

# 基于虚拟节点交换方法优化 P2P 覆盖网性能

袁瑞峰, 李振华, 陈贵海

(南京大学 计算机科学与技术系 邮编 210093)

**摘要:** 对等网络 (P2P) 通过在底层物理网上建立一个虚拟的应用层覆盖网进行工作, 覆盖网的拓扑结构通常与底层物理网很不一致, 这种不一致性严重增加了网络通信开销、降低了系统工作性能。本文提出一种称作 VNS 的虚拟节点交换的方法来动态、自适应地优化 P2P 覆盖网拓扑结构, 从而有效提高 P2P 覆盖网与物理网的一致性、提升系统工作性能。VNS 方法的优点表现在 3 个方面: (1) 优化覆盖网拓扑结构所需要的信息通过节点监听路过或收到的消息而获得, 因此信息收集不带来额外通信开销; (2) 通过给每个节点安排一个有探测作用的虚拟节点, 以辅助节点在覆盖网拓扑结构中找到其更合适的位置, 该方法是新颖的且具有很高的通用性; (3) 通过基于随机过程的理论分析和在多种拓扑结构的覆盖网上的模拟实验, VNS 方法展现出比当前代表性的提升覆盖网一致性方法 LTM 更快的优化速度、更好的稳定状态优化效果和更低的优化通信开销。

**关键词:** 对等网络 拓扑性能 虚拟节点

## Enhance overlay's performance with VNS

Yuan Ruifeng, Li Zhenhua, Chen Guihai

Department of Computer Science and Technology, Nanjing University, Nanjing, Jiangsu  
210093. China

**Abstract** - Peer-to-peer systems build up a virtual overlay above current underlying physical network. So the overlay always mismatches the physical network. This mismatch problem causes longer lookup latency, lower performance and unnecessary traffic. We propose a virtual node switching (VNS) technique that optimize the overlay's topology, and enhance overlay's performance dynamically and adaptively. There are three notable advantages of VNS: (1) VNS gets overlay information from local node's query and message transmission that leads to little traffic cost. (2) Each has a virtual as detector to help getting better location in overlay that is novel and general for application. (3) Based on Stochastic Process analysis and numbers of simulation proof, VNS shows faster optimizing speed and better result than LTM which is well-known, and meanwhile with lower traffic cost.

**Key Words:** peer-to-peer network, topology performance, virtual node

### 1. 引言

对等网络 (P2P) 通过在底层物理网上建立一个虚拟的应用层覆盖网进行工作, 覆盖网的拓扑结构通常与底层物理网很不一致, 这种不一致性严重增加了网络通信开销、降低了系统工作性能。拓扑不一致问题的解决思路可归结为如何使得节点在覆盖网拓扑结构中占据更

好的位置，即如何选取更合适的节点 ID 与邻居，从而提高 P2P 覆盖网与物理网的一致性、提升系统工作性能。

在解决拓扑不一致问题方面，现存多种不同的解决方案，这些方案主要可分为三类[2]：  
(1) 路由优化：节点发送或转发消息时，不仅由覆盖网上的距离来决定，同时也会考虑实际的路由往返时间[4][5]。路由优化方法本质上并没有改变拓扑结构，只是在路由选择时做了一些启发，因此没有从根本上解决拓扑不一致问题；  
(2) 邻居优化：通过改变节点的邻居来优化覆盖网拓扑结构，令其与物理网拓扑尽量一致。这一方法优化效果较好，但现存方法[1][3]引入的额外通信开销较大，给系统带来多余的负担；  
(3) ID 优化：在构造覆盖网时改变 ID 的选取策略，使节点 ID 同其位置信息相关联。这在一定程度上提高了拓扑一致性，但必须重新设计节点 ID 的生成方法，并且该方法由于暴露出节点位置信息，所以还带来了安全性减弱的负面影响。

