

# 基于主机标识的平面路由技术研究

孙冰<sup>1</sup>, 姚楠<sup>2</sup>, 陈佳<sup>1</sup>, 张思东<sup>1</sup>

(1. 北京交通大学电子信息工程学院下一代互联网互联设备国家工程实验室, 北京 100044;  
2. 香港科技大学, 香港)

**摘要:**随着 Internet 的发展, 现有网络将位置与身份绑定的做法, 使得终端在安全性、移动性和扩展性等方面的弊端日渐显现。提出一种新型的基于主机的平面标识路由, 这是一种基于分布式哈希表算法的平面标识路由方式, 支持变长主机标识数据通信。本方案设计的路由机制采用网状拓扑, 可以在接入路由器之间形成良好的物理连接, 为现有各类终端和网络提供了统一接入路由方案, 改善了可扩展性, 在网通层对数据包进行处理, 能够保证终端的灵活出入, 并支持终端的移动。

**关键词:**路由可扩展性; 主机标识; 平面路由

**中图分类号:** TP31

**文献标识码:** A

**文章编号:** 1673-629X(2011)05-0005-04

## Research on Techniques of Host Identifier-Based Flat Routing

SUN Bing<sup>1</sup>, YAO Nan<sup>2</sup>, CHEN Jia<sup>1</sup>, ZHANG Si-dong<sup>1</sup>

(1. National Engineering Laboratory for Next Generation Internet Interconnection Devices, School of Electronics and Information Engineering, Beijing Jiaotong University, Beijing 100044, China;  
2. Hong Kong University of Science and Technology, Hong Kong, China)

**Abstract:** Because of binding the location and identity, the drawbacks which embody the security, mobility and scalability of the terminal in the existing network have gradually become obvious with the development of the internet. This identifier-based routing scheme contributes to all kinds of the existing terminal units and networks. The design of the routing mechanism is based on a mesh topology constituted by access routers. And it can make good physical connections between the access routers, which improve the scalability of the internet significantly. In addition, processing packets via network layer guarantees the flexible accessing and exiting of terminal units, as well as the support for mobility.

**Key words:** routing scalability; host identifier; flat routing

## 0 引言

近些年来, 互联网已经成为了三网(电信网, 计算机网, 有线电视网)中发展较快的网络。随着互联网的发展, 其应用场景发生巨变, 由实验科研化向社会应用化转变, 出现了很多与体系结构密切相关的问题, 使得原来“端到端透明”的互联网核心设计原则遇到了极大的挑战。

在这种情况下, 加之由于各种不同类型终端和网络的接入, 需要为其提供一致的普遍的服务, 原有网络在可扩展性方面的弊端日渐显现, 并体现在网络安全性和移动性等诸多方面, 因此多个国家都开始投入精

力发展一种新型网络的路由方式来改善这种状况。

## 1 路由系统现状分析

### 1.1 路由可扩展性问题

路由系统是支撑整个互联网信息传输的基础。大量研究和测试表明, 目前因特网的路由表条目增长已经超过了线性速率。由于种种原因, 大量非聚合 IP 前缀被插入到路由表中, 这使得路由器需要应用更多的处理和存储资源来维护这些路由信息表项和路由更新报文<sup>[1]</sup>。当前路由条目压缩的主要手段, 就是通过 CIDR<sup>[2]</sup>来进行 IP 前缀聚合, 减少核心路由表的条目数, 但是由于各种限制因素的存在, 目前这种方式已经无法起到明显有效的作用。另外, 随着大规模的 IPv6 地址的应用和部署, 庞大的地址空间会进一步造成路由表的显著增长<sup>[3]</sup>, 极可能会超过当前路由系统的处理能力。核心路由表的膨胀会导致存储空间不足, 路由查找、路由收敛等处理变慢, 直接影响整个互联网的效

收稿日期: 2010-10-08; 修回日期: 2011-01-15

基金项目: 国家 973 重点基础研究发展规划项目 (2007CB307106, 2007CB307101); 中兴通讯产学研合作项目

作者简介: 孙冰 (1984-), 女, 天津人, 硕士研究生, 研究领域为下一代互联网; 张思东, 教授, 研究方向为通信与信息系统。

率和安全。

这些问题的产生和互联网设计中将位置和身份绑定的初始思想密不可分。在当前 Internet 中的 IP 地址,既表示主机身份又表示主机位置,地址分配往往是基于组织结构,而且相对稳定,不会根据网络拓扑改变而动态改变,这样使得单一的 IP 地址命名空间难以同时满足双重角色的要求,从而产生了路由扩展性的问题。由此可见,对 IP 地址的身份和位置双重功能进行分离,已经成为互联网面临的最重要的问题之一。

