

# 支持 Internet 实时多媒体应用的 Overlay 技术研究<sup>\*</sup>

吴云波<sup>1,2</sup> 李志蜀<sup>1</sup> 吴云海<sup>3</sup> 陈志华<sup>4</sup> 殷锋<sup>1</sup> 卢瞰<sup>1</sup> 元昌安<sup>1</sup>

(四川大学计算机学院 成都 610065)<sup>1</sup> (浙江纺织服装职业技术学院 宁波 315211)<sup>2</sup>

(云南省水电勘测设计院 昆明 650028)<sup>3</sup> (云南省昆明市气象局 昆明 650028)<sup>4</sup>

**摘要** 目前出现的大规模(large-scale)的多媒体应用对 Internet 提出了新的挑战。传统的 IP 组播(IP multicast, IPM)由于实现技术的复杂性,很难在 Internet 上推广。相反,基于叠加网络(overlay network, ON)的解决方案则受到各方面的关注。本文针对大规模实时多媒体应用需求,通过建立评价参数模型对 ON 协议进行研究。尤其对 overlay 组播(overlay multicast, OM)的 QoS 控制、可靠性以及可伸缩性机制进行了分析。与 IPM 相比,OM 虽然有性能代价(performance penalty)问题,但是 OM 能方便、灵活地利用多种高层的功能实现,基于 hop-by-hop 方式有效地提供 E2E(end-to-end)的可靠性和 QoS 保障,并且较好地支持 TCP-friendliness,具有良好的可伸缩性。此外,OM 的基础网络(infrastructure/underlay network)无关性,也使其容易在 Internet 中进行配置。以上特性使得 OM 成为支持 Internet 上大规模实时多媒体应用的最具前景的技术。

**关键词** Overlay, Internet, 实时多媒体应用, QoS, 可靠性控制, 可伸缩性

## Study on Overlay Mechanisms of Supporting Real-Time Multimedia Applications in Internet

WU Yun-Bo<sup>1,2</sup> LI Zhi-Shu<sup>1</sup> WU Yun-hai<sup>3</sup> CHEN Zhi-Hua<sup>4</sup> YIN Feng<sup>1</sup> LU Tun<sup>1</sup> YUAN Chang-An<sup>1</sup>

(School of Computer Science, Sichuan University, 24 Southern Section 1, 1st Ringroad, Chengdu 610065)<sup>1</sup>

(Department of Information Engineering, Zhejiang Textile and Fashion Institute, Ningbo 315211)<sup>2</sup>

(Yunnan Design Institute of Water Conservancy and Hydroelectric Power, Kunming 650021)<sup>3</sup>

(Kunming Meteorological Administration, West of Xihua Park, Kunming 650228)<sup>4</sup>

**Abstract** Recently, large-scale multimedia applications put a challenge on traditional Internet. Due to many drawbacks in implementing reliability and QoS control, as well as requirements in upgrading infrastructure components, IP multicast has more difficulties to be widely deployed in Internet. However, circumventing these problems via overlay networking increasingly attracts many interesting of researchers. This paper models some measures of reliability, QoS and scalability relating to large-scale real time multimedia applications. Furthermore, overlay networking protocols especially overlay multicast protocols have been analyzed referring to these issues. Comparing to IP multicast, overlay multicast may induce some delay penalty, but this shortcoming is insignificant by virtues of proper protocol design. Overlay multicast can employ upper-layer functionalities in where, thus make it feasibly and efficiently to achieve E2E reliability and QoS based on hop-by-hop. Besides, scalability as well as TCP-friendliness is guaranteed efficiently in overlay multicast. Moreover, no infrastructure support is needed, thus paving the way for widely deployment in Internet. Advantages above have made overlay multicast to be a promising techniques for large-scale real time multimedia applications in Internet.

**Keywords** Overlay, Internet, Real-time multimedia, QoS, Reliability, Scalability

## 1 引言

随着 Internet 的迅速发展,大规模的多媒体应用如在线音/视频点播、分布式文件共享、在线游戏以及联网仿真等日益膨胀,基于 one-to-many 和 many-to-many 的组播通信引起了广泛的关注。传统的 IPM 由于对路由设备的依赖,难以在 Internet 上实现跨 ASs (autonomous system) 配置。加之 QoS 和可靠性保障在 IPM 中的有效实施太过复杂,控制开销比较大。这使得 Internet 难以利用 IPM 支持大规模的多媒体组播应用。相反,作为在 Internet 上提供高层服务的 ON 越来越受到网络界的关注,相关的应用如 OM、叠加服务网(overlay service network, OSN)以及 P2P 文件共享等正在成为研究焦点。其中,OM 因其对 Internet 组播应用的有效支

持而被认为是 IPM 的替代方案。近几年来, Napster、eDonkey 和 BitTorrent 在 Internet 上获得的极大成功,越发激起人们对 OM 的研究兴趣。目前已有许多研究者对 OM 进行了研究,提出了一些富有意义的建议<sup>[2~11]</sup>。本文将针对 Internet 上大规模实时多媒体应用的特性,通过建立评价参数模型,就 QoS 控制、可靠性以及可伸缩性问题对 OM 协议进行分析研究。

