© 2014 IJIRT | Volume 1 Issue 10 | ISSN: 2349-6002

Cooperative Caching Using Dynamic Nodes for Efficient Data Access in Disruption Tolerant Network

K.Priyadharshini, Neethu Mariam Joseph,
Assistant Professor/IT, *Karpagam College of Engineering*

Abstract **- Disruption Tolerant Networks (DTNs) are characterized by unpredictable node mobility, low node density and lack of global network information. Most of research efforts in DTNs focus on data forwarding; only limited work has been made on providing efficient data access to mobile users. Here we propose a novel approach to support cooperative caching in DTNs, which provides the sharing and coordination of cached data among multiple nodes and reduces data access delay. Our basic design is to intentionally cache data at a set of Network Central Locations (NCLs), which can be easily accessed by other nodes in network. An efficient scheme which ensures appropriate NCL selection based on a probabilistic selection metric and coordinates multiple caching nodes to optimize the tradeoff between caching overhead and data accessibility. Extensive trace-driven simulations show that our approach significantly improves data access performance compared to existing schemes**.

Index Terms—**Disruption Tolerant Networks, Cooperative Caching, Network Central Locations, Data Access, Cache Replacement**

I. INTRODUCTION

Disruption Tolerant Networks (DTNs) consist of mobile devices which contact each other opportunistically. Because of low node density and unpredictable node mobility, in DTNs only intermittent network connectivity exists, and the subsequent difficulty of maintaining end- to-end communication links makes it necessary to use "carry-and-forward" methods for data transmission. Examples of these networks include groups of individuals moving in disaster recovery areas, military battled, or urban sensing applications. In those networks, node mobility is exploited to let mobile nodes carry data as relays and forward data opportunistically when contacting others user. The key problem is to determine the appropriate relay selection strategy.

Although forwarding schemes have been proposed in DTNs, there is inadequate research on providing efficient data access to mobile users, regardless of the importance of data accessibility in many mobile applications. For example, it is popular that Smartphone users can end interesting digital content from their nearby peers. In Vehicular Ad-hoc Networks (VANETs), the availability of live traffic information will be beneficial for vehicles to avoid traffic delays.

In these applications, data is only requested by mobile users at any time when they need, and requester is unaware of data locations in advance. The destination of data is unknown when data is generated. This communication paradigm differs from publish/ subscribe systems, in which data is forwarded by broker nodes to users according to their data subscriptions. Suitable network design is needed to ensure that data can be promptly accessed by requesters in such cases.

II. RELATED WORK

Research on data forwarding in DTNs originates from Epidemic routing which Foods the complete network. Some schemes do such prediction based on their mobility patterns, which are characterized by Kalman elter or semi-Markov chains. In 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 property of node contact patterns, such as the community structures and centrality, have also been also exploited for relay selection in recent social-based data forwarding schemes [9].

Data is intentionally cached at appropriate network locations with generic data and query models, 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.

III. DYNAMIC NODE SELECTION

A general caching scenario, in which each node may generate data with a globally unique identifier and finite lifetime, and may also request other data by sending queries with a finite time constraint. The cellular 3G infrastructure may usually be unavailable, or provide too limited bandwidth to transmit data traffic. Instead, mobile users rely on their opportunistic contacts for accessing data, which can be live weather report, traffic condition, or government rescue plan. Similarly, a platoon of soldiers in the military

battlefield may lose the satellite connection due to enemy attacks, and have to distribute battle reports via their opportunistic contacts. Either a data item or a query is described by a set of keywords over a keyword space, so that caching nodes can determine the appropriate data that a user is interested in. Data requesters are randomly distributed in the network. Focusing on efficiently utilizing the available node buffer to optimize the overall caching performance, which is measured by the successful ratio and delay for mobile users to access different data items.

Nodes in DTNs are well motivated to contribute their local resources for caching data, because the cached data provides prompt data access to the caching nodes themselves. Since the central nodes representing NCLs are prioritized for caching data, the closer a requester is to a central node, the sooner its queries are responded by the corresponding NCL.

