

非结构化 P2P 网络的增量式查询

朱桂明^{1,3} 郭得科² 金士尧³

(国防科学技术大学计算机学院 长沙 410073)¹ (国防科学技术大学信息系统与管理学院 长沙 410073)²
(国防科学技术大学并行与分布处理国家重点实验室 长沙 410073)³

摘 要 基于非结构化 P2P 网络的应用日益广泛,参与共享资源的节点越来越多,且呈海量增长趋势。如何在海量节点、海量资源的情况下,对同一查询语义进行增量式查询,是一个重要而颇具挑战性的问题。提出了一个通用的近似依相关性大小的非结构化 P2P 网络增量式查询算法模型,该算法模型使得针对同一查询语义的增量式查询尽力优先访问与语义最为相关的节点,从而在完全分布式的 P2P 网络中获得类似 Google^[1] 的依资源相关性排序显示及其“下一页”的功能。

关键词 P2P 网络,非结构化,增量式查询

中图法分类号 TP393 **文献标识码** A

Incremental Search for Unstructured P2P Network

ZHU Gui-ming^{1,3} GUO De-ke² JIN Shi-yao³

(School of Computer Science, National University of Defense Technology, Changsha 410073, China)¹

(School of Information System and Management, National University of Defense Technology, Changsha 410073, China)²

(National Laboratory for Parallel and Distributed Processing, National University of Defense Technology, Changsha 410073, China)³

Abstract Nowadays it is quite easy for common users to share and exchange resources on the Internet through application software based on peer-to-peer computing mode, and therefore more and more people join in peer-to-peer network to share and exchange resources. As a result, the number of peers becomes extremely large, and resources are extremely abundant and scattered. In this case, it is an important and challenging job to do incremental search to incrementally retrieve new related resources for any query. This paper presented a general incremental search model for unstructured P2P network. This general incremental search model makes incremental search try to access semantic most related nodes, and therefore it does similar searches as Google^[1] in totally distributed P2P network.

Keywords P2P network, Unstructured network, Incremental search

自 1999 年成功推出一款 P2P 文件共享系统 Napster^[2]以来,基于 Peer-to-Peer 模式的资源共享迅速成为 Internet 上最主要的资源共享方式之一。至 2006 年,Internet 上各类 P2P 应用的流量已占骨干网流量的 60% 以上^[3]。随着非结构化 P2P 网络的应用日益广泛,参与资源共享的节点日益增多,P2P 网络中的共享资源日渐丰富,且呈现海量趋势。如何在海量资源、资源完全分布在 P2P 网络各个节点中的情况下,对同一查询语义依相关性大小实现增量式查询,即获得类似 Google 的依资源相关性排序显示及其“下一页”的功能,是一个重要而颇具挑战性的问题。

1 相关工作

基于节点和资源组织与定位方法,P2P 网络主要分为两种类型:结构化 P2P 网络和非结构化 P2P 网络。结构化 P2P 网络通常采用分布哈希(DHT)技术组织节点和发布资源(如

Chord^[4],CAN^[5]和 Tapestry^[6])的元信息。结构化 P2P 网络具有较高的查询效率,然而,由于哈希方法掩盖了节点和资源的语义信息,使得查询仅能按照精确匹配的方式进行。非结构化 P2P 网络中节点之间的关系比较随意。基于查询过程中有无提示信息的使用,将非结构化 P2P 网络路由分为盲目搜索(blind search)和提示性搜索(informed search)。盲目搜索的特征是搜索过程中不使用任何提示信息,将查询消息转发给部分或全部邻居节点,直到查询满足停止条件为止,典型算法有受限广播 Gnutella^[7,8]、随机路由(RW^[7,9])算法。提示性搜索的主要思想是由资源共享节点发布共享信息,并在网络中传播和维护这些共享信息,用于指导查询消息的转发^[10]。典型的算法有 Local Index^[11],SQR^[12]和 PROSA^[13]。Local Index 对一定范围内节点的所有共享资源对象建立索引,并负责回答其范围内的查询。SQR 通过依指数衰减 Bloom Filter (EDBF)将节点共享资源的衰减 Bloom Filter 传递给邻近

