International Journal of Latest Trends in Engineering and Technology
IJLTET Special Issue- ICRACSC-2016 , pp.050-055
e-ISSN:2278-621X

FILE REDUPLICATION PROTOCOL IN ADHOC NETWORKS FOR SHARING DOCUMENTS IN P2P

AAruna Jyothi and M.L.Prasanthi

Abstract: File sharing application in mobile adhoc networks (MANETS) have attracted more and more attention in recent years. The efficiency of file querying suffers from the distinctive properties of such networks including node mobility and limited communication range and resource and intuitive method to alleviate the problem is to create file replicas in the network specifically , current file replication protocols in mobile adhoc networks have two short comings. First, they lack a rule to allocate limited resources to different files in order to minimize the average querying delay. Second, they simply consider storage as available resources for replicas. Actually a node that has a higher frequency with others provides higher availability to its files. In this paper we introduce a new concept of resource for file replication[2] which considers both nodes storage and meeting frequency. We theoretically study the influence of resource allocation on the average querying delay and derive resource allocation rule to minimize the average querying delay. We further propose distributed file replication protocol to realize the proposed rule.
Keywords: MANET, Querying delay, Distributed file replication, OFRR rule.

I. INTRODUCTION

With the increasing popularity of mobile devices, e.g, smartphones and laptops, we envision the future of MANETs consisted of these mobile devices. By MANETs, we refer to both normal MANETs and disconnected MANETs, also known as delay tolerant networks (DTNs). The former has a relatively dense node distribution in an area while the latter has sparsely distributed nodes that meet each other opportunistically. On the other side, the emerging of mobile file sharing applications motivates the investigation on the peer-to-peer(P2P)[3] file sharing over such MANETs. The local P2P file sharing model provides three advantages. First, it enables file sharing when no base stations are available (e.g., in rural areas). Second, with the P2P architecture, the bottleneck on overloaded servers in current client-server based file sharing systems can be avoided. Third, it exploits otherwise wasted peer to peer communication opportunities among mobile nodes. As a result, nodes can freely and unobtrusively access and share files in the distributed MANET environment, which can possibly support interesting applications. For example, mobile nodes can share files based on users' proximity in the same building or in a local community. Tourists can share their travel protocols, each individual node replicates files it frequently queries or a group of nodes create one replica for each file they frequently query.

Peer-to-peer Architecture

A peer-to-peer network is designed around the notion of equal peer nodes simultaneously functioning as both "clients" and "servers" to the other nodes on the network. This model of network arrangement differs from the client–server model where communication is usually to and from a central server. A typical example of a file transfer that uses the client-server model is the File Transfer Protocol (FTP) service in which the client and server programs are distinct: the clients initiate the transfer, and the servers satisfy these requests.

Unstructured networks

Overlay network diagram for an unstructured P2P network, illustrating the ad hoc nature of the connections between nodes. Unstructured peer-to-peer networks do not impose a particular structure on the overlay network by design, but rather are formed by nodes that randomly form connections to each other.(Gnutella, Gossip, and Kazaa are examples of unstructured P2P protocols).Because there is no structure globally imposed upon them, unstructured networks are easy to build and allow for localized optimizations to different regions of the overlay. Also, because the role of all peers in the network is the same, unstructured networks are highly robust in the face of high rates of "churn"—that is, when large numbers of peers are frequently joining and leaving the network.

However the primary limitations of unstructured networks also arise from this lack of structure. In particular, when a peer wants to find a desired piece of data in the network, the search query must be flooded through the network to find as many peers as possible that share the data. Flooding causes a very high amount of signaling traffic in the network, uses more CPU/memory (by requiring every peer to process all search queries), and does not ensure that search queries will always be resolved.

Furthermore, since there is no correlation between a peer and the content managed by it, there is no guarantee that flooding will find a peer that has the desired data. Popular content is likely to be available at several peers and any peer searching for it is likely to find the same thing. But if a peer is looking for rare data shared by only a few other peers, then it is highly unlikely that search will be successful.