本文提出一种称作 VNS (Virtual Node Switching) 的虚拟节点交换的方法，可以动态、自适应地优化 P2P 覆盖网拓扑结构，从而有效提高 P2P 覆盖网与物理网的一致性、提升系统工作性能。VNS 方法通过监听路过或收到的消息来收集覆盖网优化所必需的信息，通过具有辅助作用的虚拟节点来探测并找到节点在覆盖网拓扑结构中可能占据的更好位置，当节点发现覆盖网中对其来说更好的位置时，将改变自身 ID 与邻居以占据新位置。VNS 方法利用本地监听收集到的网络信息做出优化决策，包括是否进行位置改变与何时改变，因此不会带来额外的通信开销。另一方面，VNS 方法对 P2P 覆盖网的拓扑类型、节点邻居限制、消息路由方式等都没有特定的要求，因此具有很高的通用性。该方法有效结合了以上邻居优化与 ID 优化两种方法的优点，并通过本地监听与虚拟节点机制克服了以上方法的不足，提供了高效、低开销、通用性强的新颖解决方案。

本文对领域的贡献具体来说可以总结为以下几个方面：

(1) 以往解决拓扑不一致问题的各种方法多通过主动发送消息以收集所需信息，给网络带来较大的额外通信开销，而 VNS 方法优化覆盖网拓扑结构所需要的信息通过节点监听路过或收到的消息而获得，因此信息收集不带来额外通信开销；

(2) 虚拟节点的使用在 P2P 领域由来已久，但一直被用来进行平衡负载，本文第一次将虚拟节点方法用于覆盖网拓扑结构的优化，通过给每个节点安排一个具有探测作用的虚拟节点，以辅助节点在覆盖网拓扑结构中找到其更合适的位置。该方法是新颖的且具有很高的通用性；

(3) 通过基于随机过程的理论分析和在多种拓扑结构的覆盖网上的模拟实验，VNS 方法展现出比当前代表性的提升覆盖网一致性方法 LTM 更快的优化速度、更好的稳定状态优化效果和更低的优化通信开销。

以下为本文的组织方式：第 2 节给出了 VNS 方法工作机制的详细描述，第 3 节对 VNS 方法进行了理论分析与大量的实验分析结果，并同常用方法 LTM 进行了比较，第 4 节则总结了我们的工作。

## 2. VNS 方法

VNS 方法基于对拓扑结构中点集与边集的操作，实现了对拓扑网络整体性能的改进。其基本办法是，当一台计算机加入对等网络时，将同时运行两个虚拟节点，两个节点通过相互共享信息并对自身节点相关信息的比较，来决定计算机应直接连接的其他节点计算机，即主节点的邻居，通过这一方式，得到计算机在当前网络环境中更合适的位置，这一过程主要通过以下三步来实现：

### 2.1 虚拟节点的建立与运行

当一台计算机希望加入某一对等网络系统进行信息的获取与发布时，它必须在本地运行

一个覆盖网节点并通过该节点与网络中的其他计算机通信，我们将这个覆盖网节点称为主节点 (Primary Node)。同时，VNS 会在同一计算机中运行另一个 ID 与主节点不同的虚拟节点，我们称其为次节点 (Secondary Node)。因此同一台计算机中将同时运行两个不同的节点，但以上两个节点的功能是不同的：主节点与普通节点功能相同，负责计算机信息的查询与发布；而次节点不负任何信息的发布，它可以用来分担主节点的部分查询负载，以下对于两种不同的使用模式进行解释：

A) 如果协议要求节点发出多条查询消息来完成一次查询，如 Gnutella 类的无结构模型，则主节点与次节点合作，主节点发送部分查询消息，次节点辅助发送余下部分的查询消息。基于负载均衡的考虑，我们建议主节点与次节点各发送一半的查询消息，因为以上操作的目的是保证主节点与次节点可以获得等量的路由信息以进行无偏颇的优化决策。

B) 如果协议规定节点进行一次查询时仅发送一条查询消息，如 Chord 等结构化模型，则 VNS 需要主节点与次节点各自发送一条相同的查询信息。次节点发出的查询消息可以被视为主节点发出消息的冗余消息以取得更快的查询结果，同时次节点还能收集到该次查询的路由信息。

VNS 将收集到的路由信息归并到一个称为查询时延表  $LLT$  的数据集中，其中包含每个目的节点与当前节点（主节点或次节点）的查询时延。数据集  $LLT$  用于下一步计算。

