# Chapter 13
# The Information Discovery Service of Electronic Product Code Network

**Siwei Zheng and Ge Yu**

**Abstract** Aiming at the problem of low query efficiency and delay of Object Naming Service (ONS), an improved Chord algorithm is used in this paper. By introducing an improved Chord algorithm for information discovery service of EPC network, this paper classifies neighboring nodes into a group, stratifies all the nodes according to the different performance, and selects more powerful nodes as the super nodes to manage ordinary nodes in the same region. The results of simulation experiments show that the query hops and network delay are both reduced while query efficiency is improved to a certain extent. Improved Chord algorithm can effectively solve the problem of ONS query bottlenecks, thus the query efficiency in information discovery service of EPC network is improved.

**Keywords** EPC network · Object naming service · Chord algorithm · Super node · Discovery service

## 13.1 Introduction

In 1999, in support of the Uniform Code Council (UCC), the concept of the Electronic Product Code (EPC) was proposed in the automatic identification Laboratory (Auto-ID Labs) [1] established at the Massachusetts Institute of Technology (MIT), whose purpose is to form an open network where worldwide items can achieve real-time information sharing, that is EPC network [2].

EPC network consists of five parts: Electronic Product Code (EPC), Information Identification System (ID System), EPC Middleware, Information Discovery Service, EPC Information Services (EPCIS). In EPC network, the item information

S. Zheng (✉) · G. Yu
Hangzhou Institute of Service Engineering, Hangzhou Normal University, Hangzhou, China
e-mail: sijia67@163.com

W. E. Wong and T. Ma (eds.), *Emerging Technologies for Information Systems, Computing, and Management*, Lecture Notes in Electrical Engineering 236, DOI: 10.1007/978-1-4614-7010-6_13,

is randomly stored in any information server in scattered, so the key issues which EPC Network Information Discovery Service needs to solve is to how to retrieve the relevant information of specific items in the unpredictable information servers efficiently and accurately. At this time it will use Object Naming Service (ONS), which can provide a query service to EPC Information Services according to EPC code, and ultimately realize the users' access to the relevant information. But the existing ONS has some defects: ONS is a resource query service depending on the root ONS server, so when more and more query requests come, the root ONS server will be overloaded. Once the root ONS server collapses, the entire query system will face paralysis, which will severely restrict the query efficiency. But EPC network that will be built in the future will host information sharing of global items, the sheer numbers of query requests that it will face can be imagined. It can be seen, in EPC Network, an efficient information discovery service is particularly important. This paper will propose an information discovery service based on DHT network, and try to solve the disadvantage of low query efficiency and delay in the existing ONS and improve query efficiency of EPC network [3].

## 13.2 P2P Network Search Technology

### *13.2.1 DHT*

Distributed Hash Table (DHT) [4] is a data structure that can effectively achieve storage, management, and query of data in the distributed system or P2P network. Whose principle is that all the nodes in the network is constructed into a hash table, and splited into one by one fragments, then these fragments are stored in physical connected nodes according to certain rules (These rules include Chord, Pastry, CAN, etc.), and each node will maintain a fragment of the entire hash table, which is called a node routing table. In this way, the resource query of entire network is realized by one by one node routing table. This query idea using hash table has become increasingly widespread in network query.

### *13.2.2 Chord Algorithm*

Chord is a resource search algorithm proposed by the Massachusetts Institute of Technology based on DHT. Given a keyword, then the keyword is mapped to a node, usually it is denoted as (k, v), k refers to the keyword identifier, which is obtained through the hash operation of resource keyword; v is called node identifier, which is the hashed value of IP address of the node. In the compatible hash, the keywords are all stored in the position called a successor node—successor (k), which represents the first node of node identifier that is equal to k or immediately

after k. All the identifiers will be arranged in a circle, and the value follows from 0 to $2^m-1$. In order to make two nodes will not hash to the same identifier, usually m must be long enough [5, 6].

