
基于 P2P 的语义检索系统设计

The design of a P2P-based semantic retrieval system

马勇 王松 李玉卿 刘晓光 王刚

Ma Yong, Wang Song, Li Yuqing, Liu Xiao-guang, Wang Gang

(南开-百度联合技术实验室, 南开大学信息学院, 天津, 300071)

(Nankai-Baidu Joint Lab, College of Information Science, Nankai University, Tianjin, 300071)

摘要: P2P 是近年来网络研究领域的热点。当前 P2P 网络的研究多集中在文件共享的应用, 其检索机制只支持基于关键词的查询, 缺乏对语义检索的支持。本文研究内容是将语义网技术和 P2P 的优点结合起来, 建立 P2P 网络的语义检索机制。本文通过建立基于本体概念的分布式倒排索引, 使检索过程不再是关键词的精确匹配, 而是通过不同节点本体中的概念之间的语义关系的逻辑推理实现检索请求与文档在语义上的匹配。实验表明, 本文提出的结构化 P2P 网络语义检索方法, 比基于关键词精确匹配的检索方法有较高的查全率和查准率。

Abstract: In recent years, P2P network is one of hot research fields. Related works of P2P mainly focus on file-sharing application. It only supports the retrieval of the query based on keywords, and is lack of support for semantic search. In this paper, a P2P-based semantic retrieval system is presented. By using distributed inverted index based semantic network, retrieval is based on ontology located at different nodes rather than keywords. The experiments show that P2P-based semantic retrieval system has higher performance compared with retrieval based on keywords.

关键词: 面对等网; 语义网; 检索

Keywords: P2P network; Semantic network; Retrieval

中图法分类号: TP391

文献标识码: A

1. 引言

近年来, 随着 Internet 的飞速发展, 以及存储介质容量的不断扩充, 数据和信息呈现出爆炸式的增长趋势。传统的采用 C/S 模式的网络已经远远不能满足各类用户的需求。对等网 (P2P) 技术就是为适应需求的变化而提出的一种新技术。目前, P2P 网络已经被广泛应用于多个领域。典型应用包括文件共享、协同工作、分布式计算以及实时通信等。但人们要想从海量信息中检索到自己想要的信息并不容易。这种困难具体表现为: 搜索引擎的智能性不高, 基本上还是采用关键词匹配的办法; 计算机不能理解概念, 不能进行语义关联和推理。为了解决这一问题, 研究人员提出语义网的概念: 语义网即能够描述事物间的关联关系, 又包含了一部分语义信息, 有利于机器的自动处理。简单的说, 语义网就是要给 Web 上的信息加上注释——本体。其中的资源不单包含信息本身, 还包括信息的真正含义, 从而提高信息处理的自动化、智能化水平。本体为计算机理解信息提供了基础, 在描述交换信息的结构和语义等方面扮演着重要角色。

本文主要工作是研究在基于 DHT 的结构化 P2P 网络中, 如何进行有效的基于语义的检索机制。将语义网技术, 特别是本体与 P2P 系统相集成, 即充分利用结构化 P2P 网络的可扩展性, 负载均衡等优点, 又通过本体对领域知识的描述来增强对共享文档的语义描述能力, 克服现有结构化 P2P 系统支持弱语义(甚至缺乏语义)的问题。在实现文档基于本体的语义描述及节点本体映射方法的基础上, 通过构建基于本体概念的分布式索引实现 P2P 网络的语义检索。

基金项目: “863” 计划课题(2008AA01Z401); 教育部博士点基金(20070055054); 天津市科技发展计划 (08JCYBJC13000)