## 2.2 路由信息收集

主节点与次节点经过一定时间的运行，通过发起查询或转发查询消息的方式收集到一定数量的  $LLT$  数据。每个节点收集到的  $LLT$  表项数应大于该节点的邻居节点数，因为信息量少将使计算结果产生较大偏差并导致 VNS 产生错误的交换操作。我们将节点收集一次完整信息集合的时间段称为一个收集周期，将初始时间定义为  $t_0$ ，将经过第  $n$  个收集周期后的时间定义为  $t_n$ ，则可以计算出主节点与次节点每个阶段的平均查询时延  $ARTT$ ：

$$ARTT(x, t_n) = \sum LLT(x, t_n) \quad (x = p \text{ 或 } x = s)$$

我们定义另一个变量  $VAL$  来转换  $ARTT$  用于计算最终结果：

$$VAL(x, t_n) = u * ARTT(x, t_n) + (1 - u) * VAL(x, t_{n-1}) \quad (x = p \text{ 或 } x = s)$$

此处通过递归方法计算生成  $VAL$  是为了对网络的突发性波动进行缓冲，防止波动对计算产生的不良影响，其中  $u$  为权重，用来决定不同收集周期的信息的时效性，由以下公式定义：

$$u = \frac{ARTT(x, t_n)}{ARTT(x, t_n) + VAL(x, t_n)} \quad (x = p \text{ 或 } x = s)$$

最终计算结果  $Diff$  反映了同一计算机中主节点与次节点性能的差异：

$$Diff = VAL(p, t_n) - VAL(s, t_n)$$

$Diff$  越大，表明次节点的性能越优于主节点的性能，不同的 P2P 系统可以根据网络情况给出  $Diff$  的阈值  $MIN\_DIFF$ 。如果  $Diff > MIN\_DIFF$ ，表明次节点的性能明显优于主节点，并且节点交换带来的利益要高于节点交换开销，此时交换结点是有利的；否则交换结点不利于网络整体性能。

## 2.3 虚拟节点交换

节点交换的过程面临以上两种不同的情况：

a)  $Diff > MIN\_DIFF$

这种情况表明节点的交换有利于自身性能与网络整体性能的优化，因此节点直接开始进行主节点与次节点的交换。首先，主节点需要等待之前发出的所有查询都获得了查询结果，因为主节点将会在交换完成后被直接终止，而次结点将继续运行，这是防止消息的丢失对查询结果产生影响。同时在此过程中，应由主节点发出的查询消息，都转由次节点发出，所以该过程将会在一个最大查询时延周期内完成，不会出现循环等待的情况。当主次点确认所有

查询都已经返回结果后，将通知次节点对当前计算机的信息进行发布，即令次节点成为下一周期的主节点，并根据不同系统的要求发出离开系统的消息或直接离开当前覆盖网。VNS 会在主节点离开后，再次生成一个虚拟节点，作为下一周期的次节点，继续以上优化算法。

b)  $Diff < MIN\_DIFF$

这种情况表明节点的交换不利于性能的优化，当前的次节点对网络性能优化无法做出贡献，所以必须重新选择次节点。随后，节点将以相同的主节点继续 VNS 算法的优化，直到再次选择的次节点性能优于主节点。

### 3. 性能评价

本节中我们将通过基于随机过程理论的数学分析给出了 VNS 算法的优化性能，并通过实验展示了 VNS 方法的有效性，同时在相同的实验环境下与解决拓扑不一致问题的代表性方法 LTM 进行了优化性能方面的比较。

#### 3.1 理论分析

首先，我们对 VNS 算法的性能进行理论分析。VNS 算法的基本思想是，当一个节点利用 VNS 方法进行性能优化时，将在覆盖网中探寻适合的辅助节点位置，如果位置更优，则交换至该位置，否则将继续挑选的过程。该算法是纯分布式的算法，各节点的决策不会互相影响。使用随机过程方法，我们将某节点在拓扑结构中可以占据的位置进行性能排序产生一个有序序列，从而我们得出以下定理：

*定理 1. 假设使用 VNS 方法每次优化后再次选择的次节点是均匀分布的，经过  $i$  次优化后，节点可以再次选择到更优次节点的概率为  $1/(i+1)$ ，这等价于节点当前位置处于有序序列中的前  $1/(i+1)$  处。*