Each node has its own routing table (Finger Table) with m entries at most. The i-th item of the node n in the routing table is successor node s with a distance of $2^{i-1}$ at least to node n in the circle, that is, $s = \text{successor}(n + 2^{i-1}) = (n + 2^{i-1}) \bmod 2^m$, $1 \leq i \leq m$ (count modulo $2^m$). Through maintaining the routing table, each node can query any other nodes via information of some nodes in the routing table. And more close from the node in the Chord circle, more information the node will know.

## 13.3 Application of Chord Algorithm in Information Discovery Service of EPC Network

Chord algorithm has the advantage of achieving quick query resource location, but there are still some questions: (1) It does not use all the nodes discriminatively according to the differences of node performance, and makes the high-performance nodes do not be fully utilized, while the low-performance nodes are used overload, which seriously affects the efficiency of the query and the stability of the network. (2) It does not consider the actual physical topology between nodes. Chord algorithm makes each node mapped to a logical space, after the nodes are hashed, the actual physical topology will lose. It causes that the nodes which are close in the physical network may be very far apart after mapped function, and ultimately it will lead to a repeat path and take up network bandwidth.

Based on Chord algorithm, information discovery service of EPC Network takes into account the different performance and actual physical topology of the nodes. According to the difference of node performance, all the nodes are divided to super node layer (Super Node, SN) and ordinary node layer (Ordinary Node, ON). At the same time, the node's IP address is grouped by CIDR (Classless Inter-Domain Routing), and the nodes that are close in the physical address are divided into a group.

### *13.3.1 The Stratification of Nodes*

Through using the idea of Chord algorithm and regarding information servers storing related information of items as independent distributed data nodes. According to the difference of node performance, all the nodes are divided to super node layer (Super Node, SN) and ordinary node layer (Ordinary Node, ON). At the same time because there are a large number of hierarchies in the real world, the division can be match with the actual business model. Super nodes bear the task of

query and transmission, and have three tables: index record table, routing forwarding table and query cache table. Index record table stores a section of keyword index managed by super nodes and the position of the object of the key value in the area centered on super nodes. The requester can quickly locate in the related ordinary node through index record table and obtain the related address information service of data node. Routing forwarding table stores routing information between the super nodes. Query cache table stores node information of recent multiple queries [7].

### *13.3.2 The Grouping and Join in of Nodes*

Through CIDR (Classless Inter-Domain Routing), the node's IP address is grouped. With respect of traditional network protocol address, CIDR can improve the utilization of the IP address better, and can shorten the routing table. CIDR allocates network address based on variable length block, and uses the various length prefix to instead of the network number and subnet number of classification address. In the improved model, the nodes which have the same upper 16 bits of the IP address are divided into a group, and the nodes that are close in the physical address are divided into a group.

When a new node wants to add to the hierarchical network, first, make the comparison on the upper 16 bits of IP address of nodes, if the upper 16 bits already exists, the new node will join the group, and register its own resource information to the super node, and otherwise the new node will create a new group.

### *13.3.3 Resources Release*

(1) Each node in the network uses its own IP address to make hash operation and obtains a node identifier ID.
(2) First, the items to be networked use theirs own EPC codings to make hash operation, and the results are recorded as Key. This Key value maps detailed items information through Chord routing algorithm to the ordinary node that is nearest the node identifier ID.
(3) At the same time, the ordinary node will release resource information to the super node that is responsible for the management of the area. Through Chord routing algorithm, super node build the index record of Key value in the corresponding super node, which is the mapping relationship between the Key value and node Identifier ID.

