A Caching Mechanism for Improving Internet based Mobile Ad Hoc Networks Performance *

Sunho Lim† Seung-Taek Park† Wang-Chien Lee†
Guohong Cao† Chita R. Das† C. Lee Giles†‡

†Dept. of Computer Science and Engineering
The Pennsylvania State University
University Park, PA 16802 USA
{slim,separk,wlee,gcao,das}@cse.psu.edu

‡School of Information Sciences and Technology
The Pennsylvania State University
University Park, PA 16802 USA
giles@ist.psu.edu

ABSTRACT

Internet based mobile ad hoc networks (IMANETs) have several limitations to fulfill users' demands to access various kinds of information such as limited accessibility to the wired Internet, insufficient wireless bandwidth, and longer message latency. In this paper, we address the issues involved in information search and access in IMANET. A broadcast based *Simple Search (SS) algorithm* and an *aggregate caching* mechanism are proposed for improving the information accessibility and reducing average communication latency in IMANET. We evaluate the impact of caching, cache management, and access points, which are connected to the Internet, through extensive simulation. The simulation results indicate that the proposed aggregate cache can significantly improve an IMANET performance in terms of throughput and average number of hops to access data.

1. INTRODUCTION

Due to the growing interest in accessing the Internet, it is an important requirement to consider the integration of MANET with the Internet. Thus, to put the MANET technology into the context of real life, we consider an *Internet based* MANET, called IMANET [1], and investigate the problem of information search and access under this environment. An IMANET consists of a set of MTs. The MTs communicate with each other with ad hoc communication protocols. Among the MTs, some of them can directly connect to the Internet, and thus serve as AP for the rest of MTs in the IMANET. Thus, an AP is a gateway for the Internet and is assumed to have access to any information.

Although there may exist many potential applications, none of the previous work has addressed the constraints in IMANET. First, not all the MTs can access the Internet. Second, due to MTs' mobility, a set of MTs can be separated from the rest of the MTs and get disconnected from the Internet. Finally, an MT requiring multi-hop relay to access the Internet may incur a longer access latency than those which have direct access to the Internet. To address these constraints, we propose an *aggregate caching* mechanism for IMANET. The basic idea is that by storing data items in the local cache of the MTs, members of the IMANET can efficiently access the required information. Thus, the aggregated local caches of the MTs can be considered as an unified large cache for the IMANET. In addition, since information search in IMANET is different from the search engine based approach on the wired Internet, we propose a broadcast based approach, called *Simple Search (SS)* algorithm, which can be implemented on the top of existing routing protocols, to locate the requested data items. As part of the aggregate cache, a cache admission control policy and a cache replacement policy, called *Time and Distance Sensitive (TDS) replacement*, are developed to reduce the cache miss ratio and improve the information accessibility.

We conduct a simulation based performance evaluation to observe the impact of caching, cache management, and access points (which are directly connected to the Internet) upon the effectiveness of IMANET. The overall results show that the proposed methodology can relieve limitations of IMANET and improve system performance significantly.

*This research has been supported in part by NSF grants CCR-9900701, CCR-0098149, CCR-0208734, and EIA-0202007, and a grant from Ford Motor Co..

2. INFORMATION SEARCH ON IMANET

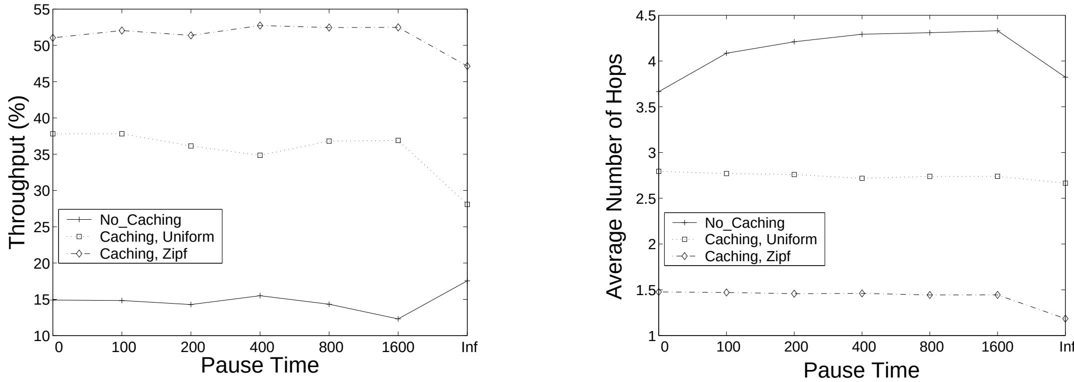
As for information access, the information from the Internet may be cached in some of the MTs within the IMANET. Moreover, any MT can be an information source. Without knowing the whereabout of information, a search algorithm is needed for IMANET as is done in the Internet. Since the concept of an aggregate cache is supported in the IMANET, data items can be received from local caches of the MTs as well as via an *access points*[1] (AP) connected to the Internet. When an MT needs a data item, it does not know exactly where to retrieve the data item from, so it broadcasts a request to all of the adjacent MTs. If an MT receives the request and has the data item in its local cache, it will send a reply to the requester to acknowledge that it has the data item; otherwise, it will forward the request to its neighbors.

