10 seconds. Therefore, if we detect two contacts for two nodes within 120 seconds, we regard these two nodes are contacted. Then, we reduce the number of contacts to 24,409 contacts.

B. Experiment Setting

The data size varies from 4MB to 50MB in the INFOCOM trace and the data size varies from 100M to 600M in the SIGCOMM trace. The contact bandwidth between nodes are assumed to be constant, 100kbps, in the mobile environment. In the experiments, the source nodes and the destination nodes are randomly selected in each round and for each experiment. All experiments are repeated for 1000 to 2000 rounds to get the average result.
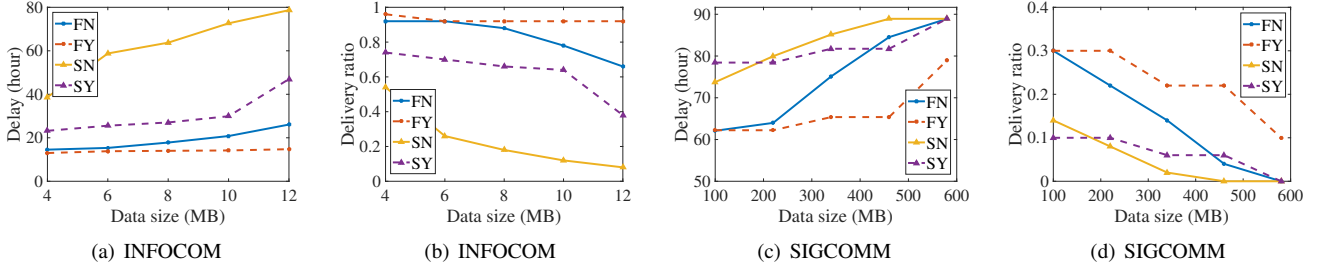
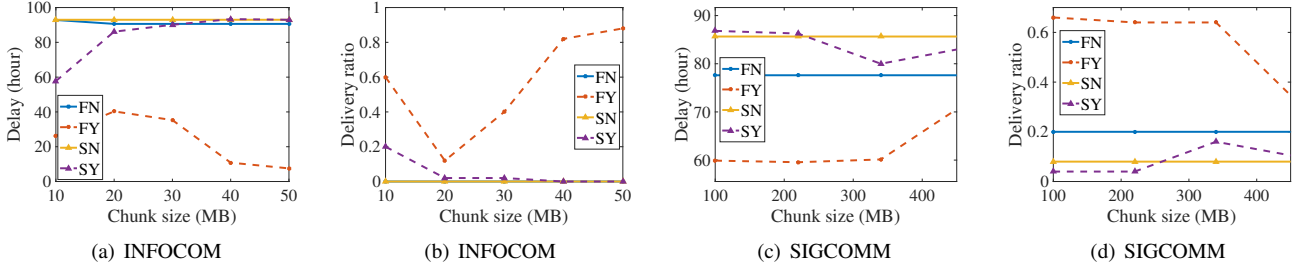Fig. 3. Performance comparison in SIGCOMM dataset.

Fig. 4. Performance comparison of different partition strategies

C. Algorithm Comparison

We propose the following 4 algorithms in different scenarios for performance comparison. They are divided into 2 categories: (1) data partitioning algorithms: the blind flooding algorithm with data partitioning, denoted as the FY algorithm, the single-copy contact-history-based algorithm [5] with data partitioning, denoted as the SY algorithm. (2) comparison algorithms: the flooding algorithm without data partitioning, denoted as the FN algorithm [4], the single-copy contact–history-based algorithm without data partitioning, denoted as the SN algorithm in the experiment. In addition, we apply the random coding in the proposed FY algorithm and conduct experiments in the scenario where there are multiple data from different sources to different destinations. They are denoted as With NC and Without NC, respectively.

D. Experimental Results

The following two subsections show the experimental results of proposed algorithms in different scenarios.

