

一种改进的非结构化 P2P 网络洪泛搜索机制

卢苇, 周韬, 邢薇薇

(北京交通大学 软件学院, 北京 100044)

摘要:非结构化 P2P 网络使用基于洪泛的查询算法来进行资源搜索。然而,这种搜索机制随着网络节点的增多,网络规模的增大,将产生大量的冗余查询消息,会导致网络流量急剧增加,引起网络拥塞。提出了一种基于转发区间的洪泛搜索机制 FIFSM(forwarding interval based flooding search mechanism),通过为消息分配不相交的转发区间,使其沿着一棵生成树的结构传播,消除了消息环路,从而避免冗余消息的产生。FIFSM 机制采用高效的网络维护策略,能够在动态环境下以较低的开销保证网络的稳定性。实验结果表明,FIFSM 机制能够降低洪泛开销,保证资源搜索的高成功率和低延迟,是一种有效的非结构化 P2P 网络资源搜索机制。

关键词:算法;计算机系统;资源优化;故障检测;容错性;网络管理;网络性能;丢包率;对等网络;可靠性分析;稳定性;时延;拓扑结构;非结构化 P2P 网络;洪泛搜索;转发区间;生成树

中图分类号:TP393

文献标志码:A

文章编号:1000-2758(2015)02-0342-09

P2P 网络是一种应用层的分布式网络,网络中各节点地位是对等的。按照资源组织与定位方法,可以将其分为结构化 P2P 网络^[1-3]和非结构化 P2P 网络^[4]。其中,由于非结构化 P2P 网络的简单性和高鲁棒性使其得到了深入的研究和广泛的应用。

以 Gnutella 为代表的非结构化 P2P 网络,是一种没有特定拓扑结构的覆盖网,通常建立在随机图上,使用基于洪泛的查询算法进行资源搜索。节点通常把资源存储在本地,也可将其备份到其他节点,以提高资源搜索的效率^[5-6]。由于没有拓扑结构上的约束,在节点频繁加入和撤离的动态环境中,非结构化 P2P 网络表现出良好的性能,并且仅需较少的维护开销。由于每一次查询都是在本地进行评估,所以它支持任意复杂类型的查询,如语义查询等。非结构化 P2P 网络采用基于洪泛的查询机制进行资源搜索。在洪泛的过程中,节点在有限的 TTL 内,不断地向所有的邻居节点转发消息,直至查询到所需的结果或 TTL 变为 0。洪泛的优点是响应时间短、覆盖范围广以及可靠性高,但洪泛会在网络中产

生大量的冗余消息,不仅增加了节点处理负担,还会占用大量的网络带宽^[7]。因此,如何在有效进行资源搜索的同时,降低冗余消息量,提高系统的可扩展性和稳定性,是非结构化 P2P 网络的一个核心问题。

针对上述问题,许多研究工作者尝试通过改进洪泛算法^[8-11],以提高非结构化 P2P 网络的可扩展性。这些方法不再盲目地向所有邻居节点转发消息,而是有策略地选择部分邻居节点转发消息。例如,随机漫步算法^[12]每次随机选择 k 个邻居节点转发消息,其搜索过程持续至查询到所需的结果或 TTL 变为 0。这样就减少了洪泛搜索产生的网络流量,但另一方面往往会产生较大的搜索延迟,并且搜索过程有可能丢失节点。另外一种优化策略是通过改变 P2P 网络结构,以达到提高搜索效率的目的,其典型如树形结构的 P2P 网络^[13-15]。在这样的系统中,由于网络结构为树形,洪泛时消息到达任意节点仅一次,从而避免了冗余消息的产生。如果树的拓扑结构是稳定的,且消息丢失率较低,则树形结构

收稿日期:2014-09-02

基金项目:国家自然科学基金(61100143、61370128、61272353)、教育部新世纪人才计划项目(NCET-13-0659)与北京高校青年英才计划项目(YETP0583)资助

作者简介:卢苇(1963—),北京交通大学教授,主要从事软件服务工程研究。

的 P2P 网络具有良好的性能。然而,在节点频繁加入和撤离的动态环境中,树的结构经常会被分割,造成大量消息的丢失。因此,为了获得良好的可靠性,树形结构 P2P 网络需要探测消息丢失并从中恢复,这会导致恢复信息的延迟。尤其当故障频繁发生时,还会引起相当大的开销,这些都会限制系统的可扩展性。树形结构 P2P 网络的另一问题是负载不均衡,树的内部节点承载了几乎所有的转发负载,而叶子节点却不分担负载。文献[16]提出,将非结构化 P2P 网络建立在结构化 P2P 网络之上,利用其结构提高洪泛的性能。文献[17]提出了一种基于生成树的洪泛算法,该算法保证遍历 N 个节点的系统只需要 $N-1$ 个消息。但由于该算法基于 chord^[3],很大程度上受其协议的限制,生成树的度数较低,消息传播延迟较大。由于 chord 邻居表中每一个表项仅维护 1 个节点,故当节点失效时,洪泛会导致大范围的节点丢失。动态环境中,其性能和可靠性高度依靠底层 chord 协议的维护能力。因此,由于协议的语义规定了覆盖网节点应如何连接,在结构化 P2P 网络上建立非结构化 P2P 网络的思想并不适合高度结构化的 DHT 协议。