证明：限于篇幅关系，详细证明请见

<http://cs.nju.edu.cn/gap/lizhenhua/resources/proofVNS.pdf>

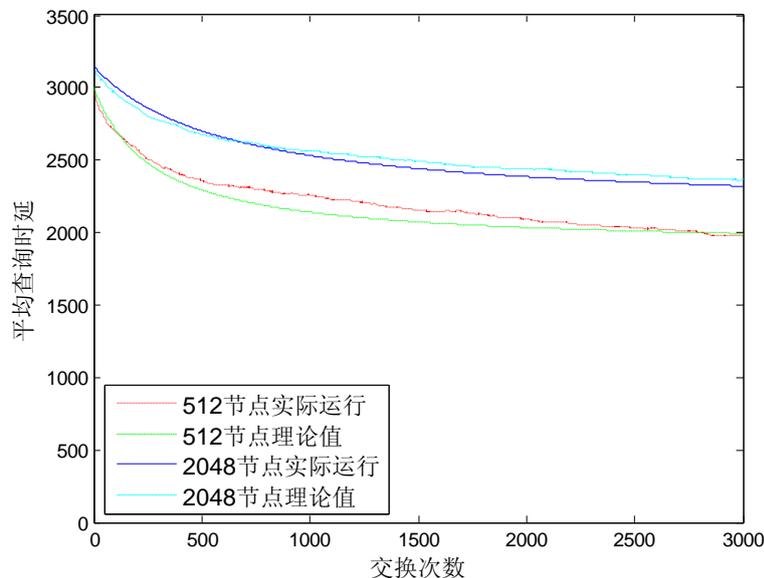


图 1. VNS 实际性能与理论值的比较

#### 3.2 实验分析

我们利用标准的拓扑网络模型，对 VNS 算法在不同节点规模情况下的优化性能进行了实验与分析。同时，在相同的实验环境下，对 VNS 算法与 LTM 算法进行了优化性能方面的比较。

图 2 展示了 VNS 对于不同节点规模的拓扑网络的优化效果，可以看出单次优化对于网络性能的影响是随着节点数目的增加而减小的，优化效果趋于收敛，最终网络状态趋于稳定。

我们以交换次数作为横坐标，这比使用系统运行时间作为横坐标要更为合理。

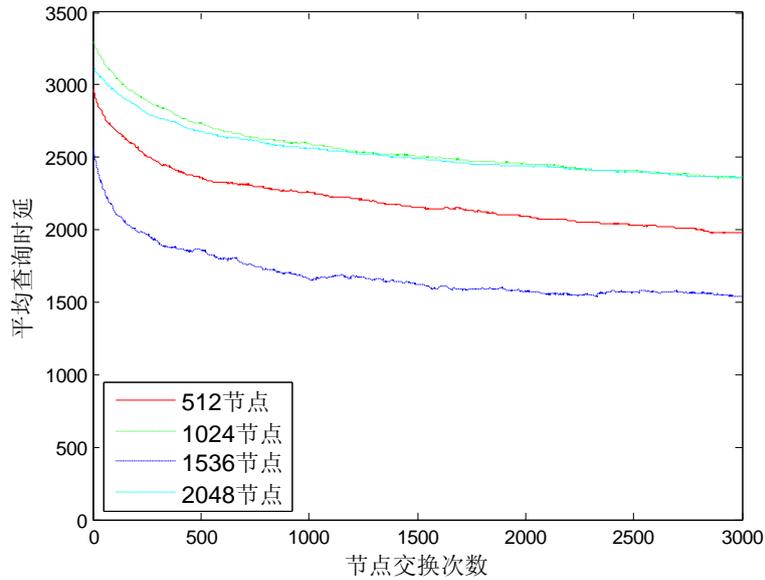


图 2. 不同节点规模下 VNS 性能的比较

我们在相同的实验环境下将 VNS 与解决拓扑不一致问题的代表性方法 LTM 进行了优化性能的比较，比较结果在图 3 中。横坐标参数的单次操作在 VNS 中为一次节点交换，在 LTM 中为一次切边与加边的操作的集合，两者都保证了节点邻居的稳定，对节点平均查询时延与查询范围不会造成额外的影响。可以看出 LTM 方法以常数速度进行优化，而优化的稳定结果较差；VNS 方法以  $O(1/x)$  的速度进行优化，其中  $x$  为节点交换次数，在性能优化速度与优化后的稳定结果方面，VNS 均明显优于 LTM 算法。