1) With/Without data partitioning: The results in Fig. 3 show experimental results of proposed four algorithms in the INFOCOM and SIGCOMM datasets. There is only one data in the experiment and the delivery deadline is 1 day. Figs. 3(a) and 3(b) show the delivery delay and the corresponding delivery ratio for different data sizes in the INFOCOM trace. In Fig. 3(a), we find that the delivery delay can be greatly reduced with data partitioning in these two data partitioning algorithms. When the data size becomes large, the average delay of the FY algorithm is only 60% of the average day of the FN algorithm. We got similar results in SIGCOMM dataset, which is shown in Fig. 4(c) and 3(d). As for the data delivery ratio, the proposed data partitioning algorithm achieves a similar performance when the data size is small. As the data size increases, the proposed data partitioning algorithm achieves a much better performance. In summary,

the performance of these algorithms follow the order: FY, FN, SY and SN. As for the performance in the SIGCOMM trace, the results are shown in Figs. 3(c) and 3(d). As for the delivery ratio, the data delivery ratio is improved at 25% at least.

2) Optimal data partition strategy: In this subsection, we verify the trade-off in different data partition strategies. Fig. 4 shows performance result in two datasets in terms of different chunk sizes. In Figs. 4(a) and 4(c), when the chunk size is really small, the delivery delay is small. The delivery delay increases a little bit and then decreases again. The reason is that when data chunk size is large, the destination node needs few data chunks to retrieve the original data. In terms of data delivery ratio, we get the similar results as shown in Figs. 4(b) and 4(d). Therefore, the data chunk size should be carefully selected to optimize the delivery ratio.

3) Impact of the network coding: One possible drawback of the data partitioning in multiple data scenario is that data partitioning will create many chunks, and thus the delivery opportunity for each chunk will decrease, especially when many data are buffered in the network. We verified such influence using 1 to 5 data in the network. Figs 5 and 6 show the impact of network coding technique in the proposed FY algorithm. Here we gradually increase the number of data generated in the network, and compare the effectiveness of network coding technique. When there are only one data, the performance improvement of network coding in INFOCOM dataset is limited as shown in Fig. 5(a). With the help of network coding, FY algorithm achieves about 6% more delivery ratio when deadline is between 4 and 15 hours. However, the advantage of network coding becomes significant when there are multiple different data as shown in Fig. 5(b). The result shows that due to the limited contact opportunity for each data, the delivery ratio of FY algorithm without network coding decreases. However, with the help of network coding technique, the delivery ratio is still similar to the scenario
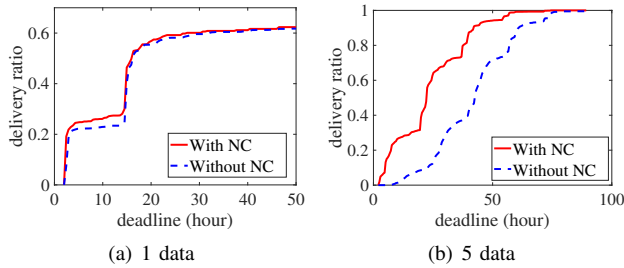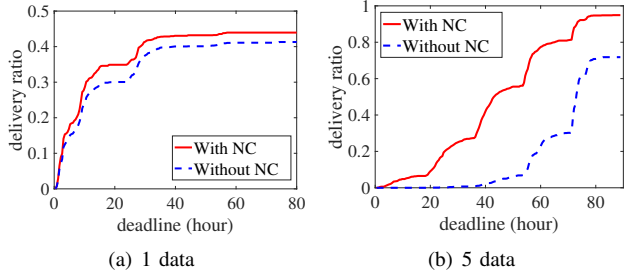
Fig. 5. Network coding in INFOCOM dataset.

Fig. 6. Network coding in SIGCOMM dataset.

where there is only one data. Another observation is that with the help of inter-data coding, each data chunk has even more opportunity, thus the delivery ratio finally reaches to almost 100%. In the SIGCOMM trace, the result is shown in Fig. 6. Fig. 4(c) shows that even there is only one data, with the help of network coding, the FY algorithm with network coding reduces the data collection problem, therefore, it achieves larger delivery ratio as shown in Fig. 6(a). When there are multiple data, the FY algorithm with network coding is much better as shown in Fig. 6(b).

