

# Proactive Replication and Search for Rare Objects in Unstructured Peer-to-Peer Networks

Guoqiang Gao, Ruixuan Li\*, Kunmei Wen, Xiwu Gu, and Zhengding Lu

Intelligent and Distributed Computing Lab, College of Computer Science and Technology,  
Huazhong University of Science and Technology, Wuhan 430074, P.R. China  
ggq@smail.hust.edu.cn, {rxli, kmwen, xwgu, zdlu}@hust.edu.cn

**Abstract.** The search efficiency problem in unstructured peer-to-peer network has not been adequately addressed so far, especially concerning search for rare objects. In this paper, we propose a proactive replication strategy to improve the search efficiency for rare objects. It uses object probing technique for peers to decide whether to establish replications for their objects or not when they join the network. This strategy can effectively increase the popularity of rare objects so as to enhance the search efficiency. We also present a rare object search algorithm. When a peer forwards a search request, forward probability is calculated according to its neighbors' degree and the number of neighbors' objects. Therefore, the search request is forwarded to the peers more likely containing target objects. Simulations show that the proactive replication strategy greatly improves the search efficiency for rare objects with moderate communication overhead. The rare object search algorithm not only improves search efficiency for rare objects, but also achieves load balance in search.

**Keywords:** peer-to-peer network, search, rare object, replication, object popularity.

## 1 Introduction

The peer-to-peer (P2P) networks can be broadly categorized into two types: structured P2P networks and unstructured P2P networks. Structured P2P networks [1] implement a Distributed Hash Table (DHT) and provide one basic operation: given a key, they map the key to a node. These systems are efficient and have a low traffic cost, but it is not clear which system is best suited to provide a reliable, load-balanced, and fault-tolerant network. Unstructured P2P networks account for millions of users dynamically connected in an ad hoc fashion, and queries, which are done through a path, are selected randomly according to a uniform distribution. They are suited to retrieve highly replicated data, but have limitations for rare information retrieval.

The typical unstructured P2P network is Gnutella [2], which implements search through using a flooding technique with a certain scope. This type of network has excessive traffic forms that account for more than 50% of network traffic [3]. Moreover, unstructured P2P has poor search efficiency, especially for rare objects. Such a

---

\* Corresponding author.

problem has been addressed by various algorithms, such as flooding, expanding ring [5] and random walks [7]. Most of these algorithms are in fact very effective for locating popular objects. However when it comes to find rare objects in unstructured P2P networks, the similar operations using these algorithms become less effective and less efficient [17]. As shown in [4], as much as 18% of all queries return no response even when results are available in the widely used Gnutella network.

This paper introduces a mechanism to address the above problems of unstructured P2P networks. Our main idea results from taking the proactive replication strategy to increase rare object proportion over unstructured P2P network and using our search scheme to improve search efficiency to the rare objects. The unstructured P2P networks such as Gnutella are power-law distribution networks. That is, the majority of the network nodes have low degrees, while a few other nodes have high degrees. Moreover, some nodes have a lot of resources, and some nodes have fewer objects. Therefore, searching the nodes with high degree and many objects is more likely to hit the target. In the search, we can take advantage of these features to forward queries to those peers more likely containing target objects in order to achieve the goal-oriented routing, which is similar to that of the structured P2P network. As we know, if an item is very rare, any query mechanism can hardly find it with limited number of steps. To alleviate this situation, we propose an object probing technique. Peers search the objects they own in the network at a certain frequency. If there are no hits, they will duplicate the objects at the terminative nodes they reach. The objective of this strategy is to increase the number of rear objects and the probability of search success. It obviously improves search efficiency for rare objects with lower popularity. The experiments show that it not only reduces the communication overhead of the search, but also minimizes the overhead of the replication.

The rest of this paper is organized as follows. Section 2 compares our strategy to related work. Section 3 presents the proposed methods. Section 4 demonstrates our algorithm through simulation. Finally, we summarize our contribution and outline items for future work.