到稿日期:2009-09-28 返修日期:2009-12-11

朱桂明(1981—),男,博士生,主要研究方向为对等计算、复杂网络,E-mail:guiming_zhu@yahoo.com.cn;郭得科(1980—),男,博士,主要研究方向为对等网络、Bloom 滤波、MIMO、无线传感器网络、无线 Mesh 网络等;金士尧(1937—),男,教授,博士生导师,主要研究方向为分布式计算、复杂系统仿真、实时系统、虚拟现实等。

节点;节点把从相同邻居链路接收到的 Bloom Filter 的“并”运算结果作为该链路的路由条目;资源搜索过程中,查询消息沿着目标对象信息量最强的方向转发。PROSA 模拟人类社会个体之间交流与合作的关系;节点之间根据彼此熟悉的程度,或知道对方的全部语义信息,或知道地方发起查询的语义,或仅仅知道对方的存在;查询路由过程中,将查询消息转发给语义最相关的邻居节点。

非结构化 P2P 网络存在一个普遍的问题是,不支持针对同一查询语义的增量式搜索。我们在先前的工作中提出了一种基于熟人关系的增量式搜索算法 IPSBSAR^[14],IPSBSAR 基于人类社会关系中的熟人关系路由机制实现了增量式搜索。本文试图不考虑非结构化 P2P 网络的具体路由算法,提出了一种通用的近似依相关性大小的非结构化 P2P 网络增量式查询算法模型,使得非结构 P2P 网络中针对同一查询语义的增量式查询尽力优先访问语义最为相关的尚未访问的节点,从而在完全分布式的 P2P 网络中获得类似 Google 依据资源相关性排序显示及其“下一页”的功能。

算法 1 rcv (Node node, Query query, ArrayList nodeslist)

```
1 /* 获取当前节点 node 发起的查询 query 对应的
2  查询会话 querySession */
3 retrieve query session querySession for query;
4 /* 将 nodeslist 中已经在会话 querSession
5  中 VisitedNodesList 的节点删除 */
6 nodeslist.delete(querySession. VisitedNodesList);
7 /* 将 nodeslist 中余下的所有节点加入到
8  会话 querySession 中的 ToBeVisitedNodesList 中 */
9 querySession.add(nodeslist);
10 /* 依据与查询 query 相关大小,
11  对 ToBeVisitedNodesList 中节点排序 */
12 querySession.sort(querySession. ToBeVisitedNodesList);
```

2 增量式查询算法模型

非结构化 P2P 网络的增量式查询的本质要求是,同一节点针对同一查询语义的多次查询请求希望返回与网络中新的语义最为相关的资源。为此,引进查询会话(session)的概念。同一查询会话的查询语义相同;每一个查询会话有两个列表:VisitedNodesList 和 ToBeVisitedNodesList。VisitedNodesList 用于保存已访问节点及其语义,ToBeVisitedNodesList 保存尚未访问节点及其语义。增量式查询算法模型的基本思想是:当查询命中时,命中节点除了向查询请求节点返回相匹配的资源或其元信息,同时还向查询请求节点返回与查询语义相关的邻居节点及其语义;查询请求节点将接收到的与查询语义相关的尚未访问的邻居节点及其语义加入到 ToBeVisitedNodesList 列表中,并按照各节点的语义与查询语义相关性的 大小依次排序。查询请求节点针对同一查询语义进行增量式搜索时,首先从该查询对应的查询会话中的 ToBeVisitedNodesList 列表中查找最为相关的节点,并将查询转发给查找到的节点;若 ToBeVisitedNodesList 列表为空,则通过 P2P 网络的一般路由算法将查询转发出去。查询请求节点接收返回节点及其语义的算法如算法 1 所示,查询请求节点进行增量式查询的算法模型如算法 2 所示。

算法 1 是查询请求节点接收返回节点及其语义的算法

rcv 伪码。算法 rcv 首先取得查询 query 的查询会话 querySession (若存在 query 对应的查询会话 querySession,则返回该查询会话 querySession,否则新建一个查询会话 querySession 返回),将 nodeslist 中存在于 VisitedNodesList 列表中的节点删除,将 nodeslist 余下的节点加入到 ToBeVisitedNodesList 中。依据查询 query 与 ToBeVisitedNodesList 中各节点语义相关性的 大小,对 ToBeVisitedNodesList 中各节点进行排序,使得越靠前的节点,其语义与 query 的相关性越大。