OM 又称为应用层组播(ALM, application layer multicast)或端系统组播(End-System Multicast)。与网络层的 IPM 不一样,OM 中参与组播会话的所有主机形成一个逻辑上的 ON。所有组播控制功能如组管理、路由以及组播树构建都在端系统中完成,并不需要基础网络的显式支持。OM 根据拓扑结构的实现方式分为:网-树结构(mesh-tree struc-

<sup>\*</sup> 863 课题资助项目课题号 2002AA144020。吴云波 博士研究生,副教授,主要方向:IPQoS, overlay networking;李志蜀 博士生导师,教授,研究方向:计算机网络与信息系统;吴云海 工程师;陈志华 工程师;殷锋 博士研究生;卢瞰 博士研究生;元昌安 博士研究生。

ture)和隐含结构(embedded/implicit structure)。网-树结构中按照 tree、mesh 建立的先后次序又可分为树优先(tree-first)和网优先(mesh-first)两种类型。Yoid<sup>[3]</sup>、Overcast<sup>[8]</sup>、ALMI<sup>[9]</sup>等属于 tree-first。这种结构的组成员先以自组织方式构建一棵树(shared tree 或者 source-specific tree),然后成员之间再建立树外的控制链路。由数据树连同其上的控制链路共同形成控制拓扑。Narada<sup>[2]</sup>、Scattercast<sup>[11]</sup>等则属于 mesh-first。这种结构的组成员首先在基础网络上形成一个网状控制拓扑 mesh。然后再按照一定的路由协议(如 DVMRP)在 mesh 中构建组播树。隐含结构的控制拓扑和数据拓扑则是同时建立的,如 CAN<sup>[4]</sup>、NICE<sup>[5]</sup>、SCRIBE<sup>[7]</sup>、Bayeux<sup>[10]</sup>等。其中,CAN、SCRIBE、Bayeux 是利用抽象的坐标空间来映像拓扑结构的,每个组成员都与该坐标空间的一个逻辑区域相对应,该区域包含有邻近成员地址的映像信息(mappings)。因此,其下一跳路由不需要网络层路由协议的支持。NICE 则是利用多层结构构建控制拓扑,而数据拓扑由前向(forwarding)规则隐含定义于其中。

## 2 OM 协议分析

Internet 上的大规模多媒体组播应用对系统的可伸缩性(scalability)、可靠性(reliability)以及 QoS 都有较高要求。以下将从这几方面对 OM 协议作进一步的分析。

### 2.1 可伸缩性(scalability)

可伸缩性反映出 OM 对组播规模的支持能力,它与协议的控制开销、需维护的状态量和可利用资源有关。为此,先建立相关的参数模型如下:

**定义 1** 若用图  $G=(N,E)$  表示基础网络, $N$  是节点集,一个节点  $\eta \in N$  代表一个路由器。 $E$  是边的集合,一条边  $(\eta_i, \eta_j) \in E$  代表基础网络的一条双向物理链路。则位于  $G$  上的叠加网络可以表示为一棵树  $o=(s,D,N_0,E_0)$ ,其中  $s$  表示源主机, $D$  是组播中的所有接收主机, $N_0 \subseteq N$  且由叠加网中的叠加链路(overlay links)相连, $E_0$  是所有叠加链路的集合, $E_0$  定义如下:

**定义 2** 一条叠加链路  $e_0=(d_s, \eta_1, \dots, \eta_n, d_r) \in E_0$ ,它由一个发送主机  $d_s$ 、若干路由器  $\eta_i$  和一个接受主机  $d_r$  顺序组成。其中, $d_s \in H_0, H_0=\{s\} \cup D, d_r \in D, \eta_i \in N_0$ 。 $e_0$  对应一个由 OM 建立的 UDP 或 TCP 连接。 $l_i(e_0)$  表示叠加链路  $e_0$  的跳数,对应路由器系列  $\eta_1, \dots, \eta_n$ 。

**性能参数<sup>[2]</sup> link stress:** 表示在一条物理链路上相同报文的重复传递次数,它反映了 ON 的负载。对于 IP M, stress 为 1。而 unicast,在最坏情形下 stress 为接收终端数。

**link stretch:** 也称为 RDP(Relative Delay Penalty),是 OM 协议的单播(unicast)路由距离与相应的 IP 单播路由(IP unicast)距离之比。在对称情形下,IPM 和 unicast 的 stretch 均为 1。link stretch 反映了 OM 引起的时延代价(delay penalty)。

**path length:** 表示 ON 中的跳数,间接描述了 link stretch。

**outdegree:** 描述数据树上节点的最大前向分组复制次数,直接反映出 stress。

**control overhead:** 表示 ON 中的非数据通信量(non-data traffic)。

**overlay cost:** 定义为 ON 中所有 overlay link 的跳数总和,即  $\forall e_0 \in E_0, \sum l_i(e_0)$ 。

其中,link stress 和 link stretch 因为依赖于底层网络拓扑很难直接计算,常常通过 path length 和 outdegree 进行分析,而 control overhead 则用进入节点的状态数计算。应用以上参数对几种典型的 OM 协议分析如下:

Narada 和 Scattercast;属于 mesh-first 拓扑。Narada 是最早的 OM 协议之一,拓扑建立时,协议首先以分布式的方式在组成员之间建立一个网状的 control 拓扑 mesh。然后,利用运行于 mesh 之上的距离向量路由协议(如 DVMRP),为每个成员建立一棵源树(source-specific tree)。这是 mesh 的生成树(spanning tree),对叶节点的扇出没有限制,最大路径长度(path length)也没有限制。mesh 成员之间需要周期性地相互交换状态信息,因此控制开销(control overheads)最大值为  $O(N^2)$ ,平均为  $O(N)$ ( $N$  是组的规模)。Narada 是为小规模组播(几十个成员)设计的。Scattercast 是对 Narada 的扩展,它引入代理机制以充分利用基础网络提供的服务。Scattercast 用类似 Narada 的方式将布置在基础网络中的代理 ScatterCastproXies(SCXs)以自组织的方式形成 mesh。然后,在 mesh 上利用传统的路由协议(以链路时延为路由参数)构建基于组播源的数据树。参加组播的成员通过邻近的代理加入组播 session。

Yoid 和 ALMI;属于 tree-first 拓扑。Yoid 也是最早的 OM 协议之一,它首先利用启发式算法为组成员构建一棵共享数据树(shared tree)。然后,树上每个成员都随机选择几个不相邻的树成员建立 link,以此形成 mesh 拓扑。mesh 主要用于共享树的恢复,其成员需要存储和维护的信息较少。Yoid 的树成员有最大出度限制(max outdegree bound),故其平均控制开销为  $O(\max \text{ deg})$ 。Yoid 没有路径长度限制。ALMI 是为对时延敏感的小规模组播应用设计的,采用集中控制方式。它首先以链路时延为参数,利用中央控制器(central controller)构成一棵最小生成树(MST, minimum spanning tree)。这是一棵源树,对 path length 没有限制。为实现 MST,所有成员相互连接形成控制 mesh,并且定期相互交换状态信息,这导致控制开销为  $O(N^2)$ 。为支持时延敏感应用,ALMI 通过设定 out degree 为常数来降低 control overhead 至  $O(N)$ , $N$  是组成员数。

CAN 和 NICE;这是两种典型的隐含结构 OM 协议。CAN 采用分布式哈希表 DHT(distributed hash table)将每个成员与一个虚拟的  $d$ -维笛卡尔空间中的点  $(x,y)$  对应,这个点拥有独立的一个小矩形或正方形相空间。这样,叠加网络的路由就转化为在笛卡尔空间各点之间的寻路问题。控制拓扑仅仅需要维持与成员相邻的  $2d$  个节点信息之间的交换,故平均控制开销为常数  $2d$ 。数据拓扑则在控制拓扑上采用洪泛(flooding)方式进行,路径长度最大值为  $O(dN^{1/d})$ 。NICE 的组成员首先构成一个分层的控制拓扑结构,所有成员都是最底层( $L_0$  层)的一部分,在每一层中成员分成一个个大小受限的簇(cluster),其范围在  $k \sim 3k$  之间, $k$  是常数。每个簇有一个中心(该中心到簇中成员的最大距离是最小的), $L_i$  层的所有中心成员又构成  $L_{i+1}$  层。若成员数  $N$ ,则最多可形成  $O(\log N)$  层,故最大路径长度为  $O(\log N)$ 。另外,按照控制拓扑协议,每个簇中的成员会定期相互交换状态更新信息,故控制开销平均为  $O(k)$ ,最大为  $O(k \log N)$ 。

以上几种协议具有一定的代表性,充分体现了 OM 三种拓扑结构的特点。类似地,利用以上参数对其他协议进行分析,可得到如下结果(表 1):

表 1 OM 协议性能

Protocol	Max. Path length	Max. out-degree	Avg. control overheads	Group size
Narada	Unbounded	User defined	$O(N)$	Small
Scattercast	Unbounded	User defined	$O(\max. \text{deg})$	Large
Yoid/HMTP	Unbounded	User defined	$O(\max. \text{deg})$	Small
CAN	$O(dN^{1/d})$	$2d$	$2d$	Large
NICE	$O(\log N)$	$k \sim 3k$	$O(k)$	Very large
Bayeux/Scribe	$O(\log N)$	$O(\log N)$	$O(\log N)$	Large
Overcast	Unbounded	Unbounded	$O(N)$	Large
ALMA	Unbounded	User defined	$O(\max. \text{deg})$	Small

由此可见, tree-mesh 协议对路径 hops 和最大出度无限制, 具有较大 stretch 和 stress。因而, 这些协议中需要增加对出度的约束功能, 以优化组播带宽和时延。embedded 协议的 average control overhead 要么趋于常数要么仅与组规模成对数关系, 表现出较好的可伸缩性 (scalability)。此外, embedded 协议的 stretch 和 stress 较小。相比之下, embedded 更适合对时延和带宽有较高要求的应用。

