Hybrid Retrieval Mechanisms in Vehicle-Based P2P Networks⋆

Quanqing Xu[1], Heng Tao Shen[2], Zaiben Chen[2], Bin Cui[1], Xiaofang Zhou[2], and Yafei Dai[1]

[1] State Key Lab for Adv Opt Commun Syst & Networks, Peking University, 100871 Beijing, China
{xqq,dyf}@net.pku.edu.cn, bin.cui@pku.edu.cn
[2] School of ITEE, The University of Queensland, Brisbane QLD 4072, Australia
{shenht,zaiben,zxf}@itee.uq.edu.au

Abstract. Mobile P2P networks have potential applications in many fields, making them a focus of current research. However, mobile P2P networks are subject to the limitations of transmission range, wireless bandwidth, and highly dynamic network topology, giving rise to many new challenges for efficient search. In this paper, we propose a hybrid search approach, which is automatic and economical in mobile P2P networks. The region covered by a mobile P2P network is partitioned into subregions, each of which can be identified by a unique ID and known to all peers. All the subregions then construct a mobile Kademlia (MKad) network. The proposed hybrid retrieval approach aims to utilize flooding-based and DHT-based schemes in MKad for indexing and searching according to designed utility functions. Our experiments show that the proposed approach is more accurate and efficient than existing methods.

1 Introduction

Mobile communication technology continues to proliferate in recent years, and mobile P2P network has been suggested as a potential solution for efficient data sharing, message transferring, information retrieval etc. As a consequence, many new applications naturally arise in telecommunication, commercial and civilian environments, such as mobile phone file sharing, commercial advertisements (ads) broadcasting [1] and traffic estimation [2], making them a focus of current research. Mobile devices (peers) in mobile P2P networks interact during physical encounters in the real world and engage in short distance wireless exchanges of data. Typically, mobile P2P networks enable direct real-time sharing of services and information among distributed peers. In contrast to wired P2P networks that are composed of static peers, mobile P2P networks are subject to the wireless bandwidth, limitations of transmission range, and highly dynamic network topology, giving rise to new challenges for research on routing, search, data

consistency maintenance, etc. The following two scenarios motivate our research on search in mobile P2P networks.

Scenario 1: Retrieving commercial ads information. A mobile P2P network can be constructed for dispatching instant and latest commercial ads. For example, staffs in a supermarket or petrol station can use their handheld devices, e.g., PDA and handphone, to publish the *instant* ads on today's promotion and discount information. Customers of interest with mobile devices can share the information with others when they are moving on roads because of the incentive mechanism. Mobile peers in an initiative mode can search other mobile peers' information under this environment to retrieve commercial ads of their interests. In such a mobile network, the influence of instant ads can be broadcasted and increased to more customers and wider areas.

Scenario 2: Retrieving traffic information. Many centralized traffic information systems have been deployed to provide traffic information of a region for mobile users, e.g. drivers, by broadcasting the information into air. In such traditional centralized mobile networks, each mobile peer has to communicate with the central server. However, many drivers have experienced the common woe of slow response or losing signals on their mobile devices when the bandwidths are most needed. Consequently, traffic information may not be updated timely and continuously, often leading to unpleasant traffic jams. In mobile P2P networks, peers can generate realtime text messages regarding the traffic and share them with other peers to understand the traffic situations just in time. If a mobile peer is heading to city, he/she can initiatively send a query to the mobile P2P network constructed by all the active mobile devices in the area to get the current traffic information of the city. In such an environment, the query is first forwarded to the mobile peers having city traffic information and then the results are routed back the querying peer. Having timely traffic information, mobile peers can avoid traffic jams and improve the traffic situations of the city.