为此,本文提出了一种改进的非结构化 P2P 网络洪泛搜索机制 FIFSM (forwarding interval based flooding search mechanism)。此机制借鉴结构化网络的策略,为每一个节点分配唯一的标识符,并建立有结构的邻居表。此外,在洪泛查询消息(下文简称查询消息)中添加 2 个字段,用于限制节点的转发范围。FIFSM 机制的实现主要包括 2 个部分:洪泛算法和网络维护策略。FIFSM 洪泛算法通过为消息分配不相交的转发区间,使其沿着一棵生成树的结构传播,避免消息环路的产生,从而避免冗余查询消息。邻居表中添加了冗余节点,消息转发过程中,如果遇到失效节点,选择冗余节点转发消息,降低节点失效造成消息丢失的可能,以提高洪泛算法的容错能力。FIFSM 采用高效的网络维护策略,通过 3 种任务周期性地维护覆盖网,保证了资源搜索的低延迟,提高了动态环境中资源搜索的成功率,使其维护开销随着节点变动率的增加而降低。

1 FIFSM 搜索机制的定义

在 FIFSM 机制中,本文定义了标识符空间、邻居表、以及前向和后向节点。

1.1 标识符空间

FIFSM 机制使用一致性哈希函数^[18]为节点分配一个 m 位的标识符, m 必须足够大,以保证 2 个节点标识符是唯一的。标识符空间是以 2^m 为模依次排列的一个标识符圆环。在标识符空间中,以任意一个点 x 为中心,把距离 x 大小为 2^{m-1} 的点称作 x 的界点,记作 $M(x)$ 。图 1 展示了以 0 为中心且 $m=3$ 的标识符空间。

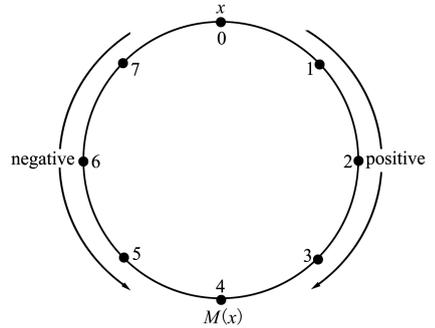


图 1 标识符空间

1.2 邻居表

在 FIFSM 机制中,对任意节点 x 都要维护 2 个邻居表,分别记录标识符空间中顺时针和逆时针方向 x 到 $M(x)$ 区间中的邻居节点。邻居表有 $m-1$ 个邻居表项,其中,第 i 个表项包含变量 start、end、interval 和 neighbors。start 为标识符空间中与节点 x 相对距离为 2^{i-1} 的标识符,end 为标识符空间中与节点 x 相对距离为 2^i 的标识符, $[start, end]$ 表示第 i 个表项所属的区间。interval 为起点是 2^{i-1} ,终点是 2^i 的区间,neighbors 为第 i 个表项的邻居节点列表,其节点数受限于节点冗余度 H 。邻居表相关变量的定义如表 1 所示。

表 1 邻居表的变量定义

Variables	Definitions
entry[i].start	$(x \pm 2^{i-1} \pm 2^m) \bmod 2^m, 1 \leq i < m$
.end	$(x \pm 2^i \pm 2^m) \bmod 2^m, 1 \leq i < m$
.interval	$[2^{i-1}, 2^i], 1 \leq i < m$
.neighbors	modes $\in [start, end]$
H	max number of neighbors

1.3 前向和后向节点

在 FIFSM 机制中,节点 x 除了要维护邻居表以外,还需要维护它的前向节点和后向节点。前向节点是节点 x 沿逆时针方向的第一个节点,记作

$predecessor(x)$ 。后向节点是节点 x 沿顺时针方向的第一个节点,记作 $successor(x)$ 。如图 2 所示,对于节点 N_0 ,它的前向节点 $predecessor(N_0) = N_7$,它的后向节点 $successor(N_0) = N_1$ 。

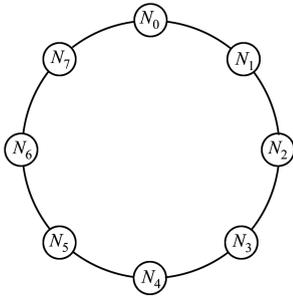


图 2 环状拓扑结构

2 FIFSM 搜索机制的洪泛算法

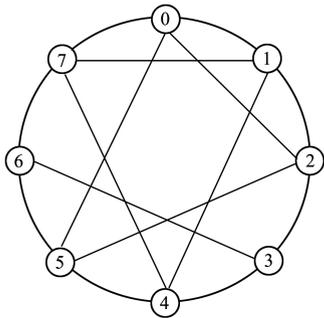
2.1 算法定义

定义 1 转发区间是标识符空间中一段连续的区间,用于限制消息的转发范围。转发区间定义为 $[LL, RL)$,其中,LL 是转发区间的左边界,RL 为转发区间的右边界。

