International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 10, October 2015
Available at http://internationaljournalofresearch.org

Efficient Data Access in Disruption Tolerant Networks Using Cooperative Caching

R.Anil Kumar #1& N.Venkatadri #2,
#1 PG Student, Dept. Of CSE, SKR College Of Engineering and Technology, Kondurusatram,Manubolu,Nellore,AP.
#2Assoc. Professor, Dept. Of CSE, SKR College Of Engineering and Technology, Kondurusatram, Manubolu,Nellore,AP

Abstract:-

Disruption tolerant networks (DTNs) are characterized by low node density, unpredictable node mobility, and lack of global network information. Most of current research efforts in DTNs focus on data forwarding, but only limited work has been done on providing efficient data access to mobile users. In this paper, we propose a novel approach to support cooperative caching in DTNs, which enables the sharing and coordination of cached data among multiple nodes and reduces data access delay. Our basic idea is to intentionally cache data at a set of network central locations (NCLs), which can be easily accessed by other nodes in the network. We propose an efficient scheme that ensures appropriate NCL selection based on a probabilistic selection metric and coordinates multiple caching nodes to optimize the tradeoff between data accessibility and caching overhead. Extensive trace-driven simulations show that our approach significantly improves data access performance compared to existing schemes.

Index Terms—Cooperative caching; disruption tolerant networks; data access; network central locations; cache replacement

2 RELATED WORK

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 [37]. In some other schemes, node contact pattern is exploited as abstraction of node mobility pattern for better prediction accuracy [4], [24], based on the experimental [7] and theoretical [5] analysis of the node contact character-istics. 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], [22], [20].

The aforementioned metrics for relay selection can be applied to various forwarding strategies, which differ in the number of data copies created in the network. While the most conservative strategy [32] always keeps a single data copy and Spray-and-Wait [31] 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 [28]. Data can be disseminated to appro-priate users based on their interest profiles [18]. Publish/ subscribe systems [36], [25] were used for data dissemina-tion, where social community structures are usuallyexploited to determine broker nodes. In other schemes [24], [2] without brokers, data items are grouped into predefined channels, and are disseminated based on users' subscriptions to these channels.

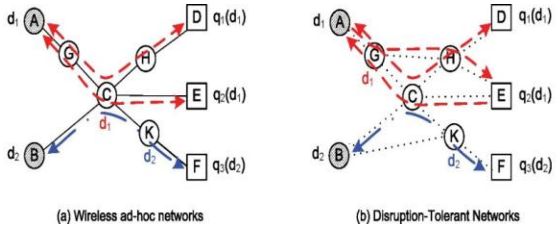
Caching is another way to provide data access. Cooperative caching in wireless ad hoc networks was studied in [35], in which each node caches pass-by data based on data popularity, so that queries in the future can be responded with less delay. Caching locations are selected incidentally among all the network nodes. Some research efforts [27], [21] have been made for caching in DTNs, but they only improve data accessibility from infrastructure network such as WiFi access points (APs) [21] or Internet [27]. Peer-to-peer data sharing and access among mobile users are generally neglected.

Distributed determination of caching policies for mini-mizing data access delay has been studied in DTNs [29], [23], assuming simplified network conditions. In [29], it is assumed that all the nodes contact each other with the same rate. In [23], users are artificially partitioned into several classes such that users in the same class are identical. In [19], data are intentionally cached at appropriate network locations with generic data and query models, but these caching locations are determined based on global network knowledge. Comparatively, in this paper, we propose to support cooperative caching in a fully distributed manner in DTNs, with heterogeneous node contact patterns and behaviors.

3 OVERVIEW

3.1 Motivation

A requester queries the network for data access, and the data source or caching nodes reply to the requester with data after having received the query. The key difference between caching strategies in wireless ad hoc networks and DTNs is illustrated in Fig. 1. Note that each node has limited space for caching. Otherwise, data can be cached every-where, and it is trivial to design different caching strategies.



