

基于 Super-P2P 的分布式知识管理模型^{*}

李志国^{1,2} 冯永¹ 钟将¹ 叶春晓¹

(重庆大学计算机学院 重庆 400030)¹ (上海宝信软件西南研发中心 重庆 400041)²

摘要 根据第二代知识管理的思想,以及现在企业、组织的分散式管理特点,论文在分析了多种 P2P 结构的基础上,将 Super-P2P 结构应用于知识管理,提出了基于 Super-P2P 结构的分布式知识管理模型,以解决集中式知识管理遇到的问题。接着,对模型的知识俘获机制以及所提供的知识服务进行了研究。通过应用案例分析,基于 Super-P2P 的分布式知识管理模型可以实现更高层次的知识共享,提供高质量、高可靠性的知识服务。

关键词 知识管理, Super-P2P, 分布式系统, 小世界

Distributed Knowledge Management Model Based on Super-P2P

LI Zhi-Guo^{1,2} FENG Yong¹ ZHONG Jiang¹ YE Chun-Xiao¹

(College of Computer Science of Chongqing University, Chongqing 400030)¹

(Southwest Research Center of Shanghai Baosight Software Corporation, Chongqing 400041)²

Abstract In terms of ideas of the second generation of knowledge management and traits of distributed management in modern enterprises or organizations etc, a variety of P2P structure is analyzed in the paper, and Super-P2P structure is applied to semantic grid. Distributed knowledge management model based on Super-P2P is presented in the paper to address problems of centralized knowledge management, and then, knowledge capture and knowledge services of the model are reported in the paper. Application case study demonstrated that the distributed knowledge management model based on Super-P2P can achieve a better knowledge sharing, and provide high-quality, high-reliability knowledge services.

Keywords Knowledge management, Super-P2P, Distributed system, Small-world

1 引言

知识管理^[1] (Knowledge Management) 指综合利用信息技术以实现知识的俘获、存储和重用。知识管理的基本目的是保存过去俘获的知识,使适当的人员在适当的时机能快捷地查询到以适当样式表示的所需知识,帮助解决工作实践中遇到的问题^[2]。2002年7月,知识管理大师 Dr. Larry Pursesak 提出了“第二代知识管理”的概念,探讨了第二代知识管理的概念和内涵,他们认为应重视非正式组织的沟通,鼓励面对面的接触,强调人与人的联系,建立信赖的环境,同时提供学习的空间,以分享并创造隐性知识。

为了更好地进行知识管理,其中最关键的环节是设计一个有效的知识库机制,以方便知识的获取、共享、存储及重用。目前国内外很多学者在这方面取得了很大的成就^[3~5]。但是设计集中式的知识库管理系统需要耗费大量的人力、物力,而且需要很长的开发周期,开发后随着各部门、人员信息的频繁变动,知识库会越来越难以维护。另外,随着经济全球化,跨国公司,虚拟经营等经济模式的转变,导致这种分散性组织中的每个个体,如各分公司、各部门、甚至个人都有可能有自己的知识库。因此,为了满足现有分散性组织的发展模式,提出一种有效的分布式知识管理模型是知识管理的发展趋势之一。这种方法有利于不同的参与者各自维护自己的知识库结构,同时又保证知识的共享和重用。

目前,peer-to-peer(对等网,简称 P2P)^[6] 技术因具有发挥闲散资源利用率的特征而得到业界的关注。P2P 改变了传统服务器/客户端的集中式服务思维模式,提出将所有节点看作对等节点(peer),它们既可以是服务的提供者,也可以是服务的接受者,拥有相同的地位。与集中式服务模式相比,P2P 结构在可扩展性方面具有天生的优势,通过增加资源或者节点的组织和功能以实现资源共享。

基于上述分析,将 P2P 技术应用于分布式知识管理,有助于解决集中式知识管理所遇到的问题,提供高质量、高可靠性的知识共享服务。

2 P2P 网络模型分析

从拓扑结构和技术上 P2P 网络可分为纯分散式 P2P 网络、混合 P2P 网络和超级 P2P(Super-P2P)网络^[7]。

2.1 纯分散式 P2P 网络

纯分散式 P2P 网络模型的拓扑结构如图 1 所示。

纯分散式 P2P 网络中没有服务器,链状的节点之间构成一个分散式网络。通过基于对等网协议的客户端软件搜索网络中存在的对等节点,节点之间不必通过服务器,可直接建立连接。这种 P2P 网络模型优点在于允许用户设定自己的规则和建立自己的网络环境;为与 Internet 合作,提供近似的即插即用特性;不仅能够在 Internet 下有效地工作,而且对于 LAN 和 Intranet 也非常有用。存在的问题^[6]:由于没有中心