作者简介: 马勇(1985—), 男, 南开大学计算机系硕士生, 从事 P2P 语义网络检索系统研究。

王松(1983—), 男, 南开大学计算机系硕士生, 从事基于本体的 P2P 语义网络检索研究。

王刚(1974—), 男, 南开大学计算机硕士生导师, 计算机学会会员, 会员号: E20-0006617M

通讯地址: 天津南开大学计算机系, 300071; Tel:022-23504780; E-mail:nkdier@hotmail.com

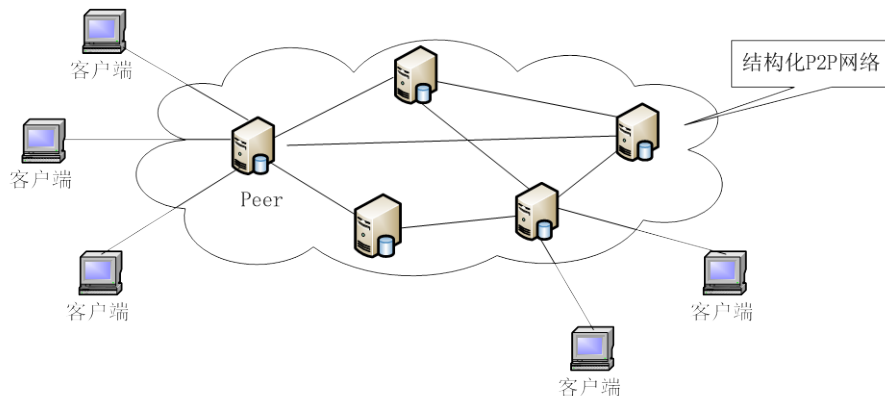
2. 相关工作

目前 P2P 系统经过了近十年的发展, 已经从第一代为中心索引拓扑和二代的无结构化体系结构发展到了现今广泛采用的第三代结构化体系结构。中心索引式的 P2P 系统如 Napster^[1], 由中心服务器存储数据索引信息, 节点通过中心服务器查询得到数据存储的节点, 然后直接和存储节点通信。二代无结构化体系结构有 Gnutella^[2], KaZaA^[3]。以 Gnutella 为例, 查询请求通过洪泛的方式广播给所有的邻居, 直到满足查询或超时, 其优点是无单点失效, 缺点是带宽消耗大, 可扩展性差。第三代结构化 P2P 系统的基本思想是采用分布式哈希表(distributed hash table, DHT)进行数据组织和检索。具体系统实现包括 Chord^[4]、CAN^[5]、Pastry^[6]、Tapstry^[7]等。同非结构化 P2P 网络相比, 基于 DHT 的结构化 P2P 网络具有很好的可扩展性、负载均衡、以及容错性, 可以应用在对可靠性和扩展性要求比较高的场合和更大规模的网络环境中。

在基于 DHT 的结构化 P2P 网络中, 利用基于关键词划分的分布式倒排索引可以实现多关键词检索, 但是这种多关键词检索方式存在两点不足: 1) 检索结果需要在不同节点之间移动, 当文档规模增大或检索关键词增多时会造成网络负载过大; 2) 检索只支持词汇之间的精确匹配, 使得检索方法只支持弱语义, 甚至缺乏语义。这是因为基于 DHT 的 P2P 网络通过对关键词的哈希变换实现资源的存储或定位。针对这种情况, 研究领域倾向于使用语义网技术, 特别是本体来解决 P2P 中出现的语义问题。文献[8][9]提出了基于本体的 P2P 系统 HyperCub。HyperCub 采用超立方的拓扑结构组织节点, 使得节点间具有较小的网络路径, 有限的网络连接, 并且避免了查询信息的盲目广播。Bibster^[10]用于在分布式的 P2P 环境下进行文献检索。Bibster 将 ACM topic hierarchy 作为系统中的全局本体, 其中的 1287 个词汇表示为概念, 用 subTopic 和 seeAlso 关系建立一个树状分类结构。每个节点根据查询主题与节点相似度决定应该将查询转发到哪些节点。但这类方法存在一个明显的共同特点, 就是系统节点必须共享并遵循一个全局本体。这同时也造成了网络不具有可移植性, 可扩展性也很差。