(a) Wireless ad-hoc networks

(b) Disruption-Tolerant Networks

Fig. 1.Caching strategies in different network environments. Data d_1 generated by node A are requested by nodes D and E, and d_2 generated by node B are requested by node F . A solid line in (a) between nodes indicates a wireless link, and a dotted line in (b) indicates that two nodes opportunistically contact each other.

The design of caching strategy in wireless ad hoc networks benefits from the assumption of existing end-to-end paths among mobile nodes, and the path from a requester to the data source remains unchanged during data access in most cases. Such assumption enables any intermediate node on the path to cache the pass-by data. For example, in Fig. 1a, C forwards all the three queries to data sources A and B, and also forwards data d_1 and d_2 to the requesters. In case of limited cache space, C caches the more popular data d_1 based on query history, and similarly data d_2 are cached at node K. In general, any node could cache the pass-by data incidentally.

However, the effectiveness of such an incidental caching strategy is seriously impaired in DTNs, which do not assume any persistent network connectivity. Since data are forwarded via opportunistic contacts, the query and replied data may take different routes, and it is difficult for nodes to collect the information about query history and make caching decision. For example, in Fig. 1b, after having forwarded query q_2 to A, node C loses its connection to G, and cannot cache data d_1 replied torequester E. Node H which forwards the replied data to E does not cache the pass-by data d_1 either because it did not record query q_2 and considers d_1 less popular. In this case, d_1 will be cached at node G, and hence needs longer time to be replied to the requester.

Our basic solution to improve caching performance in DTNs is to restrain the scope of nodes being involved for caching. Instead of being incidentally cached “anywhere,” data are intentionally cached only at specific nodes. These nodes are carefully selected to ensure data accessibility, and constraining the scope of caching locations reduces the complexity of maintaining query history and making caching decision.

3.2 Network Model

Opportunistic contacts in DTNs are described by a network contact graph GðV ; EÞ, where the stochastic contact process between a node pair i; j 2 V is modeled as an edge e_{ij} 2 E. We assume that node contacts are symmetric; i.e., node j contacts i whenever i contacts j, and the network contact graph is, therefore, undirected. The characteristics of an edge e_{ij} 2 E are determined by the properties of inter-contact time among nodes. Similar to previous work [1], [39], we consider the pairwise node intercontact time as exponentially distributed. Contacts between nodes i and j then form a Poisson process with contact rate _$_{ij}$, which is calculated in real time from the cumulative contacts between nodes i and j because the network starts. In the rest of this paper, we call the node set fj j _$_{ij}$> 0g _ V as the contacted neighbors of i.

4. Probabilistic Data Selection

The aforementioned removal of cached data essentially prioritizes popular data during cache replacement, but may impair the cumulative data accessibility. The major reason is that according to our network modeling in Section 3.2, the data accessibility does not increase linearly with the number of cached data copies in the network. More specifically, the data accessibility will increase considerably if the number of cached data copies increases from 1 to 2, but the benefit will be much smaller if the number increases from 10 to 11. In such cases, for the example shown in Fig. 10b, caching d_1 at node A may be ineffective because the popular d_1 may already be cached at many other places in the network. In contrast, removing d_6 out from the cache of node B may greatly impair the accessibility of d_6 because there may be only few cached copies of d_6 due to its lower popularity.

In other words, the basic strategy of cache replacement only optimizes the cumulative data access delay within the local scope of the two caching nodes in contact. Such optimization at the global scope is challenging in DTNs due to the difficulty of maintaining knowledge about the current number of cached data copies in the network, and we instead propose a probabilistic strategy to heuristically control the number of cached data copies at the global scope.