对已有的 OM 协议, 还可以从 overlay 体系上划分为 peer-to-peer (p2p) 和基于代理 (proxy based) 两种。基于代理的可视为是一种混合结构, 它利用了网络层提供的服务, 如 Bayeux, OMNI, Overcast 和 Scattercast。p2p 如 Narada, ALMA, CAN, NICE, Yoid 等。从这个角度对比, p2p 体系具有较好的故障冗余性 (failure redundancy)、动态性 (dynamicity) 和可伸缩性 (scalability), 配置较灵活、简单。proxy based 机制则因为利用了基础网络提供的服务而具有更高的效率。近来, 能够提供多种增值服务 (value-added) 的覆盖服务网 (overlay service network, OSN) 受到了人们的关注, 如 SON<sup>[12]</sup>、QUEST<sup>[13]</sup>、OverQoS<sup>[14]</sup> 等。OSN 采用基于代理的机制, 在 Internet 的多个 ASs 之间建立有服务质量保证的 overlay links, 提供 many-to-many 或者 one-to-many 的增值服务 (reliability)。

在此将 OM 的可靠性机制分为拥塞控制 (congestion control)、差错控制 (error-correction) 和故障恢复 (loss recovery) 三部分。与 IPM 中的可靠性控制相比, OM 有其自身特点。除了 IPM 面临的反馈风暴 (feedback implosion)、丢失路径多样性 LPM (loss path multiplicity) 以及 TCP-friendly 等问题外, 成员的对等性和自组织性使得拓扑结构变化的随机性和动态性增大, 这要求 OM 系统具有较高的鲁棒性 (robust) 和容错性 (fault tolerance)。此外, 由于存在时延代价 (delay penalty), 差错检测和错误恢复变得更为困难, 传统的基于 E2E 的 TCP 控制不能直接用于 OM 中的可靠性控制。然而, OM 可以利用端系统的智能性灵活地在高层实施多种控制机制, 提供较 IPM 更为有效、灵活的可靠性保障。

如果利用 TCP 连接建立 OM 的 overlay links, 那就很容易实现基于 hop-by-hop 的可靠性机制 (参见图 2), 其中  $w_c^o$  是窗口大小。目前具有可靠性机制的 OM 协议大都采用了这种方式, 如 Overcast、RMX<sup>[15]</sup>、ALMI、ROMA<sup>[16]</sup> 以及文 [13, 17] 的 overlay MCC 等。这是一种分而治之 (divide-and-conquer) 的思想。这种将拥塞控制和差错控制局部化的方式, 极大地提高了差错检测和恢复的速度, 降低了 E2E 时延、时延抖动以及 overlay link stress, 具有良好的 TCP-friendliness。此外, 利用基于 hop 的反馈聚集 (feedback-aggregating) 还可以有效地消除反馈风暴<sup>[9]</sup>。但是, 纯粹的 hop-by-

hop 控制并不能保障 E2E 的可靠性, 对于因 OMT 中间节点失效所引起的链路故障则仍然需要 E2E 的 ACK/NACK 机制予以保障。因此 OM 还需要一个位于 overlay 之上的可靠性机制提供 E2E 的可靠性保障。

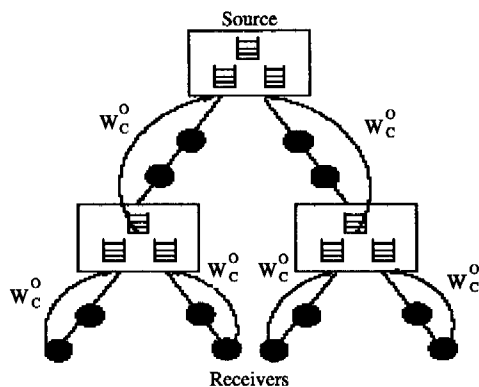


图 2 hop-by-hop OM 拥塞控制模型

已有的带可靠性控制的协议中, Overcast 中的 TCP 仅仅提供了基于 hop 的单播可靠性, 缺乏 E2E 的流控机制。ALMI 和文 [17] 应用反压机制 (backpressure) 控制发送速率, 利用带外连接 (out-of-band connection) 进行数据重传, 具有 E2E 的可靠性控制, 但是容易产生单速率控制的问题。RMX<sup>[6]</sup> 利用代理 asmod 提供针对特定应用的 E2E 可靠性机制。ROMA 将节点的上下、行端口进行部分解耦 (loosely coupled), 结合快速前向纠错反压控制, 提供多速率的 E2E 可靠性。