算法 2 void incrementalSearch(Node node, Query query, int num)

```
1 if (! localIncrementalSearch(node, query, num))
2  generalRouting(node, query, num);
3 boolean localIncrementalSearch(Node node, Query query, int num)
4 if (num ≤ 0) return true;
5 /* 获取当前节点 node 发起的查询 query 对应的查询会话
6  querySession */
7 retrieve query session querySession for query;
8 /* 如果查询会话 querySession 的 ToBeVisitedNodesList 为空,
9  则返回 false */
10 if (querySession. ToBeVisitedNodesListSize() == 0)
11  return false;
12 /* 获取查询会话 querySession 的 ToBeVisitedNodesList
13  最为相关的节点,并将其从 ToBeVisitedNodesList 列表中移出,
14  加入到 VisitedNodesList 列表中 */
15 nextPeer = querySession.ToBeVisitedNodesList.first();
16 querySession. VisitedNodesList.add(nextPeer);
17 /* 将查询 query 转发给 nextPeer 节点 */
18 forward(nextPeer, query, num);
19 /* 如果 nextPeer 没有足够相匹配资源,则迭代调用增量式查询
20  算法 */
21 if MatchedRes(nextPeer, query, num) < num
22  return localIncrementalSearch (node, query, num-MatchedRes
23  (nextPeer, query, num));
```

算法 2 给出了查询请求节点进行增量式查询的算法模型 incrementalSearch。incrementalSearch 首先通过 localIncrementalSearch 进行增量式查询。localIncrementalSearch 有一个布尔(boolean)型的返回值,用于标识增量式搜索的成功与否。当返回值为 true 时,表示成功进行了增量式查询;否则,增量式查询 localIncrementalSearch 失败。若 localIncrementalSearch 失败,则由一般的查询路由算法 generalRouting 转发查询请求。localIncrementalSearch 首先确认增量式请求的资源数目 num,若其小于或等于 0 时,表示无需再请求新的资源,从而返回 true。否则,num 大于 0,节点获取查询 query 对应的查询会话 querySession,并进一步判断 querySession 的 ToBeVisitedNodesList 是否有尚未访问的节点,若没有尚未访问的节点则返回 false。若 querySession 的 ToBeVisitedNodesList 有尚未访问的节点,则从 ToBeVisitedNodesList 列表中移出与查询 query 最为相关的节点 nextPeer,同时将 nextPeer 加入到 querySession 的 VisitedNodesList 列表中。将查询请求转发给 nextPeer。如果 nextPeer 没有足够的资源满足增量式查询请求,则以仍需请求的资源数继续调用增量式查询 localIncrementalSearch。

增量式查询算法模型需要查询命中节点在返回资源的同

时返回与查询语义相关的邻居节点;查询请求节点对于每一个查询语义 *query* 维护一个查询会话 *querySession*;查询请求节点将尚未访问的节点加入到 *querySession* 的 *ToBeVisitedNodesList* 列表中,将已访问的节点从 *querySession* 的 *ToBeVisitedNodesList* 列表中移出并加入到 *VisitedNodesList* 列表中。增量式查询沿着与查询会话 *querySession* 的 *ToBeVisitedNodesList* 语义最为相关的节点进行。当增量式查询失败时,则仍由非结构化 P2P 网络中各节点的一般路由算法转发;一般路由算法将查询消息路由到目标节点时,将目标节点及其语义和其与查询语义相关的邻居节点及其语义返回给查询请求节点,由查询请求节点通过 *rcv* 算法接收。因此,该增量式查询算法模型是一个通用的近似依相关性大小的非结构化 P2P 网络增量式查询算法模型,可与任何路由算法相结合实现增量式搜索。为了提高增量式查询的效率,要求查询请求节点保存查询过程中发现的与其具有共同兴趣的节点及其语义。