3. 基于 P2P 的语义检索系统结构

语义检索是与传统关键词检索相区别的, 对检索条件、信息组织以及检索结果显示赋予一定语义成分的一种新的检索方式。对检索请求赋予语义, 是指对用户各种输入形式(例如: 表单、关键词、或者自然语言), 通过加入特定的背景信息或者语义信息, 让计算机能够更好地理解用户输入, 使用户和计算机在对检索入口的理解上达到更大一致性; 对信息组织赋予语义, 是指在建立索引时, 不再局限于以关键词作为索引入口, 而是以语义实体、语义属性或语义关系作为切入点, 使信息检索过程具有很强的推理分析能力。可以通过两种方式对信息组织赋予语义: 一种方式是保留了传统的文档, 通过建立文档和一系列语义实体、语义属性和语义关系的联接, 给文档索引赋予语义。同时, 与关键词具有一定的权重相类似, 语义实例在文档中也具有一定的权值; 另一种方式是文档将不再是语义信息的基本单位, 取而代之的是与现实世界模型相对应的语义实体、语义属性和语义关系, 这些语义实体、语义属性和语义关系作为结构化的信息存储在特定的数据对象中, 文档检索被规约为完全形式化的概念匹配。对检索结果赋予语义, 是指检索的结果将不再是只包含检索关键词的文档, 而是在语义上扩展了检索结果空间, 包含那些语义关联的文档。这里的语义关联是指在某一领域或者应用中, 具有重大意义的与用户所感兴趣的概念之间的联系。



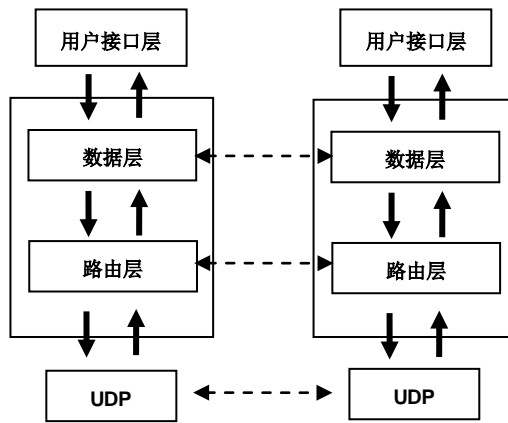


图 1 系统网络拓扑图

图 2 P2P 系统层次结构图

本文对结构化 P2P 网络基于本体的语义检索方法进行了研究。图 1 为整个系统的网络拓扑结构。系统结构采用了带有超级节点的 P2P 网络。超级节点（即图 1 中的 Peer 节点）之间基于 Chord 覆盖网协议构成了结构化 P2P 网络。每个客户端节点连接到某个超级节点，通过发现节点在概念之间的语义关系，实现检索请求在语义层次的匹配，在超级节点共享本地资源或检索系统中的资源（文档）。对检索请求的语义匹配结果通过检索关键词进行优化，将最相关的结果首先返回给用户。

4. P2P 系统的具体实现

本文所使用的 P2P 网络系统是基于结构化的 Chord 覆盖网协议实现。它的层次结构图如图 2 所示，自底向上可以分成四层：UDP 层，路由层，数据层，用户接口层。

4.1 P2P 系统层次介绍

系统的最底层是 UDP 层。节点之间的通信数据量小但频繁，因此节点之间的数据通信采用 UDP 协议。

路由层是整个 P2P 结构的核心层，它的功能主要包括 P2P 网络维护，数据转发路由，为上层提供数据发送和接收接口。网络维护包括处理节点加入和离开，以及后继列表和 finger 表的维护。

1) 路由机制。在系统中，通过哈希函数映射节点 IP 到一个 ID 空间内，关键字也映射在同样的 ID 空间内。这样的空间为 $[0, 2^m)$ (m 为标识符位数)，首尾相接形成一个环。每个节点都维护着前驱节点、后继列表和 Finger 路由表。前驱节点是指从节点开始逆时针方向遇到的第一个节点；后继列表维护着从节点开始顺时针方向遇到的 k 个活跃节点的 ID 和 IP 信息；Finger 路由表有 m 项，第 i 项记录 ID 距离至少为 2^i 的第一个活跃节点的 ID 和 IP 信息。在路由过程中，节点查询自己的后继节点列表和 Finger 表，获得离目的节点最近的前驱节点，将这个节点作为下一跳的地址，这样迭代转发最终到达目的节点。