^{*} 基金项目:浦东新区科技发展基金 PKK2005-07;国家发改委科学研究计划项目 2005-2137。李志国 博士,主研方向:知识管理、知识发现;冯永、钟将 博士;叶春晓 副教授。

管理者,网络节点难以发现;不易管理且安全性较差。此种类型的 P2P 网络如:Gnutella、FreeNet 等。

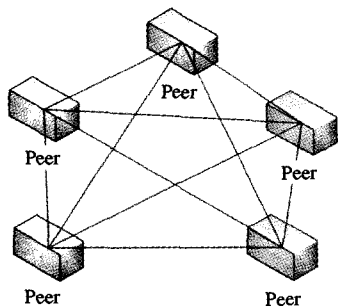


图 1 纯分散式 P2P 拓扑结构

2.2 混合 P2P 网络

混合 P2P 网络模型的拓扑结构如图 2 所示。

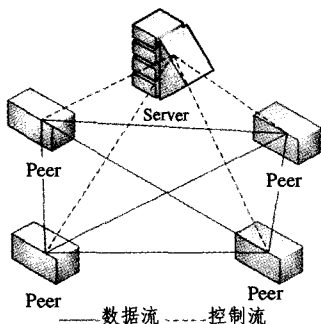


图 2 混合 P2P 拓扑结构

混合 P2P 网络各节点之间可以直接建立连接,但网络的构建需要服务器,通过集中认证,建立索引机制。然而这里的服务器仅用于辅助对等节点之间建立连接,一旦连接成功,服务器不再起作用,对等节点之间直接进行通信。这不同于 C/S 模式中的服务器,也可以认为是弱化了服务器的作用。这种 P2P 网络模型和纯分散式 P2P 网络相比,易于发现网络节点,易于管理且安全性较好,但也有类似 C/S 模式的缺陷,由于服务器瓶颈的限制,规模不易扩大,一旦中央服务器失效,系统无法正常工作。其代表是 Napster。

2.3 Super-P2P 网络

Super-P2P 网络模型的拓扑结构如图 3 所示。

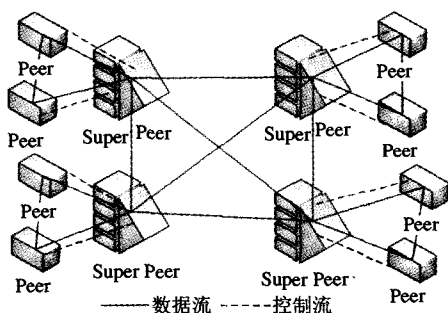


图 3 Super-P2P 拓扑结构

Super-P2P 网络中的各 peer 节点在本地 Super-peer 的控制下直接建立连接,Super-peer 一般由几台 peer 节点互联构成,主要具有 3 个职能:(1)辅助本地对等节点之间建立连接,通过集中认证,建立索引机制;(2)同时作为本地的服务提供者和服务接受者,参与通信;(3)与其它 Super-peer 节点构成对等结构,负责本地 peer 节点与其它区域 Super-peer 节点或

peer 节点的通信。

Super-P2P 网络模型具有如下优点:(1)相对纯分散式 P2P 网络,具有更强的可控性,网络节点易于发现,同时 Super-peer 节点提供了安全服务;(2)相对于混合 P2P 网络,具有更强的容错性,由于 Super-peer 节点是几台 peer 节点构成的,只要不是全部节点失效,Super-peer 节点仍然能够提供服务;(3)具有良好的可扩展性,只要适当增加 Super-peer 的规模,则整个 Super-P2P 网络的规模都可扩充。

3 基于 Super-P2P 结构的分布式知识管理模型

要实现各个分散性组织间的知识共享和复用,必须提供一套集知识检索,知识交换和共享等功能于一体的机制。论文提出了一种基于 Super-P2P 结构的分布式知识管理模型。该模型假定分布式知识管理中的每个节点都是一个独立的知识管理系统(简称为 KMS),由知识俘获引擎、控制模块、安全模块和知识库组成。整个系统强调各知识管理系统间的交互与协作,拓宽了知识管理的广度,同时各子系统间又便于开发与维护。