但是,身份和位置分离技术的关键是对现有协议体系结构核心的修改,这需要对数量庞大的终端进行协议栈的升级,需要引进一个新的复杂的主机身份标识空间,这增加了整个系统的复杂性。

### 1.2 新型网络体系下的路由技术

为了改善现有状况,国内外将开发新型信息网络作为本领域的重要研究内容。如 NewArch 计划及美国国家科学基金委员会(NSF)的 GENI、FIND 等项目都投入了大量精力开展新一代信息网络基础理论研究。我国也积极参与到下一代网络的研究中,提出了一种支持普适服务的一体化可信网络体系结构<sup>[4,5]</sup>,用于实现服务一体化、网络一体化。该体系结构允许包括固定用户<sup>[6]</sup>、移动终端、移动子网、自组网等多种类型的用户接入,各种用户享有同样的个性化业务服务,实现分布式网络资源的共享和查找服务,进行语音、数据、图像等业务的传输,确保向用户提供一致的、普遍的业务<sup>[7]</sup>。

一体化网络中引入了两种标识:接入标识 AID (Access Identifier)代表终端的公开身份信息;交换路由标识 RID (switch Routing Identifier)代表终端的位置信息,并原创性的提出 AID 与 RID 分离聚合映射理论<sup>[8]</sup>。这种身份标识与位置标识分离的技术可以克服传统互联网中由于 IP 地址具有双重性而带来的问题,能够满足新业务对移动性和安全性,以及多种接入方式等方面的需求。

一体化网络从架构上分为两个主要层次:服务层和网通层。在网通层,通过接入网和骨干网的拓扑划分,建立了接入标识和路由标识分离的基础。接入网实现各种类型的终端或者网络的接入,骨干网解决位置管理和路由技术。一体化网络虚拟接入网中的主机接入标识采用无结构化的平面标识,因而传统路由器基于网络前缀进行路由转发的机制在一体化网络虚拟接入模块当中并不适用,需要重新设计基于身份标识 AID 的一种接入网路由机制,因为 AID 本身已经没有承载关于位置的信息,那么传统的路由协议,即基于前缀的网段式路由查找方式,已经不能够适应平面标识路由机制的需求。

虚拟接入网模块见图 1。

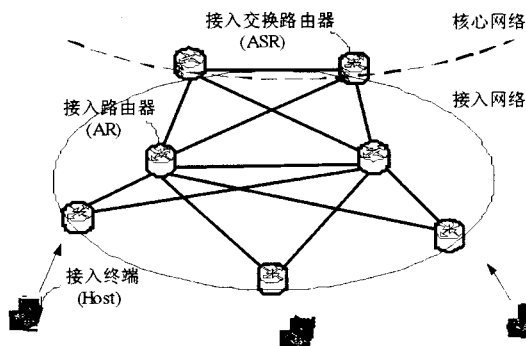


图 1 虚拟接入网模块

## 2 平面路由技术

SEATTLE<sup>[9]</sup> (A Scalable Ethernet Architecture for Large Enterprises)给出了一种在大型以太网当中应用 DHT 算法交换数据包的方法。在以太网中交换相对现有的 IP 路由网络会更加容易配置、实施和管理,但是可扩展性的优越度不如 IP 路由网络。SEATTLE 通过实施平面化的编址结构,保证了大型企业网络的即插即用,并通过最短路径路由和基于分布式哈希算法的主机信息解析机制保证系统的可用性和扩展性。但其工作仅仅停留在交换层面上,没有提供有效的路由转发机制。

ROFL<sup>[10]</sup> (Routing On Flat Labels)给出了一种在互联网中利用平面化主机标识进行编址的平面路由方法。它提出为保证在未来网络中路由体系的可扩展性,路由可以直接基于网络中主机的标识符,这样还可以消除任何路由机制依靠的网络层协议来确定网络的位置,但该方法的全网可扩展性仍有待讨论。

Compact Routing<sup>[11]</sup>对现有平面化路由方法及层次结构路由方式进行分类及总结,讨论了现有互联网路由体系中诸多需要改进的机制,并给出针对 Stretch、Path 等属性的算法评估和证明,但该工作更多的停留在理论层面。

