

# Overlay 服务网络上的服务组合模型

潘 耘<sup>1</sup>, 余镇危<sup>2</sup>, 王银燕<sup>2</sup>, 贾民政<sup>3</sup>

(1. 中国传媒大学计算机与软件学院, 北京 100024; 2. 中国矿业大学(北京)计算机系, 北京 100083;

3. 北京工业职业技术学院信息工程系, 北京 100042)

**摘 要:** 分析了在 Overlay 服务网络上进行服务组合的必要性, 指出了在进行组合时服务次序约束所带来的特定困难, 进而给出了服务次序约束的形式化定义以及带 OoS 组播路由问题的数学优化模型。借助于罚函数的方法, 将服务次序约束条件吸收到目标函数中, 从而可以用求解一般组播优化模型的算法来求解新模型。

**关键词:** Overlay 服务网络; 服务次序; 组播; 罚函数方法

## Service Composition Problem on Overlay Service Network

PAN Yun<sup>1</sup>, YU Zhenwei<sup>2</sup>, WANG Yinyan<sup>2</sup>, JIA Minzheng<sup>3</sup>

(1. Computer Science & Software Institute, Communication University of China, Beijing 100024; 2. Department of Computer, China University of

Mining and Technology-Beijing, Beijing 100083; 3. Department of Information Engineering, Beijing Polytechnic College, Beijing 100042)

**【Abstract】** This paper analyzes the necessity for combining services and the particular difficulties considering order of service constraints, and proposes the formal definition of order of service (OoS) and a programming model for multicast routing problem with OoS-constraints in the overlay network environment. Penalty-function method is employed and the OoS-constraints are integrated into the objective function, resulting in that many algorithms for general multicast model can be used to solve the proposed OoS-multicast model.

**【Key words】** Overlay service network; Order of service (OoS); Multicast; Penalty-function method

### 1 概述

网络覆盖技术(Overlay Technique)<sup>[1~4]</sup>是一种无需对基础网络层的协议进行修改即可支持新型服务的有效方式。目前, 覆盖服务网络研究主要针对具体的应用, 在假定覆盖节点都已经选定, 探讨如何在这些节点之间建立覆盖链路(Overlay Links), 以更加有效地支持提供各种覆盖应用(Overlay Applications)<sup>[2,3]</sup>, 而这些种类各异的结构往往是针对不同的特定服务而优化选择的, 对一种服务适合的结构对另外的服务就不一定是高效的。

文献[1,3]提出了通用覆盖服务网络OSN的概念, 这与作者提出的主动Overlay网络的概念<sup>[4]</sup>很类似。这种通用的OSN为各种各样不同的应用所共享; 为了实现这一目标, 文献[3]引入了覆盖代理(Overlay Brokers, OBs)的概念, 这些OBs按照“特定策略”放置在Internet之上。覆盖代理上可以定制服务, 来支持新的应用, 使得从源端到目的端的数据流可得到中间代理节点提供的特殊服务。例如一个视频分布式应用, 在代理节点上实现压缩和解压服务可降低数据传输对带宽的要求。

不同的应用需求可以在通用覆盖服务网络上组合不同的服务, 组合服务就需要按一定的次序进行组合, 例如先压缩后解压, 先加密后解密等; 而不能是相反的次序, 即在进行服务组合时需要考虑服务次序约束问题(把每一个简单服务称为服务因子, 一个复杂的服务是由若干个服务因子按照一定的服务次序组合的)。从目前发表的文章来看, 在国外也有一些大学的研究机构在进行与服务次序约束有关的研究。其中最具代表性的有 2 个: 一是UIUC大学计算机系的Jin Jingwen等提出的应用层组播服务路径发现协议<sup>[5]</sup>; 另一个是

WUSTL大学计算机系的Sumi Choi等提出的所谓可编程网络中会话配置问题<sup>[6]</sup>。在这些研究工作中, 也涉及到了服务次序约束, 但没有对其进一步的描述。

本文对服务组合问题进行了严格的形式化描述。再结合路由, 为确保端系统之间的通信按一定的服务因子次序组合服务, 研究了覆盖服务网络上进行服务组合的路由。

### 2 2 个假设

对于给定的网络和组播实例, 为了建立更加具体的数学优化模型, 还需要对服务次序约束问题进行进一步的抽象和形式化。首先, 需要 2 个假定:

(1) 每一个代理节点仅能够承担一种服务。当某个代理节点 $v$ 能够承担 $m$ ( $\geq 2$ )种服务时, 可将此节点看作是由 $m$ 个代理节点 $v_1, v_2, \dots, v_m$ 组成的完全图, 每个代理节点 $v_i$ 能够且仅能够承担这 $m$ 种服务中的 1 种, 并规定此完全图中的每条边的权值为 0(从而, 如果在原图中有节点 $u$ 与 $v$ 相邻, 则 $u$ 到此完全图的每个节点的路径长度都为 $u$ 与 $v$ 之间边的权值。可见, 这种替换跟原拓扑结构是等价的);

(2) 在一次具体的组播中, 每个代理节点可以有选择地承担或不承担它能够承担的服务, 即在模型中, 需要明确区分 2 个概念, “能够承担服务”和“承担服务”。前者指某个代理节点具有承担某种服务的能力, 但它不一定在特定的组播路由中承担此服务(或者是因为此组播不需要这种服务,

**基金项目:** 高等学校博士学科专项科研基金资助项目(20030290003)

**作者简介:** 潘 耘(1974 -), 女, 博士、副教授, 主研方向: 网络性能评价, Overlay 网络; 余镇危, 教授、博导; 王银燕, 博士生; 贾民政, 硕士

**收稿日期:** 2006-11-15      **E-mail:** panyun13@126.com

或者是这种服务已经由此组播路由上的其它节点承担);而后者是指在特定的组播路由中,某个代理节点承担了它所能提供的服务。例如图 1 所示组播路由中,假定节点 2、节点 3 都具有解压缩的功能,但是在此次组播路由中解压缩的服务仅由节点 3 来承担,节点 2 只是作为普通节点出现。

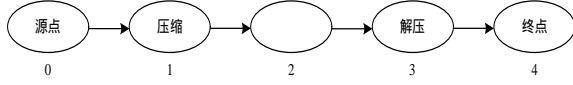


图 1 一个具体的组播路由

### 3 服务组合问题的形式化描述

首先,需要引入以下概念。

**定义 1** 服务代价函数

$$c': V \rightarrow [0, \infty), c'(v) = \begin{cases} \text{节点 } v \text{ 承担服务时的服务代价, 节点 } v \text{ 是代理节点} \\ 0, \text{节点 } v \text{ 不是代理节点} \end{cases} \quad (1)$$

和服务因子类别函数

$$s: V \rightarrow N, s(v) = \begin{cases} j, \text{节点 } v \text{ 能够承担第 } j \text{ 类服务因子} \\ 0, \text{其它} \end{cases} \quad (2)$$

显然,在第 1 个假定之下,服务因子类别函数  $s$  是定义良好的单值函数。

**定义 2** 服务因子类别集合

$$C \stackrel{\text{def}}{=} V/R = \{[v]: v \in V\} \quad (3)$$

其中  $R$  是定义在  $V$  上的二元关系

$$R \stackrel{\text{def}}{=} \{<u, v>: s(u) = s(v)\} \quad (4)$$

易证,  $R$  是等价关系,所以定义 2 有意义。

为便于描述,当  $s(v)=j$  时,可不加区分地使用“第  $[v]$  类服务因子”和“第  $j$  类服务因子”这样 2 种等价的说法;在不致引起歧义时,亦可使用“第  $v$  类服务因子”这样不太严格的说法。

**定义 3** 设  $D \subseteq V$ , 定义  $D$  上的服务次序约束  $L_D \subseteq C \times C$  集合为

$$L_D \stackrel{\text{def}}{=} \{<[u], [v]>: u, v \in D, \text{第 } [u] \text{ 类服务因子 “不滞后于” 第 } [v] \text{ 类服务因子}\} \quad (5)$$

显然,关系  $L_D$  具有自反性、反对称性和传递性,即  $L_D$  是偏序集。由于传递关系的传递闭包关系是其本身,因此,在实际应用中,一般只需给出  $L_D$  的部分序偶,然后对  $L_D$  求自反传递闭包可得到其全部偏序关系。例如,在某组播约束中,用户给出的  $L_D$  为  $\{<[1], [2]>, <[2], [3]>\}$ , 则求自反传递闭包之后  $L_D$  为  $\{<[1], [1]>, <[1], [2]>, <[1], [3]>, <[2], [2]>, <[2], [3]>, <[3], [3]>\}$  (假定该组播仅涉及第 [1]、第 [2] 和第 [3] 类服务)。另外,如果第  $[j]$  类服务是该组播所必须确保的,但是在路由中的次序无关紧要,则  $L_D$  中必须包含序偶  $<[j], [j]>$ ; 否则,在计算服务次序约束的投影时就会有遗漏(见式(11)~式(13))。最后需要说明的是,如果将此定义中的“不滞后于”改为“领先于”,则关系  $L_D$  就不具有自反性。