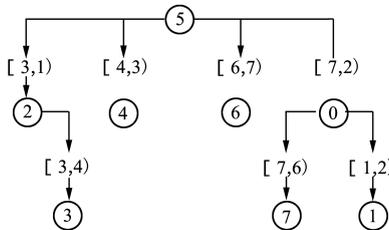
定义 2 洪泛的消息定义为 $message(Info, LL, RL)$,其中,LL 和 RL 为消息中添加的 2 个 m 位的字段,表示消息的转发区间,Info 表示消息的内容。

2.2 算法描述

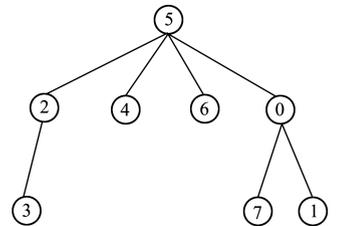
在传统的洪泛算法中,节点收到消息后,盲目地向所有邻居节点转发消息,导致消息的重复到达,产生大量的冗余消息。在 FIFSM 的洪泛算法中,节点收到消息后,仅向转发区间内的部分邻居节点转发消息,当转发区间内不存在邻居节点时,则停止转发。如图 3 所示,其中图 3a) 为节点之间的连接关系,图 3b) 展示了洪泛时消息的转发过程,图中的区间为节点的转发区间。例如,当节点 5 发起洪泛时,它选择邻居节点 2,4,6,0 发送消息,当节点 2 收到消息后,它向转发区间中的邻居节点 3 转发消息。节点 3 收到消息后,发现转发区间中不存在节点,则停止转发,其他节点与之类似。从图中可以看出,节点之间的转发区间是不相交的,且每次转发后,节点不会出现在之后的转发区间中,保证了节点上消息的不重复到达,避免了冗余信息的产生。图 3c) 展示了洪泛过程中消息的传播路径。可以看出,消息的传播路径是 1 棵生成树,消息到达任意节点仅 1 次。



a) 网络拓扑结构



b) 消息转发过程



c) 消息传播路径

图 3 洪泛生成树的生成过程

2.3 算法实现

FIFSM 的洪泛算法不仅能实现全网洪泛,也可以针对特定的范围进行洪泛,本文称之为范围洪泛。

2.3.1 发起洪泛

1) 全网洪泛

对于当前节点 x ,分别从邻居表的每一个表项中选择一个邻居节点,对于从表项 i 中选出的邻居节点 y ,把表项 i 的 start 和 end 赋予消息 m 的 LL 和 RL 字段,并向节点 y 发送消息 m 。如果表项 j 不存在

邻居节点,即 neighbors 为空,则将表项 j 所属的区间 $[start, end)$ 并入下一个节点的转发区间中。

2) 范围洪泛

对于当前节点 x ,在邻居表中找到一个目的范围 $[LL, RL)$ 内的邻居节点 y ,向它发送一个封装了 $[LL, RL)$ 的消息 m ,节点 y 收到消息 m 后就会在该范围内进行洪泛。

2.3.2 转发消息

节点 y 收到消息 m 后,对于邻居表中属于转发

区间 $[LL, RL)$ 内的表项 i , 检查其 neighbors 是否为空, 如果不为空, 则从 neighbors 中选出一个邻居节点 z , 把表项 i 的 start 和 end 赋予消息 m 的 LL 和 RL 字段, 向邻居节点 z 发送消息 m 。如果表项 i 不存在邻居节点, 即 neighbors 为空, 则将表项 i 所属的区间 $[start, end)$ 并入下一个节点的转发区间中。

2.4 算法分析

假设洪泛过程中消息的每一次转发都成功, 且不存在失效节点, 以下从节点丢失和消息冗余两方面进行算法可行性分析。

1) 节点丢失节点 x 转发消息时, 存在以下 4 种情况, 如图 4 所示。在图 4a) 中, x 的前向和后向节点都位于转发区间 $[LL, RL)$ 中, 在图 4b) 和图 4c) 中, x 的前向或后向节点位于转发区间 $[LL, RL)$ 中, 对于这 3 种情况, x 的前向和后向节点至少有一个位于转发区间中, 所以节点 x 可以把消息转发出去, 不丢失节点。在图 4d) 中, x 的前向和后向节点都不在转发区间 $[LL, RL)$ 中, 由于前向和后向节点是离 x 最近的节点, 因此 $[LL, RL)$ 内不存在其他节点, 节点 x 停止转发, 不丢失节点。综上, 该算法能够保证遍历所有的节点而不会出现节点丢失的现象。

