Advanced Query Processing in Mobile Ad Hoc Networks

Takahiro Hara
Dept. of Multimedia Engineering, Osaka University
hara@ist.osaka-u.ac.jp

1 Introduction

At the early stage of *mobile ad hoc networks (MANETs)* research, most studies focused on routing protocols to support communications among mobile nodes connected to each other by one-hop/multi-hop links. However, in many MANET applications, mobile nodes share data with each other and access data held by other mobile nodes. Typical examples are collaborative rescue operations at a disaster site, military operations, sensor networks, and exchange of word-of-mouth information in a shopping mall.

For such applications, preventing the deterioration of data availability at the point of network partitioning is a very significant issue. Data replication is the most promising solution to this potential problem [2, 4]. In our previous research, we designed effective data replication techniques in MANETs: replica relocation [2], consistency management [3], and data lookup and transmission [5].

In this research, we assume that every data request directly specifies the identifier of the target data item. However, there are many MANET applications that require more flexible and advanced queries for accessing data, such as *keyword search*, *top-k search*, and *k-nearest neighbor search*. Each of these advanced queries has particular features in terms of query processing.

2 Top-k Search in MANETs

We are currently working on efficient query processing for top-k search in MANETs [1]. In MANETs, it is important to acquire only the necessary data items, and an effective way to do this is to have each mobile node retrieve data items using a top-k query, in which data items are ordered by the score of a particular attribute and the query-issuing mobile node acquires data items with the k highest scores. In [1], we proposed a message-processing method for a top-k query to reduce traffic and maintain the accuracy of the query result, i.e., guaranteeing that data items with the k highest scores in the entire network are acquired.

In this method, each mobile node estimates data items with the k highest scores and sets a part of those scores to the *Standard Scores (SS)*. When a mobile node transmits query and reply messages, it reduces the number of candidates that are included in the top-k result by referring to the Standard Scores. Moreover, if a mobile node detects the disconnection of a radio link during the transmission of the reply message, it searches for an alternative path to transmit the reply message to the query-issuing node.

2.1 Determining the Standard Scores (SS)

In our method, each mobile node estimates data items with the k highest scores and sets N of them to its Standard Scores. Each Standard Score $B(i, j)(1 \leq j \leq N)$, calculated by mobile node M_i, guarantees that the query-issuing node acquires more than $\frac{k}{N}j(1 \leq j \leq N)$ data items.

2.2 Query processing

The procedures of the query-issuing node, M_p, and mobile nodes that receive a query message are briefly explained below. First, M_p specifies the number of requested data items k and the query condition. Then, M_p calculates the scores of its data items from the query condition using a scoring function and initializes its Standard Scores as follows:

$$B(p, j) = S(p, \frac{k}{N}j)(1 \leq j \leq N). \tag{1}$$

Here, $S(i, h)$ denotes the h-th highest score among the scores calculated by M_i. Specifically, the Standard Scores of M_p are the every $\frac{k}{N}$-th scores calculated by M_p.

Then, M_p transmits a query message with the attached Standard Scores to its neighboring mobile nodes. Each mobile host that received the query message updates the Standard Scores based on the scores of data items held by the host and forwards the revised query message to its neighbors. The reply messages of the query are sent back to M_p in the reversed routes how the query was propagated. Based on the information in the reply message, each mobile host that relays the reply determines its own threshold to guarantee that it is equal to or larger than the k-th highest score in the network. It also sends to M_p those data items it holds that have scores equal to or larger than the threshold (Figure 1).

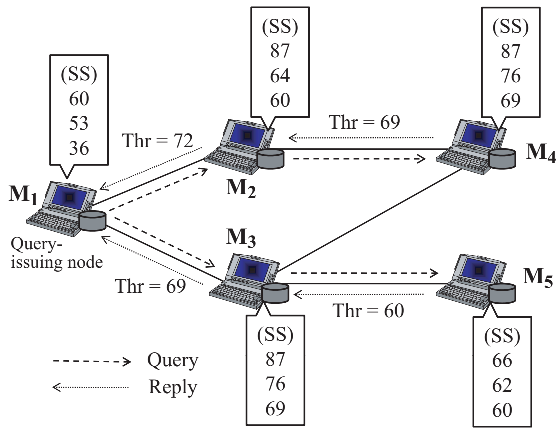

Figure 1. Query message transmission for top-k search.

By the proposed method, each mobile host can reduce the number of candidates of data items included in the top-k result, which helps to reduce the traffic in query processing.

3 Future Directions

In this section, we discuss future directions for advanced query processing in MANETs.

3.1 Top-k search

Although the method proposed in [1] can reduce the number of candidates included in the top-k result, there are several issues open to future work. First, to further reduce the traffic in query processing, we can apply an approach for predicting the distribution of scores in the network from the scores of neighbors and past query logs. Second, since the method in [1] can return the correct answer only when all mobile hosts are connected, i.e., network partitioning does not occur, we need to address the issue of how to deal with network partitioning. The most promising solution is data replication. However, we should take the features of top-k query into account when designing the replication protocol.

3.2 Keyword search

When documents such as Web documents are shared among mobile hosts, keyword search is used to find documents of interest. Unlike general keyword search for Web documents in the Internet, in MANETs we cannot assume the availability of a server with substantial computational resources to make an index (like that of Google) for ranking documents for the specified keyword in MANETs. Alternatively, in MANETs, each mobile host makes an index for its own documents, and connected mobile hosts collaboratively work to process keyword searches. Moreover, to cope with network partitioning, we should consider effective replication protocols for both documents shared among mobile hosts and their indexes.

Furthermore, when sharing documents of hypertext, e.g. Web documents, it is common for mobile users to successively access multiple documents connected by hyperlinks. In such a case, it is sometimes not desirable to return only a document that is the most relevant to the keyword specified by the user; instead, it may be preferable to return a document from which many other documents can be reached by navigating through hyperlinks to increase the user's satisfaction level.

3.3 k-nearest neighbor search

To achieve a k-nearest neighbor (kNN) search in MANETs, we need to design a method for the type of query processing specialized for kNN search. It is assumed that such a query processing method would be similar to that for a top-k search; therefore, we could design it by following the query processing method for top-k search.

On the other hand, replication methods should be different from those for top-k search. This is because in top-k search, data items with higher scores are involved in top-k searches more frequently than those with lower scores, i.e. their access frequencies are higher. However, in kNN search, data items with higher scores are not necessarily involved in query results more frequently, since the query points specified by the users change depending on the queries. In other words, such a search is not simply accessing data items with higher scores. Therefore, we need other approaches specialized for kNN search.

References

[1] R. Hagihara, M. Shinohara, T. Hara, and S. Nishio: "A message processing method for top-k query for traffic reduction in ad hoc networks," *Proc. MDM 2009*, to appear, 2009.

[2] T. Hara, "Effective replica allocation in ad hoc networks for improving data accessibility," *Proc. IEEE Infocom 2001*, pp.1568-1576, 2001.

[3] T. Hara, and S.K Madria: "Consistency management strategies for data replication in mobile ad hoc networks," *IEEE Trans. on Mobile Computing*, to appear, 2009.

[4] P. Padmanabhan, L. Gruenwald, A. Vallur, M. Atiquzzaman, "A survey of data replication techniques for mobile ad hoc network databases," *VLDB Journal*, Vol.17, No.5, pp.1143-1164, 2008.

[5] M. Shinohara, T. Hara, and S. Nishio: "A data transmission method using multicast in mobile ad hoc networks," *Proc. MDM 2009*, to appear, 2009.