A) NCL Selection Metric

i) Trace-based Validation

The practical applicability of the fore mentioned NCL selection metric is based on the heterogeneity of node contact patterns, such that nodes in DTNs differ in their popularity and few nodes contact many others frequently. These traces record contacts among users carrying mobile devices in conference sites and university campuses. The mobile devices, including Mica2 sensors or smart phones, are distributed to users participating into the experiment. Devices equipped with Bluetooth interface periodically detect their peers nearby, and a contact record is made when two devices move close to each other. Devices equipped with Wi-Fi interface search for nearby Wi-Fi Access Points (APs) and associate themselves to APs with the best signal strength. A record of contact is done when two devices are associated to the same AP. The detected contacts are recorded in the local storage of mobile devices. After the experiment ends, these devices are called back so that the recorded contacts are processed and analyzed.

Calculate pair wise node contact rates based on their cumulative contacts during the entire trace. Inappropriate values of make Ci close to 0 or 1. Instead, values of Tare adaptively determined in different traces to ensure the differentiation of NCL selection metric values of nodes. T is set as 1 hour for the two Info com traces, 1 week for the MIT Reality trace, and 3 days for the UCSD trace.

ii) Global Selection

When global network knowledge about the shortest opportunistic paths among mobile nodes and pair wise node contact rates are available, central nodes representing NCLs can be selected sequentially by the network administrator before data access.

Let Ncl denote the set of selected central nodes; every time the node with the highest metric value is selected as the next central node, until the required K central nodes are selected.

By doing so we ensure that the selected central nodes will not be clustered on the network contact graph. The parameter is determined by the average node contact frequency. This parameter is generally trace-dependent for different DTN traces. A network warm-up period is reserved for nodes to collect information and calculate their pair wise contact rates, and central nodes are selected after the warm-up period ends. Data is unable to be intentionally cached at the NCLs during the warm-up period. Instead, data is incidentally cached by nodes in the network, for data access. In particular, each requester sends queries directly to the data source and caches the received data locally for responding other pass-by queries in the future.

Each node is only notified about the identifiers of central nodes, this notification is cost-effective without producing noticeable communication overhead, even in cases where the central nodes frequently change. Note that the central nodes are selected due to their popularity in the network, rather than their storage or computation capabilities. Therefore, we assume that the central nodes have similar capabilities in computation, data transmission and storage with other nodes in DTNs.

B) Caching Scheme

The idea is to intentionally cache data at a set of NCLs which can be promptly accessed by other nodes. The scheme consists of the following two components:

- When a data source generates data, it pushes the data to central nodes of NCLs which are prioritized to cache data. One copy of data is cached at each NCL node. If the caching buffer of a central node is full, another node nearer to the central node will cache the data. Decisions are automatically made based on buffer conditions of nodes involved in the pushing process.
- Requester multicasts a query to central nodes of NCLs to pull the data, and a central node forwards the query to the caching nodes. Number of data copies is returned to the requester.

i) Caching Location

Whenever a node generates new data, *S* pushes the data to NCLs by sending a data copy to each central node representing a NCL. The opportunistic path weight to the central node as relay selection metric for such data forwarding and a relay forwards data to another node with a higher metric than itself. This "Compare-and-Forward" strategy has been widely used in the literature for efficient data forwarding. On opportunistic path, this strategy probabilistically ensures that each forwarding reduces the remaining delay for data to be delivered to the central node. For newly generated data, the initial caching locations are automatically determined during the forwarding process based on node buffer conditions. The caching locations are then dynamically adjusted by cache replacement according to query history. In general, data is forwarded to and cached at central nodes. Here the forwarding process only stops when the caching buffer of the next relay is full, and data is cached at the current relay in such cases.

ii) Cache Replacement

For each data item in the network, the locations where it is cached are dynamically adjusted via cache replacement. This replacement is based on data popularity, and generally places popular data nearer to the central nodes of NCLs. Traditional cache replacement strategies such as LRU, which removes the least-recently used data from cache when new data is available, are ineffective due to its over-simplistic consideration of data popularity. Greedy-Dual-Size calculates data utility by considering data popularity and size simultaneously, but cannot ensure optimal selection of cached data.