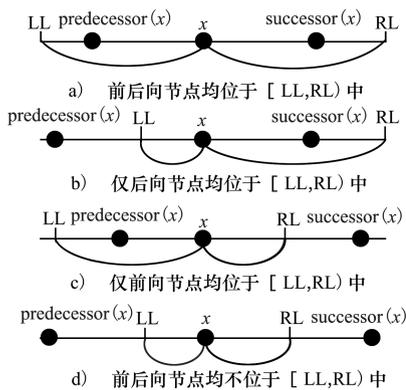


图 4 消息转发的情况

2) 消息冗余

该算法保证了节点之间具有不相交的转发区间, 且被访问过的节点不会出现在之后的转发区间中, 从而保证了每个节点收到且仅仅收到 1 次相同的消息, 不会产生冗余消息。

3 FIFSM 搜索机制的网络维护

P2P 网络的环境是高度动态的, 需要在频繁加入和撤离节点时保证网络的稳定性。网络维护主要

包括两部分: 节点的加入和撤离以及邻居表的维护。

节点 n 加入网络时, 需要通过外部机制联系到一个在线的节点, 由该节点在全网洪泛一个节点连接请求。如果一个节点能够接受节点 n 作为邻居节点, 则该节点就会直接发送回复消息给节点 n , 最后它们对方添加到各自的邻居表中。在 FIFSM 搜索机制中, 当节点 y 收到节点 x 的连接请求时, 在如下 2 种条件之一发生的情况下, 节点 y 会主动连接节点 x : ① 邻居表项的节点数没有达到节点冗余度 H ; ② 节点 x 是节点 y 的前向或后向节点。

而节点 n 撤离网络时, 会发送“leave”消息给它的邻居节点, 当这些节点收到“leave”消息时, 便将节点 n 从其邻居表中删除。

3.1 邻居表维护

在动态环境下, 邻居表中可能出现失效节点、前向和后向节点不一致以及空邻居表项, 会降低 FIFSM 搜索机制的性能。为此, 本文提出了 3 种任务, 分别解决相应问题。其中, fault detector 负责探测失效节点并修复邻居表, SP fixer 负责维护前向和后向节点的一致性, connect task 负责为节点增加新的连接。

1) fault detector

节点的故障撤离会导致邻居表中失效节点的产生, 为了探测失效节点并修复邻居表, fault detector 周期性地 (60 s) 向邻居节点发送“Are you there!”的探测消息, 然后等待回复, 如果在几次询问后仍未收到某个节点的回复, 便向它发送“leave”消息, 并把该节点从邻居表中移除。

2) SP fixer

节点加入网络后, 由于网络的动态性, 节点之间维护的前向和后向节点可能不一致。在图 5a) 中, 节点 a, b, c 依次排列在节点标识符空间中, 节点 a 维护的后向节点是 c , 而节点 c 维护的前向节点是 b 而不是 a , 即节点 a 的后向节点和节点 c 的前向节点不一致。这种情况下, 节点 a 会在范围内发起洪泛, 当节点 b 收到消息后, 发现节点 a 是自己的前向节点, 则主动与节点 a 建立连接, 这样节点 a 的后向节点和节点 b 的前向节点都实现了更新。同理, 图 5b) 展示了节点 c 的前向节点的维护过程。

3) connect task

在动态的网络中, 节点的频繁撤离会造成节点邻居表中空表项的产生, 即该表项的邻居节点数为 0, 不利于查询消息的高效转发, 增加洪泛搜索的延

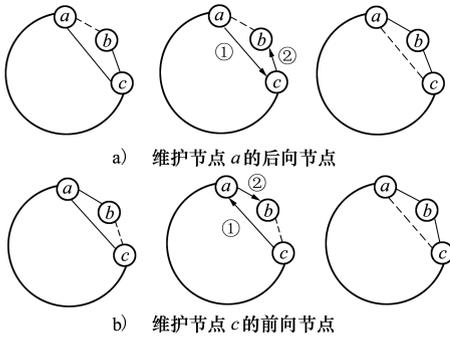


图 5 前向和后向节点的维护策略

迟。为了降低洪泛搜索的延迟,需要为邻居表增加连接。connect task 会定期地检查邻居表中的每一个表项,并根据表项的邻居节点数进行相应处理。

1) 当邻居节点数为 0 时,connect task 会主动在该表项所属的区间 [start, end] 内洪泛。同时,接受来自其他节点的连接请求。

2) 当邻居节点数不为 0 且小于 H 时,connect task 不会主动洪泛。但它仍然接受来自其他节点的连接请求。

3) 当邻居节点数达到 H 时,connect task 既不主动洪泛,也不接受来自其他节点的连接请求。

4 仿真实验

为了评估 FIFSM 搜索机制的性能,本文在 OMNET++^[19] 仿真平台上,建立了 FIFSM 搜索机制的仿真模型(以下简称为 FIFSM 模型),并进行了一系列的实验。实验中,我们采用大小为 2^{15} 的标识符空间。