3.1 Super-P2P 分布式知识管理系统结构

基于 Super-P2P 的分布式知识管理系统结构如图 4 所示。

如图 4 所示,Super-P2P 分布式知识管理系统由多个知识管理系统互联构成 Super-P2P 结构,每个知识管理系统主要具有如下角色:(1)整个结构中的一个 Super-peer 节点;(2)本地用户的知识服务提供者;(3)其它地区 Super-peer 节点的知识服务提供者;(4)其它地区 Super-peer 节点的知识服务接受者;(5)对本地用户节点有一定的控制作用,如增加、删除用户节点;(6)其它地区 Super-peer 节点的候选者;(7)维护管理本地知识库。

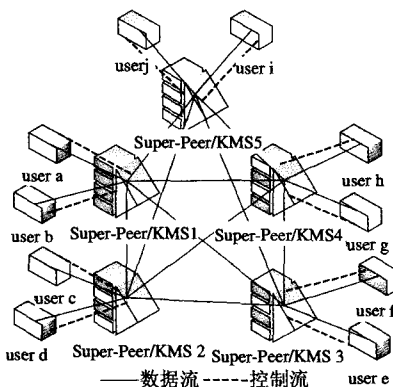


图 4 Super-P2P 分布式知识管理系统结构

相对于一般的 Super-P2P 结构,具有如下区别:(1)本地一般用户不再作为 peer 节点,只接受来自本地知识管理系统的服务。这样有两个好处,其一是减轻了一般用户的负担;其二是增强了一般用户的安全性。(2)知识管理系统构成 P2P 结构,互为知识服务提供者和接受者。由于知识管理系统具有较强的服务功能、良好的容错性和可靠性以及增强的安全服务措施,因此使得整个分布式系统能够更可靠、更安全地提供知识服务。(3)一旦知识管理系统失效,那么距离最近的知识管理系统充当本地的知识服务提供者,进一步保证了整个分布式知识管理系统提供知识服务的能力。

3.2 Super-P2P 分布式知识管理系统的知识俘获机制

已经有研究成果表明利用小世界现象构造算法可以提高

P2P 系统中的搜索效率^[8]。

Watt 和 Strogatz 提出的模型(简称 WS 模型)^[9]是一种常用的 Small-World 模型: n 个节点分布在一个圆环上,初始状态时,每个节点有 k 个连接,分别连向最近的 k 个点。然后,依次调整各节点的连接,以概率 P 随机地改变连接的终端,但避免连向节点本身。

记 $D(i, j)$ 为节点 i 和 j 之间的最短距离,平均路径距离 L 的计算公式如下:

$$L = \frac{1}{n(n-1)/2} \sum_{1 \leq i, j \leq n} D(i, j)$$

当时, $P \approx 0, L \rightarrow n/2k$ 此时网络拓扑呈规则状态;当 $0.001 < P < 0.01$ 时, $L \rightarrow \ln n / \ln k$, 此时节点不仅与相邻节点存在连接,还与远距离节点建立了少数的快捷连接,正是这些快捷连接有效缩短了 L ,使整个网络呈现出 Small-World 特征。当搜索消息沿小世界环路传递时,其平均路径长度为 $O(n)$ 。

利用 Small-World 的思想在 Super-P2P 分布式知识管理系统中的知识俘获过程描述如下:

Step1: 搜索时请求节点首先根据要搜索的知识建立知识标识 K_ID , 检索本地知识管理系统 A 的知识库, 如果匹配则 A 返回结果; 否则 A 到路由表寻找距离其最近的直接连接节点 B;

Step2: B 检索本地的知识库, 如果有匹配项, 则返回搜索结果; 否则 B 返回本地搜索失败信息;

Step3: A 同时向其它所有直接连接节点发送搜索请求, 若搜索成功, 则返回结果;

Step4: 否则 A 向其它所有已建立快捷连接的远程节点发送搜索请求, 若搜索成功, 则返回结果; 否则, 向用户返回搜索失败信息。

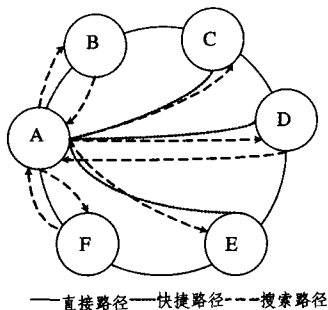


图 5 基于 Small-World 的 Super-P2P 分布式知识俘获

图 5 是搜索原理示意图。如图 5 所示, A 节点首先向距离最近的直接相连节点 B 发出搜索请求, B 返回失败信息; A 继续向直接相连节点 F 发出搜索请求, F 返回失败信息; A 同时向其已建立快捷连接的节点 C、D 和 E 发出搜索请求, 最后 D 返回成功结果。

