

Efficient Transmission of Data by a Sequence of Clusters for a Disruption Tolerant Network

Soumya Patil

PG Student, Dept of CSE,
AMCEC Bangalore
Karnataka, India

Mrs, Ramya

Asst. Professor, Dept of CSE,
AMCEC, Bangalore
Karnataka, India

Basavaraj Patil

Asst. Professor, Dept. of ISE
S.D.M. Institute of Technology
Ujire, Karnataka, India

Abstract— Presently wireless networks are deployed in various extreme environments where they suffer from different levels of link disruptions depending on the severity of the operating conditions. Disruption/Delay tolerant networks (DTN), are designed with node density, node mobility, and less global network information. The major research work in DTNs is focused on data forwarding, but there are many challenges in providing an efficient access the data to a mobile user. In this paper, we put forward a new method to enhance the caching in DTNs, which is needed for sharing the cached data for multiple nodes and decreasing the delaying access. Here the cached data in group of nodes, a center node can be access the data from all neighbor nodes in the network. We propose an efficient method which ensures that proper selection of center node based on a probabilistic selection metric. The simulations prove that our approach extensively improves data access performance compared to existing schemes.

Key Words —DTN, cooperative caching, CN.

I. INTRODUCTION

Disruption tolerant networks (DTNs) composed of mobile devices that are connected to each other. Due to the less node density and random node mobility, the connectivity is irregular between the nodes in DTNs and is also difficult for maintaining the end-to end communication links for data transmission. Popular examples of such irregularly connected networks (ICNs) conditions include:

- 1) Exotic Media Networks (EMNs) interconnecting extraterrestrial nodes (e.g. satellites, deep space probes, landers, orbiters, etc.) which are periodically communicate with base (i.e. terrestrial) nodes using high latency Radio Frequency (RF) transceivers.
- 2) Wireless Sensor Networks (WSNs), Mobile Wireless Sensor Networks (MWSNs) and Sensor/Actuator Networks (SANs) deployed in extreme regions (e.g. Amazons, deep volcanic or underwater areas, etc.), which consist of low power sensor nodes that, for the purpose of energy saving, periodically switch between active/sleep modes and are unable to provide continuous interaction with a data-collecting server.
- 3) Mobile Ad-Hoc Networks (MANETs) typically consisting of nodes (e.g. GPSs, PDAs, Cellular Phones, Tracking devices, Laptops, etc.) placed over incessantly moving objects (e.g. vehicles, moving individuals, animals, etc.). Communication in these networks is regularly intermittent either due to nodes going out of communication range of each other or obstructing obstacles or node destructions as is the case in Battlefield Wireless Military Networks (BWMNs) [11]. In such networks, the mobility node is broken to let

mobile nodes carry data as relays and forward data frequently when get in touch with others. The main problem is “How to determine the appropriate relay selection strategy”.

Although forwarding schemes have proposed DTNs [4], [1], [13], there is limited research work on providing efficient data access to mobile users. Example, Smartphone Users can find important digital content from their nearby users. In vehicular ad-hoc networks (VANETs), the live traffic information is necessary for vehicles to avoid traffic delays.

In these applications, data are available only whenever mobile users requested and when they are in need and the requestor does not know the position of data in advance. The destination of data is, hence, not known when data are generated.

This communication paradigm differs from publish/subscribe systems in which data are forwarded by broker nodes to users according to their data subscriptions. Appropriate network design is needed to ensure that data can be promptly accessed by requesters in such cases.

A common technique used to improve data access performance is caching, i.e., to cache data at appropriate network positions based on query history, so that queries in the future can be responded with less delay. Although cooperative caching has been studied for both web-based applications [15] and wireless ad hoc networks, to allow sharing and coordination among multiple caching nodes, it is difficult to be realized in DTNs due to the lack of persistent network connectivity. First, the opportunistic network connectivity complicates the estimation of data transmission delay, and furthermore makes it difficult to determine appropriate caching positions for reducing data access delay. This difficulty is also raised by the incomplete information at individual nodes about query history. Second, due to the uncertainty of data transmission, multiple data copies need to be cached at different positions to ensure data accessibility. The difficulty in coordinating multiple caching nodes makes it hard to optimize the tradeoff between data accessibility and caching overhead.