The basic idea is to probabilistically select data to cache when the problem in (12) is solved by a dynamic programming approach. More specifically, if data d_i are selected by the dynamic programming algorithm, it has probability u_i to be cached at node A. This algorithm is described in detail in Algorithm 1, where GetMaxð**S**; S_AÞ calculates the maximal possible value of the items in the knapsack via dynamic programming, and SelectDataðd_{imax} Þ determines whether to select data d_{imax} to cache at node A by conducting a Bernoulli experiment with probability u_{imax} . Such probabilistic selection may be iteratively conducted multiple times to ensure that the caching buffer is fully utilized. By proposing this probabilistic strategy, we still prioritize the popular data with higher utility during the caching decision, but also enable the data with less popularity to have nonnegligible chance to be cached.

Algorithm 1. Probabilistic Data Selection at node A among the data set **S**:

```
$i_{min}$ ¼ argmin$_i$f$s_i$ j $d_i$ 2 S; $x_i$ ¼¼ 0g
while S6¼ ;&&$S_A$>$s_{i_{min}}$ do
    $V_{max}$ ¼ GetMax(S, $S_A$)
    $S^0$¼ S
    while $S^0$¼6 ;&&$V_{max}$>0 do
        $i_{max}$ ¼ argmax$_i$f$u_i$ j $d_i$ 2 $S^0$g
        if SelectData ($d_{i_{max}}$ )¼¼ true &&$V_{max\_}s_{i_{max}}$ then
            $x_{i_{max}¼}$ 1
```

9. $\mathbf{S}$ ¼ $\mathbf{S}$ n $d_{i_{max}}$
10. S_A ¼ S_A _ $s_{i_{max}}$, V_{max} ¼ V_{max} _ $s_{i_{max}}$
11. $\mathbf{S}^0$ ¼ $\mathbf{S}^0$ n $d_{i_{max}}$
12. i_{min} ¼ argmin$_i$fs_i j d_i 2 $\mathbf{S}$; x_i ¼¼ 0g

5. PERFORMANCE EVALUATION

We evaluate the performance of our proposed caching scheme by comparing it with the following schemes:

. No Cache, where caching is not used for data access and each query is only responded by data source.

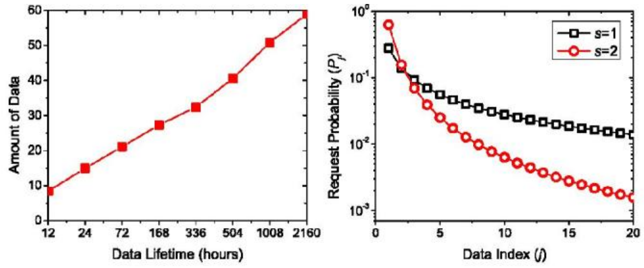
. Random Cache, in which every requester caches the received data to facilitate data access in the future.

. CacheData [35], which is proposed for cooperative caching in wireless ad hoc networks, and lets each relay in DTNs cache the pass-by data based on their popularity.

Bundle Cache [27], which packs network data as bundles and makes caching decision on pass-by data by considering the node contact pattern in DTNs, so as to minimize the average data access delay.

Cache replacement algorithms are proposed in Cache-Data and Bundle Cache, and will also be used in our evaluations. For Random Cache, LRU is used for cache replacement. The following metrics are used for evaluations. Each simulation is repeated multiple times with randomly generated data and queries for statistical convergence:

. Successful ratio, the ratio of queries being satisfied with the requested data. This ratio evaluates the coverage of data access provided by our proposed caching schemes.



(a) Amount of network data (b) Data request probabilities

Fig. 2. Experiment setup.

- . Data access delay, the average delay for getting responses to queries.

- . Caching overhead, the average number of data copies being cached in the network.

5.1Experiment Setup

Our performance evaluations are performed on the In-focom06 and MIT Reality traces. In all the experiments, central nodes representing NCLs are globally selected before data and queries are generated. The first half of the trace is used as warm-up period for the accumulation of network information and subsequent NCL selection, and all the data and queries are generated during the second half of trace.