2) 网络维护和 Finger 表刷新。P2P 网络动态性高，节点的加入和离开频繁。为了维护网络的稳定，节点定时发送稳定数据包检测第一后继节点是否存在，后继节点收到检测包后要回送应答包；后继节点没有收到前驱节点的检测包，则认为前驱节点失效，前驱节点没有收到后继节点的应答包，则认为后继节点失效。节点发现后继节点失效后，将第一后继节点从后继列表中删除，然后向新的第一后继节点发送稳定数据包。同时为了保证 Finger 表的有效性，节点也需要定时的刷新 Finger 表项。节点定时的选择 Finger 表中的一项，查询它最新的后继节点。

3) 节点的加入。新节点加入时必须先向网络中已知的一个节点发送加入请求；节点接收到加入请求后将它转发到新节点的第一后继节点，第一后继节点接收到加入请求后，将新节点作为自己的前驱节点，同时向新节点回送应答自己的前驱节点和后继列表信息；新节点得到应答消息后，计算自己的前驱节点和后继列表，这样新节点完成加入。

4) 节点的离开。节点的离开有两种情况：主动离开和失效。节点主动离开时，它向前驱节点发送离开请求，前驱节点收到请求后，将节点从自己的后继列表中删除；节点失效时，不会发送离开请求，系统通过稳定检测也能很快恢复。

数据层是数据的存储和管理层。数据层负责接收和处理用户的发布、查询请求。数据的存储采用轻量级的数据库 BerkeleyDB。它接收到发布请求后，从中提取出相关信息，存入到数据库中，并通过路由层向

发起者回送应答；它接收查询请求后，在数据库查询得到结果，然后将结果通过路由层发送给发起者。

用户接口层提供的接口有两个：发布信息 and 查询信息。发布信息时，先通过数据层向目的节点发送发布请求，然后等待数据层接收到的发送成功应答。查询信息时，先通过数据层向目的节点发送查询请求，然后等待数据层接收到的查询结果。这样，发布信息和查询信息就像在本地执行的一样。

4.2 数据冗余设计

现有 P2P 系统中面临的一个重要问题是：节点频繁的加入和离开，导致整个系统的高动态性，数据的可用性大大的降低。为了提高数据可用性，我们在数据层加入了数据冗余设计。

常用的数据冗余方法可以分成两种：副本冗余和容错编码冗余。副本冗余就是在整个系统中放置一定数目的完全副本，其中一个副本失效后，可以通过其它的副本恢复。副本放置的方式有很多种，最常见的是基于后继列表的放置方式。相关的研究分析表明^[11]，在副本冗余中，数据的可用性与节点的生存时间分布有关；节点频繁的加入会降低副本的可用性。容错编码冗余就是将要存储的数据切分成 M 块，然后利用容错编码将其编码成 N 块 ($N > M$)，通过其中任意的 M 块都能恢复原始数据。

我们采用了两种方法结合的方法，先通过容错编码，将编码后的数据发送到网络中；然后对编码后的数据采用副本冗余，将副本放置在节点后继列表中。主节点接收到发布数据的请求后，除了将编码数据存放在本节点，还要把数据的副本发送给后继列表中的节点。如图 3 所示，源数据被切分层 M 块，经过容错编码成 N 块；编码后的数据块通过哈希运算后，发送到 P2P 网络中的节点，以图 3 中编码后的最后一块为例，它被发送到主节点 (Δ 所示节点)，同时被备份到主节点的后继节点列表上 (\circ 所示节点)。这样主节点失效后，后继节点能继续提供服务，提高了数据可用性。