In this paper, we propose a new method to address the challenges and to proficiently support the cooperative caching in DTNs. The basic plan is to collect the cached data at a set of center located nodes, each of which corresponds to a set of neighboring mobile nodes, which can be easily accessed by nearest nodes in the network. Each Center node is represented as Central Node (CN) which has high priority in the network and is given preference for caching data. Due to the restricted caching buffer of central nodes, multiple nodes are involved for caching. We make sure that popular data are always

cached nearer to the central nodes through dynamic cache replacement method based on query transaction. Our work as follows:

- a) We develop an efficient approach to CN (Center Node) selection in DTNs based on probabilistic selection parameters. The selected CNs is having high priority for the correct response to user queries with lesser time in network storage and transmission.
- b) We propose a data transfer method to organize multiple caching nodes for replying to the user queries.
- c) We reduce the transactions to minimize the average number of cached data copies between data accessibility and caching overhead in the network.
- d) We propose necessity-based cache replacement methods which dynamically check the cache positions depending on query history and our method achieves good tradeoff between the data accessibility and data access delay.

II. LITERATURE SURVEY

Research on data forwarding in DTNs originates from Epidemic routing [34], which floods the entire network. Some later studies focus on proposing efficient relay selection metrics to approach the performance of Epidemic routing with lower forwarding cost, based on prediction of node contacts in the future. Some schemes do such prediction based on their mobility patterns, which are characterized by Kalman filter [8] or semi-Markov chains. In some other schemes, node contact pattern is exploited as abstraction of node mobility pattern for better prediction accuracy [4], based on the experimental [7] and theoretical [5] analysis of the node contact characteristics. The social network properties of node contact patterns, such as the centrality and community structures, have also been also exploited for relay selection in recent social-based data forwarding schemes [9].

The parameter for relay selection can be applied to different forwarding strategies, which vary in the number of data copies created in the network. While the most conservative strategy always keeps a single data copy and Spray-and-Wait holds a fixed number of data copies, most schemes dynamically determine the number of data copies. In Compare-and-Forward [12], a relay forwards data to another node whose metric value is higher than itself. Delegation forwarding [13] reduces forwarding cost by only forwarding data to nodes with the highest metric. Data access in DTNs, on the other hand, can be provided in various ways. Data can be disseminated to appropriate users based on their interest profiles [16]. In other schemes [2] without brokers, data items are grouped into predefined channels, and are disseminated based on users' subscriptions to these channels.

Caching is another way to provide data access. Cooperative caching in wireless ad hoc networks was studied in, in which each node caches pass-by data based on data popularity, so that queries in the future can be responded with less delay. Caching positions are selected incidentally among all the network nodes. Some research efforts have been made for caching in DTNs, but they only improve data accessibility from infrastructure network such as WiFi access points (APs) or Internet. Peer-to-peer data sharing and access among mobile users are generally neglected.

Distributed determination of caching policies for minimizing data access delay has been studied in DTNs, assuming simplified network conditions. In [16], it is assumed that all the nodes contact each other with the same rate. The users are artificially partitioned into several classes such that users in the same class are identical, but these caching positions are determined based on global network knowledge. So in this paper, we propose novel approach to support cooperative caching in a fully distributed manner in DTNs, in heterogeneous network.

III. CENTRAL NODE

The selection of CNs based on probabilistic parameter by evaluating the data transmission delay among nodes in DTNs and by validating the availability of such parameters in practice depending on the type of node connection pattern in realistic DTN traces. Then we propose the practical approach of selecting as CN based on availability of network information.

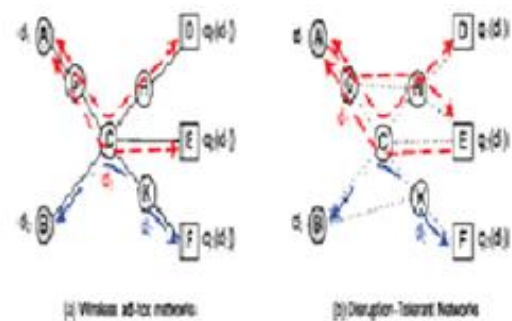


Fig.1. Caching strategies in different network environments. Data d1 generated by node A are requested by nodes D and E, and d2 generated by node B are requested by node F. A solid line in (a) between nodes represents a wireless link, and a dotted line in (b) represents that two nodes opportunistically contact each other.