Improving previous work by proposing a probabilistic cache replacement strategy, which appropriately selects the data to be cached and heuristically balances between the cumulative data accessibility and access delay.

C) NCL Load Balancing

First, the central nodes cache the most popular data in the network and respond to the frequent queries for these data. Second, the central nodes are also responsible for broadcasting all the queries they receive to other caching nodes nearby. However, such functionality may quickly consume the local resources of central nodes which include their battery life and local memory. In addition, would like the caching schemes to be resilient to failures of central nodes. In this section, we focus on addressing this challenge, and propose methods which efficiently migrate the functionality of central nodes to other nodes in cases of failures or resource depletion. In general, the methods we present in this section can be used to adjust the deployment of central nodes at run-time, such as adding or removing central nodes according to up-to-date requirements on caching performance.

i) Adjustment of Caching Locations

After a new central node is selected, the data cached at the NCL represented by the original central node needs to be adjusted correspondingly, so as to optimize the caching performance. This movement is achieved via cache replacement when caching nodes opportunistically contact each other. Each caching node at the original NCL re-calculates the utilities of its cached data items with respect to the newly selected central node. In general, these data utilities will be reduced due to the changes of central nodes, and this reduction moves the cached data to the appropriate caching locations which are nearer to the newly selected central node.

Evaluate the caching scheme performance by comparing it with the following schemes:

- **No Cache**- where no caching is used for data access and each query are only responded by data source.
- **Random Cache**-Every requester caches the received data to facilitate data access in the future.
- **Cache Data**-which is implemented 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**-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 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.
- **Data access delay**- average delay for getting responses to queries.
- **Caching overhead**- average number of data copies being cached in the network

ii) Effectiveness of Cache Replacement

The cache replacement strategy is compared with the traditional replacement strategies including FIFO and LRU. It is also compared with Greedy-Dual-Size which is widely used in web caching. Use MIT Reality trace for such evaluation. FIFO and LRU leads to poor data access performance due to improper consideration of data popularity. Data access delay of FIFO and LRU also becomes much longer when increases. Greedy-Dual-Size performs better than FIFO and LRU due to consideration of size and data popularity, but it is unable to ensure optimal cache replacement decision.

D) Result Analysis

The performance evaluations are performed on the Incofom06 and MID Reality traces, In all the experiments, central nodes representing NCL are globally selected before data and queries are generated. The nearest 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.

For each region created one node is selected as an Network Central Location. The node with high priority is selected as an NCL node. The NCL node is used to send the requested data to all other nodes in the region.