3 理论分析

(1) 增量式查询

本文提出的非结构化 P2P 网络的增量式查询算法模型,查询命中节点在返回相匹配的资源时所返回的与查询语义相关的邻居节点及其语义,查询请求节点将接收到的与查询语义相关的尚未访问节点保存在该查询对应的查询会话中的 *ToBeVisitedNodesList* 列表中。当针对同一查询语义进行增量式查询时,首先从该查询对应的查询会话中的 *ToBeVisitedNodesList* 列表中选择与语义最为相关的节点转发查询;当 *ToBeVisitedNodesList* 列表为空,增量式查询失败时,将查询由一般路由算法转发。该增量式查询算法模型总是试图访问新的未被该查询节点访问过的语义相关节点以获得新的资源,从而实现增量式查询。

(2) 效率

该增量式算法模型能否高效地实现增量式搜索,取决于非结构化 P2P 网络中语义相关节点之间的连接关系及查询路由算法。如果网络中任何一个节点均存在到其它所有语义相关节点的通路,查询路由算法能够高效地查找任何一个语义相关节点,那么增量式查询将通过顺序遍历该查询会话中的 *ToBeVisitedNodesList* 列表中的节点,而获得新的语义相关资源,从而具有极高的查询效率。

4 实验

我们在 Peersim^[16] 上实现了增量式查询模型,并将其分别与 PROSA, RW, SQR 和 Local Index 查询路由算法相结合,实现增量式查询。模拟过程中,PROSA, RW, SQR 和 Local Index 查询路由算法的查询请求节点均保存向其返回相关资源对象的节点及其语义,但不改变 SQR 和 Local Index 算法的原有信息场。PROSA 算法的查询请求节点以随机方式选择一个邻居节点转发,中间节点保存查询消息的发起节点及查询语义,并将查询消息尽可能转发给与语义最为相关的邻居节点;SQR 路由算法始终将查询消息沿着成员资格信息量最大的方向传递;Local Index 路由算法中各节点根据其存储的一定范围内节点的资源对象元信息转发查询请求,而对于当前节点没有匹配的资源对象元信息的查询请求,将其随

机转发给当前节点的任一邻居节点。

定义资源占有率为节点本身所具有的与查询相关的资源数占有与查询相关的资源数的比率,用于衡量节点进行增量式查询时获取新资源的能力。实验过程为,在表 1 所给参数的基础上,时间[0,50]内所有节点在各个时间点上以各自的兴趣点作为查询语义进行增量式查询,在时间[50,100]内随机选择 20 个节点,各节点随机选择其兴趣外的兴趣点作为查询语义进行增量式搜索。实验结果如图 1 所示。图 1 给出了同一时刻所有查询节点的资源占有率的均值。

表 1 实验参数

参数	值	参数	值
节点个数	1000	Max TTL	10
初始节点平均度数	4	k-RW 的 k 值	1
初始拓扑	random	Bloom Filter 比特位数	6000
兴趣总数	20	哈希函数的个数	16
每个节点的兴趣个数	1	衰减指数 d	1.2
节点每个兴趣对应的资源个数	1	各节点语义 Bloom Filter 的最大传递距离	4
每次查询请求的资源个数	3		

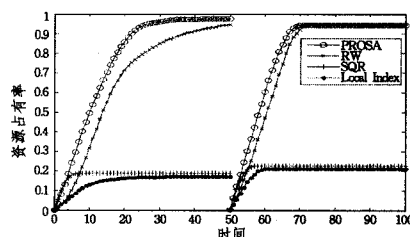


图 1 与不同路由算法相结合的增量式查询模型实验结果对比图

图 1 中,增量式查询模型与不同路由算法相结合,产生的增量式查询效果不尽相同。在[0,50]和[50,100]内,对于 PROSA 算法,在各查询节点进行少量增量式查询后,各节点的平均资源占有率即达到了较高的水平。对于 RW 算法,增量式查询模型的增量式查询效果略逊于 PROSA 算法,远好于 SQR 和 Local Index 算法。对于 SQR 和 Local Index 算法,各查询节点的平均资源占有率维持在 0.2 左右。PROSA 和 RW 算法的查询请求节点以随机路由方式转发查询请求,保存向其返回相关资源对象的节点及其语义,更易使得网络中所有语义相关节点形成直接连通子图。相对 RW 算法而言,PROSA 算法由于中间节点保存查询消息的发起节点及查询语义而获得更好的性能。SQR 和 Local Index 算法由于查询请求始终优先转发给周围最为相关的节点,使得查询消息限入了近乎为同样搜索路径上进行资源搜索的困境,很难使网络中所有语义相关的节点形成直接连通子图,从而 SQR 和 Local Index 算法的资源占有率达到 0.2 左右时再无法获得新的相关资源。