3.1 Selection of Central Node(CN)

The methods for identifying the required K CNs in practice based on the center node selection parameters. K is defined as a parameter determined by the network performance requirements. The network information regarding the pair wise node communication counts and shortest distanced paths between the mobile nodes are important to evaluate the metric values of each mobile node. The maintenance of this network information is costlier in DTNs because of the lack of constant connectivity between each node. Therefore, we assume that the complete information is globally available for the selection of center node then we define the distributed CN selection methods are approximate with global selection results and then processed on each individual node in an independent manner.

3.1.1 Global Selection

When global network information about the pair wise node connected values and minimal opportunistic paths between the mobile nodes are available, central nodes representing CNs can be identified in order before data access by the network admin. Let INC denotes the set of selected central nodes, every time the node in V and INC with the highest value is

selected as the next central node, until the required K central nodes are selected. The center node C_i is calculated (1) using V and INC as below, i.e.

$$C_i = \frac{1}{|V \setminus N_C|} \cdot \sum_{j \in V \setminus N_C} p_{ij}(T). \quad \text{-----(1)}$$

Then the selected central nodes will not be grouped based on the network communication graph. The parameter traces, T is calculated by the average node communication frequency for different DTN traces.

A t time set for collecting the information and to calculate their pairwise communication values. Then central nodes are selected after the set time t ends. Data which are not capable to be caching at the CNs during the set up time by nodes in the network the data are incidentally cached for data access. Every request sends queries directly to the data source, and caches the received data locally for communicating with the other near-by queries in the future.

3.1.2 Distributed Selection

When global network information is unavailable a node manages the information regarding the pair wise communication values and finds the shortest paths to other nodes via opportunistic contacts.

IV. CACHING SCHEME

We define our new cooperative caching method; purposely cache data at a set of CNs are accessed by other nodes. Our method consists of the following three components:

1. When a data is generated by data source it is added to nodes of CNs which are preferred to cache data. One copy of data is cached at each CN. If the caching buffer is full, another node near to the central node will cache the data. These processes are done based on buffer circumstances of each node involved in the adding process.
2. A requester multicasts a query to central nodes of CNs to receive the data and a central node processed the query to the caching nodes. Multiple data copies are sent back to the requester, which optimizes the tradeoff between data accessibility and transmission cost by controlling the number of returned data copies.
3. Whenever two different caching nodes communicate and make sure that important data are cached nearer to central nodes utility-based cache replacement method performed. We cache more number of important data copies to optimize the data access delay and also cache the less important data to make sure that the overall data accessibility.

4.1 Caching Position

Whenever a node S generates new data, S adds the data to CNs by sending a copy of data to each node representing a CN. The optimized path weights is used as relay selection parameter for data forwarding and then relay forwards data to another node with a higher value of itself. This ‘‘Compare-and-Forward’’ strategy has been widely used for efficient data forwarding. According, the opportunistic path, ensures that each forwarding reduces the remaining delay for data to be delivered to the central node. For the data generated recently the initial caching positions are automatically identified during

the forwarding process based on node buffer conditions. The caching positions are then dynamically adjusted by cache replacement according to query history.

Data is forwarded to and cached at central nodes. This forwarding process only stops when the caching buffer of the next relay is full and data are cached at the current relay in different cases. The identification of caching position is explained in figure 2, where the bold lines represent opportunistic communication links used to forward data, and the dashed lines represent data forwarding stopped by node buffer constraint. Central node C1 is able to cache data, but data copies to C2 and C3 are stopped and cached at relays R24 and R33 respectively, because neither C2 nor R34 has enough buffers to cache data. The caching position at a CN may not be the communicated neighbor of a central node like in the case of nodes R33. From this it is easy to find that the set of caching nodes at each CN forms a connected subgraph of the network communication graph at any time during data access. This property necessarily facilitates the delivery of user queries to the caching nodes.

4.2 Queries

Suppose that any node may request data and the data requesters are randomly distributed in the network. A requester multicast a query within a given time constraint to all the nodes to pull data and existing multicast schemes in DTNs can be exploited for this reason. After receiving the query a central node immediately responds to the requester with data if whether it is cached locally or it has broadcasted the query to the nodes nearby. This process is explained in Fig. 3. While the central node C1 is able to return the cached data to R immediately, the caching nodes A and B only reply to R after they receive the query from central nodes C2 and C3, respectively. The broadcasting of query ends when query expires. Each time when the caching node caches data at CNs maintains updated information about query history and used in cache replacement.