In this paper, we study the problem of search in mobile P2P networks, where each published message is a document. We use the boolean model (i.e. exact match) to retrieve information in mobile P2P networks. In traditional structured P2P networks, Distributed Hash Table (DHT) [3] based data retrieval techniques, given a key, the query will be routed to a specific peer, which is responsible for storing the value associated with the key. However, due to the peers' frequent mobility, the DHT-based retrieval technique alone is not expected to perform equally well in mobile P2P networks due to high DHT maintenance cost. In unstructured P2P networks, the flooding-based technique is the most popular data retrieval one. Flooding entails message processing at every peer. However, it is expensive in terms of communication cost and computational resources in mobile P2P networks. To achieve efficient retrieval, we propose a Hybrid Retrieval approach (HR), which is automatic and economical. In the mobile network, a region partition service can be deployed to be known to all participating peers since vehicle-based peers have the GPS function. The region covered by the network is divided into subregions that construct a mobile Kademlia (MKad)

network. HR is designed to utilize flooding-based and DHT-based schemes for quick search.

The main contributions of this paper include: First, we present a novel mobile DHT based on subregions: MKad and its maintenance mechanisms. Second, we propose an effective and efficient search approach named Hybrid Retrieval (HR) in mobile P2P networks. HR is self-adaptive to choose either flooding-based scheme or DHT-based scheme for efficient indexing and searching by analyzing their expected search cost according to cost functions in mobile P2P networks. Third, we confirm the effectiveness and efficiency of our methods by conducting an extensive performance measured by search recall, query latency, query routing efficiency.

The remainder of this paper is organized as follows: Section 2 reviews the related work. In Section 3, we propose a mobile DHT based on Kademlia. We present a hybrid retrieval approach in Section 4. In Section 5, we show the experimental results. Finally, Section 6 concludes the paper with a discussion about our future work.

2 Related Work

There are few existing works on search in mobile P2P networks. The Geographic Hash Table (GHT) system [4] has been initially developed for data storage in sensor networks. Its benefits would be that a responsible cell may be empty without any negative influences, where a perimeter used for storing the location information would simply be outside of the responsible cell. Thus, the robustness is increased. In contrast, GHT may increase complexity of the scheme in combination with a higher network load for maintaining the information in the system. GHT stores information in those peers that form a perimeter around a point. This point in Hierarchical Location Service (HLS) [5] would be the center of the responsible cell. The peer closest to this point is responsible for regularly sending the information to the other members of the perimeter. If the closest peer fails, another member of the perimeter will sent an update after a timer has expired and a new closest peer can be elected. The geographic location system (GLS) [6] is a scalable location service that performs the mapping of a node identifier to its location. GLS can be combined with geographic forwarding to implement unicast.

Rybicki et al. [2] proposed a new paradigm to implement traffic information systems, using an infrastructure-based P2P network consisting of vehicles. This approach has several advantages over traditional VANET-based systems. In addition, it might also be possible for both VANET and P2P based systems to coexist and complement each other. Wolfson et al. [7] proposed a novel search algorithm called Rank-Based Broadcast (RBB) based on device proximity for discovery of local resources in mobile P2P networks. Eriksson et al. [8] presented Cabernet for delivering data to moving vehicles, which uses WiFi access points encountered during drives for network connectivity and provides a beneficial way to use the WiFi networks from moving vehicles.

3   A Novel Mobile DHT

3.1   Preliminaries

Most file queries are for highly replicated objects in P2P networks [9]. However, queries for rare objects are also substantial [10]. Both studies correctly reflect different aspects of the Zipfian distributions. The head of the Zipfian popularity distribution was shown in [9], and hence the query requests are measured based on the objects that match the top 50 queries seen. On the contrary, [10] focuses on the long tail of the distribution. Individual rare objects in the tail may not be requested frequently. However, these queries represent a substantial fraction of the workload, and are worth optimizing. The popularity distribution of a file-sharing workload is flatter than what we would expect from a Zipfian distribution [11]. Therefore, both frequent and rare objects are important for retrieval.