## 2 Related Work

To find an object, a node queries its neighbors in unstructured P2P networks. The most typical query method is flooding, where the query is propagated to all neighbors within a certain radius. Flooding introduces duplicate messages and duplicate queries, which create pure overhead. Since flooding has inherent limitations, some improved methods have been proposed, such as expanding ring [5]. Although this approach has less message duplication than overall flooding, it does produce many extra traffic messages, especially for querying rare objects. Studies [7], [8] have shown random walk to be significantly more efficient than flooding, which forwards a query message to some randomly chosen neighbors at each step until the object is found. The query cache strategy [6] can also reduce the communication overhead, but shortens the length of the queries. With this strategy, nodes cache the results of the successful queries so that the same subsequent queries can quickly hit the target through the cache. This mechanism not only improves the search efficiency, but also reduces the communication overhead. The search efficiency of these improvements is

significantly increased for those hot resources or more replicated resources, but the efficiency remains small for rare objects.

Published measurements of unstructured P2P networks such as Gnutella show that they have power-law degree distributions [9]. This distribution reflects the existence of a few nodes with very high degree and many with low degree. Adamic [10] proposed a high degree seeking strategy, in which a search request is forwarded to the high degree node because they have more neighbors. However, this method easily leads to information gathering, and harms the distribution of the search. Based on power-law characteristics, we improve the high degree seeking strategy through the probability forward. Our scheme tries to balance the load of each node in the search with the similarity search efficiency.

Replication will increase search efficiency of popular objects through storing the object at the requester or all nodes along searching path when a search is successful. Presently, replication strategy is mainly passive, as this method cannot help rare objects. Cohen [12] studied a replication strategy to improve search efficiency in unstructured P2P network in which target objects establish replication with the square root of the number of visits. This strategy can greatly improve the search efficiency of hot resources, but for rare objects the efficiency of the search has not improved much.

When searching for rare objects, the researchers put forward the strategy of object replication. When a peer joins a P2P network, it actively installs its object's references at a set with  $O(\gamma n^{1/2})$  of randomly selected peers [13]. This strategy can improve the searching efficiency for rare objects, but it can introduce a high communication cost. Moreover, it is hard to estimate the network peer number  $n$ , especially for dynamical P2P networks. Krishna et al. [11] uses Supernode-Constrained Random Walk (SCRW) to search rare objects with two-hop index replication, and proposes three two-hop index replication strategies: full replication, square-root replication and constant replication. In SCRW, nodes always forward the query to one of it randomly selected supernode neighbors. Although this strategy can improve the search efficiency for rare objects, it needs the P2P network with supernode and will incur heavy overheads. Shen et al. [16] put forward an adaptive replication strategy. Compared to existing algorithms, this strategy requires the replication popularity to be relatively low. To achieve better search efficiency, the replication location is not random, but on the so-called network routing node. It can improve the search hit rate, and reduce the search latency as well.

In this paper, we propose a proactive replication strategy which can replicate the rare objects at the nodes of P2P network, and also propose a rare object search algorithm. By doing this, we can not only enhance the search efficiency for rare objects, but also reduce the unnecessary communication overhead for replication. In the next section, we will introduce our strategy in detail.

### 3 Replication and Search Strategy

To improve the search efficiency of rare objects, we must replicate the rare objects and improve the searching algorithm. In this section, we will first describe the object popularity and the proactive replication (PR) strategy for replicating the rare objects, and then discuss the rare objects searching (ROS) algorithm.

### 3.1 Object Popularity and Object Value

Object popularity is the proportion of the number of the nodes with an object to the network size. Object popularity is very important, and it directly determines the efficiency of search in unstructured P2P networks.

Let  $N$  represent the number of nodes in an unstructured P2P network, and an object  $R$  in this network has  $r$  copies, then object  $R$ 's popularity is refined as  $P=r/N$ . In unstructured P2P network, the job of search algorithms is an exploratory attempt. Therefore, the popularity of target objects directly affects the search efficiency. If an object  $R$  is searched at a peer in the P2P network, then the search hit probability is  $P$ . Using standard random walk as the search algorithm, the maximum number of hops is  $k$ , and then the search hit probability  $Pr$  for object  $R$  is refined as  $Pr(R)=kP$ .