4 Super-P2P 分布式知识管理系统的知识服务

Super-P2P 分布式知识管理系统的知识服务体系如图 4 所示。服务被分成 3 个层次:

(1) 存储层: 在本地语义网络节点中存储领域知识和其它语义网络节点的知识索引。

(2) 服务支持层: 提供知识访问服务、知识表示服务、语义映射服务以及知识发现与匹配服务。整个知识服务层运行在 XML+RDF+Ontology 之上, XML+RDF+Ontology 构成知识管理系统之间以及知识管理系统和用户之间相互理解的基础。

础。

(3) 用户层: 与服务支持层通过基于 OWL 的查询和操作消息来进行通信, 它从用户端接受知识查询请求, 将其表示为 OWL 消息模式, 然后通过知识服务支撑环境发送消息给知识服务处理器。

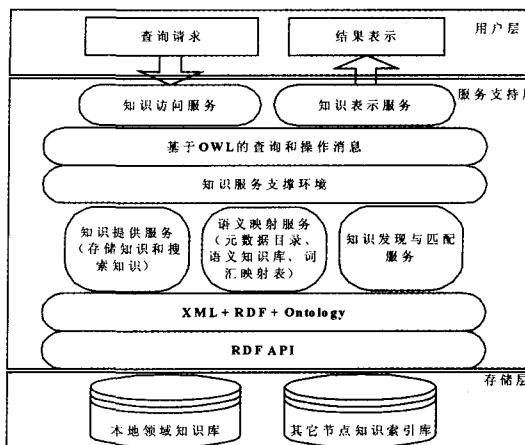


图 6 基于 Super-P2P 语义网络的知识服务体系

4.1 知识访问与表示服务

用户使用知识访问服务接口向本地知识管理系统发送请求, 并使用知识表示服务显示从本地知识管理系统返回的结果。知识访问与表示服务是知识管理系统的高层服务。

4.2 知识提供服务

知识提供服务从知识访问服务中获取请求, 然后转换为 OWL 格式, 并在本地知识库中搜索本地存储的知识, 将得到的结果返回给用户。用户调用知识表示服务显示知识。

4.3 语义映射服务

将不同的局部本体映射为全局本体非常重要, 这个服务被称为语义映射服务, 它在背景知识和以下组件的基础上使用本体映射引擎执行映射服务^[10]。

(1) 元数据目录: 包括一些信息源的描述, 如语义、存储路径、类型和供应者等。

(2) 语义知识库: 包括理解本体概念和属性所需要的知识, 如同义词、中英文比较等。

(3) 词汇映射表: 为了获取一致的属性含义和数据描述视图, 有必要定义一个全局词汇和本地 RDF 的相同语义描述表。

4.4 知识发现与匹配服务

当用户向知识管理系统发送请求时, 知识发现与匹配服务利用 3.2 节描述的方法进行知识发现与匹配。

5 应用案例

假设一个大型软件公司各个分公司都分别建立了各自的知识管理系统, 采用 Super-P2P 结构形成分布式知识管理系统, 现总公司需要获得各个分公司关于某段时间内某一类型软件产品在各个地区的销售情况, 那么整个分布式系统的处理机制如下:

Step1: 首先知识提供服务从知识访问服务中获取用户请求, 转换为 OWL 格式, 并提供产品属性和产品销售属性, 产品属性主要包括产品类型、研发部门、产品名称、研发日期等, 产品销售属性包括销售日期、销售地点、销售量等;

(下转第 196 页)

件为 $(\theta_{fp} > \pi/2$ 或 $\theta_{fp} < -\pi/2)$ 且 $(\omega_{fp} > \pi/2$ 或 $\omega_{fp} < -\pi/2)$ 且 $d_{fp} > V_r$ 。



图6 人工鱼集群行为的动画实现画面

4 实现效果

下面的画面是采用 Microsoft Visual C++ .net 作开发平台,辅助使用 3DSMax 6.0 作为建模工具,用 Direct3D9 作为渲染工具,创作出来的动画效果。