3.2   Region Partition

For an administrative region covered by a mobile P2P network, we divide it into a number of subregions based on system requirements in a recursive manner for maintenance. The region is first divided into two half subregions (R_e and R_w) based on the north/south direction (i.e. longitude), where R_e and R_w are respectively represented by 1 and 0; and then for R_e and R_w, they are also divided into two half subregions (R_n and R_s) based on the east/west direction (i.e. latitude), where R_n and R_s are respectively represented by 1 and 0 as well. The above procedure that is shown in Fig. 1 is recursively processed until the differences in longitude and latitude of a subregion are both less than given thresholds LO and LA. Consequently, the whole administrative region is divided into multiple geographical subregions, which form the network topology. Each subregion is represented by a unique region ID, and each peer uses this embedded service to keep the location information of all subregions in the whole network. Since the network topology consists of geographical subregions, each subregion is responsible for a set of keys of data objects. The keys of data objects are distributed among the subregions, such that each *key* is mapped to a subregion R. This key-to-subregion mapping is known to all peers by using a variety of Kademlia [3] named Mobile Kademlia (MKad). Note that the key-to-subregion mapping is used in MKad, while the key-to-peer mapping is utilized in the original Kademlia.

In MKad, each subregion has a unique ID and each key is a 160-bit identifier. Every document is associated with a home subregion, where the document is initially stored. To assign ⟨*key*, *value*⟩ pairs to particular subregions, MKad relies on a notion of distance between a subregion ID and a key identifier, where *key* is a 160-bit identifier of a word or document and *value* is a word or document itself. Given a subregion ID R and a *key*, MKad defines the distance between them as their bitwise exclusive or (XOR) interpreted as an integer, $d(R, key) = R \bigoplus key$. Since XOR is unidirectional, it ensures that all lookups for the same key converge along the same path, regardless of the originating peer. Thus, caching ⟨*key*, *value*⟩ pairs along the lookup path alleviates the load of hot spots. In MKad, there are four subregions with respect to a object: query subregion,
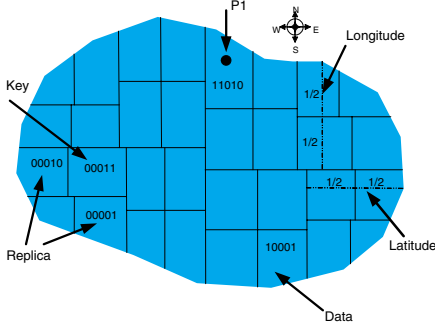
Fig. 1. The procedure of region partition ($k = 2$, k is the number of closest subregions to the key)

Algorithm 1. Maintaining the *MKad* network

Input: $\mathcal{R}$ is a set of regions
Output: the MKad network
// *Publish rare words*
1 **foreach** *subregion* $R \in \mathcal{R}$ **do**
2   Gather new words $\mathcal{W}$ of R based on gossip
3   **foreach** *word* $w \in \mathcal{W}$ **do**
4   R publishes the word w

// *Drop indexes of popular words*
5 **foreach** *subregion* $R \in \mathcal{R}$ **do**
6   R count $\#requested$ of each word
7   **foreach** *word* $w \in \mathcal{W}$ **do**
8   **if** $w's\#requested \geq \Theta_{th}$ **then**
9   Drop the popular word w's index

key subregion, replica subregion and home subregion in our proposed approach. For example, they are respectively regions: 11010, 00011, {00010, 00001} and 10001 as shown in Fig. 1.

3.3 Aggregation for a Word

To obtain the aggregation for a word w in a subregion, there are two possible approaches. One is gossip-based approach, where peers exchange their current aggregates with their neighbors till the aggregates almost converge to the global values; the other is hierarchy-based approach, where peers come into being a hierarchy and pass the aggregates in a bottom-up manner along the hierarchical path. Hierarchy-based approach requires a super peer mechanism, which is hard to be implemented in mobile P2P networks. Therefore, we utilize gossip-based approach to obtain w's aggregation.