4.1 静态评估

为了评估 FIFSM 洪泛算法的有效性和容错能力,本文进行了 2 组实验。实验开始后,所有的节点依次加入网络,并在实验过程中保持在线状态。

4.1.1 FIFSM 洪泛算法的有效性

为了把 FIFSM 洪泛算法与传统的洪泛算法进行对比,本文参考 Gnutella 0.4 协议,在 OMNET++ 平台上建立了 Gnutella 的仿真模型,并在相同的条件下进行实验。实验中,我们设置不同的网络规模,节点数 N 从 2^3 到 2^{14} 依次增长,选择不同的节点发起洪泛,观察是否所有的节点都被访问到了,并统计产生的消息数。图 6 为 FIFSM 模型和不同度数的

Gnutella 模型中洪泛遍历所有节点产生的消息数对比实验结果,其中,度数指 Gnutella 中每个节点的连接数。从图中可以看出,在相同的网络规模下,Gnutella 模型产生了比 FIFSM 模型更多的消息,并随着网络规模和度数的增大而大幅度增加,而 N 个节点的 FIFSM 模型中仅产生 $N-1$ 个消息。这是由于传统的洪泛算法会产生大量冗余消息,而 FIFSM 洪泛算法不产生任何冗余消息。

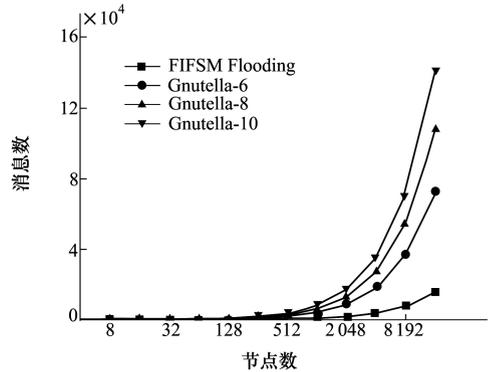


图 6 洪泛遍历所有节点的消息数

4.1.2 FIFSM 洪泛算法的容错能力

在 FIFSM 洪泛算法中,查询消息携带了接收节点的转发区间,查询消息的丢失会导致该转发区间内节点的丢失。为了评估 FIFSM 洪泛算法在故障环境下的容错能力,本文定义 α 和 β , 分别代表链路故障下的丢包率和失效节点率。实验中,我们设置 $N = 1\ 000$ 和 $4\ 000$ 。

实验中,我们设置丢包率 α 在 $0.1\% \sim 0.6\%$ 范围内, $\alpha = 0.1\%$ 表示 1 000 个消息会有 1 个消息丢失。节点加入网络后,我们让每个节点都发起 1 次洪泛,统计节点被访问到的次数,并计算丢失的节点数。图 7 为不同 α 对应的节点丢失率曲线。

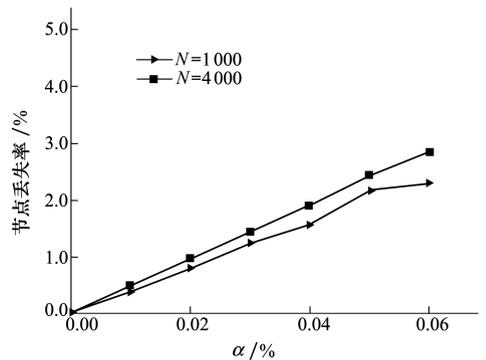


图 7 链路故障下洪泛的节点丢失率

从图中可以看出,随着 α 的增加,节点丢失率呈线性增长,而且随着节点数的增加,节点丢失率也随之增加。可以看出,消息的丢失对搜索的成功率有着很大的影响,实际网络中,可以采用 TCP 作为传输层协议,保证消息传输的可靠性,避免链路故障造成的消息丢失的现象。

为了评估 FIFSM 洪泛算法对节点故障的容错能力,实验中,令节点以 β 的比率故障离开,造成节点失效的情况,我们统计 30 s 内若干次洪泛后丢失的节点数。图 8 展示了 $N=1\ 000$ 的网络,在不同的 H 和 β 下节点的丢失率曲线。

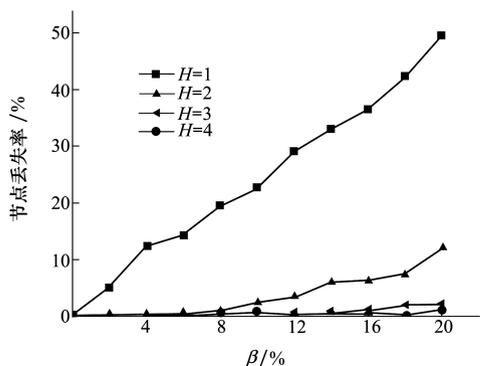


图 8 节点故障下洪泛的节点丢失率

从图中可以看出,当 $H=1$ 时,节点的丢失率随着 β 的增大而迅速增加,但随着 H 的增大,节点丢失率显著降低。当 $H=4$ 时,节点的丢失率几乎为 0。这是因为邻居表采用了冗余节点机制,每一个表项的邻居节点数可以达到 H 。当节点转发消息的时候,如果遇到失效节点,可以选择其他的冗余节点转发消息。假设节点失效概率为 p ,则 1 次转发失效的概率为 p^H 。因此随着 H 的增大,FIFSM 洪泛算法的容错能力也随之提高。