图7 鱼群逃逸动画实现画面

其中,图6被圈中的的是不满足集群条件的人工鱼。图7

被圈中的的是捕食者—鲨鱼。

结束语 本文从复杂系统的角度研究了人工鱼群的自组织行为,建立了鱼群的多 Agent 系统模型,并在此基础上建立了鱼群的觅食、逃逸模型。但是只简单地考虑到影响个体鱼集群的几个因素,事实上还有很多生理、心理、遗传等因素需要考虑,同时人工鱼群作为整体,还存在着个体间通信、合作捕食关系等。这些都将是后续需要研究的内容。随着复杂系统和基于认知的人工鱼高级行为规划研究的深入,鱼群自组织行为研究的理论基础也将不断丰富,在这些理论的指导下我们将完善鱼群系统,生动地展现海洋鱼群运动、生活。



图8 鱼群觅食动画实现画面

参考文献

- 1 Shaw E. Fish in Schools. *Natural History*, 1975,84(8):40~46
- 2 班晓娟,曾广平,涂序彦. 基于自学习的人工鱼感知系统设计与实现. *电子学报*, 2004, 32(12)
- 3 Reynolds C W. Flocks, Herds, and Schools: A Distributed Behavioral Model in Computer Graphics, 21(4) (SIGGRAPH 87 Conference Proceedings) pages 25~34. <http://www.red.com/cwr/boids.html>
- 4 Tu Xiaoyuan. *Artificial Animals for Computer Animation: Biomechanics, Locomotion, Perception, and Behavior*, dissertation for doctor degree, 1999

(上接第 186 页)

Step2:知识提供服务进行如下处理。1)根据知识提供服务提供的产品属性值和产品销售属性值,在本地知识库中搜索本地存储的知识,然后将查询结果以一定的格式反馈给用户。2)如果本地反馈的信息不能满足用户的需求,则调用知识发现与匹配服务,通过论文提出的知识俘获机制获取相关知识,并在各个局部本体中进行匹配;

Step3:根据语义映射服务提供的元数据目录、语义知识库和词汇映射表,更新各个不同的局部本体与全局本体之间的映射关系,同时更新本地的其它节点知识索引库,为今后同类知识的需求和获取提供便利;

Step4:知识表示服务将最终查询结果,以一定的格式反馈给用户。

以上过程描述了 Super-P2P 分布式知识管理模型在具体应用中的处理方式,其中一些处理细节,以及知识反馈给用户的表示形式,目前已经有篇文章对其进行研究和论述,不属于论文的研究重点,因此不进行详细的论述。

结论 为了更好地促进知识的交流和共享,论文将 P2P 技术应用于知识管理,提出一种基于 Super-P2P 结构的分布式知识管理模型,以解决集中式知识管理遇到的问题,并对模型提供的知识服务进行了研究和论述。通过应用案例分析,

Super-P2P 分布式知识管理模型促进了分散性组织间的知识交流和知识共享,为企业分布式知识管理系统的构建提供了一种行之有效的解决方案。

参考文献

- 1 段淳林,曹洲涛. 重构企业的知识管理[J]. *经济师*, 2004, 2: 157~158
- 2 沈洁,罗建利. 基于多 Agent 系统的分布式知识管理研究[J]. *系统工程理论与实践*, 2006, 1(1): 42~47
- 3 罗炜,统秉枢,田凌. 协同知识管理中利用共享本体建立产品状态模型[J]. *计算机辅助设计与图形学学报*, 2004, 2: 191~196
- 4 李飞,高济. OKMF: 一个基于本体论的知识管理系统框架[J]. *计算机辅助设计与图形学学报*, 2003, 12: 1538~1543
- 5 Abdulmajid H M, Lee S P. An ontology-based knowledge model for software experience management[J]. *Journal of Knowledge Management Practice*, May 2004
- 6 Dreamtech software group. Peer to peer application development: cracking the code[M]. Wiley Publishing, 2001
- 7 Fiorano Software, Whitepaper: Super-Peer Architectures for Distributed Computing[Z]. <http://www.fiorano.com/whitepapers/superpeer.pdf>, 2001
- 8 黄道颖,黄建华,庄雷,李祖鹏. 基于主动网络的分布式 P2P 网络模型[J]. *软件学报*, 2004, 7(15): 1081~1089
- 9 Watts D J, Strogatz S H. Collective dynamics of small-world networks[J]. *Nature*, 1998, 393(6): 440~442
- 10 黄屹,顾进广,陈莘萌,陈和平. 网络环境下基于本体的知识库模型研究[J]. *武汉大学学报(理学版)*, 2005, 51(5): 603~608