When a peer is moving out of a subregion, its aggregation results and cached documents are transferred to some peers within the subregion which have the following features: 1) low mobility speed; 2) close to the center of the subregion, or moving towards the center; 3) enough cache space. A peer with low mobility and locating near the subregion's center is probably selected to store the aggregates, since it will leave the subregion in the near future with a low probability. In our proposed scheme, gossip-based aggregation has two functions: identifying rare words and publishing new words into MKad. We also differentiate mobile peers by their movements, which could be *intra-subregion* movement or *inter-subregion* movement. For an intra-subregion movement, a peer moves only within the same subregion, thus the overheads are trivial. For an inter-subregion movement, on the contrary, a peer moves out of its initial subregion to a neighboring subregion. Peers need to check their positions periodically to detect an inter-subregion movement.

3.4 MKad Network Maintenance

For each peer in mobile P2P networks, its storage space is limited and thus it is not able to store excessive objects. On the other hand, it is also expensive

to maintain too many popular objects in the MKad network due to the limited bandwidth. Popular objects are expected to be dropped from the MKad network if they can be easily and quickly found by flooding-based techniques. Here we use gossip-based aggregation to decide if a word is popular or rare. We present an algorithm shown in Algorithm 1 for maintaining the MKad network. For new words, they are often rare. Thus they are published in the MKad network (lines 2-4). Given a word, if its number of request is not less than Θ_{th} where Θ_{th} is a threshold, the word is regarded as a popular one. Then, the indexes of popular words are dropped from the MKad network (lines 5-9).

4   Hybrid Retrieval

For information retrieval in mobile P2P networks, generally two search schemes can be applied: flooding-based scheme and DHT-based scheme. We first present their cost models, followed by the hybrid retrieval algorithm which can automatically select the more efficient scheme for a query word.

For the flooding-based scheme, the total expected search cost for a query word w is given in [12]: $M^{\mathbf{u}}(w) = \sum_{i=1}^{N} C_{u_i} Pr(X > u_{i-1})$, where a search strategy with TTL values $\mathbf{u} = [u_1, u_2, \cdots, u_N]$, X is the object w's location, C_{u_i} is the search cost of with TTL value u_i, and $u_0 = 0$ is assumed. C_{u_i} was mentioned in [13] that the number of messages incurred with a TTL value of u_i is roughly $u_i + \beta u_i^2$, where β is a constant with respect to the network parameters. The search policy that minimizes this cost reads as following:

$$F(w) = M^{\mathbf{u}^*}(w) = \arg\min_{\mathbf{u} \in U} \sum_{i=1}^{N} C_{u_i} Pr(X > u_{i-1}) \qquad (1)$$

where U denotes the set of all admissible search strategies (TTL sequences), i.e., all vectors $\mathbf{u}$ such that $u_i < u_{i+1}$ for all $1 \le i \le N - 1$. Formula 1 can be solved backward in time using standard dynamic programming techniques [12].

For DHT-based scheme, the search cost can be approximately modeled as below:

$$D(w) = \frac{\|R_w - R_o\|}{\gamma} + \sum_{R_i \in \mathbb{R}} \frac{\|R_i - R_o\|}{\gamma} y(R_i, w) \qquad (2)$$

where R_o is the originating subregion for the query w, R_w is the subregion that is responsible for the index of w according to MKad, $y(R_i, w)$ is the boolean function returns 1 if R_i publishes w, or 0 otherwise, γ is the transmission range, and $\|.\|$ represents the distance between two subregions.

Based on a simple flooding algorithm presented in Algorithm 2, Algorithm 3 outlines our method - HybridRetrieval. When a peer sends a query q which may contain multiple words, the peer generates a query message containing the following two fields: 1) its identity, and 2) the set of query words. If the search cost of flooding-based scheme is not greater than that of DHT-based scheme (line 5), it floods the query with a given TTL L to determine if any of its neighbors have query answers (line 5). If not, the requesting peer can use the MKad protocol

Algorithm 2. Flooding	**Algorithm 3.** HybridRetrieval

Algorithm 2. Flooding

Input: P_o is the query originator, R is a subregion, q is a query, L is a TTL
Output: search results $\mathcal{S}$
// R_o is the originating subregion
1   **if** R *is not* R_o **then**
2   **foreach** *peer* $P \in R$ **do**
3   Retrieve results $\mathcal{S}$ for q from P