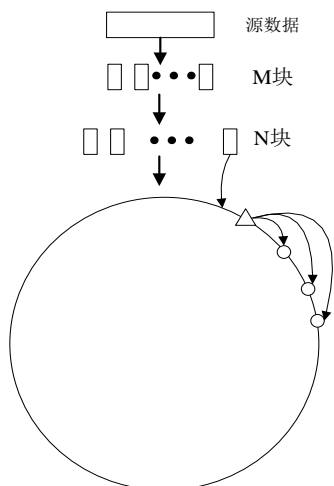


图 3 数据冗余设计图

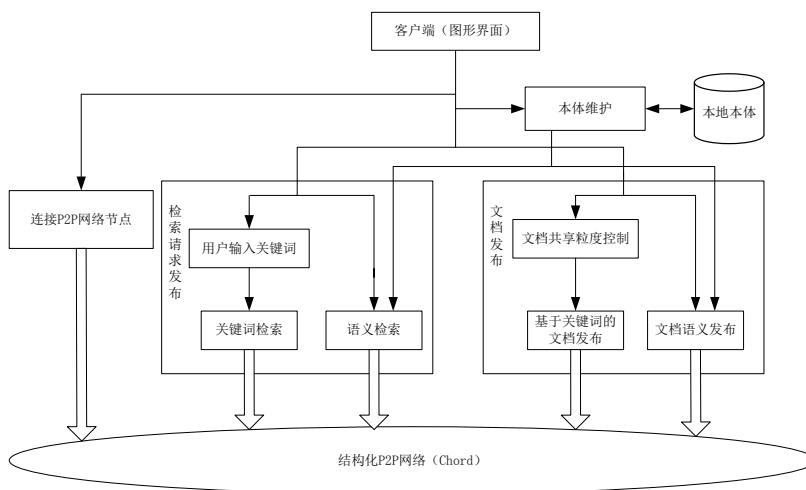


图 4 语义检索系统结构图

5. 语义检索系统的实现

在本文的系统中，语义检索系统是作为 P2P 网络的上层应用来实现的，如图 4 所示。语义检索系统它可以划分成本体操作，文档发布，检索请求发布三个模块。客户端程序在连接 Chord 节点服务器之后可以进行文档的发布和文档的检索。发布和查询文档的最大的特点是增加了语义发布和语义检索功能。它们的基础之一就是对信息组织赋予语义：即在建立索引时，不再局限于以关键词作为索引入口，而是以语义实体、语义属性和语义关系作为切入点。为此，本文提出以本体概念作为索引入口，通过分布式哈希表在结构化 P2P 网络中构建分布式倒排索引，使内容相近的文档对应的文档描述信息（在本文即为文档描述符）通过概念的哈希被存储到同一个节点，形成在语义上的聚集。同时，在目标节点存储概念对应的文档描述信息，而不仅仅是文档列表，那么可以直接在节点本地完成文档的匹配，无需再通过其它的网络节点，节省了网络带宽。

系统实现的关键技术之一是使用了文档描述模型来对文档进行标注，它包括文档语义模型（语义层）和文档统计模型（语法层）。文档语义模型 $SM = \{C, L\}$ ，C 是表示文档主题，表明文档涉及的主要领域知

识，即表示文档所属的知识分类，L 表示文档语义描述，描述了文档的内容，由本体概念和关系构成。这样通过引入领域本体对文档内容进行描述，展现了文档的语义特征；文档统计模型是应用向量空间模型（VSM），通过关键词对文档内容的描述，描述了文档在语法方面的信息，表示了文档的统计特征。同时系统应用到了文献[12]提出的 LMATCH 本体映射方法，对异构本体进行了同构操作，实现了异构本体检索的互操作。

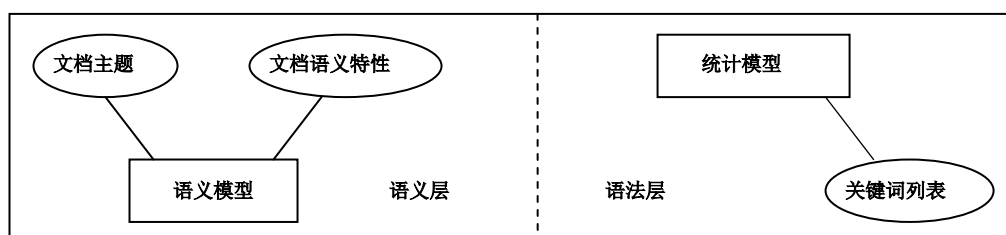


图 5: 文档描述模型

5.1 本体操作

针对本体操作本文提出了两种方案：第一，应用普林斯顿大学的 WordNet^[13]进行关键词的发布和检索。通过系统调用 WordNet 词库,应用上下位关系和整体部分关系对文档发布时的描述关键词以及检索时的关键词进行概念的提取，以 Synset 集概念作为最终关键词进行相关的发布和检索操作。第二，由专业领域内专家应用本体描述语言 OWL^[14]构建 RDF^{[15][16][17]}本体文件，应用相关的语义推理机推理出概念之间的联系（这里使用的是 Pellet^[18]），形成一棵本体语义树结构，用户可以根据自己理解的某文档的语义，在已构建本体树中选择语义最贴近的节点概念作为该文档的语义描述信息，以它作为发布和检索的关键词。