### *13.3.4 Resources Query*

Once the query request is issued, nodes in the network will raise a query based on Chord routing algorithm:

(1) Once nodes receive the user's query of Key = Hash (EPC encoding), firstly they will research the node that is closest to the Key value in the ordinary nodes in their own area. Then they send the received query request to that request node.
(2) If the resource needed is not found in the ordinary node layer, the node will continue to send query request to the super node in its own area. When this super node receives the query request, it will continue to forward query request to other super nodes.
(3) When a message is sent to a super node, it will initiate a query in the area where the super node is. The super node will first check whether there is an index matched with the Key value in its own query cache table and index record table, if there is, then it will return the resource address of the source query node. Otherwise, the super node will send query request to another super node which is closest to the Key value.
(4) After the super node receives the resource query request, if the matched information with the Key value can be found in the index record table, the query request node will receive the query request issued by the super node. Otherwise, it will return a query failed information.
(5) According to returned query information, the user will directly query related resources of items from nodes storing item information, and obtain detailed item information.
(6) After be successfully queried, the super node in the area where the query is issued will record the Key value of this time and the node information. The benefit of the query is that when other nodes query the same Key value again, the system does not need to find information in the network of other areas, and can directly obtain the Key value and corresponding node's IP address through query cache table. It reduces query hops, and improves the query efficiency. Figure 13.1 shows the model of information discovery service based on Chord algorithm.

## 13.4 Experiments and Analysis

The paper generates the simulation by NS2, and makes a performance compare within ONS query algorithm, Chord routing algorithm and improved Chord routing algorithm used in EPC network from two aspects of average query hops and average query delay respectively.

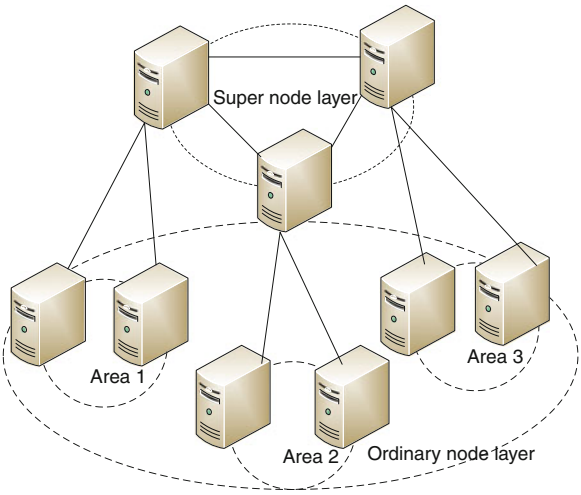


**Fig. 13.1** Information discovery service based on Chord algorithmChord algorithm

The experiment adopts 5 different node numbers (2000, 4000, 6000, 8000, 10000), and each node stores 20 documents. Experiment will be carried out 10 times for each set of values.

As can be seen from Fig. 13.2, with the growth of the number of nodes, the average query hops of Chord algorithm and improved Chord algorithm are also growing, but they do not dramatically increase, and remain in a relatively stable range. The average query hops of Chord algorithm increase in the proportion of $O(\ln N)$ with the increase of the number of nodes $N$. Improved Chord algorithm curve is more flat than Chord algorithm curve, and its average query hops is few. This is because that the node division according to the physical position makes nodes in the same network segment are divided in the same group, and the hierarchical Chord model makes that each super node is responsible for ordinary nodes