4   **else**
5      $\mathbb{P} = \{P_o\}$
6   **while** $L -- > 0$ **do**
7         $\mathbb{P}$=Get peers from $\mathbb{P}$ with flooding-based scheme
8   **foreach** *peer* $P \in \mathbb{P}$ **do** goto 3

9   Return gossip-based aggregation results to P_o

Algorithm 3. HybridRetrieval

Input: P_o is the query originator, q is a query
Output: search results $\mathcal{S}$
1   **foreach** *word* $w \in q$ **do**
2   **if** w *is indexed in* $MKad$ **then**
3   q's key subregion $R_{key} = R_w$
4   break

// w is the 1st word of q indexed in MKad
5   **if** $F(w) \leq D(w)$ **then** Flooding(P_o, R_o, q, L)
6   **else**
7   Achieving the home regions $\mathcal{R}_{home}$ via R_{key}
8   **foreach** $R_h \in \mathcal{R}_{home}$ **do**
9   Routing the query q to R_h
10   Flooding(P_o, R_h, q, 1)

to get a subregion (R_{key}) of query word according to the query (line 7). The home subregion of w is determined by searching the subregion whose region ID is the closest to the key of w. Peers that locate outside the home subregion drop the query message without further processing. Flooding is restricted within a subregion for quick query response and savings in network bandwidth (line 10).

A peer that routes the query message towards the destination subregion checks its location in the message's header to determine if it is within that subregion. The first peer that receives the query message inside the destination subregion floods the message within the subregion to locate the peers holding the documents including the requested key words. Each peer in that subregion processes the query message to determine if it has the requested documents. When the documents are located, the response is sent back to the original requesting peer and the query process expires.

5   Experimental Studies

5.1   Experimental Setup

For the simulations, a discrete event simulator NS-2 (*http://www.isi.edu/nsnam /ns/*) is employed with the IEEE 802.11 MAC layer. The NS-2 simulation model simulates peers moving in an unobstructed plane. GPSR [14] is used as the wireless routing protocol. We have modified it to provide routing to subregions instead of specific destinations by forwarding the packet towards the subregion and using the flooding-based scheme inside the subregion.

To measure the performance of our search schemes, we simulate the algorithms on various mobile network topologies by varying the number of peers from 100 to 1000, mean speeds from 1 to 20 m/s, and motion regions from 1500m × 1500m to 4000m × 4000m. The default parameters: # of peers is 400, # of document objects is 10,000, transmission range of a peer is 250m, motion region is 2000m × 2000m, mean speed is 10m/s, length or width of a region is 250m.

5.2   Data Set

We utilize TREC data from the 2GB Web track (WT2G) to simulate documents (ads or traffic information), and their associated words as their descriptive tags.

The top n words within each document are associated with the tags that are valid in describing the document. The words used to describe a document are drawn from its corresponding Web document. It makes the simulations more realistic. We use word frequency to simulate the strength of description that a word has in a document. The relative popularities of documents are arbitrarily assigned according to a Zipf distribution.

5.3   Retrieval Models

A series of experiments are conducted to study several combinations of routing query messages and search in mobile P2P networks. In our experiments, we used simulations to evaluate HR and compared its performance to three other search models in vehicle-based mobile P2P networks. The same partition service of a region is utilized in the four models, which is known to all participating peers.

The four search models read as follows: 1) **Flooding**: Search using the flooding-based scheme alone. 2) **GLS+MKad**: Grid Location Service (GLS) [6] builds an ad hoc network's area by using a fixed grid. Each peer maintains 1- and 2-hop neighbor lists. The neighbor lists are built from piggyback messages and indicate a peer's location as well as other parameters like the peer's 1-hop neighbors. In addition, an appropriate cell size is carefully selected regarding the transmission range of the peers. Each peer in a cell knows about all other peers in the same cell. 3) **HLS+MKad**: Hierarchical Location Service (HLS) [5] divides the area of a mobile network into smaller areas called cells and assign each peer a set $\mathbb{S}$ of these cells. Position updates and requests are sent to (possibly different) subsets of $\mathbb{S}$. The selection of the subset depends on a hierarchical grouping of the cells of $\mathbb{S}$ and the position of the peer which computes the subset. The intersection of two subsets computed for the same peer is non-empty in HLS. 4) **HR**: All the subregions construct the MKad network. HR chooses either flooding-based scheme or DHT-based scheme for efficient indexing and searching by analyzing their expected search cost according to cost functions in mobile P2P networks.