To make the standard random walk search successfully for object  $R$ , we need to ensure that  $Pr(R)$  is equal to 1. According to Equation (1), it can be seen that the object  $R$  can be searched successfully if the number  $r$  of object  $R$  in the P2P network is up to a certain standard.

$$Pr(R) = 1 \Rightarrow r = \frac{N}{k} \quad (1)$$

In unstructured P2P networks, the object popularity directly affects the search efficiency, especially for rare objects. Therefore, we can build replication based on the object popularity. How to obtain the object popularity will be discussed in the next section.

Although the object popularity can be used as a basis for replication, some rare objects in the P2P network are valueless. For example, the node owner accidentally shares a private image, but it has no value to other users in the network. The overhead of replication can be reduced through the introduction of object value. The object  $R$ 's value  $V_R$  at a time 't' in Equation (2) is refined to:

$$V_R(t+1) = \mu V_R(t) + \eta \frac{R_q(t)}{N_q(t)} * 100 \quad (2)$$

where  $0 \leq \mu, \eta \leq 1$ , the larger the value of  $\mu$ , the more importance of the historical requests to  $V_R$ , while the larger the value of  $\eta$ , the more importance of the recent requests to  $V_R$ .  $R_q(t)$  is the number of requests received for the object  $R$  and  $N_q(t)$  is the total number of requests received by the node at a time 't'. We can take  $V_R$  as another basis for replication.

### 3.2 Proactive Replication

From the analysis above, we know that we can improve object  $R$  search hit rate through increasing the popularity of object  $R$ . However, in a distributed environment, each node only has the partial information, so to obtain the object popularity is very difficult. Therefore, it is not certain which objects should be replicated and how much copies should be made. The existing rare objects replication strategy is to replicate all objects to build the same number of replication, although this can improve the hit rate of the search algorithm, it also brings unnecessary communication overhead.

To address this issue, this paper presents a simple solution: When a node joins a P2P network, the node initiates a random walk to search its own objects, and the hit

can be defined by success probe or failure probe. The node  $i$ 's proactive probe function for object  $R$   $proactiveProbe(i, R)$  is refined in Equation (3).

$$proactiveProbe(i, R) = \begin{cases} 1 & i \text{ successfully search } R \\ -1 & \text{else} \end{cases} \quad (3)$$

The node establishes a copy for those objects who fails to probe. The replication method is replicating the object in the last node during the proactive probe. The basic idea of this strategy is this: If one fails to find the source, the same thing may happen to the others. This strategy is not to obtain an accurate popularity of an object, but whether the popularity is less than the minimum value with which search can be done successfully. In order to ensure the minimum objects popularity and which object can be searched successfully, each node must pass on a regular probe to maintain objects popularity. The node for the next probe time for object  $R$  is refined in Equation (4).

$$T(x) = g \times 2^l, l = \begin{cases} m+x & m+x > 0 \\ 0 & \text{else} \end{cases} \quad (4)$$

where  $g$  is configuration interval value,  $x$  is the return value of  $proactiveProbe(i, R)$ ,  $m$  is the last  $l$  values and the initial value of  $m$  is 0.

By defining the above probe time, we can effectively reduce the number of probes. In the meantime, the network maintains the minimum popularity of objects through the joint efforts of the network nodes. The detail cost analysis will be described in the experimental section.

The biggest cost of the proactive replication comes from the regular probe. To reduce this cost, object value can be introduced in regular probe of proactive replication. However, it is very difficult to compute the visit frequency of an object  $R_q$  in the distributed environment. In this article, we introduce a simple mechanism: Each object owner only computes the visit frequency of this object and its replication. Since the node's own object value  $V_R$  is 0, when it joins the network, in order to enhance the popularity of valuable objects, the node must be allowed to run a certain number of times before the factor of object value is introduced into regular probe. When the node's running times is greater than a predefined threshold  $Tp$ , and the object value of  $V_R$  is less than a certain threshold  $Vt$ , the probe is not triggered, and its next probe time is set to  $T(1)$ .

Node  $i$ 's proactive replication algorithm for object  $R$  is shown in Algorithm 1.