另外, OM 可以灵活地利用高层的控制功能来实现多种可靠性控制。RON<sup>[18]</sup> 是一种基于测量 (infrastructure/ measurement-based) 的方案, 它采用绕道机制 (detouring path) 选取性能相似的替代故障路径, 以此提供传输可靠性。文 [19] 的 MPOT 模型 (multi-path overlay transmission) 是一种混合 IPM/OM 方式。它利用 overlay 旁路机制改善 IPM 的拥塞控制性能。当 IPM 树发生拥塞时, MPOT 立即在收、发端之间建立多条带数据缓存功能的 overlay 链路分流数据。而当拥塞消除时, MPOT 会选用有效的 IP 链路进行传输。这样可以提高组播拥塞时的吞吐量。文 [20] 建议的 AOM (adaptive overlay multicast) 则是利用 E2E 时延与 RTT (round trip time) 作为参数构建 OMT, 引入备份路径 (backup path), 以提高 OMT 的容错性, 适应大数据文件和收端数量较多的应用, 如 Internet audio/video 广播。文 [21] 应用启发式算法构造 max-min 速率优化的 OM 树, 在细粒度上支持公平的多速率 OM。同样, ROMA 通过对节点上下行端口的部分解耦提供多速率支持。文 [17] 则通过对 OMT 分支节点的下行链路进行解耦, 以此提高 OM 的可伸缩性。可见, 以灵活多样的高层控制机制来实现 OM 的可靠性, 这正是 OM 的优势所在。

为支持 E2E 的 QoS 应用, 网络必须采取综合的控制策略, 包括: 接纳控制 (admission control)、QoS 约束路由、链路共享的分组调度技术以及资源预留和管理等。支持 QoS 的组播在 Internet 中具有广泛的应用前景, 越来越受到网络界的关注。QoS 组播主要是路由问题, 即在 QoS 约束下的组播寻路问题, 分为最优化问题和性能界约束问题。最优化问题就是寻找对应 QoS 度量的最优路径; 而性能界约束问题就是寻找大于对应 QoS 度量 (如带宽) 或小于对应 QoS 度量 (如时延) 的一条路径, 即在满足性能界要求的集合中选择一个解。QoS 组播问题的形式化描述如下:

**定义 3** 用带权重的无向图  $G=(N,E,w)$  表示网络拓扑,  $N$  是节点集,  $E$  是边集, 一条边  $(\eta_i, \eta_j) \in E$  连接节点  $\eta_i$  和  $\eta_j$  的一条双向物理链路, 简记为  $(i, j)$ 。从节点  $\eta_i$  到  $\eta_j$  的一条路径则记为  $[i, j]$ 。  $w(i, j) \geq 0$  是边  $(i, j)$  的权函数, 它与 QoS 度量参数相对应。权函数反映了网络的状态。对于加性参数 (additive, 如时延、时延抖动、分组丢失率等), 一条路径的某个 QoS 参数值就等于该路径上所有链路的对应权值之和。而对于凹性参数 (concave, 如带宽和策略标记 policy flags 等), QoS 参数值则等于路径上对应权值的最小 (最大) 值。

**定义 4** 路径时延表示路径上所有链路时延之和, 记为  $d[i, j] = \sum_{(i,p) \in [i,j]} d(i, j)$ ; 路径带宽表示路径中链路的最小带宽, 记为  $b[i, j] = \min_{(i,p) \in [i,j]} b(i, j)$ 。

设  $M \subseteq N$  是包含一个组播成员的节点集, 源节点是  $s$ , 接收节点是集合  $R = M - \{s\}$ 。组播树  $T$  是覆盖  $M$  中所有节点的  $G$  的子图  $T=(N_T, E_T)$ , 其中  $R \subseteq N_T \subseteq N, E_T \subseteq E$ 。  $T$  中还包括有非组播成员的中  $R \subseteq N_T \subseteq N, E_T \subseteq E$ 。  $T$  中还包括有非组播成员的中继节点。将接收节点  $r \in R$  的 E2E 时延要求记为  $D(r)$ , 带宽要求记为  $C(r)$ , 则组播约束路由问题即为: 对于每一个接收节点  $r \in R$ ,  $T$  上的路径  $[s, r]$  必须满足  $d[s, r] \leq D(r)$  以及  $b[s, r] \geq C(r)$ 。链路开销 (link cost) 表示该链路支持的所有接收节点中的最大带宽需求, 记为  $c(i, j) = \max\{C(r) | (i, j) \in [s, r]\}$ 。组播树的总开销是树上所有链路开销的总和, 记为  $C(T) = \sum_{l \in E_T} C(l)$ 。因此, Steiner (Steiner Minimum Tree, SMT) 问题就是寻找对确定的源的最小开销组播树, 即  $C(SMT) = \min_{T_l \in T} C(T_l)$ 。