5.2 Data Generation

Each node periodically checks whether it has generated data which has not expired yet. If not, the node determines whether to generate new data with probability p_G. Each generated data have finite lifetime uniformly distributed in range ½0:5T; 1:5T &, and the period for data generation decision is also set as T . In our evaluations, we fix p_G ¼ 0:2, and the amount of data in the network is, hence, controlled by T , as illustrated in Fig. 12a for the MIT Reality trace. Similarly, data size is uniformly distributed in range ½0:5s_{avg}; 1:5s_{avg}&, and caching buffers of nodes are uniformly distributed in range ½200; 600 Mb&. s_{avg} is adjusted to simulate different node buffer conditions.

Note that in this section, we compare the performance of our proposed schemes with the existing work. When T is large, indicating long intercontact time among mobile nodes in the network, our experimental setup increases the data lifetime accordingly. In this way, we ensure nonnegligible caching performance in the network and furthermore comprehensive performance comparisons. We could rea-sonably infer that the comparison results we have in this section will still hold, when the average intercontact time in the network is reduced and enables efficient access on data with shorter lifetime.

6.CONCLUSION:

In this paper, we propose a novel scheme to support cooperative caching in DTNs. Our basic idea is to intentionally cache data at a set of NCLs, which can be easily accessed by other nodes. We ensure appropriate NCL selection based on a probabilistic metric; our approach coordinates caching nodes to optimize the tradeoff between data accessibility and caching overhead. Extensive simula-tions show that our scheme greatly improves the ratio of queries satisfied and reduces data access delay, when being compared with existing schemes.

REFERENCES

[1] J. Burgess, B. Gallagher, D. Jensen, and B. Levine, “MaxProp: Routing for Vehicle-Based Disruption-Tolerant Networks,” Proc. IEEE INFOCOM, 2006.

[2] 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.

[3] P. Cao and S. Irani, “Cost-Aware WWW Proxy Caching Algorithms,” Proc. USENIX Symp. Internet Technologies and Systems, 1997.

[4] 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.

[5] 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.

[6] E. Daly and M. Haahr, “Social Network Analysis for Routing in Disconnected Delay-Tolerant MANETs,” Proc. ACM MobiHoc, 2007.

[7] 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.

[8] 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.

[9] 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.

[10] V. Erramilli, A. Chaintreau, M. Crovella, and C. Diot, “Delegation Forwarding,” Proc. ACM MobiHoc, 2008.

[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] K. Fall, “A Delay-Tolerant Network Architecture for Challenged Internets,” Proc. ACM SIGCOMM Conf. Applications, Technologies, Architectures, and Protocols for Computer Comm., pp. 27-34, 2003.

[15] L. Fan, P. Cao, J. Almeida, and A. Broder, “Summary Cache: A Scalable Wide-Area Web Cache Sharing Protocol,” IEEE/ACM Trans. Networking, vol. 8, no. 3, pp. 281-293, June 2000.

[16] M. Fiore, F. Mininni, C. Casetti, and C.F. Chiasserini, “To Cache or Not to Cache?” Proc. IEEE INFOCOM, pp. 235-243, 2009.

[17] 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.

[18] W. Gao and G. Cao, “User-Centric Data Dissemination in Disruption Tolerant Networks,” Proc. IEEE INFOCOM, 2011.

[19] W. Gao, G. Cao, A. Iyengar, and M. Srivatsa, “Supporting Cooperative Caching in Disruption Tolerant Networks,” Proc. Int’l Conf. Distributed Computing Systems (ICDCS), 2011.

[20] W. Gao, Q. Li, B. Zhao, and G. Cao, “Multicasting in Delay Tolerant Networks: A Social Network Perspective,” Proc. ACM MobiHoc, pp. 299-308, 2009.

[21] Y. Huang, Y. Gao, K. Nahrstedt, and W. He, “Optimizing File Retrieval in Delay-Tolerant Content Distribution Community,” Proc. IEEE Int’l Conf. Distributed Computing Systems (ICDCS), pp. 308-316, 2009.

[22] P. Hui, J. Crowcroft, and E. Yoneki, “Bubble Rap: Social-Based Forwarding in Delay Tolerant Networks,” Proc. ACM MobiHoc, 2008.

