

|                                     |   |
|-------------------------------------|---|
| P.O.Box 8718, Beijing 100080, China | Journal of Software, Sept. 2004,15(9):1375-1384                     |
| E-mail: jos@iscas.ac.cn             | ISSN 1000-9825, CODEN RUXUEW, CN 11-2560/TP                         |
| http://www.jos.org.cn               | Copyright © 2004 by The Editorial Department of Journal of Software |

## PeerIS:基于Peer-to-Peer的信息检索系统

凌波, 陆志国, 黄维雄, 钱卫宁, 周傲英

[Full-Text PDF](#) [Submission](#) [Back](#)

凌波<sup>1,2</sup>, 陆志国<sup>1</sup>, 黄维雄<sup>3</sup>, 钱卫宁<sup>1</sup>, 周傲英<sup>1</sup>  
<sup>1</sup>(复旦大学 计算机科学与工程系, 上海 200433)  
<sup>2</sup>(中国浦东干部学院, 上海 200233)  
<sup>3</sup>(新加坡国立大学 新加坡-麻省理工联盟, 新加坡)

作者简介: 凌波(1974—), 男, 江西赣州人, 博士, 主要研究领域为P2P计算, 分布式数据库, 信息检索, 信息经济; 陆志国(1978—)男, 硕士, 主要研究领域为P2P计算, 数据库系统实现技术; 黄维雄(1973—), 男, 马来西亚人, 博士, 主要研究领域为数据库性能, P2P计算, 基于Internet的应用; 钱卫宁(1976—), 男, 博士, 主要研究领域为P2P计算, 海量数据挖掘, 流数据管理与分析; 周傲英(1965—), 男, 博士, 教授, 博士生导师, 主要研究领域为对等计算, Web数据管理, 数据挖掘, 流数据管理与分析.

联系人: 凌波 Phn: +86-21-65643024, E-mail: lingbo@fudan.edu.cn, <http://www.cs.fudan.edu.cn/wpl>

Received 2003-11-29; Accepted 2004-02-03

### Abstract

In this paper, the emerging P2P computing is first briefly introduced, including its distinct features, potential merits and applications, and the problems from which the existing P2P-based data sharing systems are suffering are further point out. To address these problems, the concept of P2P-based information retrieval is proposed, which can not only exploit the potential merits of P2P to overcome the problems of traditional information retrieval systems (e.g., lacking of scalability), but also achieve fully semantic retrieval and sharing in the context of P2P systems. Based on the ideology, PeerIS, a P2P-based information retrieval system is developed. Then, the architecture of PeerIS and its peers' components are presented. The key issues of implementation are described, including communication, semantics-based self-reconfiguration, query processing and self-adaptive routing mechanisms, are also described. Finally, an experimental study is used to verify the advantages of PeerIS.

Ling B, Lu ZG, Ng WS, Qian WN, Zhou AY. PeerIS: A peer-to-peer based information retrieval system. *Journal of Software*, 2004,15(9):1375~1384.

<http://www.jos.org.cn/1000-9825/15/1375.htm>

### 摘要

介绍了对等计算(peer-to-peer,简称P2P)的特征、潜在优势和应用范围,指出了当前P2P数据共享系统存在仅支持弱语义(甚至缺乏语义)和粗粒度(文件水平)共享等局限性.针对这种现状,提出了基于P2P的信息检索,既可充分发掘P2P技术的潜在优势,克服传统信息检索系统的可伸缩瓶颈等问题,又可实现P2P数据共享系统语义丰富和细粒度的信息检索与共享;并开发出PeerIS:基于P2P的信息检索系统.描述了PeerIS的整体构架与节点的内部结构;重点阐述了PeerIS的通信机制、自配置机制、查询机制以及自适应路由机制等实现关键技术;并用实验证明了PeerIS的优异性.

基金项目: Supported by the National Nature Science Foundation of China under Grant No.60373019 (国家自然科学基金); the High Education Doctorial Subject Research Program of MoE of China under Grant No.20030246023 (国家教育部博士点基金资助); the Science and Technology Commission of Shanghai Municipal Government of China under Grant No.03DZ15028 (上海市科委重大研究项目)

### References:

[1] Ng WS, Ooi BC, Tan KL. BestPeer: A self-configurable peer-to-peer system. In: Chrysanthis PK, ed. Proc. of the 18th ICDE. San Jose: IEEE Computer Society Press, 2002. 272.

[2] Anderson d. Peer-to-Peer: Harnessing the Benefits of a Disruptive Technology. O'Reilly & Associates, Inc., 2001. 67~76.

[3] ICQ Home Page. <http://www.icq.com/>

[4] Groove Home Page. <http://www.groove.net>

[5] Napster Home Page. <http://www.napster.com/>

[6] Gnutella Development Home Page. <http://gnutella.wego.com/>

[7] Druschel P, Rowstron A. PAST: A large-scale persistent peer-to-peer storage utility. In: Elphinstone K, ed. Proc. of the HotOS VIII. Schoss Elmau: IEEE Press, 2001. 65~70.

[8] Rowstron A, Druschel P. Storage management and caching in PAST: A large-scale persistent peer-to-peer storage utility. In: Chateau Lake Louise, ed. ACM SOSP 2001. Banff: ACM Press, 2001. 188~201.

[9] Kalnis P, Ooi B, Papadias D, Tan K. An adaptive peer-to-peer network for distributed caching of olap results. In: Ramakrishnan R, ed. ACM SIGMOD. Madison: ACM Press, 2002. 25~36.

[10] Ng WS, Ooi BC, Tan KL, Zhou A. Peerdb: A p2p-based system for distributed data sharing. In: Dayal U, ed. Proc. of the 19th ICDE. Bangalore: IEEE Computer Society Press, 2003. 633~644.

[11] Salton G, Lesk ME. Computer evaluation of indexing and text processing. Journal of the ACM, 1968,15(1):8~36.

[12] Yang Y, Liu X. A re-examination of text categorization methods. In: Callan J, ed. Proc. of the ACM SIGIR. Berkeley: ACM Press, 1999. 42~49.