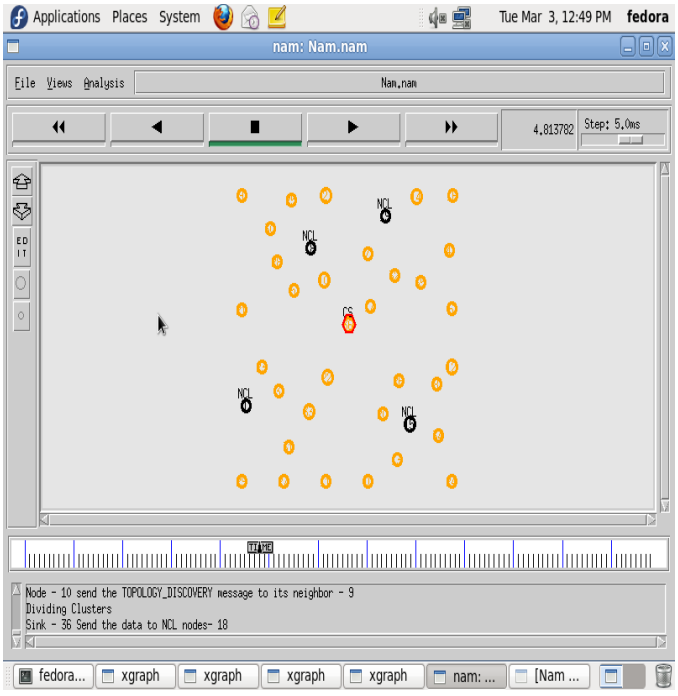

Figure 1.1 NCL Selections

Each region is identified by the colour given to it. The new NCL is selected due to the lower energy of current NCL.The NCL which lost its energy is identified by the dot (.) attached to the node.

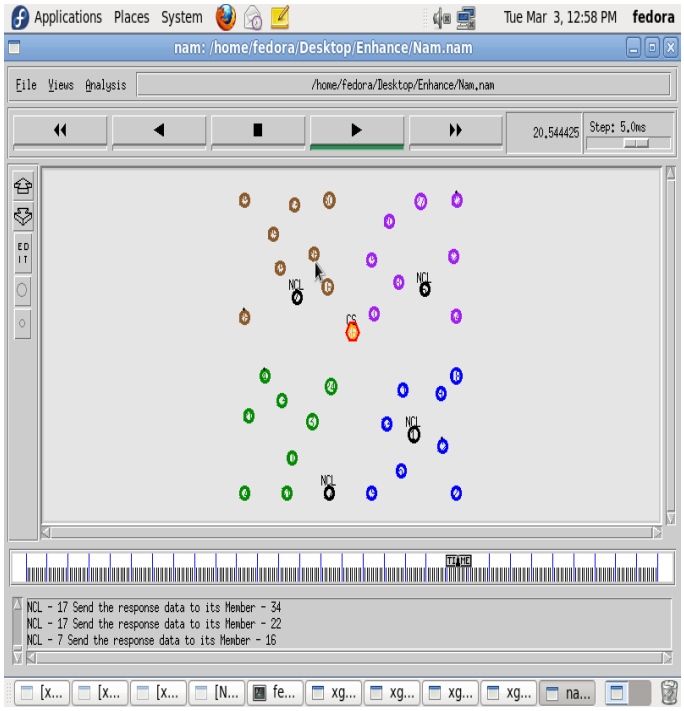

Figure 1.2 Changing NCL nodes

The newly selected Network Central Location node starts passing the requested data to all other nodes in the region. Likewise when the current NNCL loses its energy another NCL is selected dynamically and the process is continued.

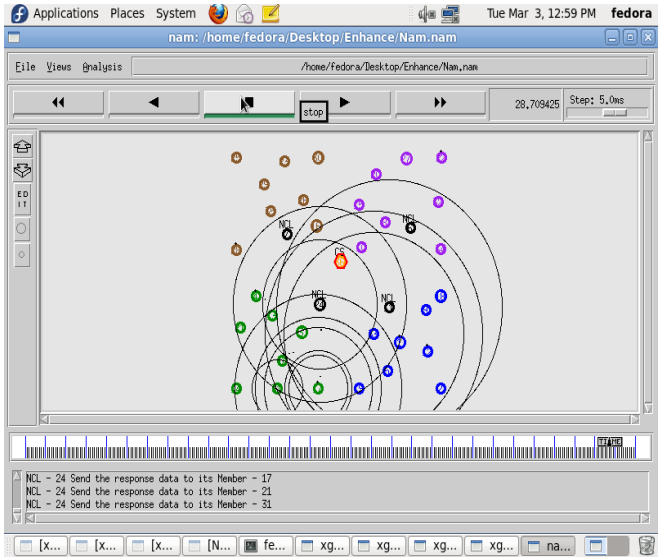

Figure 1.3 Message Passing

IV. CONCLUSION

The presence of different types of meta-data allows us to think in terms of a meta-data space, where coordinates are combinations of space, time, and tag values. From this point of view, a message with its associated metadata can be considered as "covering" parts of the meta-data space. This abstraction is important because it allows us to make decisions about priorities in message forwarding based on characteristics of a message's content and its ability to cover more or less of the meta-data space. For example, a photo of a geographic area for which no data has yet been received could be considered more valuable (and thus worthy of higher delivery priority, if authentic) than a message containing a photo of locations for which data has already been received. Messages that arrive at the cluster are stored in a message database.