实验结果显示,与 PROSA、RW 路由算法相结合的增量式查询模型,在进行少量增量式查询后,即获得了网络中绝大多数与语义相关的资源,达到了较好的效果。为使增量式查询一般算法模型高效地执行,查询路由算法必须能够快速使得网络中所有与语义相关的节点形成直接连通子图。

结束语 非结构化 P2P 网络增量式查询是一个重要而颇具挑战性的问题,本文提出了一个通用的近似依相关性大小的非结构化 P2P 网络增量式查询算法模型,该算法模型使得针对同一查询语义的增量式查询尽力优先访问语义最为相

关的节点,从而在完全分布式的 P2P 网络中获得类似 Google 的依据资源相关性排序显示及其“下一页”的效果。

参 考 文 献

- [1] Google[EB/OL]. <http://www.google.com>
- [2] Napster[EB/OL]. <http://www.napster.com>
- [3] Suh K, Figueiredo D R, Kurose J, et al. Characterizing and Detecting Skype-Relayed Traffic[C] // Proc. IEEE INFOCOM. Washington DC; IEEE, 2006; 1-12
- [4] Stoica I, Morris R, Karger D R, et al. Chord: A Scalable Peer-to-Peer Lookup Service for Internet Applications[J]. IEEE/ACM Transactions on Networking, 2003, 11(1): 17-32
- [5] Ratnasamy S, Francis P, Handley M, et al. A Scalable Content-Addressable Network[J]. Proc. ACM SIGCOMM. New York; ACM, 2001, 31(4): 161-172
- [6] Zhao B Y, Kubiatowicz J D, Joseph A D. Tapestry: A Fault-Tolerant Wide-Area Application Infrastructure[J]. Computer Communication Review, 2002, 32(1): 81
- [7] Tsoumakos D, Roussopoulos N. Analysis and comparison of P2P search methods [C] // Proc. of the 1st international conference on Scalable information systems. New York, USA; ACM, 2006; 25
- [8] Chawathe Y, Ratnasamy S, Breslau L, et al. Making Gnutella-Like P2P Systems Scalable[C] // Proc. ACM SIGCOMM. New York, USA; ACM, 2003; 407-418
- [9] Lv Q, Cao P, Cohen E, et al. Search and Replication in Unstructured Peer-to-Peer Networks[C] // Proc. of the 16th international conference on supercomputing. New York, USA; ACM, 2002; 84-95
- [10] Zhang Yi-ming, Lu Xi-cheng, Zheng Qian-bing, et al. PST: An efficient search algorithm for large-scale P2P systems[J]. Journal of Software, 2008, 19(6): 1473-1480
- [11] Yang B, Garcia-Molina H. Efficient Search in Peer-to-Peer Networks[C] // Proceedings of the 22nd IEEE International Conference on Distributed Computing Systems (ICDCS). Washington DC; IEEE, 2002
- [12] Abhishek K, Jun (Jim) X, Ellen W Z. Efficient and Scalable Query Routing for Unstructured Peer-to-Peer Networks[C] // Proc. of the IEEE INFOCOM. Washington DC; IEEE, 2005; 1162-1173
- [13] Carchiolo V, Malgeri M, Mangioni G, et al. Emerging Structures of P2P Networks Induced by Social Relationships[J]. Computer Communications, 2008, 31(3): 620-628
- [14] 朱桂明, 金士尧, 郭得科. IPSBSAR: 一种基于熟人关系的增量式 P2P 搜索算法[J]. 计算机研究与发展, 2009, 46(8): 1260-1269
- [15] Jelasity M, Montresor A, Jesi G P. The Peersim Simulator [CP/OL]. [2008-05-10]. <http://peersim.sf.net>
- [16] Lee I W C, Fapojuwo A O. Modeling Wireless TCP Connection Arrival Process[C] // GLOBECOM. Press by IEEE, 2006
- [17] Chen L, Jiang Z J, Zhu Y F. Design and Logical Analysis of Authenticated Key Exchange Protocol[C] // WiCOM. 2008
- [18] Maneerat N, Varakulsiripunth R. Composition Method of Communication System Specifications in Asynchronous Model and Its Support System[C] // ICON. 2001
- [19] Orset J M, Alcalde B, Cavalli R. An EFSM-based Intrusion Detection System for Ad Hoc Networks[C] // ATVA. 2005
- [20] Zhu L Z, Zhang L W, Li Q. Modeling and Evaluating of Net Protocol Based on Timed Colored Petri Net[C] // Proc. of Modeling and Simulation in Electronics. Chengdu, 2008