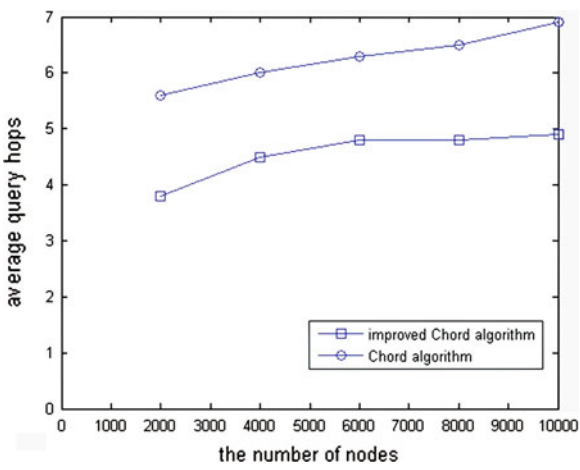


**Fig. 13.2** Comparison of average hops on two models

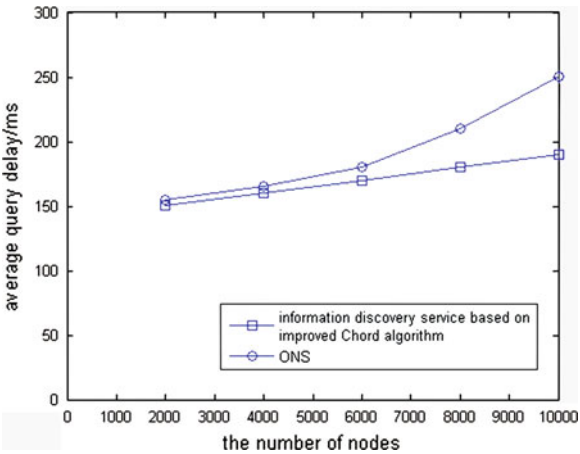


**Fig. 13.3** Comparison of average delay on two models

of the area, so access speed between nodes in an area becomes fast, and it reduces query hops, finally query efficiency has been significantly improved to a certain extent.

As can be seen from Fig. 13.3, the average query delay of information discovery service based on improved Chord algorithm is around 150 ms. With the increase of the number of nodes, the value is in a flat range. This is because when the nodes issue the query request, they are no longer dependent on the central node, but exchange data among the nodes, and grouped node routing reduces forwarding times of the entire network, so the number of nodes have few influence on the query time. The average query delay of ONS gradually increases as the number of nodes increase, this is because the query mechanism of ONS depends on the root servers, once the query request is too large, the request will wait in line, at the same time, ONS will experience two address resolution processes of ONS returning the URI address and DNS returning the IP address, which results in a delay of the query.

## 13.5 Conclusion

Aiming at the problem of overload in ONS root server, this paper presents an improved Chord algorithm of information discovery service. Improved information discovery service has the following advantages: (1) facing the information resolution of a large number of items, it has high performance in load balancing, avoiding the bottleneck problem of root ONS. And average query delay has reduced by 25 % than ONS. It demonstrates that it can meet the massive query needs of EPC network. (2) The nodes are grouped according to the physical topology, discovery algorithm only forwards information among the super nodes, and super nodes and neighboring ordinary nodes constitute an autonomous area,

which all improve query efficiency to some extent. (3) Query cache table in the super node is also conducive to quickly locate to the previous query node information. The items information in the EPC network are related to the commercial interests and privacy, so security protection mechanisms play an important role in information discovery service of EPC network. The next step is to study the security problem of information discovery service.

## References

1. AutoID Labs homepage [EB/OL]. http://www.autoidlabs.org/ (2010)
2. Xiaoli, D, Baoping, Y: ONS service in EPC network. Microelectron. Comput. **22**(2), 17–21 (2005)
3. EPCglobal Inc.V.1.0.1. Object Naming Service (ONS) Standard [S] (2008)
4. Stoica, I., et al.: Chord: a scalable peer-to-peer lookup protocol for Internet applications. IEEE/ACM Trans Netw. **11**(1), 17–32 (2003)
5. Liu H, Cheng L: Research on naming service architecture of internet of things based on DHT. Appl. Res. Comput. **6**(28), 2327–2329 (2011)
6. Yong, H: The summarizing of Chord and improvement algorithms in P2P systems. Comput Knowl. Technol. **8**(1), 38–71 (2006)
7. Liu, D: Study on the information discovery service of Internet of Things based on P2P [D]. Zhengzhou Univercity, Zhengzhou (2011)
8. Minghui, J.: Hierarchical bidirectional Chord. International conference on educational and information technology (ICEIT2010), pp. 486–489 (2010)