## 3 接入网平面路由系统设计

利用 SEATTLE 的设计基础,设计了接入网主机标识平面路由。这是一种基于分布式哈希表算法 (DHT) 的平面标识路由方式,是一种接入网支持变长主机标识数据通信的方法。

DHT 采用分布式检索的路由算法<sup>[12]</sup>,已成为国际结构化对等网络 (Peer-to-Peer, P2P) 研究和应用的热点。其采用的 Chord<sup>[13]</sup>是 2001 年 MIT 提出的一种分布式查找算法,其核心思想就是要解决在 P2P 应用中遇到的基本问题,即如何找到存有特定数据的节点。

采用这种设计方法有如下优势:首先,可以为现有各类终端和网络提供统一的普遍的接入路由方案,改善了网络的可扩展性;其次,该路由方案能够使终端在网络中灵活的加入或退出,并支持终端在接入网的范围内移动。

一体化网络接入网的平面路由机制见图 2。

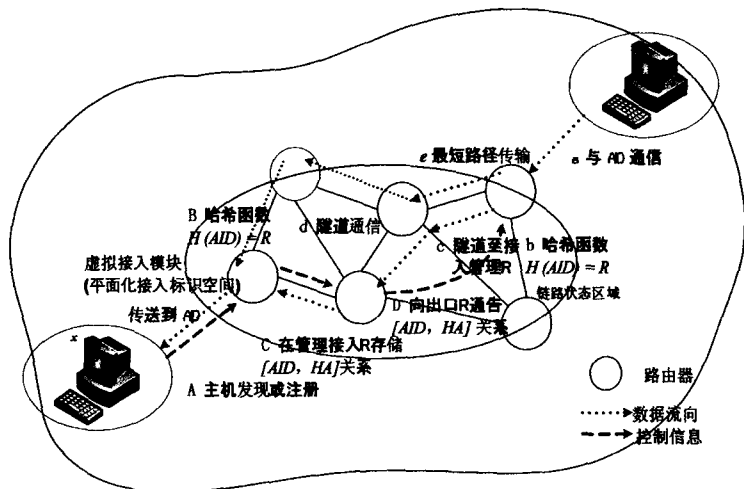


图2 一体化网络接入网的平面路由机制

接入网主机标识平面路由方案与网络拓扑的部署有紧密的关系。不同拓扑在路由机制的设计和评估中占有决定性的因素。本方案设计的路由机制采用接入路由器构成较具代表性的全网状拓扑连接的虚拟接入模块。这样,接入路由器之间可以形成良好的物理连接,为接入路由器之间的通信提供有力支持。

该路由方案概述如下:

### 3.1 注册过程

主机首次进入接入网或者从其他位置移动到新的网关接入路由器时,需要由当前网关接入路由器代理主机进行主机-位置信息注册(见图3)。注册过程如下:

- (1) 主机向网关接入路由器发送自身接入标识 AID。
- (2) 网关接入路由器接收到主机发送的首次接入信息, 将该主机 AID 保存在自身管理的主机列表及本地路由表中。
- (3) 网关接入路由器通过哈希函数  $H$  生成管理接入路由器位置信息, 即  $H(AID) = R$ 。并将主机-位置对应关系 [主机接入标识, 网关接入路由器] ([Host Access ID, Hosting Access Router]) 发送至管理接入路由器  $R$ , 进行注册。

(4)管理接入路由器 R 接收到主机-位置对应关系[主机接入标识,网关接入路由器]([Host Access ID,Hosting Access Router]),将该对应关系保存在自身的管理关系列表当中,并为发送注册信息的网关接入路由器回复 ACK 确认信息。

