

非结构化P2P系统的路由算法

侯孟书, 卢显良, 周旭, 詹川

(电子科技大学计算机科学与工程学院 成都 610054)

【摘要】针对非结构化对等网络中洪泛搜索机制的路由盲目性问题,提出了一种利用节点积累的经验指导节点传播查询的路由算法。在该路由算法中,节点记录自己关注的主题和满足主题的邻居节点,建立主题与邻居节点的对应关系表,当节点收到查询后节点就利用该表来指导节点选择查询路由,以便更快地找到查询结果。仿真试验和性能分析表明该算法有效地减少了查询带来的网络流量,提高了查找成功率。

关键词 对等网络; 洪泛; 查询; 文件共享

中图分类号 TP393 文献标识码 A

New Route Algorithms of Unstructured P2P Systems

HOU Meng-shu, LU Xian-liang, ZHOU Xu, ZHAN Chuan

(School of Computer Science and Engineering, UEST of China Chengdu 610054)

Abstract The problem of the unstructured peer-to-peer system is the scalability of the system due to network traffic cost. This paper presents a new route algorithm of unstructured peer-to-peer system with routing indices. In the algorithm, a peer records the subject which it query about in the past and who can answer the subject. When peers receive a new query, it can forward the query to a selected peer which can answer the subject in the past. In this way the peer reduces the blind search. In the end, we test our design through simulations and the result shows that our design routes the query effectively, and reduces the network traffic greatly, improve the searching success ratio.

Key words peer-to-peer network; flooding; query; file sharing

在P2P (Peer-to-Peer)系统中,所有节点在功能上是对等的,既可以是客户机又可以是服务器,称为对等机(SERVER-client, servant)。通常, P2P系统分为非结构化(Unstructured)P2P系统、结构化(Structured)P2P系统和松散结构化(Loosely Structured)P2P系统三类。这三类结构各有优缺点,而非结构化P2P系统应用最为广泛。但是,由于非结构化P2P系统采用洪泛(Flooding)搜索机制,查询从一个节点以广播方式传播到其他节点,直到查找到查询结果,导致每次查询产生大量的网络流量,对网络造成很大的负担,影响了非结构化P2P系统的可扩展性。针对以上问题,本文提出了一种新的带有索引机制的路由(Random Walk with Routing Index, RWRI)算法。

1 P2P系统

在P2P系统中,每个节点都与系统中的其他节点相连接,系统中所有节点及其链接形成一个覆盖在物理网络上的虚拟网络,该虚拟网络称为一个覆盖网(Overlay Network)。在覆盖网中任意两个节点间可以通信,消息通过网络的边进行传播,节点之间完全平等。覆盖网中的每个节点拥有共享数据集,这些共享数据集通常以文件的形式存在,但并不局限于文件。查询采用适合于共享数据集的方式,如果共享数据集以文件

的形式存在,则可以对文件属性查询,也可以对文件内容查询。

对于非结构化P2P系统中洪泛机制带来的网络流量问题,目前主要从下面3个方面解决。

1) 网络拓扑。KaZaA和Morpheus使用超节点(Supernode)对系统中的节点区别对待,让计算能力强、存储容量大、高带宽、低延迟的节点成为超节点,充当一个区域的服务器,从而解决非结构化P2P系统的扩展性问题。文献[1]将兴趣相同或相似的节点进行聚类(Clustering),这样查询易于在聚类中满足,减少查询传播的次数,从而减少网络流量。

2) 查询消息的跳数(Time-To-Live, TTL)。文献[2]使用逐渐增加TTL的值代替固定的TTL值,文献[3]使用本地索引和TTL相结合的方式。

3) 搜索机制。文献[2]提出了用随机漫步来代替洪泛机制,文献[4]根据Gnutella中节点分布呈幂律(Power Law)特性,提出将查询请求路由到聚集度高的节点,以此来克服随机漫步的盲目性。文献[5]采用改进的宽度优先算法和智能搜索机制。文献[3]提出了有向宽度优先算法,通过某些策略选择路由的节点。文献[6]使用多种路由索引来选择路由的节点。

为方便以后的叙述,首先引入几个概念:

1) 查询的源节点,即发出查询的节点。在非结构化P2P系统中,为了限制查询消息传播的时间,源节点指定了查询消息的TTL,在Gnutella中,TTL一般设定为7。