**Algorithm 1.** Node  $i$ 's proactive replication algorithm for object  $R$

```

Procedure proactiveReplication(i, R, t)
Begin
  If (t > Tp and  $V_R(t) < Vt$ ) Then
    After T(1), call proactiveReplication(i, R) again;
  Else
    k = proactiveProbe(i, R);
    If (k = -1) Then
      Replicate R in the last peer;
    End If
    After T(k), call proactiveReplication(i, R) again;
  End If
End

```

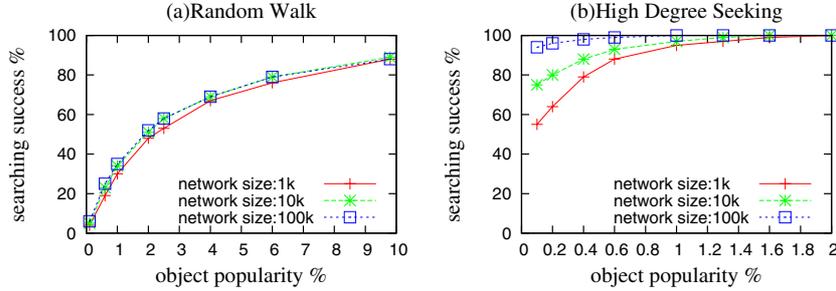


Fig. 1. The RW and HDS hit rate with different object popularity

### 3.3 Analyze of Search Efficiency

Random walk (RW) significantly reduces the network overhead, and the search hit rates are as well reduced. In the same hit rate, taking high degree seeking (HDS) strategy requires a lower popularity of objects than random walk. To analyze search efficiency, we design simulation experiment with different network sizes for RW and HDS. Fig. 1(a) depicts a diagram for the search hit rate and object popularity using the RW under different networks, and Fig. 1(b) describes the efficiency of HDS. Fig. 1 and 2 shows that the search hit rate of RW only reaches about 90% under 10% objects popularity, whereas HDS only requires 0.5% objects' popularity in the same search hit rate. At the same time, the correlation of the search hit rate and the objects popularity do not change in different network sizes. As we discussed earlier, the higher popularity objects have a higher hit rate of search. In subsequent experiments, this paper will only use a relatively large network size, because it is resembled to the actual network.

Since random walk is essentially a complete blind detection, it is ideal that the object popularity is relatively high, which is not suitable to our strategy characterized by the minimum object popularity. The high degree seeking search only forwards search to neighbor nodes with relatively large degree, because these nodes have more neighbors and are more likely to contain the target resources. However, there is also the issue of information gathering when high degree seeking and high degree nodes may become a routing node, which results in network bottlenecks. Furthermore, there is the problem of fixed routing for high degree seeking. If we always choose high degree neighbor as the forwarding node, the search routing is fixed, which could lead to some nodes in the network not being covered. Therefore, we need a search algorithm that can utilize some characteristic of P2P networks such as the degree of nodes to improve efficiency, meanwhile, the algorithm is able to avoid some problems of arising from itself such as the heavy load of the high degree nodes.

### 3.4 Rare Object Search

In order to more effectively search for rare objects, this paper proposes a search scheme which combine with the degree and the object number of neighbors to obtain forward decision. Experimental results show that this strategy has good search efficiency for rare objects. Each node in the network keeps the object number  $r$  and the

degree  $k$  of its neighbors.  $j_r$  and  $j_k$  represent the object number and the degree of neighbor node  $j$ , respectively. When a node  $i$  forwards the query to neighbor  $j$ , the forwarding probability  $Pr_{ij}$  and  $Pk_{ij}$  respectively based on the object number and the degree of neighbor are defined by Equation (5).

$$Pr_{ij} = \frac{j_r}{\sum_{p \in K} p_r}; Pk_{ij} = \frac{j_k}{\sum_{p \in K} p_k} \quad (5)$$

where  $K$  is the set of  $i$ 's neighbors,  $p$  is the node of the set  $K$ , and  $p_r$  and  $p_k$  represent the object number and the degree of the node  $p$ , respectively.