5.2 文档发布

文档的发布可以根据需求选择基于 WordNet 方式发布或是基于构建的本体文件进行发布。前者，把提取的语义集合概念以及本地地址整合成元数据信息，以得到的语义集合概念作为哈希关键字，把元数据信息发布到网络中去；后者，在本体模型中选择描述该文档的主题和语义信息，加上本地地址整合成元数据，以文档的主题为哈希关键字，再把元数据发布到网络中去

5.3 检索请求发布

文档的检索可以选择基于 WordNet 方式检索或基于构建的本体文件进行检索。前者，把所要查找的信息发布到网络中，其中是以所要查找的关键字在 WordNet 中提取的 Synset 集合概念作为哈希关键字在分布式结构化 P2P 网络中进行查找；后者，需要根据本体描述需求文档的主题、包含的语义信息、关键字等信息，以文档的主题为哈希关键字，把所查找的信息发布到网络中。最终网络返回给用户结果列表，用户可以选择相应的项目进行文档下载。

6. 实验测试

本节对本文给出的基于本体的 P2P 网络语义检索方法进行了仿真实验。实验是在一台 PC 机上模拟完成的。其中以文件表示节点来模拟 P2P 网路系统，整个系统由 100 处理能力相同的节点组成。实验中，使用康奈尔大学和华盛顿大学关于科学艺术和工程的课程分类作为节点本体^[19]，并在其中增加了连接关系，如“teachBy”、“teachFor”、“prerequisiteWith”等。每个本体具有不同的概念分类层次，分别包含 124 个和 137 个概念。以本体概念名称作为关键字，以其哈希结果数值 hash (concept) 作为文件名构建文本文件（称为节点文本文件）。每个文本文件表示 P2P 网络系统中的一个节点。如果多个概念的哈希结果数值相同，那么表明它们属于同一节点，并将对应的数据存放于同一个文本文件中。

实验对本文提出的基于本体的结构化 P2P 网络语义检索方法 semantic search 与原始的基于关键词倒排索引的结构化 P2P 网络多关键词检索方法 original search，以及文献[20]提出的混合倒排索引的结构化 P2P 网络多关键词检索方法 hybrid search 进行了比较。

实验的目标是检验语义检索方法的检索结果质量和对网络资源(特别是带宽资源)的占用情况。为此，实验采用了查全率和查准率以及检索流量作为评价标准^[21]。查全率是指检索完成时返回给用户的结果中，

相关文档数量（即文档描述信息的数量）与系统中总的相关文档数量的比例。查全率越高就越能满足用户的需求，特别是当用户不能明确地定义检索需求时，高的查全率能保证用户找到满足要求的答案。查准率是指检索完成时，相关文档数量与检索返回的全部文档数量的比例。查准率表明了检索的有效性。检索流量主要是指检索结果在网络中不同节点之间的移动造成的网络流量，通过计算与检索请求相匹配的文档描述信息的数量与描述信息的字节数的乘积得到。

实验中使用不同个数的关键词构建查询，三种检索方法的查全率和查准率的比较如图 6 和图 7 所示^[12]，其中对 semantic search 检索的语义匹配结果未进行优化。从图中可以看出 semantic search 方法具有较高查全率和查准率，并且关键词个数对查全率和查准率的影响较小。original search 和 hybrid search 方法完全是通过关键词的字面信息匹配而获得结果，作为其检索结果的文档必须包含检索关键词，因此，检索关键词的数量会影响检索结果，也就是说检索关键词越多，检索的查全率和查准率越高。但是，而 semantic search 方法通过对相同或相近主题的文档进行抽象的概念层次上的匹配而获取结果，并不是局限于关键词的字面信息。因此，其检索结果的查全率和查准率要高于 original search 和 hybrid search。对于 semantic search，决定其基于语义的检索结果的主要因素是表示检索语义描述的描述逻辑表达式。而检索关键词数量越多，相应的描述逻辑表达式会越复杂，并且越能明确表示期望文档的语义。因此，检索结果的查全率和查准率越高。