GLS and HLS are both position-based routing approaches and do not support search in mobile P2P networks. Thus, we integrate the MKad function into them and use it to get the relevant peers for a given query. Furthermore, GLS and HLS are utilized to route between a response peer and a requesting peer.

5.4   Evaluation Methodology

We adopt three performance metrics to evaluate retrieval in mobile P2P networks, i.e., retrieval accuracy, query latency and query routing efficiency.

Recall For automatically-generated queries, it is expensive for us to acquire relevance judgments in mobile P2P networks. Instead, we used the retrieval results from a single large collection as the baseline, whereafter measuring how well the P2P network could reproduce this baseline. The single large collection is the subset of the experimental test data set used to define peer contents, and agreement is measured over all the data retrieved for each query. Although this methodology is not perfect, it is reasonable because distributed retrieval systems are not better than the "single collection" baseline. Accuracy is measured with
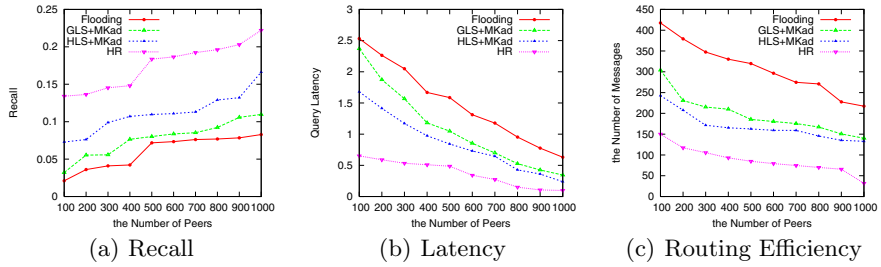
(a) Recall (b) Latency (c) Routing Efficiency

Fig. 2. Effect of Peer Densities

forms of set-based *recall* ($R = \frac{|A \cap B|}{|B|}$), where A is the set of the retrieval data in P2P networks, B is the set of the retrieval data using the subset collection of the single test data set. $R(q)$ captures the fraction of relevant data a retrieval algorithm is able to identify and present to the user.

Query Latency. The query latency is the elapsed time for a query to get its response. It records the overall query processing time.

Query Routing Efficiency. Query routing efficiency is usually measured by the average number of query messages routed for each query in mobile P2P networks.

5.5   Experimental Results and Analysis

Experiment 1: Effect of Peer Densities We consider the effects of peer densities on query recall, latency and routing efficiency of different methods. Fig. 2(a), 2(b) and 2(c) show that HR yields the best performance on query recall, latency and routing efficiency. HR uses the proposed utility functions and carefully selects query techniques: flooding and MKad, which can significantly improve the three metrics when the query is initialized and forwarded between peers. HR uses the flooding-based technique to search popular words and uses GPSR to return retrieval results. On the other hand, HR uses MKad to locate rare words and uses the flooding-based scheme to find them only in single subregions.

However, GLS and HLS only use MKad to locate words. In the static P2P networks, the query recall of DHT techniques is relatively high. However, the performance drops greatly because mobile P2P networks are highly dynamic. On the other hand, GLS and HLS achieve a much worse query performance than HR because they produce much more update packets leading to higher network load and network congestion in the end. Moreover, there is a tradeoff between the maintenance cost for the MKad network and the query effectiveness regarding GLS and HLS.