Based on the above analysis, the high degree seeking can improve the search efficiency, but it will bring the load problem. The probability  $Pk_{ij}$  defined here is similar to the high degree seeking. In order to reduce the load of high degree nodes, we introduce the probability  $Pr_{ij}$  based on the object number of nodes. A node cannot contain all information of its neighbors' resources. For example, in some P2P applications, the node builds the index based on terms of its documents. The index may be very big, so the nodes cannot store it each other. Therefore, the number of objects, which may be the number of terms in some applications, can be used as an important factor in guiding query route.  $Pr_{ij}$ , like  $Pk_{ij}$ , can lead the query to those power nodes which are more likely to contain the target objects. The experiments show that this strategy is available. Node  $i$  forwards query to neighbor  $j$  with the forward probability defined by Equation (6).

$$P_{ij} = \lambda Pr_{ij} + (1 - \lambda) Pk_{ij}, \quad 0 < \lambda < 1 \quad (6)$$

where  $\lambda$  is an adjustable parameter with the initial value being 0.5. The parameter  $\lambda$  can be automatically adjusted in the running system. For example, since search success is due to the factor of objects, you can increase the value of  $\lambda$ . The above forwarding strategy not only finds rare objects more effectively, but also enables search distribution. This strategy solves the problem of fixed routing in high degree forward and will be conducive to explore new objects.

In the rare object search strategy, the node  $i$ 's forwarding algorithm is shown as Algorithm 2.

**Algorithm 2.** Node  $i$ 's forwarding algorithm

```

Procedure Forward (i)
Begin
  For (each neighbor j in peer i);
    Calculate  $Pr_{ij}$ ;
    Calculate  $Pk_{ij}$ ;
  End For
  For (each neighbor j in peer i)
    Calculate  $P_{ij}$ ;
    Forward query to neighbor j according  $P_{ij}$ ;
  End For
End

```

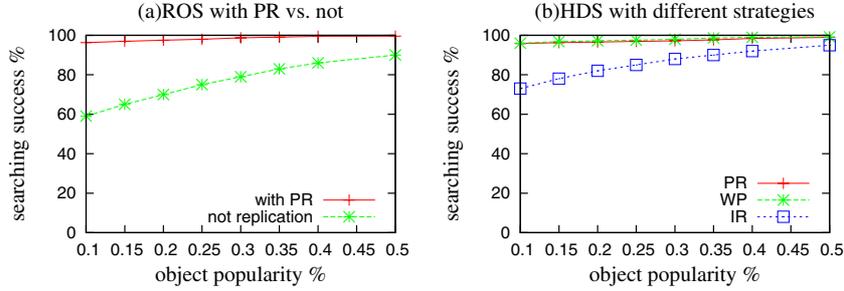


Fig. 2. Effectiveness of Proactive Replication

## 4 Experimental Evaluation

In this section, we analyze the efficiency of proactive replication (PR) strategy and rare object search (ROS) algorithm presented in this paper through the experiment. In order to obtain a more realistic experimental environment, the size of the node in network is in the magnitude of 100,000. Taking PeerSim [14] as the simulation-driven kernel, the initial topology is generated by Brite [15], and the distribution of network resources is defined by Zipf. In the simulations, we set  $Tp$  as 20 cycle times,  $Vt = 100$ ,  $\mu = \eta = 1$ ,  $g = 2$  cycle times,  $\lambda = 0.5$ , and don't change the value of these parameters for simplicity.

### 4.1 Effectiveness of Proactive Replication

To assess the efficiency of proactive replication strategy and rare object search algorithm, we run rare object search (ROS) algorithm proposed in this paper in the network which proactive replication strategy is applied to and the network without any replication strategies. Fig. 2 depicts the correlation of object popularity and search hit rate in the networks in which ROS is run. Experiments show that the search efficiency of ROS has been greatly improved as proactive replication strategy is applied in the P2P network. The object popularity of the horizontal axis in Fig. 2(a) represents the system initialization popularity of the object itself, rather than the object popularity after using proactive replication strategy. This is to visually show the efficiency of the search for rare objects. The rare object search algorithms forward queries to those power nodes, so it has high search efficiency with the proactive replication strategy which provides the necessary popularity of objects for successful search. Therefore, our proactive replication strategy is available to improve the search efficiency for rare objects according to the experimental results.