考虑到剩余带宽和缓存空间, 相关的 QoS 路由问题是链路优化 (link-optimization) 和链路约束 (link-constrained) 路由。如寻找具有最大瓶颈带宽的路径 (优化路由) 或最大带宽超过给定值的路径 (约束路由)。考虑到时延、时延抖动以及网络开销 (cost) 的路由问题, 则是路径/树优化 (path/tree-optimization) 和路径/树约束 (path/tree-constrained) 路由, 前者寻找代价最小的路径 (指组成路径的所有链路的总代价), 后者如寻找时延代价满足要求值的路径。其他 QoS 路由问题不外是这四种基本问题的组合。其中, 不相关参数的多约束路径问题 (multi-constrained path, MCP) 和路径约束路径优化问题 (path-constrained path-optimization, PCPO), 这两者已证明是 NP-complete 的<sup>[22]</sup>。另外, 最小开销组播树 (the least-cost multicast tree) 问题即 Steiner tree problem 是寻找总开销最小的组播树, 这是 NP-complete 问题。还有, 约束 Steiner 树 (constrained Steiner tree problem) 问题是寻找时延受限的最小开销树, 也是 NP-complete 问题<sup>[23]</sup>。

与单播 (unicast) QoS 路由相比, 组播面临的困难要大得多。首先, 分布式多媒体应用的 QoS 需求往往涉及时延、时延抖动、带宽 QoS 以及分组丢失率等多个参数, 是多约束优化问题, 其中许多是 NP-complete 问题。其次, 组播路由由协议比较复杂, 既要收集、维护网络和组成员的状态信息, 又要维护树结构的动态拓扑 (如组播树重构 (tree rearrangement) 和树/核迁移 (tree/core migration)), 管理成员的变化。QoS 计算的引入, 进一步增加了协议设计的复杂性。由于 QoS 相关信息具有动态特性, 增加了控制开销。而且在不确定状态信息下, 如何实现满足 QoS 约束的组播路由和树无疑更为困难。有关 IPM 的 QoS 问题已有许多文献进行了研究<sup>[24~27]</sup>, 文[28, 29]对此作了很好的总结。

与 IPM 相比, OM 能够灵活地应用高层的功能来支持 Internet 上的 QoS 应用。文[30]利用 ON 节点中的 in situ 组

件, 以基于测量的方法 (measure/infrastructure-based) 提供广域网范围的低时延保证。in situ 由运行于用户层的几个后台程序 (user-layer daemons) 模块组成, 包括测量模块、路径模块和路由模块。路径计算采用了时延后退算法 (delay-regression), 可以得到与下层网络分布无关的近似最小时延解。文[31]通过设置服务综合代理 (service composition agents) 组建 ON。OMT 的构建采用启发式路由算法, 并用仿真退火算法对其进行优化。此外, 还引入了应用层接入控制 (application level admission control) 保证连接的 QoS 要求。文[32]利用基于探测 (probe-based) 的资源预留机制, 结合分组的 Differ-Serv 标记提供不同的 QoS 保障。文[33]则提出一种相关代价路由算法 (correlation cost routing) 在 ON 中进行多路径选择, 以支持多速率传输。

另一方面, 一种基于 ON 在 Internet 上提供多种增值服务的覆盖服务网 OSN 受到广泛关注。OSN 可以方便地在 Internet 上实现有偿的 QoS 服务。这正成为 ON 研究的一个热点。有关的协议如: OverQoS 采用 CLVL (controlled loss virtual link) 技术, 以 overlay 方式为 Internet 提供 QoS 服务, 如不同的速率分配、统计带宽和差错率 (loss rate) 要求。SON 用 overlay 方式提供能满足用户不同 QoS 需求的 Internet 增值 (value-added) 服务, 使用 pipe SLAs 连接两个服务网关 (service gateway) 提供带宽保证。QUEST 类似 SON, 利用 SON 节点 (service overlay network node) 和 SLAs (service level agreements) 提供 QoS 保障和负载均衡机制。QRON<sup>[34]</sup> 则是利用 OBs (overlay broker) 建立 overlay 服务网 (overlay service network, OSN), 提出了两个 QoS 路由算法进行 OBs 之间选路, 即类最短路径算法 (modified shortest-distance path, MSDP) 和比例带宽最短路径算法 (proportional bandwidth shortest path, PBSP)。

**总结** 底层网络无关性使得 ON 很容易在 Internet 中进行配置。而且, 借助高层的充分支持, ON 能够以多种方式灵活地提供有效的 E2E 质量控制和可靠性保证。这使得基于 ON 方式的 OM 在支持 Internet 的大规模 P2P 文件共享、流媒体应用以及实时多媒体组播应用方面, 具有许多 IPM 无法比拟的优越性。OM 能方便、灵活地利用多种高层的功能实现, 基于 hop-by-hop 方式有效地提供 E2E (end-to-end) 的可靠性和 QoS 保障, 并且较好地支持 TCP-friendliness, 具有良好的可伸缩性。有关 ON 和 OM 的研究还有待进一步深入, 尤其是关于可靠性、安全性机制和 QoS。以下课题值得今后做进一步研究:

- 1) 构建高效的、面向非特定应用的 OM 系统, 并具有可靠性、QoS 保障机制以及良好的可伸缩性;
- 2) ON 的安全性机制;
- 3) 无线网络中支持流媒体应用的 OM 系统以及可靠性、QoS 机制。