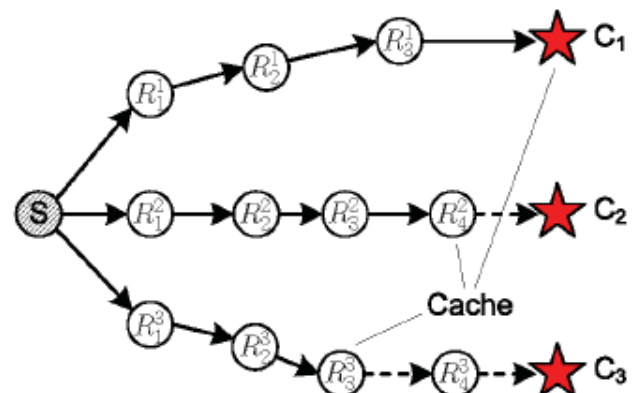


Fig.2. Determining caching position at CNs.

4.3 Probabilistic Response

Multiple data copies are replied to the requester from CNs to ensure that the requester receives data before query expires. However, only the first data copy received by the requester is useful, and all the others are essentially useless and waste network resources. The major challenge for solving this problem arises from the intermittent network connectivity in

DTNs. First, it is difficult for caching nodes to promptly communicate with each other, and hence, the optimal number of data copies returned to the requester cannot be determined in advance. Second, a relay carrying a data copy does not know the positions of other data copies being returned, and therefore cannot determine whether the requester has received data. In this section, we propose a probabilistic scheme to address these challenges and optimize the tradeoff between data accessibility and transmission overhead. Our main idea is that, after received the query, a caching node probabilistically decides whether to return the cached data to the requester. Different strategies are used for this decision, according to the availability of network contact information.

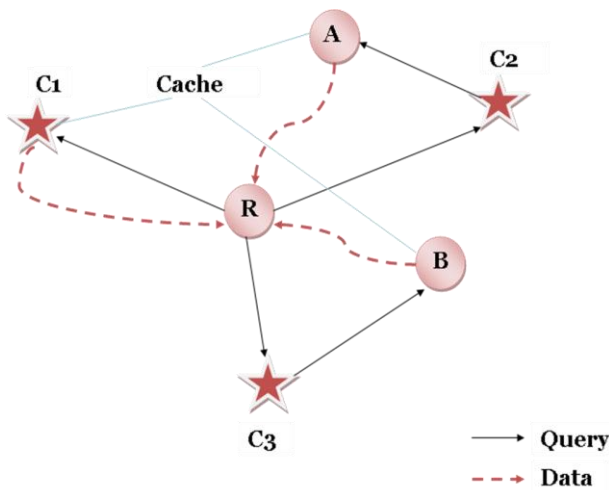


Fig. 3. Pulling data from the CNs.

4.4 Cache Replacement

In the network for each data element, the positions where it is cached are dynamically adjusted via cache replacement. Cache replacement scheme is based on data popularity and places popular information nearest to the nodes of CNs. Earlier cache replacement methods like LRU removes the recently used data from cache when new information is available and ineffective due to consideration of important data. Greedy Dual Size [6] evaluates the data utility by considering important data and size concurrently, but cannot be sure optimal selection of cached data. We improved the early work by defining a new probabilistic cache replacement strategy which appropriately selects the data to be cached and heuristically balances between the cumulative data accessibility and access delay.

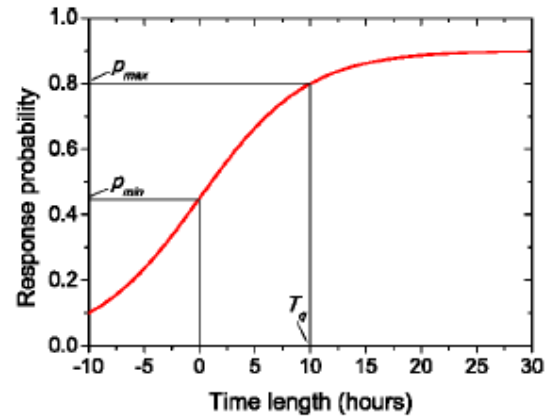


Fig. 4. Probability for deciding data response

V. ARRANGEMENT OF CACHING POSITIONS

When a new central node is identified, then it is selected as CN, the data cached is represented by the original central node. The needs are adjusted correspondingly, so as to optimize the caching performance. After the processing of central node C1 will be migrated to C3 then the nodes A, B, and C near C1 are not considered as good positions for caching data anymore is as shown in figure 5. Instead of that the cached data at the nodes needs to be migrated to nodes nearest to C3. This migration is achieved through cache replacement mechanism when caching nodes communicating each other. Each caching node at the original CN recalculates the utilities of its cached data elements w.r.t the newly selected CN. These data utilities will be decreased by the changes of central nodes and reduction moves the cached data to the exact caching positions that are nearest to the newly identified CN. These Changes in central nodes and its caching positions major affect on caching performance as shown in Fig. 5. Therefore the performance deprivation will be gradually removed over time by the cache replacement mechanism. Then we will evaluate the impact on real DTN traces.