Table 1. The amount of message in different replication strategies

	WP	PR
Average message processed in each node	2162	1391
Factor	1	1.55

Due to object probe, the proactive replication strategy will produce some additional communication overhead compared to other algorithms. In order to analyze these overhead, we collect the messages of objects probe and replication in the P2P network with proactive replication strategy. The average number of each node objects is 10 in the experiment. We take the replication strategy with probability (WP) proposed by Ronaldo [13] to conduct a comparative analysis. WP replicates  $\gamma n^{1/2}$  copy for each object of node,  $n$  is the number of network nodes. WP has the query success rate at 85% when  $\gamma=1$ . The success rate of ROS reaches 95% after using PR strategy. To be fair, we take  $\gamma$  as 1 in the experiment. Table 1 shows the average amount of messages processed by the two kinds of replication strategies in each node in the simulation. The proactive replication proposed in this paper reduces the communication traffic by a factor of 1.55 compared to the existing WP algorithms. Obviously, the traffic of PR algorithm will be increased with the simulation cycles. However, if the object probe cycle interval is properly handled, we can maintain low communication overhead and in the meantime we can achieve a reasonable popularity of objects.

To analyze the performance of PR, we introduce another replication strategy: two-hop index replication (IR) similar to the strategy proposed in [11]. In IR, each node sends its index to all of its two-hop neighbors in its routing table. Fig. 2(b) shows the efficiency of high degree seeking (HDS) strategies with three different replication strategies. As you can see, IR has the least effective, the efficiency of PR and WP for rare objects is similar. However, the proactive replication strategy generates less communication overhead, so the proactive replication strategy PR is superior to the replication strategy with probability WP in overall performance.

## 4.2 Effective of Rare Object Search

We can see, from the Fig. 2, the search efficiency of rare object search (ROS) presented by this paper is similar to the high degree seeking (HDS) for searching rare objects. However, HDS forward routing is fixed, which can easily lead to information gathering; while the ROS uses the probability of forwarding, which can improve the problem of load balancing posed by query. The system with high degree forward strategy may lead the high degree nodes to become a bottleneck. To alleviate the load of high degree nodes, we can firstly run random walk and HDS after search failure to improve HDS. We call this strategy as improved HDS (IHDS). Although IHDS could improve the problem of load, it increases latency for rare objects, and brings additional communication overhead as well. Therefore, taking ROS as search scheme can not only ease the network traffic load, but also lessen the search delay of rare objects. Fig. 3(a) describes the search delay of ROS and IHDS for rare objects in P2P network with PR strategy. The vertical axis shows the maximum number of hops for successful search to target objects. IHDS firstly use random walk to search an object. However, random walk generally fails with the specified TTL value for rare objects with low popularity. Therefore, IHDS has high latency overhead.