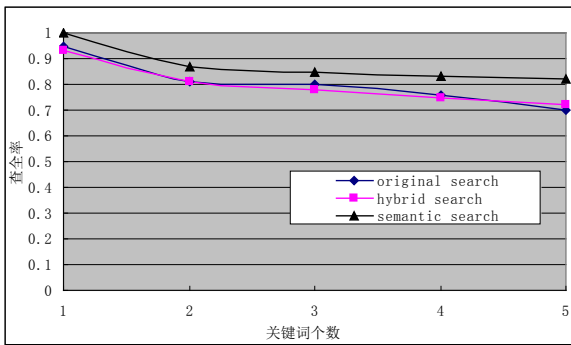


图 6 三种检索方式查全率比较

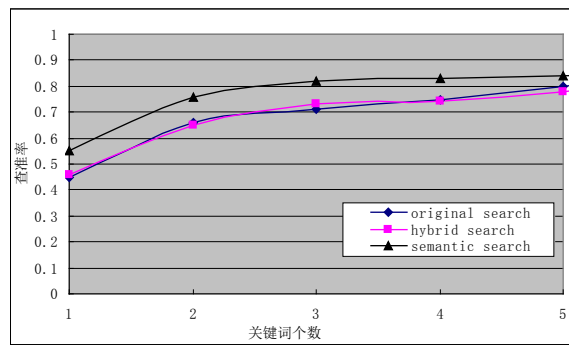


图 7 三种检索方式查准率比较

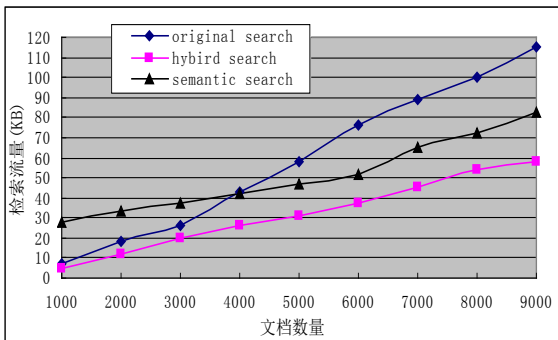


图 8 不同检索方式产生的检索流量与检索关键词个数的关系系统

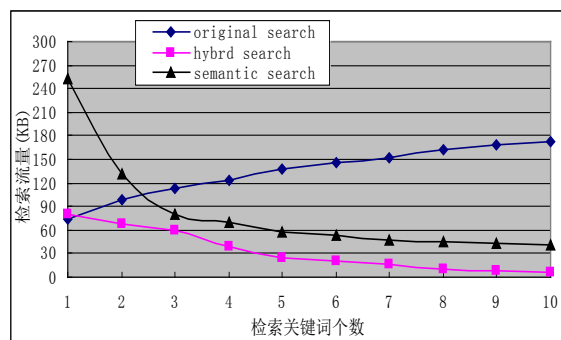


图 9 不同检索方式产生的检索流量与中文档数量的关系

同时，我们也对三种检索方法在检索过程中产生的网络流量进行对比分析，如图 8 和图 9 所示^[12]。从图中可以看出 original search 在检索过程中造成较大的检索流量，并且其检索流量随检索关键词个数的增加而增大。造成这种现象的主要原因是：original search 为了获取检索结果，需要在不同节点之间对单个关键词的匹配结果进行“与”操作。hybrid search 和 semantic search 方法检索流量比较接近。这是因为：二者的倒排索引中不仅包含了文档标识，同时还增加了其它文档描述信息，对于 hybrid search 增加了文档关键词集合，对于 semantic search 在关键词集合的基础上又增加了文档语义信息，因此在目标节点即可获取到与检索请求匹配的最终结果，避免了中间匹配结果在不同节点之间移动。