Obviously, the flooding-based technique is the worst on three metrics among the four approaches because of three general reasons: 1) it is ineffective and inefficient to return the query answers; 2) it is less effective for locating rare items: 28.1% of all queries receive 5 or fewer results, and 12.3% of queries receive no results in this experiment, although it is highly effective for locating popular words, which are retrieved in large quantities. On the other hand, although it is
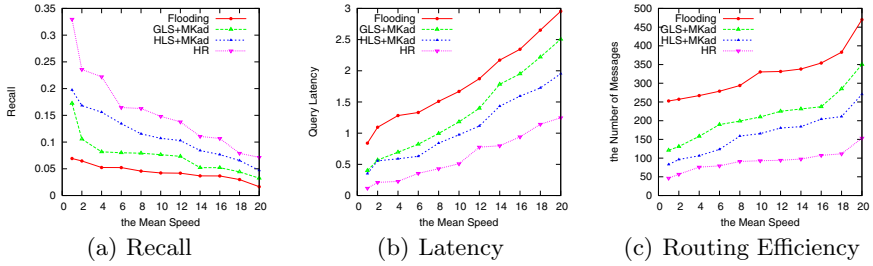
Fig. 3. Effect of Mobility Speeds

highly efficient for locating popular words, where the queries have good lookup time, it has bad returning time; 3) it is less efficient for locating rare objects and the results have poor response time. For queries that return a single result, the first result arrives after 7.2 seconds on average. For queries that return 5 or fewer results, 5.6 seconds elapsed on average before receiving the first result in this experiment. However, HR uses the designed utility functions to select the query techniques for improving the query recall, latency and routing efficiency. As the peer density increases, HR can construct stabler MKad network to search rare words and locate much more popular words with smaller TTL.

Experiment 2: Effect of Mobility Speeds To evaluate the scalability of HR with respect to mobility speeds, we increase the mobility speeds from 1m/s to 20m/s. We still evaluate three metrics: query recall, latency and routing efficiency of four approaches. The results presented in Fig. 3(a), 3(b) and 3(c) indicate that HR still achieves significantly better query performance than the other three methods.

It is more difficult for the flooding-based mechanism to locate rare words and return retrieval results at higher speed. The network load in GLS and HLS increases with growing peer speed because more updates are sent that lead to worse query routing efficiency. On the other hand, a higher load produces more collisions resulting in re-queries and in unwanted delays in GLS and HLS. Facing to high peer speed, HR still has better query performance than the other three methods because it is based on subregions besides the proposed cost functions. On the other hand, HR construct the MKad network only for rare objects, which is effective and efficient at high speed. On the contrary, both GLS and HLS are not effective and efficient to construct the MKad network for all the objects at high speed.

Experiment 3: Effect of Motion Regions with a Fixed Peer Density We consider the effect of motion regions with a fixed peer density (100 peers per square kilometer) on query performance of four methods. Fig. 4(a), 4(b) and 4(c) show the results of query recall, latency and routing efficiency. Similarly, HR can retrieve significantly better query performance than previously proposed methods. Clearly, the flooding method is the worst among the four approaches, which does not need to be elaborated any further.

To improve the query performance, the proposed utility functions in HR help to carefully select query techniques: flooding and MKad. Moreover, HR uses
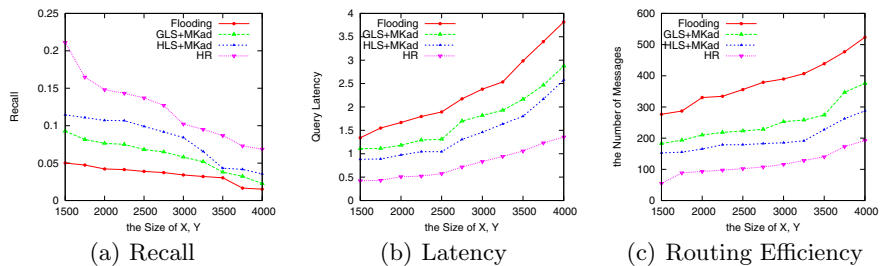
(a) Recall   (b) Latency   (c) Routing Efficiency

Fig. 4. Effect of Motion Regions with a Fixed Peer Density

GPSR to locate response peers and return query results to a requesting peer. Since GPSR makes greedy forwarding decisions using the only information about a router's immediate neighbors, both query latency and query efficiency are improved further. However, GLS and HLS only use MKad to locate words. Besides previous explanations in experiment 1 and 2, their query messages in a larger region must be forwarded to requesting peers and response ones via a longer distance, which is a more difficult task in mobile P2P networks with fast movers.