服务次序约束偏序关系  $L_D$  也可以用  $L_D$  的关系矩阵表示如下:

$$A_{L_D} = (l_{ij})_{n \times n}, n = |V|, l_{ij} = \begin{cases} 1, <[v_i], [v_j]> \in L_D \\ 0, \text{others} \end{cases} \quad (6)$$

另外,对于一棵具体的组播树  $T$ , 其中的代理节点既可以选择承担相应的代理服务,也可以选择不承担服务,故需要一个服务标志向量来表明这一点。所以,引入定义 4。

**定义 4** 组播树  $T$  所对应的服务标志向量定义为

$$\overline{a}_T = (a_1, a_2, \dots, a_{|V|}), a_j = \begin{cases} 1, v_j \in V_T \text{ 且承担服务} \\ 0, \text{others} \end{cases} \quad (7)$$

并定义矩阵

$$A_T = (t_{ij})_{n \times n}, n = |V|, t_{ij} = \begin{cases} a_i \cdot a_j, \text{在 } T \text{ 中 } v_i \text{ 是 } v_j \text{ 的祖先} \\ 0, \text{others} \end{cases} \quad (8)$$

**定义 5** 为了刻画组播树  $T$  对服务次序约束  $L_D$  的违背(violation)程度,还需要定义如下矩阵:

$$A_v = A_{L_D}^T \cdot * A_T = (w_{ij})_{n \times n} = \begin{cases} 1, <v_i, v_j> \in L_D, \text{在 } T \text{ 中 } v_i \text{ 是 } v_j \text{ 的祖先} \\ 0, \text{others} \end{cases} \quad (9)$$

其中:  $A_{L_D}^T$  为  $A_{L_D}$  的转置矩阵的闭包; 2 个同型矩阵的  $*$  运算定义为对应元素分别相乘构成的新矩阵。定理 1 保证了该定义的正确性。

**定理** 在组播树  $T$  中, 节点对  $<u, v>$  违背服务次序约束, 当且仅当在  $A_v$  中第  $u$  行第  $v$  列元素  $w_{uv} = 1$ 。

**证明** 在组播树  $T$  中, 节点对  $<u, v>$  违背服务次序约束, 意味着: 节点  $u$  和  $v$  均在  $T$  中承担服务, 并且在  $T$  中  $u$  是  $v$  的祖先节点(即  $t_{uv} = 1$ ), 但是节点  $u$  所承担的服务却“滞后于”节点  $v$  所承担的服务(即  $<[u], [v]> \in L_D^C$ )。对于任意的二元关系  $R$ , 其逆关系  $R^C$  定义为:  $xR^C y$ , 当且仅当  $yRx$ 。显然, 关系  $R$  的关系矩阵的转置就是关系  $R^C$  的关系矩阵。所以,  $<[v], [u]> \in L_D$ , 即  $l_{vu} = 1$ 。故  $w_{uv} = l_{vu} * t_{uv} = 1$ 。

反之, 如果在组播树  $T$  中, 节点对  $<u, v>$  不违背服务次序约束, 则有: (1) 节点  $u$  或者节点  $v$  不在  $T$  中, 此时  $t_{uv} = 0$ ; (2) 节点  $u$  和  $v$  均在  $T$  中, 但是节点  $u$  不是  $v$  的祖先节点, 此时  $t_{uv} = 0$ ; (3) 节点  $u$  和  $v$  均在  $T$  中, 并且  $u$  是  $v$  的祖先节点, 但是  $u$  或者  $v$  不承担代理服务, 此时  $a_u = 0$  或者  $a_v = 0$ , 则  $t_{uv} = a_u * a_v = 0$ ; (4) 节点  $u$  和  $v$  均在  $T$  中,  $u$  是  $v$  的祖先节点, 并且  $u$  和  $v$  均承担代理服务(此时  $t_{uv} = 1$ ), 但是它们所承担的服务之间没有违背服务次序约束, 即: 要么它们所承担的服务之间无次序约束(此时  $l_{uv} = l_{vu} = 0$ ); 要么  $u$  所承担的服务“不滞后于”  $v$  所承担的服务(此时  $l_{uv} = 1$ , 则由偏序关系的反对称性知  $l_{vu} = 0$ ), 从而,  $w_{uv} = l_{vu} * t_{uv} = 0$ 。