7. 结束语

本文提出了一种在结构化 P2P 网络上基于本体的语义检索方法：通过引入本体来捕获词汇的语义以及词汇之间的语义关系，并且建立基于本体概念的分布式倒排索引。这样，检索过程不再是关键词的精确匹配，而是通过不同节点本体中的概念之间的语义关系的逻辑推理实现检索请求与文档在语义上的匹配。实验表明，本文提出的结构化 P2P 网络语义检索方法，比基于关键词精确匹配的检索方法有较高的查全率和查准率。

在后续工作中，我们将着眼于建立各种面向专业领域的本体，利用这些本体实现更为精确的语义检索系统。工作的难点在于，如何引导和帮助领域专家准确地建立相关的本体，并真正发挥出基于本体的语义检索能力。

参考文献

- [1] The Napster Homepage. <http://www.napster.com>,2002.
- [2] The Gnutella Homepage. <http://gnutella.wego.com>,2002.
- [3] KaZaA file sharing network, <http://www.kazaa.com/>, 2002.
- [4] Stoica I, Morris R, Karger D, et al. Chord: a scalable peer-to-peer lookup service for Internet applications.[C] In: Proceedings of the ACM SIGCOMM '01,149-160 2001.
- [5] Ratnasamy S, Francis P, Handley M, et al. A scalable content-addressable network. [C] In ACM SIGCOMM.161-172, 2001.
- [6] Rowstron A and Druschel P. Pastry: Scalable decentralized object location and routing for large-scale peer-to-peer systems.[C] In: Proceedings of 18th IFIP/ACM International Conference on Distributed Systems Platforms (Middleware'01). 329-350, 2001.
- [7] Zhao B Y, Kubiawicz J, et al. Tapertry: An infrastructure for wide-area fault-tolerant location and routing.[R] U.C. Berkeley Technical Report UCB/CSD-01-1141, April, 2001.
- [8] Debajyoti Mukhopadhyay , Sounak Chakravorty , Sudarshan Nandy. An Algorithm for Hyper-cube-based Ontology Construction .In: 10th International Conference on Information Technology , 2007
- [9] Habib Rostami , Jafar Habibi , Emad Livani. Semantic partitioning of peer-to-peer search space .In:Computer Communications 32 .619-633,2009
- [10] Haase P, Broekstra J, Ehrig M, et al. Bibster-A semantic-based bibliographic peer-to-peer system. In: McIlraith SA, Plexousakis D, Harmelen FV, eds. Proc. of the ISWC 2004, Berlin, Springer-Verlag, 2004. 122~136
- [11] Guangping Xu, Gang Wang, Jing Liu: A Hybrid Redundancy Approach for Data Availability in Structured P2P Network Systems. [C] In: Pacific Rim Dependable Computing(PRDC) .195-199,2007
- [12] 马文辉.P2P 网络基于异构本体的语义检索机制的研究. [D] 天津：南开大学,2008
- [13] Kruse P M, Naujoks A, Roesner D, et al. Clever search: A wordnet based wrapper for internet search engines. In: Proceedings of the 2nd GermaNet Workshop, 2005
- [14] Dean M, Schreiber G, Bechhofer S, et al. OWL web ontology language reference.[S] W3C recommendation, 2004.
- [15] <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/> ,0:00~24:00 download
- [16] <http://www.w3.org/TR/2004/REC-rdf-schema-20040210/>,0:00~24:00 download
- [17] <http://www.w3.org/TR/2004/REC-owl-features-20040210/>,0:00~24:00 download
- [18] Bijan Parsia, Evren Sirin. Pellet: An OWL DL reasoner.[C] In: Proceedings of the Third International Semantic Web Conference Poster. 51-53, 2004.
- [19] <http://pages.cs.wisc.edu/~anhai/wisc-si-archive/>,0:00~24:00 download
- [20] Tang C Q, Dwarkadas S. Hybrid global-local indexing for efficient peer-to-peer information retrieval.[C] First Symposium on Networked Systems Design and Implementation. 16-16 ,2004.
- [21] Wenhui Ma, Guangping Xu, Gang Wang, Jing Liu. Detecting Semantic Mapping of Ontologies with Inference of Description

Logic.[C] In: Proceedings of the Ninth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD).393~398,2008.