2) 查询的目标节点,即满足查询的节点。目标节点将查询的结果返给源节点。

3) 查询的中间节点,即消息从查询的目标节点到查询的源节点所经过的所有节点。

4) 路由索引表(Routing Index Table, RIT),即本地节点查询关键字索引表。假设覆盖网的拓扑图如图1所示,节点A的RIT如表1所示,其中行表示关键字,列表示节点A的邻居节点。表1中 K_i 为节点A过去一段时间内所查询的关键字,其中 (K_i, B) 的值为 m ,表示通过节点B成功查询到 K_i 的次数为 m 。例如 (K_1, B) 的值为4,表示通过节点B成功查询到 K_1 的次数为4; (sum, B) 的值为14,表示通过节点B成功查询到所有关键字的总和为14。这里“通过节点B成功查询到 K ”有两种情况,一是节点B本身是查询的目标节点,二是节点B是查询的中间节点。

5) 本地索引表(Local Index Table, LIT),对本地节点的文档按主题关键字建立的倒排索引。

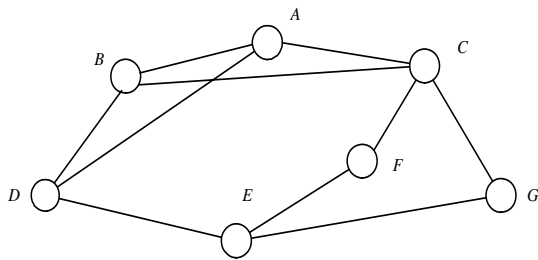


图1 P2P网络拓扑图

表1 节点A的RIT

	B	C	D
sum	14	16	6
K_1	4	7	1
K_2	3	2	3
K_3	6	3	1
K_4	1	4	1

2 路由索引表的建立和维护

2.1 LIT的建立和维护

当节点首次加入非结构化P2P系统时,节点对本地的文档建立LIT表,LIT表的内容主要包括文件属性,如文件名、文件作者、文件创建时间等。当有新的共享文档加入到共享目录时,节点更新自己的LIT表,以反映本节点最新的共享文档。

LIT的作用有两个:一是节点加入非结构化P2P系统时,更新邻居节点的路由索引表RIT;二是节点收到查询后,用于快速检索本地是否有查询的结果。

2.2 RIT的建立和维护

节点A加入非结构化P2P系统的过程如下:首先,节点A和节点B取得联系,当B同意与A建立链接后,B将自己的邻居节点C、D告诉A;其次,A试图与B的邻居节点建立链接,当A分别与C、D建立链接后,将自

己的LIT表传递给B, C, D; 最后, B, C, D根据节点A的LIT表更新各自的RIT表。由于负载平衡等原因, B可能不同意与A建立链接, 这种情况下B就推荐一个邻居节点给A。当节点A刚加入系统时, 由于A没有发出查询, 所以A的RIT表并不知道A主要查询那些主题的内容, 也无法对A收到的查询给出指导, 这时A的路由算法采用随机漫步, 即从A的邻居节点中随机选取几个节点来投递查询。当A多次查询后, A的RIT表不断记录A查询的主题, 并通过检查查询命中消息, 建立主题与满足查询主题的邻居节点的对应关系, 这种情况下, 节点A就可以通过RIT表来选择查询的路由, 例如当A收到主题为KI的查询时, 根据A的RIT表, 可知该查询应该投递到节点C, 因为在历史上, 节点B满足查询主题KI的次数为4, 而节点D满足查询主题KI的次数为1。当节点A收到查询后, 对节点A的RIT表进行如下维护:

- 1) 如果消息是来自节点A, 则将查询的关键字加入节点A的RIT表。
- 2) 如果消息是来自其他节点的查询结果, 假设查询结果来自节点B, 则更新节点A的RIT表, $(K, B) = (K, B) + 1$; $(sum, B) = (sum, B) + 1$ 。

3 路由算法

根据前述的路由算法原理, 现给出RWRI的具体算法:

Routing Select

Input Query Q , a client C

Output Query answers, Query forwards

For Match (Q , LIT. key [i])

Send local answers for Q to C

If StopCondition () return

RouteSelectNode = Φ

For Match (Q , RIT. Key [i])

If Max(Q , N) RouteSelectNode.Add(N)

If RouteSelectNode == Φ

RouteSelectNode.Add(Max(sum N))

RouteSelectNode.Add(RandomNode())

Send Q to RouteSelectNode

If StopCondition() return

对RWRI算法的说明如下:

- 1) 算法输入的参数是查询 Q 和 Q 的源节点, 输出的是查询结果和查询传播的路由节点;
- 2) 函数Match(Q , T)用于查询表 T 中的项是否和 Q 匹配, 集合RouteSelectNode用于记录查询投递的节点;
- 3) 当查询 Q 到达时, 不论 Q 来自节点自身还是其他节点, 算法RWRI都要运行, 以便返回结果给 Q 的源节点, 当查询 Q 的结束条件不满足时, 将 Q 路由到最有可能获得结果的节点;
- 4) 无论RIT表中是否有与 Q 匹配的项, 都随机选取几个邻居节点加入到要路由的节点集合中, 这样的处理是基于以下的考虑: 一是避免每次投递都集中在少数几个节点, 即避免这少数几个节点的负载过重而产生系统热点(Hot Pot); 二是当投递到Max(Q , N)节点查询失效时, 下次仍然将同样的查询投递到该节点是不合适的, 加入随机选取邻居节点, 可以增加查询的召回率(Recall), 使RWRI算法更加健壮。