The impact of CN load balancing scheme defined on caching performance. The evaluation results saves g ¼ 100 Mb by setting K ¼ 8 for the Reality trace.

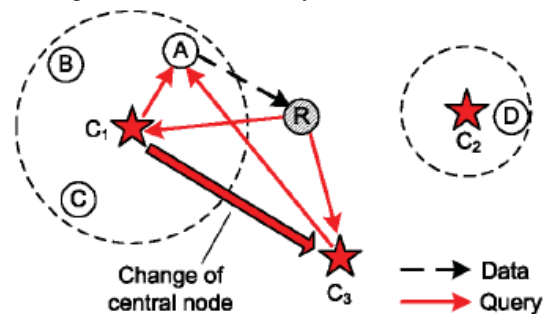


Fig. 5. Load balancing

The CN checks whether to migrate the functionality to another node with a probability p time-time continuously. The time for making decision is to be 10% of the trace length and the evaluation results with different values of p on the reality

trace. When the central nodes change, the existing caching positions become inappropriate and therefore the successful ratio of data access is decreased. The reduction can be up to 40% when the lifetime of data is short but significantly to 10% when there is longer time for the queries to be forwarded to the caching nodes. The impact of CN load balancing is determined by the frequency of the changes of central nodes. The impact of CN load balancing on the caching performance is largely determined by the specific network condition and data access pattern.

VI. SIMULATION RESULTS

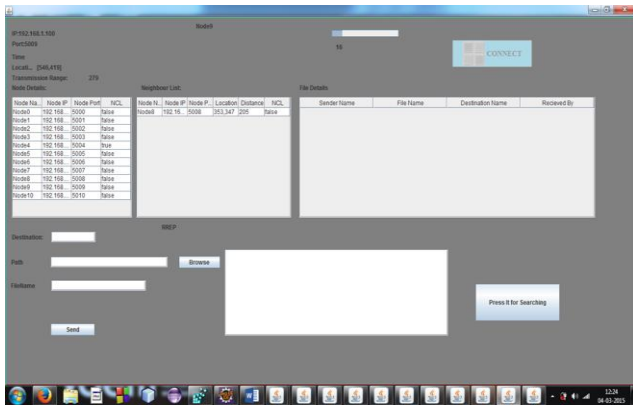


Fig. 6 CN is true

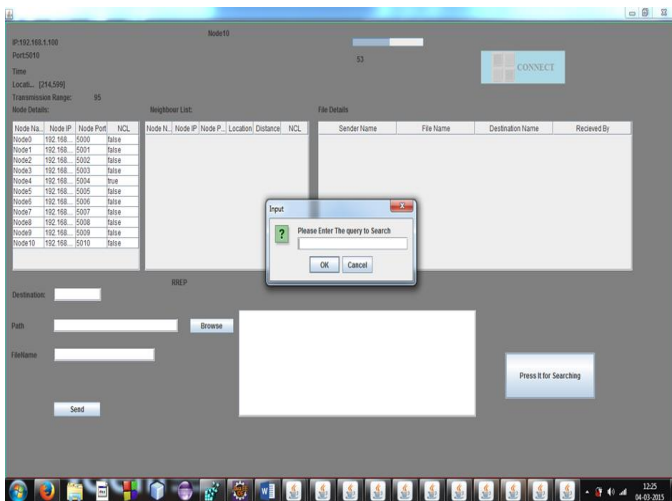


Fig.7. Enter the query to search

VII. CONCLUSION

We proposed new method to maintain cooperative caching in DTNs. Our intentional is to cache data at a set of CNs, which can be easily accessed by nodes. We make sure that the exact CN selection is done based on a probabilistic parameter; our method coordinates caching nodes to optimize the time between data accessibility & caching overhead. The trace driven simulations show that our method is improved; the ratio of queries satisfied and reduces data access delay compared with existing schemes.