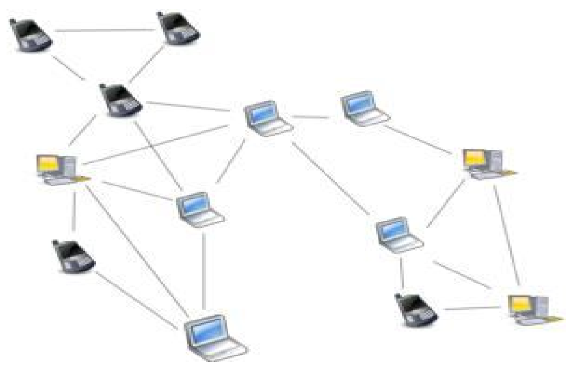

Fig1.1:Unstructured P2P

Structured networks

Overlay network diagram for a structured P2P network, using a distributed hash table (DHT) to identify and locate nodes/resourcesIn structured peer-to-peer networks the overlay is organized into a specific topology, and the protocol ensures that any node can efficiently search the network for a file/resource, even if the resource is extremely rare.This enables peers to search for resources on the network using a hash table: that is, (key, value) pairs are stored in the DHT, and any participating node can efficiently retrieve the value associated with a given key.
However, in order to route traffic efficiently through the network, nodes in a structured overlay must maintain lists of neighbors that satisfy specific criteria. This makes them less robust in networks with a high rate .More recent evaluation of P2P resource discovery solutions under real workloads have pointed out several issues in DHT-based solutions such as high cost of advertising/discovering resources and static and dynamic load imbalance.

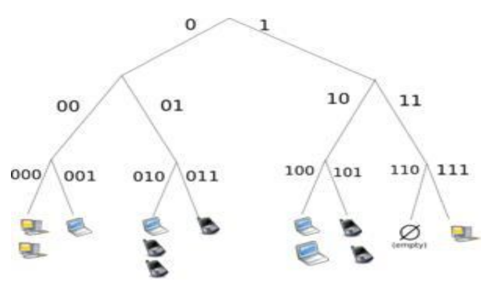

Fig 1.2:Structured p2p

EXISTING SYSTEM

- In the former, redundant replicas are easily created in the system, thereby wasting resources. In the latter, though redundant replicas are reduced by group based cooperation, neighboring nodes may separate from each other due to node mobility, leading to large query delay.

- There are also some works addressing content caching in disconnected MANETs/ DTNs for efficient data retrieval or message routing. They basically cache data that are frequently queried on places that are visited frequently by mobile nodes. Both the two categories of replication methods fail to thoroughly consider that a node's mobility affects the availability of its files.

DISADVANTAGES OF EXISTING SYSTEM

- Node mobility, limited communication range and resource, have rendered many difficulties in realizing such a P2P file sharing system.

- Broadcasting can quickly discover files, but it leads to the broadcast storm problem with high energy consumption.

- In spite of efforts, current file replication protocols lack a rule to allocate limited resources to files for replica creation in order to achieve the minimum average querying delay, i.e., global search efficiency optimization under limited resources.

They simply consider storage as the resource for replicas, but neglect that a node's frequency to meet other nodes (meeting ability in short) also influences the availability of its files. Files in a node with a higher meeting ability have higher availability.

II. PROBLEM DEFINITION

- In this paper, we introduce a new concept of resource for file replication, which considers both node storage and node meeting ability. We theoretically study the influence of resource allocation on the average querying delay and derive an optimal file replication rule (OFRR) that allocates resources to each file based on its popularity and size. We then propose a file replication protocol based on the rule, which approximates the minimum global querying delay in a fully distributed manner.

- We propose a distributed file replication protocol that can approximately realize the optimal file replication rule with the two mobility models in a distributed manner.

- Our experiment and simulation results show the superior performance of the proposed protocol in comparison with other representative replication protocols.

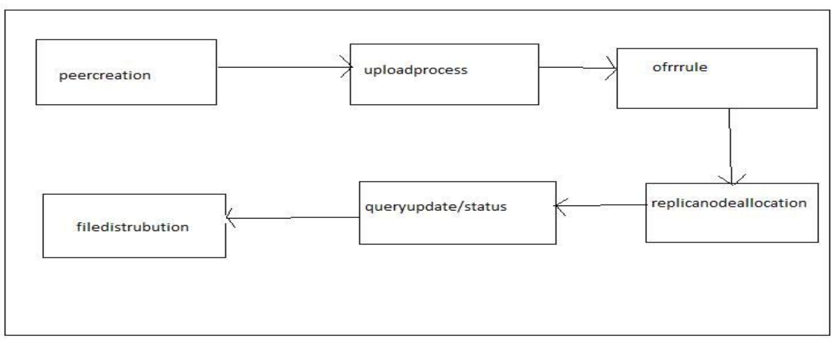

Fig 2 : Block Diagram

III. PROPOSED ALGORITHM

Proposed algorithm is a framework of four step process. Steps are discussed below

Step 1: Optimal File Replication with the RWP Model
Step 2: Community-Based Mobility Model
Step 3: Meeting Ability Distribution
Step 4: Design of the File Replication Protocol

IV. IMPLEMENTATION AND RESULTS MODULES

- Optimal File Replication with the RWP Model
- Community-Based Mobility Model
- Meeting Ability Distribution
- Design of the File Replication Protocol

MODULES DESCSRIPTION

➢ **Optimal File Replication with the RWP Model**

- In the RWP model, we can assume that the inter-meeting time among nodes follows exponential distribution. Then, the probability of meeting a node is independent with the previous encountered node.