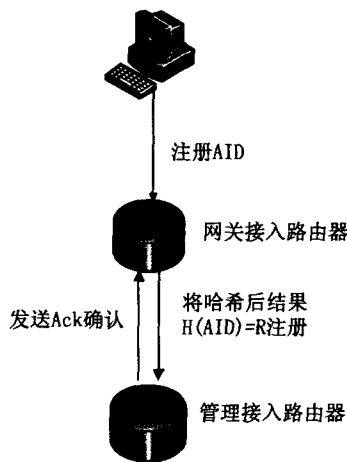


图 3 主机注册过程

### 3.2 转发过程

接入路由器在接收到任意数据包时需要做如下处理完成路由转发：

- 流向  
制信息
- (1)若接收到的数据包来自用户接口,则跳转到(2);若接收到的数据包来自路由器接口,则跳转到(3)。
- (2)以数据包的目的 AID 为索引,查找本地路由表,若命中,则转发至相应接口,转发过程结束;若未命中,则通过哈希函数  $H$  生成管理接入路由器位置信息, $H(AID) = R$ ,并将数据包以隧道方式转发至管理接入路由器  $R$ 。转发过程结束。
- (3)若数据包来自隧道,则跳转到(4);若数据包不是来自隧道,以数据包目的 AID 为索引,查找本地路由表,若命中,则转发至相应接口;若未命中,则丢弃该数据包。转发过程结束。

(4)将隧道数据包解封装,以数据包原 AID 为索引,查询本地路由表,若命中,则转发至相应接口;若未命中,则查询管理关系列表,若命中,则以隧道方式转发至相应网关接入路由器,若未命中,则以隧道方式转发至接入交换路由器(Access Switching Router, ASR)。转发过程结束。

### 3.3 退出过程

(1) 当某接入路由器从网络中退出时,通过链路状态路由协议使接入网络中其他接入路由器从邻居列表中清除该路由器,并调整其间的路由可达性和逻辑连接状态维护;

(2)接入网络中其他接入路由器将退出的路由器节点从本机路由器邻居列表中清除；

(3) 接入网络中其他接入路由器遍历自身的本地管理主机-位置关系列表,将原本由退出节点负责管理的主机-位置对应关系重新注册到退出路由器节点的后继路由器上。

#### 4 仿真结果分析

选取北京交通大学作为一个虚拟接入网,使用其 IPV4 网的地址段(59.64.x;202.112.x;192.168.x),将所有的 IPV4 前缀均化为 32bits 的全平面标识符,在这样全平面的接入标识环境下,网络中心的拓扑图为依据,设定网络中存在的路由器节点个数(假设所有可路由的节点均参与到平面路由当中),应用哈希算法 SHA-1 对 IP 地址进行处理,测量每个路由器能分到的路由条目(见图 4)。

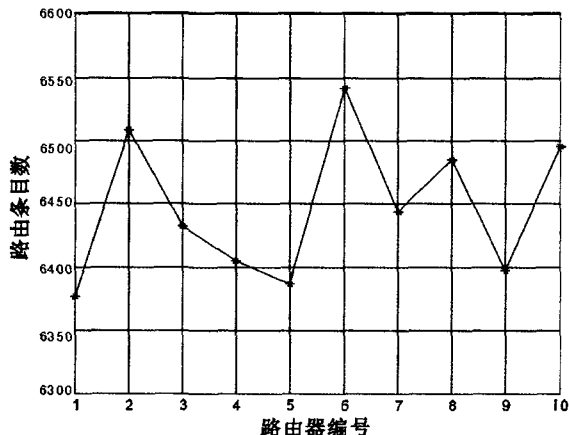


图 4 路由器存储路由条目的分布

可知,各个路由器需要存储的路由条目个数差别不大,平均为 6400 dests / router,负载较为均衡。

以上情况是假设接入网内的接入路由器 ID 等距分布,在实际情况下,路由器 ID 可能不会均匀分布,每个路由器的存储条目会根据其 ID 值而变化。现变换路由器 ID 间距,路由器 1 的间距依次变大,路由器 6 的间距次第减小,观察其路由条目数,测试结果见图 5。