Cooperative caching in Disruption Tolerant Network provides quick data access to mobile users. It reduces data delay efficiently and it takes important step towards fast data access with complete avoidance of data delay in the future.

REFERENCES

[1] Fall, K. (2003). A Delay-Tolerant Network Architecture for Challenged Internets.In Proceedings of ACM SIGCOMM 2003.

[2] Seth, A., Kroeker, D., Zaharia, M., Guo, S., and Keshav, S. (2006). Low-cost Communication for Rural Internet Kiosks Using Mechanical Backhaul.In Proceedings of ACM MOBICOM 2006.

[3] Chaintreau, A., Hui, P., Crowcroft, J., Diot, C., Gass, R., Scott, J. (2006). Impact of Human Mobility on the Design of Opportunistic Forwarding Algorithms.In Proceedings of INFOCOM 2006.

[4] Palen, L. and Liu, S. (2007). Citizen Communications in Crisis: Anticipating a Future of ICT-Supported Participation, In Proceedings of ACM CHI 2007.

[5] Palen, L., Hiltz, S., and Liu, S. (2007). Online Forums Supporting Grassroots Participation in Emergency Preparedness and Response. Communications of the ACM, 50 (3) (Mar. 2007): 54-58.

[6] Liu, S., Palen, L., Sutton, J., Hughes, A., Vieweg, S. (2008). In Search of the Bigger Picture: The Emergent Role of On-Line Photo-Sharing in Times of Disaster. In Proceedings of ISCRAM 2008.

[7] Vieweg, S., Palen, L., Liu, S., Hughes, A., Sutton, J. (2008). Collective Intelligence in Disaster: An Examination of the Phenomenon in the Aftermath of the 2007 Virginia Tech Shootings. In Proceedings of ISCRAM 2008.

[8] Hughes, A. and Palen, L. (2009). Twitter Adoption and Use in Mass Convergence and Emergency Events. In Proceedings of ISCRAM 2009.

[9] Hughes, A., Palen, L., Sutton, J., Liu, S., and Vieweg, S. (2008) "Site-Seeing" in Disaster: An Examination of On-Line Social Convergence. In Proceedings of ISCRAM 2008.

[10] H. Dubois-Ferriere, M. Grossglauser, and M. Vetterli Age matters: efficient route discovery in mobile ad hoc networks using encounter ages. Proc. MobiHoc, pages 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. In Proceeding of MobiSys. ACM, 2008.

[12] W. Gao, G. Cao, A. Iyengar, and M. Srivatsa. Supporting cooperative caching in disruption tolerant networks. In Proceedings of Int'l Conf. on Distributed Computing Systems (ICDCS), 2011.

Priyadharshini.K M.E in Computer Science and Engineering from Anna University, 2013 and B.E in Information Technology from Avinashilingam University, 2011. The area of interests includes Computer networks and security.

She is working as Assistant Professor in the department of Information Technology, Karpagam College of Engineering. She worked as Assistant Professor in the department of Computer Science and Engineering, Karpagam Institute of Technology.

Ms. Priyadharshini is a member in professional societies like IAENG and theRED.

Neethu Mariam Josephholds M.Tech in Network and Internet Engineering from Karunya University, 2013 and B.Tech in Information Technology from Anna University, 2011. The area of interests includes Cloud Computing, Computer networks and security.

She is working as Assistant Professor in the department of Information Technology, Karpagam College of Engineering. She worked as Assistant Professor in the department of Computer Science and Engineering, Karpagam University.

Ms. Neethu isa member in professional societies like IAENG and theRED.