REFERENCES

- [1] A. Balasubramanian, B. Levine, and A. Venkataramani, "DTN Routing as a Resource Alposition Problem," Proc. ACM SIGCOMM Conf. Applications, Technologies, Architectures, and Protocols for Computer Comm., pp. 373-384, 2007.
- [2] C. Boldrini, M. Conti, and A. Passarella, "ContentPlace: Social-Aware Data Dissemination in Opportunistic Networks," Proc. 11th Int'l Symp. Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM), pp. 203-210, 2008.
- [3] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "Web Caching and Zipf-Like Distributions: Evidence and Implications." Proc. IEEE INFOCOM, vol. 1, 1999.
- [4] J. Burgess, B. Gallagher, D. Jensen, and B. Levine, "MaxProp: Routing for Vehicle-Based Disruption-Tolerant Networks," Proc. IEEE INFOCOM, 2006.
- [5] H. Cai and D.Y. Eun, "Crossing over the Bounded Domain: From Exponential to Power-Law Inter-Meeting Time in MANET," Proc. ACM MobiCom, pp. 159-170, 2007.
- [6] P. Cao and S. Irani, "Cost-Aware WWW Proxy Caching Algorithms," Proc. USENIX Symp. Internet Technologies and Systems, 1997.
- [7] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of Human Mobility on Opportunistic Forwarding Algorithms," IEEE Trans. Mobile Computing, vol. 6, no. 6, pp. 606-620, June 2007.
- [8] P. Costa, C. Mascolo, M. Musolesi, and G. Picco, "Socially Aware Routing for Publish-Subscribe in Delay-Tolerant Mobile Ad Hoc Networks," IEEE J. Selected Areas in Comm., vol. 26, no. 5, pp. 748-760, June 2008.
- [9] E. Daly and M. Haahr, "Social Network Analysis for Routing in Disconnected Delay-Tolerant MANETs," Proc. ACM MobiHoc, 2007.
- [10] H. Dubois-Ferriere, M. Grossglauser, and M. Vetterli, "Age Matters: Efficient Route Discovery in Mobile Ad Hoc Networks Using Encounter Ages," Proc. ACM MobiHoc, pp. 257-266, 2003.
- [11] J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden, and H. Balakrishnan, "The Pothole Patrol: Using a Mobile Sensor Network for Road Surface Monitoring," Proc. ACM Sixth Ann. Int'l Conf. Mobile Systems, Applications and Services (MobiSys), 2008.
- [12] V. Erramilli, A. Chaintreau, M. Crovella, and C. Diot, "Diversity of Forwarding Paths in Pocket Switched Networks," Proc. Seventh ACM SIGCOMM Conf. Internet Measurement (IMC), pp. 161-174, 2007.
- [13] V. Erramilli, A. Chaintreau, M. Crovella, and C. Diot, "Delegation Forwarding," Proc. ACM MobiHoc, 2008.
- [14] M. Fiore, F. Mininni, C. Casetti, and C.F. Chiasserini, "To Cache or Not to Cache?" Proc. IEEE INFOCOM, pp. 235-243, 2009.
- [15] W. Gao and G. Cao, "On Exploiting Transient Contact Patterns for Data Forwarding in Delay Tolerant Networks," Proc. IEEE Int'l Conf. Network Protocols (ICNP), pp. 193-202, 2010.
- [16] J. Reich and A. Chaintreau, "The Age of Impatience: Optimal Replication Schemes for Opportunistic Networks," Proc. ACM Fifth Int'l Conf. Emerging Networking Experiments and Technologies (CoNEXT), pp. 85-96, 2009.
- [17] J. Zhao, P. Zhang, G. Cao, and C. Das, "Cooperative Caching in Wireless P2P Networks: Design, Implementation, and Evaluation," IEEE Trans. Parallel & Distributed Systems, vol. 21, no. 2, pp. 229-241, Feb. 2010.
- [18] H. Zhu, L. Fu, G. Xue, Y. Zhu, M. Li, and L.M. Ni, "Recognizing Exponential Inter-Contact Time in VANETs," Proc. IEEE INFOCOM, 2010.
- [19] Wei Gao, Guohong Cao, Arun Iyengar and Mudhakar Srivatsa, "Cooperative Caching for Efficient Data Access in Disruption Tolerant Networks" IEEE transactions on mobile computing, vol. 13, no. 3, March 2014.