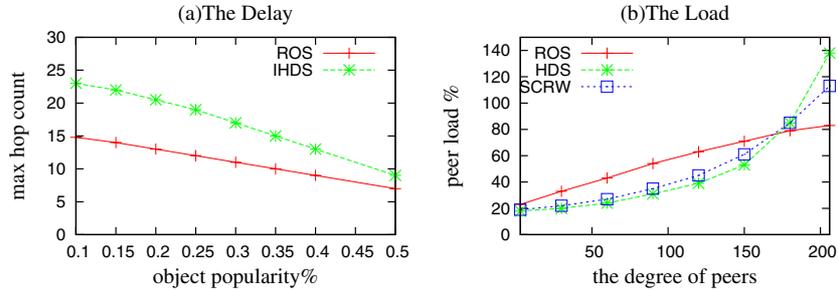


Fig. 3. Effectiveness of Rare Object Search

In order to carry out the analysis about load, we design an experiment to discuss the load balance of nodes with different degrees as follows: The network topology is generated by Brite [15], the minimum degree of nodes is 3, the maximum degree is 206, and the degree distribution of nodes obeys power-law distribution. Fig. 3(b) is the simulated results of ROS, HDS and SCRW (Supernode-Constrained Random Walk) proposed in [11] in which nodes always forward the query to one of its randomly selected supernode neighbors. As you can see, the load of ROS is balanced, and the load of the node with the greatest degree is lower than 100%. While the high load of the HDS and SCRW is mainly concentrated in the high degree nodes, the utilization of the nodes with small degree is relatively lower. Therefore, ROS algorithm presented in this paper not only has a better search efficiency, but also improves the load balance of nodes compared to existing algorithms.

### 4.3 Performance in Churn

The proactive replication can withstand the high rate of nodes dynamics. Churn arises from continued and rapid arrival and failure (or departure) of a large number of peers in the P2P network. We evaluated the efficiency of the proactive replication in unstructured P2P network in churn. Experiment results verified the resilience of the proactive replication in churn. In the simulation, node join and voluntary departures are modeled by a Poisson process with a mean rate of  $R$ , which ranges from 0.05 to 0.5. A rate of  $R=0.05$  corresponds to one node joining and leaving every 20 seconds on average. Fig. 4(a) plots the average lookup path length versus node join/leave rate. We can see that the proactive replication has the best effect because the cycle probes of the nodes with PR which ensure the objects popularity.

Churn not only appears in the network, but also exists in the content. We take another experiment to analyze the dynamicity of the content (popularity changing over time). In the simulation, the set of search objects varies with time period, that is, the popular objects in a time period will become unpopular in another time period. We can adjust  $\mu$  in equation (2) to increase the replication of popular rare objects. The proactive replication ensures high resistance to content churn and this is evident from Fig. 4(b) in which the system generates a good success rate.

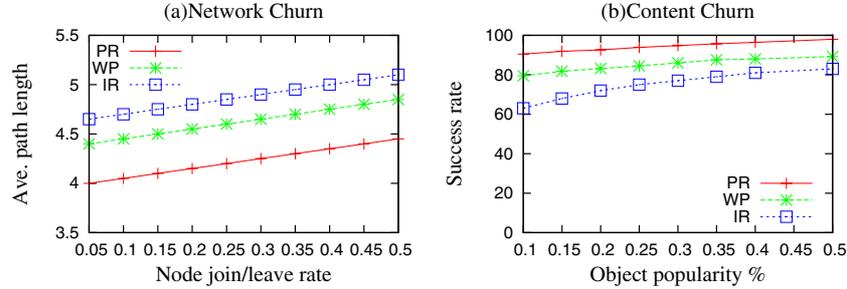


Fig. 4. Performance in churn

## 5 Conclusions

This paper presents a new proactive replication strategy for rare objects, as well as proposes a rare object search algorithm through the improvement of the high degree seeking strategy. Through the proposed algorithms, we can greatly improve the search efficiency for rare objects, and keep the communication cost associated with search relatively low. Experiments showed that there is a significant improvement in efficiency for rare objects after using proactive replication strategy, and the communication overhead brought by search is lower than other replication strategies. We design the comparative experiment with other search algorithms, and the results show that the search algorithm in this article has better efficiency for rare objects, and ensures load balance of search. In order to reduce unnecessary communication overhead, the replication strategy proposed by this paper is the algorithms of the least amount replication, which can decide whether to create a replication through the probe of objects. This will result in the communication overhead in probe. However, experiments have shown that the overhead is less than replication strategy of a constant amount. As a result, it will be interesting to see if the communication overhead brought by probing can be further reduced by fine-tuning the interval between detection cycles. The dynamical adjustment for the proportion of degree and objects in the rare object search algorithm is also a place worth exploring. The future work will focus on the cycle of object probing and the optimization of search strategy.

**Acknowledgments.** This work is supported by National Natural Science Foundation of China under Grant 60873225, 60773191, 70771043, National High Technology Research and Development Program of China under Grant 2007AA01Z403, Natural Science Foundation of Hubei Province under Grant 2009CDB298, Wuhan Youth Science and Technology Chenguang Program under Grant 200950431171, Open Foundation of State Key Laboratory of Software Engineering under Grant SKLSE20080718, Innovation Fund of Huazhong University of Science and Technology under Grant Q2009021.