- Therefore, we define the meeting ability of a node as the average number of nodes it meets in a unit time and use it to investigate the optimal file replication. Specifically, if a node is able to meet more nodes, it has higher probability of being encountered by other nodes later on.

- A node's probability of being encountered by other nodes is proportional to the meeting ability of the node. This indicates that files residing in nodes with higher meeting ability have higher availability than files in nodes with lower meeting ability. So we take into account both meeting ability and storage in measuring a node's resource.

- When a replica is created on a node, it occupies the memory on the node. Also, its probability of being met by others is decided by the node's meeting ability. This means that the replica naturally consumes both the storage resource and the meeting ability resource of the node.

➢ **Community-Based Mobility Model**

- In this module, we conduct the analysis under the community-based mobility model. We consider each node's satisfying ability. It is defined as a node's ability to satisfy queries in the system and is calculated based on the node's capacity to satisfy queries in each community.

- In this model, since nodes' file interests are stable during a certain time period, we assume that each node's file querying pattern (i.e., querying rates for different files) remains stable in the considered period of time. Then, the number of nodes in a community represents the number of queries for a given file generated in this community. As a result, a file holder has low ability to satisfy queries from a small community.

- Thus, we integrate each community's fraction of nodes into the calculation of the satisfying ability.

➢ **Meeting Ability Distribution**

- We measured the meeting ability distribution from real traces to confirm the necessity to consider node meeting ability as an important factor in the resource allocation in our design.

- For each trace, we measured the meeting abilities of all nodes and ranked them in decreasing order. We see that in all traces, node meeting ability is distributed in a wide range. This matches with our previous claim that nodes usually have different meeting abilities. Also, it verifies the necessity of considering node meeting ability as a resource in file replication since if all nodes have similar meeting ability, replicas on different nodes have similar probability to meet requesters, and hence there is no need to consider meeting ability in resource allocation.

➢ **Design of the File Replication Protocol**

- We propose the priority competition and split file replication protocol (PCS). We first introduce how a node retrieves the parameters needed in PCS and then present the detail of PCS.

- In PCS, each node dynamically updates its meeting ability and the average meeting ability of all nodes in the system. Such information is exchanged among neighbor nodes.

- We introduce the process of the replication of a file in PCS. Based on OFRR, since a file with a higher P should receive more resources, a node should assign higher priority to its files with higher P to compete resource with other nodes. Thus, each node orders all of its files in descending order of their Ps and creates replicas for the files in a top-down manner periodically.

- The file replication stops when the communication session of the two involved nodes ends. Then, each node continues the replication process for its files after excluding the disconnected node from the neighbor node list. Since file popularity, Ps, and available system resources change as time goes on, each node periodically executes PCS to dynamically handle these time-varying factors. Each node also periodically calculates the popularity of its files (qj) to reflect the changes on file popularity (due to node querying pattern and rate changes) in different time periods. The periodical file popularity update can automatically handle file dynamism.