Based on the idea described above, we propose an information search algorithm, called *Simple Search (SS)*, to determine an information access path to the MTs with cached data of the request or to appropriate APs. The decision is based on the arriving order of acknowledgments from the MTs or APs. Once the MT receives the requested data item, it triggers the cache admission control procedure to determine whether it should cache the data item.

3. AGGREGATE CACHE MANAGEMENT

In IMANET, caching data items in the local cache helps in reducing latency and increasing accessibility. If an MT is located along the path in which the request packet travels to an AP, and has the requested data item in its cache, then it can serve the request

[1]The access point here is a logical notation. An AP equipped with appropriate antennas can directly communicate with the Internet through wireless infrastructures including cellular base stations, and Low Earth Orbit (LEO) or geostationary (GEO) satellites.



(a) Throughput

(b) Latency

Figure 1: Throughput and latency as a function of pause time.

without forwarding it to the AP. In the absence of caching, all requests should be forwarded to the appropriate APs. Since the local caches of the MTs virtually form an aggregate cache, a decision as to whether to cache the data item depends not only on the MT itself, but also on the neighboring MTs. Therefore, we propose a cache admission control and a cache replacement algorithm.

Cache Admission Control: When an MT receives the requested data item, a cache admission control is triggered to decide whether it can cache this item. In this paper, the cache admission control allows an MT to cache a data item based on the distance of other APs or MTs, which have the requested data item. If the MT is located within Γ hops from them, then it does not cache the data item; Otherwise it caches the data item. The cached data items can be used by closely located MTs. Therefore, the same data items are cached at least Γ hops apart. Here, Γ is a system parameter.

The primary idea is that, in order to increase accessibility, we try to cache as many data items as possible, while trying to avoid too many duplications. Although caching popular data items aggressively in closer MTs helps to reduce the latency, in this work, we give more weight to the data accessibility than to access latency. A rationale behind this is that it is meaningless to reduce access latency when a set of MTs is isolated from other MTs or the AP, and they can not access any interested data items. Instead of waiting until the network topology changes, it is better for the MTs to have even high probability of finding the requested data items. Since Γ value enables more distinct data items to be distributed over the entire cache due to admission control, more data items can be accessible and thus the overall data accessibility is increased.

Cache Replacement Policy: A cache replacement policy is required when an MT wants to cache a data item, but the cache is full, and thus it needs to victimize a data item for replacement. Two factors are considered in selecting a victim. The first factor is the distance (δ), measured by the number of hops away from the AP or MTs, which has the requested data item. Since δ is closely related to the latency, if the data item with a higher δ is selected as a victim, then the latency would be high. Therefore, the data item with the least δ value is selected as the victim.

The second factor is the access frequency of data items. Due to mobility of the MTs, the network topology may change frequently. As the topology varies, the δ values become obsolete. Therefore, we use a parameter (τ), which captures the elapsed time of the last updated δ. The τ value is obtained by $\frac{1}{t_{cur}-t_{update}}$, where t_{cur} and t_{update} are the current time and the last updated time of δ for the data item, respectively. If τ is close to 1, δ has recently been updated. If it is close to 0, the updated gap is long. Thus, τ is used as an indicator of δ to select a victim. In this paper, we suggest a *Time and Distance Sensitive (TDS)* replacement policy based on these two factors.

4. PERFORMANCE EVALUATION

In order to evaluate the efficiency of the proposed schemes, we developed a model which is similar to [1]. We assume that an MT can not only connect to the Internet but also can forward a message for communication with other MTs via wireless LAN (e.g. IEEE 802.11). We have done extensive simulation to analyze various performance metrics. Here we include a subset of the results due to space limitation. For additional results, please refer to [2].

In Figure 1(a), data accessibility is greatly improved when we use the aggregate cache. Throughput is increased more than twice compared to the no cache case. With caching, there is a high probability of the requested data items being cached in the MT's local cache or at other MTs'. Even though a set of MTs is isolated from the AP, in contrast to the no cache case, they still try to access the cached data items among them. Note that almost 200% improvement is achieved compared to the no cache case when data access pattern follows Zipf distribution. Figure 1(b) shows the effect of the aggregate cache on the average latency. Since a request can be satisfied by any one of the MTs located along the path in which the request is relayed to the AP, unlike to the no cache case, data item can be accessed much faster. As expected, latency is reduced by more than 50% with caching. The results demonstrate the effectiveness of aggregate caching schemes.

5. CONCLUSIONS

We proposed an aggregate caching scheme, which includes a broadcast based search and a cache management technique to improve the communication performance of IMANET. Simulation results indicated that the proposed aggregate cache can significantly improve IMANET performance.

6. REFERENCES

[1] M. S. Corson, J. P. Macker, and G. H. Cirincione, "Internet-Based Mobile Ad Hoc Networking," in *IEEE Internet Computing*, July–August 1999, pp. 63–70.

[2] S. Lim, S. Park, W. Lee, G. Cao, C. R. Das, and C. L. Giles, "An Aggregate Caching for Internet based Ad Hoc Networks," Tech. Rep. CSE–02–017, Dept. of Computer Science and Engineering, The Pennsylvania State University, University Park,PA 16802, Oct 2002.