4 实验结果

为了评价RWRI算法, 建立了仿真程序来模拟P2P环境, 网络拓扑使用BRITE生成^[7], 并且与另外两个泛洪(Flooding, FLD)算法和随机漫步算法(Random Walk, RW)进行了比较。

- 1) FLD算法: 在随机网络拓扑上, 使用洪泛算法来定位查询结果。
- 2) RW算法: 在随机网络拓扑上, 使用随机漫步算法来定位查询结果。

实验首先对比了三种算法产生的网络流量随返回结果数的变化情况, 实验结果如图2所示。图中横坐标

R 表示请求返回的结果数, 纵坐标 F 表示产生的网络流量(单位为Byte)。从图中可以看出RWRI产生的网络流量明显少于FLD和RW产生的网络流量。其次, 对比了三种算法查找成功率随消息 TTL 变化的情况, 实验结果如图3所示。图3中横坐标表示 TTL 的值, 纵坐标表示查找的成功率。从图中可以看出, RWRI的查询成功率在 TTL 较小的时候明显高于其他两种算法, 这是因为RWRI减少了查询的盲目性。随着 TTL 的增加, 三种算法的查找成功率趋于一致。

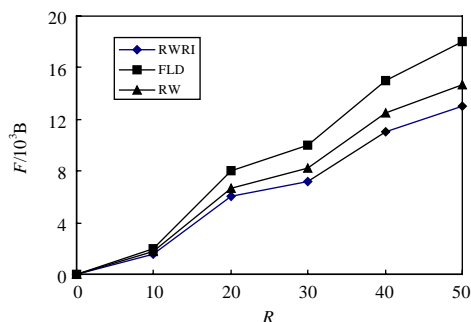


图2 结果数与产生的网络流量关系图

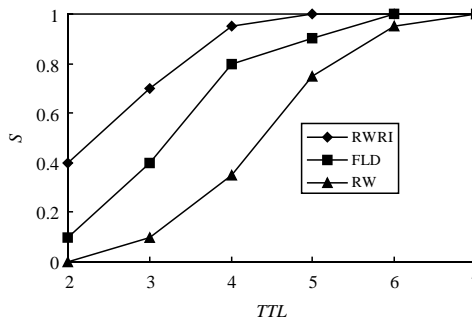


图3 查找成功率随 TTL 的变化图

5 结论

非结构化P2P系统采用洪泛机制定位查询结果, 往往造成大量的网络流量, 影响了非结构化P2P系统的可扩展性。本文提出了一种新的查询路由算法RWRI, 该算法通过记录节点关注的主题和满足该主题的目标节点, 将过去的经验保留下来, 指导节点路由选择。实验结果表明, RWRI能有效地减少查询带来的网络流量, 提高查找的成功率。

参 考 文 献

- [1] Ng C H, Sia K C. Peer clustering and firework query model[C]. In Proc. of 11th World Wide Web Conference, Hawaii, USA, 2002
- [2] LV Q, Cao P, Cohen E, et al. Search and replication in unstructured peer-to-peer networks[C]. In: Proceedings 16th ACM International Conference on Supercomputing(ICS'02) New York, 2002
- [3] Yang B, Garcia-Molina H. Efficient search in peer-to-peer networks[C]. In Proc. of ICDCS'02, Vienna, Austria, 2002
- [4] Admaic L A, Lukose R M, Puniyani A R, et al. Search in power-law networks [EB/OL]. <http://www.parc.xerox.com/istl/groups/iea/papers/plsearch/index.html>, 2004 - 03 - 05
- [5] Kalogeraki V, Gunopulos D, Zeinalipour-Yazti D. A local search mechanism for peer-to-peer networks[C]. In: Proc. of the Eleventh International Conference on Information and Knowledge Management, Virginia, USA, 2002
- [6] Crespo A, Garcia-Molina H. Routing indices for peer-to-peer systems[C]. In Proceedings International Conference on Distributed Computing Systems, Arizona, USA, 2002
- [7] Medina A, Lakhina A, Mattaf I, et al. BRITE: An approach to universal topology generation[C]. In Proceedings of the International Workshop on Modeling, Analysis and Simulation of Computer and Telecommunications System-MASCOTS'01, Cincinnati, Ohio, 2001

编辑 熊思亮