- [21] GNU gprof[EB/OL]. <http://www.cs.utah.edu/dept/old/text-info/as>
- [22] Carofiglio G, Caretto M. Route Stability in MANETs under the Random Direction Mobility Model[J]. IEEE Transaction on Mobile Computing, 2009, 8(2): 1167-1179
- [23] Coll H, Lloret J, Francisco J S. Does NS2 really simulate MPLS network? [C] // International Conference on Autonomic and Autonomous Systems, 2008
- [24] Bendov O. On the Validity of the Longley-Rice (50,90/10) Propagation Model for HDTV Coverage and Interference Analysis[C] // Proc. of NAB Broadcast Engineering Conf. Las Vegas, NV, 1999
- [25] Andersson S, Ryden T. Maximum likelihood estimation of a structured MMPP with applications to traffic modeling[C] // ITC Specialist Seminar. Monterey, 2000
- [26] 陈果. 基于遗传算法的 ARMA 模型定阶新技术[J]. 机械工程学报, 2005, 45(1): 41-45
- [27] Suarez C A H, Parra O S, Martinez L F P. Traffic Model Based on Time Series to Forecast Traffic Future Values Within a WIFI Data Network[C] // Wireless Communications, Networking and Mobile Computing Conference. 2008
- [28] Chen C, Xu Y, Zhang L. Some Remarks on ON/OFF Network Traffic[C] // Proc. of Workshop on Power Electronics and Intelligent Transportation System. Guangzhou, 2008
- [29] Jeong H D J, Lee J S R, McNickle D, et al. Suggestions of efficient self-similar generators[J]. Simulation Modelling Practice and Theory, 2007, 15(3): 328-353
- [30] Hosking J R M. Fractional differeneing [J]. Biometrika, 1981; 165-176
- [31] Kulkarni V, Marron J, Smith F. A cascaded on-off model for TCP connection traces [R]. UNC/OR/TR01-1. University of North Arolina, 2001
- [32] Feldmann W, Gilbert A C, Willinger W. Data-network as cascades; investigating the multifractal nature of Internet WAN traffic[C] // ACM SIGCOMM'98 Conference. Vancouver, BC, Canada, 1998
- [33] Lee I W C, Fapojuwo A O. Modeling Wireless TCP Connection Arrival Process[C] // GLOBECOM. Press by IEEE, 2006
- [34] Chen L, Jiang Z J, Zhu Y F. Design and Logical Analysis of Authenticated Key Exchange Protocol[C] // WiCOM. 2008
- [35] Maneerat N, Varakulsiripunth R. Composition Method of Communication System Specifications in Asynchronous Model and Its Support System[C] // ICON. 2001
- [36] Orset J M, Alcalde B, Cavalli R. An EFSM-based Intrusion Detection System for Ad Hoc Networks[C] // ATVA. 2005
- [37] Zhu L Z, Zhang L W, Li Q. Modeling and Evaluating of Net Protocol Based on Timed Colored Petri Net[C] // Proc. of Modeling and Simulation in Electronics. Chengdu, 2008
- [38] GNU gprof[EB/OL]. <http://www.cs.utah.edu/dept/old/text-info/as>
- [39] Carofiglio G, Caretto M. Route Stability in MANETs under the Random Direction Mobility Model[J]. IEEE Transaction on Mobile Computing, 2009, 8(2): 1167-1179
- [40] Coll H, Lloret J, Francisco J S. Does NS2 really simulate MPLS network? [C] // International Conference on Autonomic and Autonomous Systems, 2008
- [41] Bendov O. On the Validity of the Longley-Rice (50,90/10) Propagation Model for HDTV Coverage and Interference Analysis[C] // Proc. of NAB Broadcast Engineering Conf. Las Vegas, NV, 1999

(上接第 46 页)