证毕。

于是, 设  $\Lambda = \{T: T \text{ 是网络 } G \text{ 的组播树}\}$ , 对  $\forall T \in \Lambda$ , 可定义  $T$  对服务次序约束  $L_D$  的违背函数如下。

**定义 6** 组播树  $T$  对服务次序约束  $L_D$  的违背函数定义为

$$v: \Lambda \rightarrow [0, 1], v(T) = \frac{1}{|AR|^2} \sum_{i=1}^n \sum_{j=1}^n w_{ij} \quad (10)$$

即  $A_v$  中非 0 元素之比例。

另外, 为了在组播树中确保某类服务因子, 还需要引入以下概念。

**定义 7** 称向量

$$\overline{p}_{L_D} = (x_1, x_2, \dots, x_k), k = |C|, x_j = \begin{cases} 1, [v_j] \in C_{L_D} \\ 0, \text{otherwise} \end{cases} \quad (11)$$

和

$$\overline{p}_T = (y_1, y_2, \dots, y_k), k = |C|, y_j = \begin{cases} 1, [v_j] \in C_T \\ 0, \text{otherwise} \end{cases} \quad (12)$$

分别为服务次序约束  $L_D$  和组播树  $T$  到服务因子类别集合  $C$  上的投影向量, 其中

$$C_{L_D} \stackrel{\text{def}}{=} \{[u]: [u] \in C, \exists [v] \in C (<[u], [v]> \in L_D \vee <[v], [u]> \in L_D)\} \quad (13)$$

和

$$C_T \stackrel{\text{def}}{=} \{[u]: [u] \in C, u \in V_T\} \quad (14)$$

分别为服务次序约束  $L_D$  和组播树  $T$  到服务类别因子集合  $C$  上的投影。

显然, 当  $\overline{p}_{L_D} = \overline{p}_T$  时, 组播树  $T$  确保了服务约束集合  $L_D$  中

出现的每类代理服务。

综上所述,带服务次序约束的覆盖服务网络模型可表示为带权图  $G=(V,E,c;s,c')$ , 其上的一个单源组播路由实例可以描述为三元组  $M=(source,D,L_D)$ , 其中  $source \in V$  为组播源点,  $D \subseteq V$  为组播终点集,  $L_D$  为服务次序约束集合, 需要求解的问题是: 寻找一棵以  $source$  为根、包含  $D$  中所有节点、满足服务次序约束  $L_D$  的 Steiner 树, 并使其组播代价(节点间的通信代价和承担代理服务的节点的服务代价之和)达到最小化, 即

$$\begin{cases} \min_{T \subseteq G} f = \sum_{e \in E_T} c(e) + \sum_{j=1}^n c'(v_j) * a_j \\ s.t. v(T) = 0, \\ \overline{P_{L_D}} = \overline{P_T}, \\ T \text{ 是以 } source \text{ 为根的 } G \text{ 的子树, 且 } D \subseteq V_T \end{cases} \quad (15)$$

其中,  $a_j, v(T), \overline{P_{L_D}}, \overline{P_T}$  的定义见式(8)~式(14)。

采用罚函数方法, 可以将此模型中的 2 个等式约束化去, 得

$$\begin{cases} \min_{T \subseteq G} f = \sum_{e \in E_T} c(e) + \sum_{j=1}^n c'(v_j) * a_j + P_1(v(T)) + P_2(\|\overline{P_{L_D}} - \overline{P_T}\|) \\ s.t. T \text{ 是以 } source \text{ 为根的 } G \text{ 的子树, 且 } D \subseteq V_T \end{cases} \quad (16)$$

其中,  $\|\cdot\|$  可以根据具体应用取所需的范数, 在这里, 取 2-范数; 罚函数  $P_1$  和  $P_2$  可根据求解(式(3)~式(16))方法的不同而选取, 原则是  $P_1(0) = P_2(0) = 0$ , 并且  $\exists M > 0$  使得当  $x > M$  时,  $P_1(x) \rightarrow \infty, P_2(x) \rightarrow \infty$ 。

(上接第 24 页)

滑模型进行预测。表 1 是各种预测方法误差评价指标的比较。

表 1 不同方法选取  $\alpha$  的 SSE 对比

	$\alpha = 0.6$	$\alpha_s =  C_s $	optimize $\alpha$
SSE(v)	597.219 5	665.406 6	596.885 2
SSE(q)	651.072 7	669.745 2	643.544 0
SSE(o)	222.047 1	217.591 6	221.279 3

从表 1 可以看出, 当  $\alpha$  取值 0.6 时的 MARE 和 MAXARE 都较小。当  $\alpha$  自适应时的平滑预测误差指标较之  $\alpha$  不变时还要更好。