VI. Conclusion

In this paper, we observe limited contact duration in opportunistic mobile networks. Unlike existing works that use data replication schemes to improve the performance, we propose a data partitioning forwarding algorithm to utilize plenty short-contact duration in networks to minimize the data delivery delay and maximize the data delivery ratio within the deadline. A blind flooding algorithm with data chunk replication and analyze the optimal homogeneous data partitioning strategy is proposed and discussed. Extensive experiments on realistic traces show that our distributed scheme outperforms existing replication schemes and achieves better performance than schemes without partitioning, which demonstrates the necessity of data partitioning in opportunistic mobile network, i.e., more than 25% delivery ratio improvement.

References

[1] C. V. N. I. Cisco, "Global mobile data traffic forecast update, 2013–2018," *white paper*, 2014.

[2] http://www.wi-fi.org, 2016.

[3] M. Seufert, G. Darzanos, I. Papafili, R. Łapacz, V. Burger, and T. Hoßfeld, "Socially-aware traffic management," in *Socioinformatics-The Social Impact of Interactions between Humans and IT*. Springer, 2014, pp. 25–43.

[4] A. Vahdat, D. Becker *et al.*, "Epidemic routing for partially connected ad hoc networks," 2000.

[5] V. Erramilli, M. Crovella, A. Chaintreau, and C. Diot, "Delegation forwarding," in *Proceedings of the ACM MobiHoc*, 2008.

[6] Y. Wang, W.-S. Yang, and J. Wu, "Analysis of a hypercube-based social feature multipath routing in delay tolerant networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 9, pp. 1706–1716, 2013.

[7] W. Gao, Q. Li, B. Zhao, and G. Cao, "Social-aware multicast in disruption-tolerant networks," *IEEE/ACM Transactions on Networking*, vol. 20, no. 5, pp. 1553–1566, 2012.

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

[9] A.-K. Pietil"ainen and C. Diot, "Dissemination in opportunistic social networks: The role of temporal communities," in *Proceedings of the ACM MobiHoc*, 2012.

[10] https://en.wikipedia.org/wiki/Coupon_Collector's_problem, 2016.

[11] S.-Y. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, 2003.

[12] P. Ostovari, J. Wu, and A. Khreishah, "Network coding techniques for wireless and sensor networks," *The Art of Wireless Sensor Networks*, pp. 129–162, 2014.

[13] Z. Li, Y. Liu, H. Zhu, and L. Sun, "Coff: contact-duration-aware cellular traffic offloading over delay tolerant networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 11, pp. 5257–5268, 2015.

[14] Z. Lu, X. Sun, and T. La Porta, "Cooperative data offloading in opportunistic mobile networks," 2016.

[15] N. Wang and J. Wu, "Opportunistic wifi offloading in a vehicular environment: Waiting or downloading now?" in *Proceedings of the IEEE INFOCOM*, 2016.

[16] W. Chang, J. Wu, and C. C. Tan, "Enhancing mobile social network privacy," in *Proceedings of the IEEE GLOBECOM*, 2011.

[17] N. Wang and J. Wu, "Optimal cellular traffic offloading through opportunistic mobile networks by data partitioning," in *Proceedings of the IEEE ICC*, 2018.

[18] W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in delay tolerant networks: a social network perspective," in *Proceedings of the ACM MobiHoc*, 2009.

[19] N. Wang and J. Wu, "A general data and acknowledgement dissemination scheme in mobile social networks," in *Proceedings of the IEEE MASS*, 2014.

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

[21] P. Hui, J. Crowcroft, and E. Yoneki, "Bubble rap: Social-based forwarding in delay-tolerant networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 11, pp. 1576–1589, 2011.

[22] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: an efficient routing scheme for intermittently connected mobile networks," in *Proceedings of the ACM SIGCOMM workshop*, 2005.

[23] G. P. Perrucci, F. H. Fitzek, and J. Widmer, "Survey on energy consumption entities on the smartphone platform," in *Proceedings of the IEEE VTC*, 2011.

[24] S. Bandyopadhyay, E. J. Coyle, and T. Falck, "Stochastic properties of mobility models in mobile ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 11, pp. 1218–1229, 2007.

[25] X. Zhang, G. Neglia, J. Kurose, and D. Towsley, "On the benefits of random linear coding for unicast applications in disruption tolerant networks," in *Proceedings of the IEEE WiOpt*, 2006.