4.2 动态评估

为了评估动态环境下 FIFSM 机制的搜索成功率、维护开销以及搜索延迟,根据文献[20]的策略,本文建立了动态的网络环境。我们在网络中随机指定一部分节点为“保留节点”,它们在实验的过程中不离开网络,始终保持在线状态。其他的节点则为“非保留节点”,这些节点在实验开始后,以 0.5 的概率选择是否加入网络。这样,网络初始化后,非保留节点有大约一半处于在线状态,一半处于非在线状态。在实验过程中,每隔一段时间,实验中设置为 60 s,“非保留节点”以概率选择是否从在线状态转

换到非在线状态,反之亦然,这样我们就建立了一个变动率为 λ 的动态网络。本文针对不同的 λ 值进行实验,实验的运行时间设置为 3 600 s。

4.2.1 FIFSM 的搜索成功率

为了评估动态环境中 FIFSM 机制的搜索成功率,实验设置 SP fixer 工作周期为 20 s,实验开始后,网络中的在线节点每 10 s 发起 1 次洪泛,“保留节点”将记录被访问到的次数,同时,我们还将记录所有节点的总洪泛次数,于是就可以计算出实验过程中每个“保留节点”的命中率,即访问次数占总洪泛次数的百分比。通过对所有“保留节点”的命中率取平均值,则近似得到在当前 λ 的动态环境下洪泛的节点命中率。图 9 展示了 $N=1\ 000$ 和 4 000 时, λ 在 $[0,0.3]$ 范围内的动态环境下,洪泛时节点的命中率曲线。从图中可以看出,随着 λ 的增大,节点命中率逐渐降低,但均保持在 99.9% 以上,并且不随网络规模的增大而剧烈变化,保持了动态环境下的高度可靠性。

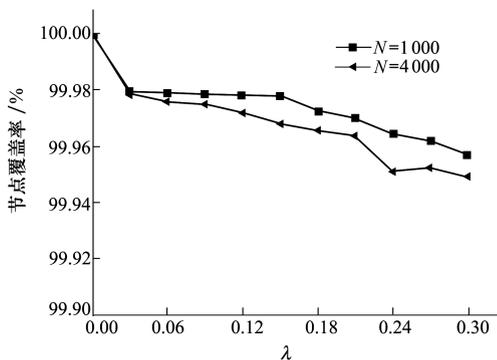


图 9 动态环境下节点的命中率

4.2.2 FIFSM 的维护开销

为了评估 FIFSM 的网络维护开销,实验中,我们主要统计节点平均每分钟收到的控制消息数,这里的控制消息主要指 SP fixer 和 connect task 产生的控制消息。

设 SP fixer 工作周期为 20 s,connect task 工作周期为 60 s,节点个数,图 10 为不同 λ 下的控制消息数曲线。从图中可以看出,SP fixer 产生的控制消息数保持在 7 个左右,这是因为 SP fixer 仅在当前节点的前向和后向节点区间之内洪泛控制消息,所以消息不会被大范围转发。此外,connect task 产生的控制消息数则随着 λ 的增加而逐渐减小,因为当节点加入网络的时候,会在全网进行洪泛,与网络中

的节点建立连接,增加其邻居表的饱和度。虽然节点的离开可能会导致其他节点空路由表项的产生,但由于每个表项的节点数最多可以达到节点冗余度 H ,降低空表项产生的可能,因此随着 λ 的增大,节点的加入越多,connect task 洪泛的次数也随之减少。综上,在动态环境中,FIFSM 的维护开销随着节点变动率 λ 的增加而降低,大大减少了网络维护开销。