## References

1. Ratnasamy, S., Francis, P., Handley, M., Karp, R.M., Shenker, S.: A scalable content-addressable network. In: The 2001 annual conference of the Special Interest Group on Data Communication (SIGCOMM 2001), pp. 161–172. ACM Press, San Diego (2001)
2. Gnutella\_protocol, [http://gnunet.org/papers/gnutella\\_protocol.pdf/](http://gnunet.org/papers/gnutella_protocol.pdf/)
3. Ripeanu, M.: Peer-to-Peer Architecture Case Study: Gnutella Network. In: The 1st International Conference on Peer-to-Peer Computing (P2P 2001), pp. 99–100. IEEE Press, Linköping (2001)
4. Loo, B.T., Hellerstein, J.M., Huebsch, R., Shenker, S., Stoica, I.: Enhancing p2p file-sharing with an internet-scale query processor. In: The thirtieth International Conference on Very Large Data Bases (VLDB 2004), pp. 432–443. Morgan Kaufmann, Toronto (2004)
5. Lv, Q., Cao, P., Cohen, E., Li, K., Shenker, S.: Search and replication in unstructured peer-to-peer networks. In: The 16th Annual ACM International Conference on Supercomputing (ICS 2002), pp. 84–95. ACM Press, New York (2002)
6. Yin, Z., Jin, H., Zhang, C., Yuan, Q., Zhao, C.: Adaptive Query-Caching in Peer-to-Peer Systems. In: The 2nd IFIP International Conference on Network and Parallel Computing (NPC 2005), pp. 97–104. Springer Press, Beijing (2005)
7. Gkantsidis, C., Mihail, M., Saberi, A.: Random Walks in Peer-to-Peer Networks. In: The 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2004). IEEE Press, HongKong (2004)
8. Gkantsidis, C., Mihail, M., Saberi, A.: Hybrid search schemes for unstructured peer-to-peer networks. In: The 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2005), pp. 1526–1537. IEEE Press, Miami (2005)
9. Adamic, L.A., Lukose, R.M., Puniyani, A.R., Huberman, B.A.: Search in Power-Law Networks. *Phys. Rev. E* 64, 046135 (2001)
10. Adamic, L.A., Lukose, R.M., Huberman, B.A.: Local search in unstructured networks. In: Bornholdt, S., Schuster, H.G. (eds.) *Handbook of Graphs and Networks*. Wiley-VCH, Berlin (2003)
11. Puttaswamy, K.P.N., Sala, A., Zhao, B.Y.: Searching for Rare Objects using Index Replication. In: The 27th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2008), pp. 1723–1731. IEEE Press, Phoenix (2008)
12. Cohen, E., Shenker, S.: Replication strategies in unstructured peer-to-peer networks. In: The 2002 annual conference of the Special Interest Group on Data Communication (SIGCOMM 2002), pp. 177–190. ACM Press, Pittsburgh (2002)
13. Ferreira, R.A., Ramanathan, M.K., Awan, A., Jagannathan, S.: Search with Probabilistic Guarantees in Unstructured Peer-to-Peer Networks. In: The 5th International Conference on Peer-to-Peer Computing (P2P 2005), pp. 165–172. IEEE Press, Konstanz (2005)
14. PeerSim, <http://peersim.sourceforge.net/>
15. Brite, <http://www.cs.bu.edu/brite/>
16. Shen, H.: EAD: An Efficient and Adaptive Decentralized File Replication Algorithm in P2P File Sharing Systems. In: The 8th International Conference on Peer-to-Peer Computing (P2P 2008), pp. 99–108. IEEE Press, Aachen (2008)
17. Qiao, Y., Bustamante, F.E.: Structured and Unstructured Overlays under the Microscope: A Measurement-based View of Two P2P Systems That People Use. In: The 2006 USENIX Annual Technical Conference (USENIX 2006), pp. 341–355. USENIX Press, Boston (2006)