## 参 考 文 献

- 1 Deering S, Cheriton D. Multicast Routing in Datagram Internetworks and Extended LANs. In: ACM Transactions on Computer Systems, May 1990
- 2 Chu Y-H, Rao S G, Zhang H. A Case for End System Multicast. In: Proc. of ACM SIGMETRICS, June 2000
- 3 Francis P. Yoid: Extending the Multicast Internet Architecture. 1999, White paper, <http://www.aciri.org/yoid/>
- 4 Ratnasamy S, Francis P, Handley M, Karp R, Shenker S. A scalable content-addressable network. In: Proc. of ACM Sigcomm, Aug. 2001
- 5 Banerjee S, Bhattacharjee B, Kommareddy C. Scalable application layer multicast. In: Proc. of ACM Sigcomm, August 2002

(下转第 129 页)

Coverage 和 precision 分别从推荐的广泛性和精确性方面对推荐方法进行衡量,而 F-测度则是两者的结合。由公式(3)可见,无论忽视 coverage 和 precision 中的哪一个,都将造成 F-测度的降低。

一般来说,具有较高覆盖率和准确率则说明推荐模型具有较好的性能。我们不妨把采用上述算法 1 和算法 2 的推荐模型 AdaptiveSite 叫做 A 模型,把那些采用只基于使用挖掘算法的推荐模型叫做 B 模型,把在使用挖掘基础上考虑了内容相关性(但没经过 frame 裁剪)的推荐模型叫做 C 模型。由于 A 模型既考虑了用户的访问路径,又考虑了基于页面内容的用户兴趣,只从候选推荐集中选出高度感兴趣(由推荐因子反映)的推荐页面(经过 frame 裁剪的),不像 B 模型的一般做法是基于关联规则取出所有后续页面(包括导航页),因此 A 模型推荐集中集中反映用户兴趣的页面数  $|R-LS|$  接近 B 模型,但推荐集页面数  $|RS|$  却较 B 模型少。所以, A 模型在覆盖率和 B 模型相近的情况下,准确率要高于 B 模型。这样, A 模型就达到了推荐页面不是很多,但大部分是用户感兴趣的目的。C 模型也是从候选集中选出用户高度感兴趣的推荐页面,但页面中用户不感兴趣的 frame 没经过裁剪。这样, A 模型相对 C 模型来说,降低了网络传输流量,加快了响应速度。