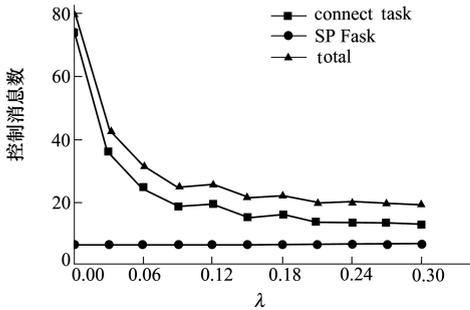


图 10 动态环境中 FIFSM 的维护开销

4.2.3 FIFSM 的搜索延迟

为了评估 FIFSM 机制在动态环境中的搜索延迟,我们分别在 $\lambda=0.15, 0.25$ 的动态环境中进行实验。设 connect task 的工作周期为 60 s,实验开始后,在线节点在每个固定的时间间隔(10 s)后发起一次洪泛,统计查询消息经历的跳数,实验结果如表 2 所示。表 2 的第一列为实验运行的参数,包括节点变动率 λ 、节点数 N 和节点冗余度 H 。第二列至第五列为不同跳数范围内包含的消息数占总消息数的比例。最后一列为所有消息的平均跳数。实验结果表明,在 $N=1\ 000$ 和 $4\ 000$ 的不同 λ 的动态网络中,消息的跳数主要集中在 8 跳以内,平均跳数均小于 5 跳。对比实验 $\langle 0.15, 1\ 000, 1 \rangle$ 和 $\langle 0.15, 1\ 000, 4 \rangle$ 以及 $\langle 0.15, 4\ 000, 1 \rangle$ 和 $\langle 0.15, 4\ 000, 4 \rangle$ 的结果可知,在相同的 N 和 λ 下,随着 H 的增大,消息的平均跳数也随之减少。综上,FIFSM 机制在动态环境中具有较低的洪泛搜索延迟。

表 2 FIFSM 的洪泛搜索延迟

$\langle \lambda, N, H \rangle$	% of 1-4 hops	% of 5-8 hops	% of 9-12 hops	% of >12 hops	avg hop count
$\langle 0.05, 1\ 000, 1 \rangle$	50.48	49.26	0.27	0	4.13
$\langle 0.15, 1\ 000, 1 \rangle$	51.60	47.93	0.46	0	4.10
$\langle 0.25, 1\ 000, 1 \rangle$	51.91	47.70	0.38	0	4.10
$\langle 0.05, 1\ 000, 4 \rangle$	76.15	23.85	0	0	3.67
$\langle 0.15, 1\ 000, 4 \rangle$	77.85	22.15	0	0	3.63
$\langle 0.25, 1\ 000, 4 \rangle$	79.29	20.71	0	0	3.63
$\langle 0.05, 4\ 000, 1 \rangle$	35.71	63.94	0.36	0	4.95
$\langle 0.15, 4\ 000, 1 \rangle$	38.71	60.87	0.42	0.002	4.91
$\langle 0.25, 4\ 000, 1 \rangle$	40.64	58.97	0.39	0.001	4.88
$\langle 0.05, 4\ 000, 4 \rangle$	50.05	49.91	0.03	0	4.43
$\langle 0.15, 4\ 000, 4 \rangle$	53.04	46.96	0.01	0	4.36
$\langle 0.25, 4\ 000, 4 \rangle$	53.61	46.39	0	0	4.34

5 结 论

本文提出了一种改进的非结构化 P2P 网络洪泛搜索机制-FIFSM,其主要内容和贡献包括:(1)提出了在非结构化 P2P 网络中使用一致性哈希函数为节点分配标识符,建立结构化的邻居表的优化策略,并在邻居表中增加冗余节点,提高洪泛算法在节点失效时的容错能力。(2)提出并实现了基于转发区间的洪泛算法,避免了冗余消息的产生,降低了洪泛搜索的代价。另外,增加了范围洪泛的功能,以降低特定范围内搜索的开销。(3)提出了高效的网络维护机制,其维护开销随着网络动态程度的增加而降低。实验结果表明,FIFSM 机制不仅降低了非结构化 P2P 网络基于洪泛的资源搜索的开销,提高了非结构化 P2P 网络的可扩展性,并且能够在动态环境中保持高度的稳定性和可靠性。未来的工作将考虑节点之间的实际物理距离,建立与底层物理网络更加匹配的覆盖网拓扑,利用节点之间的邻近性,进一步降低 FIFSM 洪泛搜索的消息延迟,合理利用网络带宽。

参考文献:

- [1] Ratnasamy S, Francis P, Handley M, et al. A Scalable Content-Addressable Network[C]//Proceedings of the 2001 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, New York: ACM Press, 2001: 149-160
- [2] Rowstron A I T, Druschel P. Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems[C]//Proceedings of the IFIP/ACM International Conference on Distributed Systems Platforms Heidelberg, Springer-Verlag, 2001
- [3] Stoica I, Morris R. Chord: A Scalable Peer-to-Peer Lookup Service for Internet Applications[J]. ACM Sigcomm Computer Communication Review, 2001, 31(4):149-160
- [4] Clip2com. The Gnutella Protocol Specification. <http://rfc-gnutella.sourceforge.net/Development>, 2010
- [5] Cohen E, Shenker S. Replication Strategies in Unstructured Peer-to-Peer Networks[J]. ACM Sigcomm Computer Communication Review, 2002, 32(4):177-190
- [6] Morselli R, Bhattacharjee B. Efficient Lookup on Unstructured Topologies[J]. IEEE Journal of Communications, 2007, 25(1): 62-72
- [7] Ritter J. Why Gnutella Can't Scale. No, Really. <http://www.darkridge.com/~jpr5/doc/gnutella.html>, 2001
- [8] 朱桂明,郭得科,金士尧,等. 基于副本复制和 Bloom Filter 的 P2P 概率路由算法[J]. 软件学报,2011, 22(4):773-781
Zhu Guiming, Guo Deke, Jin Shiyao, et al. P2P Probabilistic Routing Algorithm Based on Data Copying and Bloom Filter[J]. Journal of Software, 2011, 22(4):773-781 (in Chinese)
- [9] Kitamura H, Fujita S. A Biased k-Random Walk to Find Useful Files in Unstructured Peer-to-Peer Networks[C]//2009 International Conference on Parallel and Distributed Computing, Applications and Technologies, 2009: 210-216
- [10] 叶培顺. 非结构化 P2P 网络的一种改进搜索算法[J]. 计算机与现代化, 2013(12): 44-47
Ye Peishun. Improved Search Algorithm for Unstructured P2P Network[J]. Computer and Modernization, 2013(12): 44-47 (in Chinese)
- [11] 马文明,孟祥武,张玉洁. 面向非结构化 P2P 网络的双向随机漫步搜索机制[J]. 软件学报, 2012, 23(4): 894-911
Ma Wenming, Meng Xiangwu, Zhang YuJie. Bidirectional Random Walk Search Mechanism for Unstructured P2P Network[J]. Journal of Software, 2012, 23(4): 894-911 (in Chinese)
- [12] Gkantsidis C, Mihail M, Saberi A. Random Walks in Peer-to-Peer Networks[C]//IEEE Conference on Computer Communications, 2004:120-130
- [13] Jagadish H V, Ooi B C, Vu Q H. BATON: a Balanced Tree Structure for Peer-to-Peer Networks[C]//Proceedings of the 31st International Conference on Very Large Data Bases, 2005
- [14] Hu Z. Improved Algorithm of Unstructured P2P Network Topology Structure[C]//2009 International Symposium on Intelligent Ubiquitous Computing and Education, 2009: 358-361
- [15] 杨亚,宋俊德. 一种适合异构 P2P 网络的树形结构覆盖层[J]. 高技术通讯, 2009, 19(3): 230-236
Yang Ya, Song Junde. TSOHEN: a Tree Structure Overlay for Heterogeneous P2P Networks[J]. Chinese High Technology Letters, 2009, 19(3): 230-236 (in Chinese)
- [16] Castro M, Costa M, Rowstron A. Should We Build Gnutella on a Structured Overlay? [J]ACM Sigcomm Computer Communication Review, 2004, 34(1):131-136
- [17] Elansary S, Alima L O, et al. Efficient Broadcast in Structured P2P Networks[J]. Peer-to-Peer Systems II, 2003, 2735: 304-314
- [18] 林雅榕,侯整风. 对哈希算法 SHA-1 的分析和改进[J]. 计算机技术与发展, 2006, 16(3):124-126
Lin Yarong, Hou Zhengfeng. Analysis and Improvement to Algorithm of SHA-1[J]. Computer Technology And Development, 2006, 16(3):124-126 (in Chinese)
- [19] Melamed R, Keidar I. Araneola: A Scalable Reliable Multicast System for Dynamic Environments[J]. Journal of Parallel and Distributed Computing, 2008, 68(12):1539-1560

An Improved Flooding Based Search Mechanism in Unstructured P2P Network

Lu Wei, Zhou Tao, Xing Weiwei

(School of Software Engineering, Beijing Jiaotong University, Beijing 100044, China)

Abstract: In the unstructured P2P networks, the flooding-based search algorithm is used to search resources; however, with increasing nodes and network scale, flooding-based search will produce large amount of redundant query messages, which will lead to heavy traffic and congestion of the network. We propose a Forwarding Interval based Flooding Search Mechanism (FIFSM). By assigning a disjoint forwarding interval to each message, they spread along a spanning tree to avoid message loops, thus eliminating redundant messages. The efficient network maintenance strategy is presented in FIFSM; this ensures the stability of the network in dynamic environment at a very low cost. Experimental results and their analysis show preliminarily that FIFSM, as an efficient search mechanism in unstructured P2P network, can reduce flooding overhead and achieve high success rate of resource search and low latency.

Key words: algorithms; computer system; cost reduction; fault detection; fault tolerance; network management; network performance; packet loss; peer to peer networks; reliability analysis; stability; time delay; topology; unstructured P2P networks; flooding search; forwarding interval; spanning tree

《工程索引》数据库 (Ei Compendex database) 收录的第一作者为王伟论文的 2006 年以后的 Ei 主题词达到 10 个

Ei Compendex database 收录的第一作者为王伟的论文是《西北工业大学报》第 33 卷第 1 期(2015 年 2 月出版)的第 1 篇, Ei Compendex 的编号是 20150900580151, 10 个 2006 年以后(含 2006 年)出现的 Ei 主题词是: aeroelasticity, angular momentum, angular velocity, dynamic models, elastic deformation, flexible wings, momentum, nonlinear analysis, rigid wings, unmanned aerial vehicles(UAV)。

胡沛泉
2015 年 4 月