6 Conclusions and Future Work

In this paper, we gave two scenarios to motivate our research on search in mobile P2P networks. And then we proposed hybrid retrieval policies to utilize flooding-based and MKad-based schemes for quick search in mobile P2P networks. Our experiments show our approach yields better performance. In particular, the query response time and the number of messages per query are reduced substantially without losing accuracy. We believe that possible directions to future work include some research topics such as extending our method with cooperative cache in mobile P2P networks, evaluating the MKad network's performance and maintenance cost, improving the simple flooding algorithm, and examining the impact of diverse duplication mechanisms in our retrieval model.

References

1. Chen, Z., Shen, H.T., Xu, Q., Zhou, X.: Instant advertising in mobile peer-to-peer networks. In: ICDE (to appear, 2009)
2. Rybicki, J., Scheuermann, B., Kieß, W., Lochert, C., Fallahi, P., Mauve, M.: Challenge: peers on wheels - a road to new traffic information systems. In: MOBICOM, pp. 215–221 (2007)
3. Maymounkov, P., Mazières, D.: Kademlia: A peer-to-peer information system based on the xor metric. In: Druschel, P., Kaashoek, M.F., Rowstron, A. (eds.) IPTPS 2002. LNCS, vol. 2429, pp. 53–65. Springer, Heidelberg (2002)
4. Ratnasamy, S., Karp, B., Yin, L., Yu, F., Estrin, D., Govindan, R., Shenker, S.: Ght: a geographic hash table for data-centric storage. In: WSNA, pp. 78–87 (2002)
5. Kiess, W., Fussler, H., Widmer, J., Mauve, M.: Hierarchical location service for mobile ad-hoc networks. SIGMOBILE Mob. Comput. Commun. Rev. 8(4), 47–58 (2004)

6. Li, J., Jannotti, J., Couto, D.S.J.D., Karger, D.R., Morris, R.: A scalable location service for geographic ad hoc routing. In: MOBICOM, pp. 120–130 (2000)
7. Wolfson, O., Xu, B., Yin, H., Cao, H.: Search-and-discover in mobile p2p network databases. In: ICDCS, p. 65 (2006)
8. Eriksson, J., Balakrishnan, H., Madden, S.: Cabernet: vehicular content delivery using wifi. In: MOBICOM, pp. 199–210 (2008)
9. Chawathe, Y., Ratnasamy, S., Breslau, L., Lanham, N., Shenker, S.: Making gnutella-like p2p systems scalable. In: SIGCOMM, pp. 407–418 (2003)
10. Loo, B.T., Huebsch, R., Stoica, I., Hellerstein, J.M.: The case for a hybrid p2p search infrastructure. In: Voelker, G.M., Shenker, S. (eds.) IPTPS 2004. LNCS, vol. 3279, pp. 141–150. Springer, Heidelberg (2005)
11. Gummadi, P.K., Dunn, R.J., Saroiu, S., Gribble, S.D., Levy, H.M., Zahorjan, J.: Measurement, modeling, and analysis of a peer-to-peer file-sharing workload. In: SOSP, pp. 314–329 (2003)
12. Chang, N.B., Liu, M.: Revisiting the ttl-based controlled flooding search: optimality and randomization. In: MOBICOM, pp. 85–99 (2004)
13. Baryshnikov, Y., Coffman, E., Jelenkovic, P., Momcilovic, P., Rubenstein, D.: Flood search under the california split rule. Operations Research Letters 32(3), 199–206 (2004)
14. Karp, B., Kung, H.T.: Gpsr: greedy perimeter stateless routing for wireless networks. In: MOBICOM, pp. 243–254 (2000)