**总结** 本文对自适应站点 Web 设计方法研究的现状做了分析,并给出了结合使用挖掘和内容挖掘的自适应站点设计算法,构造了一个实现该算法的自适应站点模型 AdaptiveSite,分析了它的推荐质量。从推荐质量分析可以看出,基

于上述算法的自适应站点模型具有较好的推荐优化能力。

## 参考文献

- 1 Fracca F M, Lanzi P L. Recent Developments in Web Usage Mining Research. Data Warehousing and Knowledge Discovery, LNCS(2737), 2003. 140~150
  - 2 Mobasher B, Cooley R, Srivastava J. Creating adaptive sites through usage-based clustering of URLs. In: Proc. Knowledge and Data Engineering Exchange Workshop, IEEE, 1999. 19~25
  - 3 郭新涛,梁敏,阮备军,等. 挖掘 Web 日志降低信息搜索的时间费用[J]. 计算机研究与发展, 2004, 41(10): 1737~1747
  - 4 Kamdar T. Creating adaptive web servers using incremental web log mining; [Master's thesis]. Computer Science Department, University of Maryland, Baltimore County, 2001
  - 5 Zaiane O R. Web usage mining for a better web-based learning environment. In: Proceedings of Conference on Advanced Technology for Education, 2001. 450~455
  - 6 Zhu Jiahua, Hong Jun, Hughes J G. PageCluster: Mining conceptual link hierarchies from Web log files for adaptive Web site navigation. Transactions on Internet Technology, ACM, 2004, 4(2): 185~208
  - 7 Pazzani M J, Billsus D. Adaptive Web Site Agents. Autonomous Agents and Multi-Agent Systems, 2002, 5(2): 205~218
  - 8 Cooley R, Mobasher B, Srivastava J. Grouping Web Page References into Transactions for Mining World Wide Web Browsing Patterns. In: Proc. Knowledge and Data Engineering Exchange Workshop, IEEE, 1997. 2~9
  - 9 张银奎,廖丽,宋俊,等译. 数据挖掘原理[M]. 北京:机械工业出版社, 2003
  - 10 Mobasher B, Dai H, Luo T, et al. Improving the effectiveness of collaborative filtering on anonymous web usage data. In: Proceedings of ACM Workshop on Web Information and Data Management, 2001. <http://maya.cs.depaul.edu/~mobasher/papers/itwp01.pdf>
- (上接第 121 页)
- 6 Chawathe Y, McCanne S, Brewer E. RMX: Reliable Multicast for Heterogeneous Networks. In: Proc. of IEEE INFOCOM, 2000
  - 7 Castro M, Druschel P, Kermarrec A-M, Rowstron A. SCRIBE: A large-scale and decentralized application-level multicast infrastructure. IEEE Journal on Selected Areas in communications (JSAC), 2002
  - 8 Jannotti J, Gifford D K, et al. Overcast: Reliable Multicasting with an Overlay Network. In: Proc. of USENIX Symposium on Operating Systems Design and Implementation, Oct. 2000
  - 9 Pendarakis D, et al. ALMI: An Application Level Multicast Infrastructure. In: Proc. of the 3rd USENIX Symposium on Internet Technologies and Systems, March 2001
  - 10 Zhuang S Q, Zhao B Y, Joseph A D, Katz R, Kubiawicz J. Bayeux: An architecture for scalable and fault-tolerant wide-area data dissemination. In: Eleventh International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV 2001), 2001
  - 11 Chawathe Y. Scattercast: An Architecture for Internet Broadcast Distribution as an Infrastructure Service; [Ph. D. Thesis]. University of California, Berkeley, Dec. 2000
  - 12 Duan Z, Zhang Z, Hou Y T. Bandwidth provisioning for service overlay networks. presented at the SPIE ITCOM Scalability and Traffic Control in IP Networks (II)'02, Boston, MA, July 2002
  - 13 Gu X, Nahrstedt K, Chang R H, Ward C. QoS-Assured Service Composition in Managed Service Overlay Networks. In: Proc. IEEE 23rd International Conference on Distributed Computing Systems, Providence, May 2003
  - 14 Subramanian L, Stoica I, Balakrishnan H, Katz R H. Over QoS: Offering Internet QoS using overlays. presented at the HotNet-Workshop, Princeton, NJ, Oct. 2002
  - 15 Chawathe Y, McCanne S, Brewer E A. RMX: Reliable Multicast for Heterogeneous Networks. In: Proc. of IEEE INFOCOM, 2000
  - 16 Kwon Gu-In, Byers J. ROMA: Reliable Overlay Multicast with Loosely Coupled TCP Connections. In: Proc. of INFOCOM, 2004
  - 17 Urvoy-Keller G, Biersack E. A Congestion Control Model for Multicast Overlay Networks and its Performance. In: Proc. of NGC, 2002
  - 18 Andersen D, Balakrishnan H, Kaashoek F, Morris R. Resilient Overlay Networks. In: Proc. of ACM Symposium on Operation Systems Principles, 2001
  - 19 Pi Renjie, Song Junde. Multi-Path Transmission based on Overlay Network. In: Proc. of AINA, 2004
  - 20 Wu Shuju, et al. Improving the Performance of Overlay Multicast with Dynamic Adaptation. In: Proc. of IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas, NV, Jan. 2004
  - 21 Cui Yi, Xue Yuan, Nahrstedt K. Max-min Overlay Multicast, Rate Allocation and Tree Construction. In: Proc. of IEEE International Workshop on QoS(IwQoS'04)
  - 22 Garey M R, Johnson D S. Computers and Intractability: A Guide to the Theory of NP-completeness. Freeman, San Francisco, 1979
  - 23 Salama H F, Reeves D S, Viniotis Y. Evaluation of multicast routing algorithms for real-time communication on high-speed networks. IEEE Journal on Selected Areas in Communications, April 1997, 15(3): 332~345
  - 24 Faloutsos M, Banerjee A, Pankaj R. QoS-MIC: Quality of Service sensitive Multicast Internet protocol. In: ACM SIGCOMM, Vancouver, Canada, Sept. 1998. 144~153
  - 25 Chen Shigang, Nahrstedt K. A QoS-Aware Multicast Routing Protocol. In: Proc. of IEEE INFOCOM 2000, Tel-Aviv, Israel, March 2000. 1594~1603
  - 26 Henrique L, Costa et al. Hop-by-Hop Multicast Routing Protocol. In: ACM SIGCOMM, Aug. 2001. 249~259
  - 27 Kuipers F, van Mieghem P. MAMCRA: A Constrained-Based Multicast Routing Algorithm. Computer Communications, 25(8): 801~810
  - 28 Aaron S, Manimaran G. A Survey of QoS Multicasting Issues. IEEE Communications Magazine, June 2002
  - 29 Bin W, Jennifer C H. Multicast Routing and Its QoS Extension: Problems, Algorithms, and Protocols. IEEE Network, Jan./Feb. 2000
  - 30 Nageswara S V R. Overlay Networks of In Situ Instruments for Probabilistic Guarantees on Message Delays in Wide-Area Networks. IEEE JSAC, 2004, 32(1)
  - 31 Sun Chang W, Simon R. End-host multicasting in support of distributed real-time simulation systems. In: Proc. of the 37th Annual Simulation Symposium (ANSS'04)
  - 32 Bianchi G, Belfari-Melazzi N, et al. QUASIMODO: Quality of Service Aware Multicasting over DiffServ and Overlay Networks. IEEE Network, Jan./Feb. 2003
  - 33 Ma Z, Shao H-R, Shen C. A New Multi-path Selection Scheme for Video Streaming on Overlay Networks. In: Proc. of IEEE Intl. Conf. on Communications, June 2004 (ICC 2004, TR2004-064)
  - 34 Li Zhi. QRON: QoS-Aware Routing in Overlay Networks. IEEE JSAC, 2004, 22(1)