图 1 是流量实测数据与采用固定权重和自适应权重进行平滑预测的结果比较。

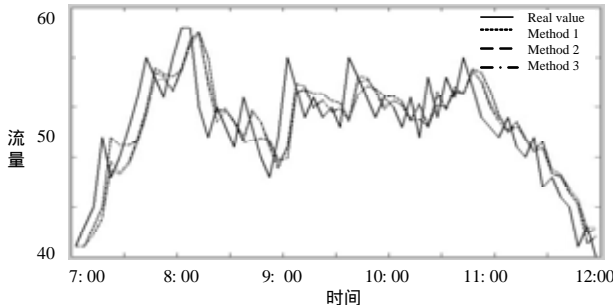


图 1 实测流量与 3 种不同预测算法流量预测比较

#### 4 结语

本文的结果使困扰指数平滑应用的初值选取、参数确定、预测偏差等问题获得了较完整的解决。所建立的动态指数平滑优化模型, 由于最终的  $\min SSE$  是借助原始的时间序列对  $\alpha$  进行自适应优选后得到的, 加之平滑中每一项的权重随着  $\alpha$

#### 4 结束语

本文研究了在覆盖服务网络上进行带 OoS 约束的路由问题, 建立了 OoS 问题的框架和模型, 还显得有些复杂, 探讨对 OoS 问题的更加简洁的描述模型始终是有意义的。另一方面, 由于 OoS 问题本身的复杂性, 提出的算法还远不是最优的, 因此设计新的更加高效的算法是需要进一步研究的内容。

#### 参考文献

- Li Zhi, Mohapatra P. Impact of Topology on Overlay Routing Service[Z]. [http://www.ieee-infocom.org/2004/Papers/09\\_1.PDF](http://www.ieee-infocom.org/2004/Papers/09_1.PDF).
- Andersen D H, Balakrishnan H, Kaashoek M F, et al. Resilient Overlay Networks[C]//Proceedings of the 18<sup>th</sup> ACM Symposium on Operating Systems Principles, Banff, Canada. 2001-10: 131-145.
- Li Zhi, Mohapatra P. QRON: QoS-aware Routing in Overlay Networks[J]. IEEE Journal on Selected Areas in Communications, 2004, 22(1): 29-40.
- 潘耘. 主动 Overlay 网络及其关键技术的研究[D]. 北京: 中国矿业大学(北京)研究生院, 2003-12.
- Xu D, Nahrstedt K. Finding Service Paths in a Media Service Proxy Network[C]//Proceedings of the ACM/SPIE Conference on Multimedia Computing and Networking. 2002-01: 171-185.
- Choi S, Turner J, Wolf T. Configuring Sessions in Programmable Networks[J]. Computer Networks, 2003, 41(2): 269-284.

的变化都在不停地自动调整, 因此新模型及其算法不仅没有改变指数平滑法的基本特征, 而且对时间序列的适应能力更强。当需要预测成百上千个项目时, 这一特性具有特别重要的意义。此外, 当数据结构发生变化, 初始规定的平滑系数不再适应时, 可以自动改正。本文所提供的方法无疑可以推广至高次及差分-指数平滑模型中。

#### 参考文献

- ITS American. National Intelligent Transportation Systems Program Plan: A Ten-year Vision[Z]. 2002. <http://www.itsa.org/>.
- Sun Hongyu, Liu Henry X. Short-term Traffic Forecasting Using the Local Linear Regression Model[Z]. 2003. [www.itsa.uci.edu/its/publications/jour.html](http://www.itsa.uci.edu/its/publications/jour.html).
- 刘静, 关伟. 交通流预测方法综述[J]. 公路交通科技, 2004, 21(3): 82-85.
- 王正武, 黄中祥. 短时交通流预测模型的分析与评价[J]. 系统工程, 2003, 21(6): 97-100.
- 杨兆升. 城市交通流诱导系统理论与模型[M]. 北京: 人民交通出版社, 2000.
- 张忠平. 指数平滑法[M]. 北京: 中国统计出版社, 1996.
- “预测方法和技术的应用研究”课题组. 中国预测技术发展研究[J]. 预测, 1991, (5): 10-17.
- 徐大江. 预测模型参数的指数平滑估计法及其应用的进一步研究[J]. 系统工程理论与实践, 1992, 12(2): 25-37.
- Berger J O. Statistical Decision Theory and Bayesian Analysis[M]. 2<sup>nd</sup> ed. New York: Springer-Verlag, 1996.