[23] S. Ioannidis, L. Massoulie, and A. Chaintreau, “Distributed Caching over Heterogeneous Mobile Networks,” Proc. ACM SIGMETRICS Int’l Conf. Measurement and Modeling of Computer Systems, pp. 311-322, 2010.

[24] V. Lenders, G. Karlsson, and M. May, “Wireless Ad Hoc Podcasting,” Proc. IEEE Fourth Ann. Comm. Soc. Conf. Sensor, Mesh and Ad Hoc Comm. and Networks (SECON), pp. 273-283, 2007.

[25] F. Li and J. Wu, “MOPS: Providing Content-Based Service in Disruption-Tolerant Networks,” Proc. Int’l Conf. Distributed Computing Systems (ICDCS), pp. 526-533, 2009.

[26] S. Martello and P. Toth, Knapsack Problems: Algorithms and Computer Implementations. John Wiley & Sons, 1990.

[27] M.J. Pitkanen and J. Ott, “Redundancy and Distributed Caching in Mobile DTNs,” Proc. ACM/IEEE Second Workshop Mobility in the Evolving Internet Architecture (MobiArch), 2007.

[28] I. Psaras, L. Wood, and R. Tafazolli, “Delay-/Disruption-Tolerant Networking: State of the Art and Future Challenges,” technical report, Univ. of Surrey, 2010.

[29] 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.

[30] S.M. Ross, Introduction to Probability Models. Academic, 2006.

[31] T. Spyropoulos, K. Psounis, and C. Raghavendra, "Spray and Wait: An Efficient Routing Scheme for Intermittently Connected Mobile Networks," Proc. ACM SIGCOMM Workshop Delay-Tolerant Networking, pp. 252-259, 2005.

[32] T. Spyropoulos, K. Psounis, and C. Raghavendra, "Efficient Routing in Intermittently Connected Mobile Networks: The Single-Copy Case," IEEE/ACM Trans. Networking, vol. 16, no. 1,
pp. 63-76, Feb. 2008.

[33] B. Tang, H. Gupta, and S.R. Das, "Benefit-Based Data Caching in Ad Hoc Networks," IEEE Trans. Mobile Computing, vol. 7, no. 3,
pp. 289-304, Mar. 2008.

[34] A. Vahdat and D. Becker, "Epidemic Routing for Partially Connected Ad Hoc Networks," Technical Report CS-200006, Duke Univ., 2000.

[35] L. Yin and G. Cao, "Supporting Cooperative Caching in Ad Hoc Networks," IEEE Trans. Mobile Computing, vol. 5, no. 1, pp. 77-89, Jan. 2006.

[36] E. Yoneki, P. Hui, S. Chan, and J. Crowcroft, "A Socio-Aware Overlay for Publish/Subscribe Communication in Delay Tolerant Networks," Proc. 10th ACM Symp. Modeling, Analysis, and Simulation of Wireless and Mobile Systems (MSWiM), pp. 225-234, 2007.

[37] Q. Yuan, I. Cardei, and J. Wu, "Predict and Relay: An Efficient Routing in Disruption-Tolerant Networks," Proc. ACM MobiHoc,
pp. 95-104, 2009.

[38] J. Zhao, P. Zhang, G. Cao, and C. Das, "Cooperative Caching in Wireless P2P Networks: Design, Implementation, and Evalua-tion," IEEE Trans. Parallel & Distributed Systems, vol. 21, no. 2,
pp. 229-241, Feb. 2010.

[39] 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.

Guide Profile:

Mr.N.Venkatadri Was Born In Andra Pradesh, India. He is Working as Asso.Prof.,M.TechDeptment of CSE, SKR College Of Engineering and Technology, KonduruSatram, manubolu,Nellore.

Student Profile:

R.AnilKumar.was born in AndraPradesh,India. He recevivedB.tech Degree from SVUCETirupati, Chittoor(DT). I am pursuing M.Tech Degree in CSE from JNTU Ananthapur.