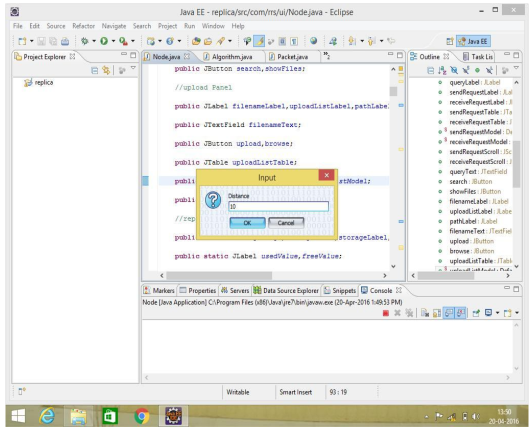

Fig 4.1 :Creation Of Node

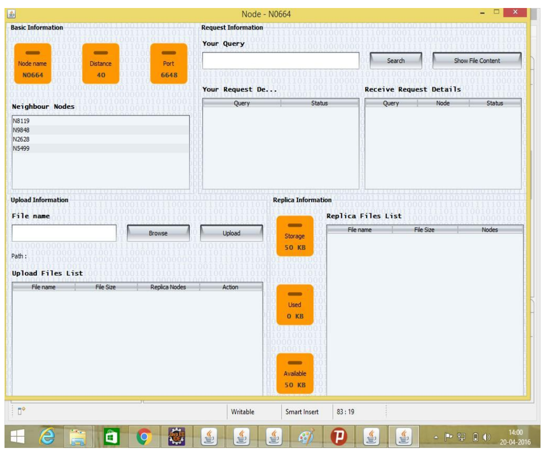

Fig 4.2: Minimum number of Nodes

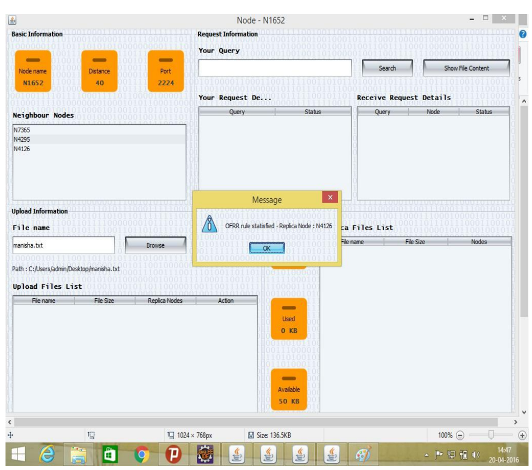

Fig 4.3: OFRR Rule

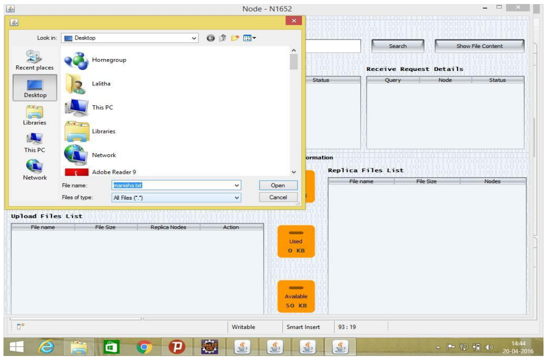

Fig 4.4: File Upload

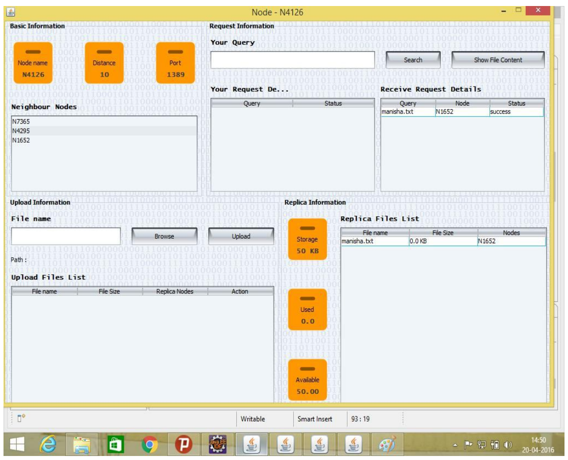

Fig 4.5: Replica Files List

V.CONCLUSION

We investigated the problem of how to allocate limited resources for file replication for the purpose of global optimal file searching efficiency in MANETs. Unlike previous protocols that only consider storage as resources, we also consider file holder's ability to meet nodes as available resources since it also affects the availability of files on the node.We first theoretically analyzed the influence of replica distribution on the average querying delay under constrained available resources with two mobility models, and then derived an optimal replication rule that can allocate resources to file replicas with minimal average querying delay.

Finally, we designed the priority competition and split replication protocol (PCS) that realizes the optimal replication rule in a fully distributed manner. Extensive experiments on both GENI testbed, NS-2, and event-driven simulator with real traces and synthesized mobility confirm both the correctness of our theoretical analysis and the effectiveness of PCS in MANETs.

REFERENCES

[1] C. Palazzi and A. Bujari, "*A Delay/Disruption Tolerant Solution for Mobile to Mobile File Sharing*," Proc. IFIP/IEEE Wireless Days, 2010

[2] J. Zheng, J. Su, K. Yang, and Y. Wang, "*Stable Neighbor Based Adaptive Replica Allocation in Mobile Ad Hoc Networks*," Proc. Int'l Conf. Computational Science (ICCS),2004.

[3] Kang Chen, Student Member, IEEE and Haiying Shen, Senior Member, IEEE, "*Maximizing P2PFile Access Availability in Mobile Ad Hoc Networks though Replication for Efficient FileSharing*", IEEE TRANSACTIONS ON COMPUTERS, VOL. 64, NO. 4,APR ,2015.

[4] L. Kleinrock, Queueing Systems, Volume II*: Computer Applications*. John Wiley & Sons, 1976.

[5] T. Hara and S.K. Madria, "*Data Replication for Improving Data Accessibility in Ad Hoc Networks*," IEEE Trans. *Mobile Computing*, vol. 5, no. 11, pp. 1515-1532, Nov. 2006.

[6] 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.

[7] Y. Huang et al., "*Optimizing File Retrieval in Delay-Tolerant Content Distribution Community*," Proc. Int'l Conf. Distributed Computing Systems (ICDCS), 2009.