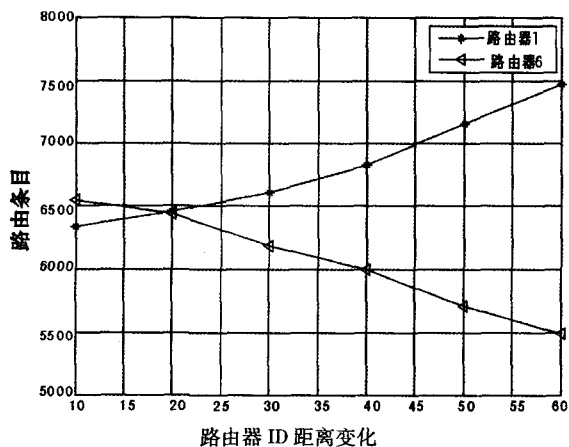


图 5 路由器 ID 距离变化后路由条目的分布

可以看出,路由器标识符的选取机制会对路由表条目的分配造成一定影响,当路由器 ID 距离较近时,可以存储的路由条目则较少,反之,则存储条目较多。

传统基于前缀的路由方式中,核心路由表的前缀数量大概维持在 30 ~ 40 dests / router,而平面化后每

个路由器所负责的路由条目远远大于这个值,因此在今后的工作中可以考虑平面标识路由与传统前缀路由共存,优化该路由机制。

在现有的仿真环境中,附加上其他外围机制,包括数据包发送、路由表生成、路由表更新等操作。当外围节点开始加入这个网络,触发其接入路由器的消息处理机制,完成节点的注册过程。然后在接入网中转发信令消息数据包,每经过一跳路由器,便加入相应的路由信息,记录该数据包在网络中所经历的路径,见图 6。

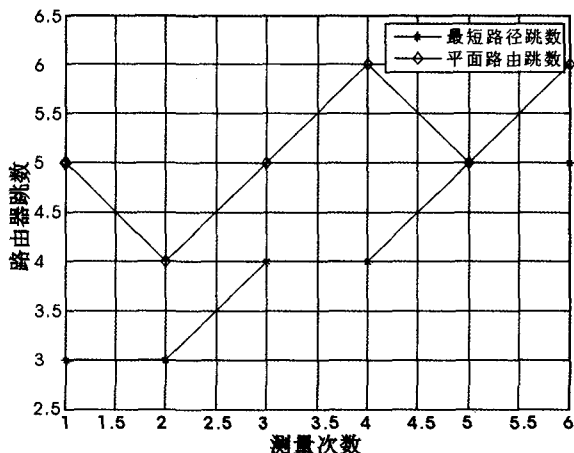


图 6 数据包经历路由跳数与传统最短路径比较

由图 6 可以看出,虽然数据包所经历的跳数不小于最短路径,但差别不大,相比平面路由所带来的优势,可以忽略不计。

#### 5 结束语

文中阐述了基于主机标识的平面路由方式,并进行了仿真测试。该设计克服了现有网络的路由可扩展性和移动性差的问题,能够适应虚拟接入网中平面化路由机制。在今后的工作中,可以考虑将平面化路由和基于前缀的平面路由相结合,使该设计得到优化。

#### 参考文献:

- [1] 中国互联网网络发展状况统计报告[R]. 中国互联网络信息中心(CNNIC),2010.
- [2] Verkaik P, Broido A, claffy KC, et al. Beyond CIDR Aggregation[R]. [s.l.]: CAIDA, 2004.
- [3] Meyer D, Zhang L, Fall K. Report from the IAB Workshop on Routing and Addressing[S]. IETF RFC 4948, 2007.
- [4] 张宏科, 苏伟. 新网络体系基础研究——一体化网络与普适服务[J]. 电子学报, 2007(4): 593-598.
- [5] 董平, 秦雅娟. 支持普适服务的一体化网络研究[J]. 电子学报, 2007(4): 599-606.
- [6] 范岩, 张思东. 身份与位置分离映射机制下固定子网接入的研究[J]. 计算机技术与发展, 2010, 20(10): 30-34.

(下转第 12 页)

表 4 Lymph 数据集中 normal 类的实验结果

| Percent | Recall | Precision | F-measure |
|---------|--------|-----------|-----------|
| 50%     | 0      | 0         | 0         |
| 100%    | 0.5    | 1         | 0.667     |
| 150%    | 0.5    | 0.5       | 0.5       |
| 200%    | 0.667  | 1         | 0.8       |
| 250%    | 0.75   | 1         | 0.857     |
| 300%    | 1      | 1         | 1         |

表 5 Lymph 数据集中 fibrosis 类的实验结果

| Percent | Recall | Precision | F-measure |
|---------|--------|-----------|-----------|
| 50%     | 0      | 0         | 0         |
| 100%    | 0      | 0         | 0         |
| 150%    | 0.75   | 1         | 0.857     |
| 200%    | 0.8    | 1         | 0.889     |
| 250%    | 0.9    | 1         | 0.947     |
| 300%    | 1      | 1         | 1         |

## 4.2 结果分析

通常认为影响稀有类分类的重要因素是数据分布的不平衡性,也就是说对于稀有类问题,普通的分类算法往往失效,但文中实验结果表明,数据分布的不平衡性影响稀有类分类的一个因素,在特定的类比率下,使样本规格变大,普通的分类算法往往也可以取得很好的分类结果。