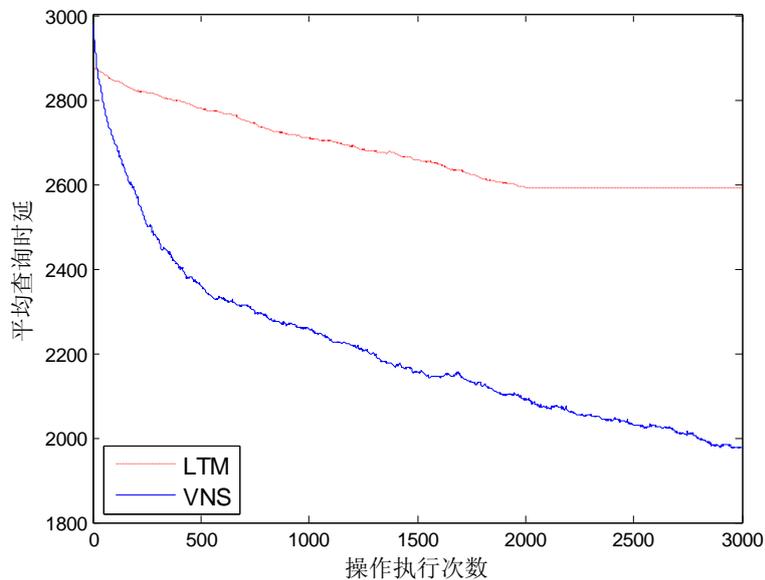


图 3. 1000 节点情况下的比较

## 4. 总结

本文提出一种称作 VNS 的虚拟节点交换的方法来动态、自适应地优化 P2P 覆盖网拓扑结构，从而有效提高 P2P 覆盖网与物理网的一致性、提升系统工作性能。该方法利用本地监听与虚拟节点机制克服了以往常用方法的不足，提供了高效、低开销、通用性强的新颖解决

方案。我们通过基于随机过程的理论分析和在多种拓扑结构的覆盖网上的大量模拟实验，证实了 VNS 方法比当前代表性的提升覆盖网一致性方法 LTM 具有更快的优化速度、更好的稳定状态优化效果和更低的优化通信开销。我们下一步的工作主要集中在不暴露节点位置信息的前提下，充分利用 VNS 方法与节点位置信息，使得节点的 ID 的排序与节点位置趋向一致，进而提出更有效的解决拓扑不一致性问题的解决方案。

## 致谢

本文受到以下项目的资助：

国家自然科学基金项目：实用化对等网络技术的研究（60573131）

江苏省自然科学基金前期预研项目：新型 P2P 计算技术的基础研究（BK2005208）

## 参考文献

- [1] Yunhao Liu, Xiaomei Liu, Li Xiao, Lionel M. Ni, Xiaodong Zhang. 2004. Location-aware Topology Matching in P2P Systems. INFOCOM'04, 2220-2230.
- [2] Krishna P. Gummadi, Ramakrishna Gummadi, Steven D. Gribble, Sylvia Ratnasamy, Scott Shenker, Ion Stoica. 2003. The Impact of DHT Routing Geometry on Resilience and Proximity. SIGCOMM'03, Karlsruhe, 381-394.
- [3] T.Q. Qiu, G. Chen, M Ye, Edward Chan and Ben Zhao, Towards Location-aware Topology in Both Unstructured and Structured P2P Systems, International Conference on Parallel Processing (ICPP2007), Xi'an, China, September 10 - 14, 2007.
- [4] I. Abraham and D. Malkhi. Compact routing on euclidian metrics. In PODC, July 2004.
- [5] Jonathan Ledlie, Michael Mitzenmacher, Margo Seltzer, Peter Pietzuch. Wired Geometric Routing. In IPDPS '07.
- [6] Sylvia Ratnasamy, Mark Handley, Richard Karp, and Scott Shenker. Topologically-Aware Overlay Construction and Server Selection. In Proceedings of the INFOCOMM, 2002.

作者简介：

袁瑞峰(1982- )，男，硕士研究生，研究方向：对等网络，分布式系统仿真与性能测量。

李振华(1983- )，男，硕士研究生，研究方向：对等网络，分布式流媒体处理。

陈贵海(1963- )，男，教授，博士生导师，研究方向：无线网络计算，对等网络，并行计算。