## 5 结束语

文中对稀有类分类问题进行了研究,分析了影响稀有类分类问题的因素,探讨了稀有类分类的评估标准。针对影响稀有类分类的一个因素:样本规格的大小进行研究,在同等类分布比率下,改变样本规格的大小,在 weka 平台下进行实验,得到数据集中稀有类的 recall、precision 和 F-measure 值。实验结果表明,在特定的类比率下,使样本规格变大,普通的分类算法往往也可以取得很好的分类结果。同时也说明,数据分布的不平衡性只是影响稀有类分类的一个因素,即使数

据分布的极不平衡,通过增加样本中稀有类实例的数目(类比率不变),也可以提高稀有类分类的各个指标。

文中的实验基于 Rotation Forest 进行,下一步的工作是,扩大实验的范围,基于多个集成分类器,例如:Bagging、Boosting 等,对比多个集成分类器下的实验结果。

## 参考文献:

- [1] Han J, Kanber M. 数据挖掘:概念与技术[M]. 范明,孟小峰,译. 北京:机械工业出版社,2001.
- [2] Sun Yanmin, Kamel M S, Wong A K C, et al. Cost-sensitive boosting for classification of imbalanced data[J]. Patter Recognition, 2007, 40: 3358-3378.
- [3] 刘艳霞, 职为梅, 杨亮. 稀有类分类问题研究[J]. 微型机与应用, 2005(6): 54-56.
- [4] 职为梅, 范明. 稀有类分类问题探讨[J]. 计算机技术与发展, 2010, 20(7): 250-253.
- [5] 职为梅, 范明. 利用基本显露模式两阶段分类稀有类[J]. 微机发展(现更名:计算机技术与发展), 2005, 15(12): 44-47.
- [6] Visa S, Ralescu A. Issues in mining imbalanced data sets - a review paper[C]//in Proceedings of the Sixteen Midwest Artificial Intelligence and Cognitive Science Conference. [s. l.]: [s. n.], 2005: 67-73.
- [7] Kuncheva L I, Rodriguez J J. An experimental study on Rotation Forest ensembles[C]//MCS 2007, Lecture Notes in Computer Science. Berlin: Springer, 2007: 459 - 468.
- [8] 范明, 魏芳. 挖掘基本显露模式用于分类[J]. 计算机科学, 2004, 31(增刊): 207-309.
- [9] Fan H, Ramamohanarao H. Bayesian Approach to Use Emerging Patterns for Classification[C]//Proc. of 14th Australasian Database Conference. [s. l.]: [s. n.], 2003.
- [10] Rodriguez J J, Kuncheva L I, Alonso C J. Rotation Forest: a new classifier ensemble method[J]. IEEE Trans. Pattern Anal. Mach. Intell., 2006, 28(10): 1619-1630.
- [11] Blake C, Merz C. UCI repository of machine learning databases [M]. Irvine, CA: University of California, Department of Information and Computer Science, 1998.

(上接第 8 页)

- [7] 张宏科. 实现一体化网络中普适服务的方法: CN, 2006 10169727.8[P]. 2006.
- [8] 张宏科. 一种实现一体化网络服务的体系结构: CN, 200510134579.1[P]. 2005.
- [9] Kim C, Caesar M, Rexford J. Floodless in SEATTLE: A Scalable Ethernet Architecture for Large Enterprises[C]// SIGCOMM'08. Seattle, Washington, USA: [s. n.], 2008.
- [10] Caesar M, Condie T, Kannan J. ROFL: Routing on Flat Labels [C]// SIGCOMM'06. Pisa, Italy: [s. n.], 2006.
- [11] Krioukov D, Claffy K, Fall K, et al. On compact routing for

the Internet[J]. ACM SIGCOMM Computer Communication Review (ccR), 2007, 37(3): 41-52.

- [12] Ratnasamy S, Shenker S. Routing algorithms for DHT: Some open questions [C]// Proceedings of IPTPS. Boston, MA: [s. n.], 2002: 278-293.
- [13] Stoica I, Morris R, Liben-Nowell D. et al. Chord: A Scalable Peer-to-Peer Lookup Protocol for Internet Applications[J]. Proceedings of the 2001 conference on applications, technologies, architectures, and protocols for computer communications, 2001, 11(1): 149-160.