

多通道并行传输中基于收益最大化的流量分配方法^{*}

郑艳伟^{1,2}, 倪宏², 邓浩江², 刘磊²

(1. 中国科学院研究生院, 北京 100049; 2. 中国科学院声学研究所 国家网络新媒体工程技术研究中心, 北京 100190)

摘要: 针对已有的多通道并行传输中的流量分配方法不区分业务、也未考虑资源收益的问题, 分析了业务参数模型, 引入业务的收益作为多通道并行传输中流量分配的效益评价函数, 通过将流量分配映射为带宽分配建立了基于收益最大化的多通道并行传输流量分配模型。基于边际收益递减法则, 提出了一种边际收益优先的多业务流多通道并行传输流量分配算法, 该算法先根据边际收益优先的贪心策略获取最大化收益的流量分配方案, 然后通过调整业务在不同通道间的流量来保证业务的最低质量要求, 最后通过典型业务实例分析和仿真实验验证了算法的有效性。与基于带宽比例策略的流量分配算法相比, 所提出的方法能够在保证业务质量的前提下, 实现带宽资源的优化配置, 提高运营商的收益。

关键词: 多通道并行传输; 流量分配; 收益最大化; 业务质量

中图分类号: TP393

文献标志码: A

文章编号: 1001-3695(2012)11-4260-05

doi:10.3969/j.issn.1001-3695.2012.11.066

Traffic distribution method for concurrent multipath transfer based on maximum revenue

ZHENG Yan-wei^{1,2}, NI Hong², DENG Hao-jiang², LIU Lei²

(1. Graduate University of Chinese Academy of Sciences, Beijing 100049, China; 2. National Network New Media Engineering Research Center, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: In order to solve the problem that existing traffic distribution methods for concurrent multipath transfer do not distinguish services or consider the revenue of resource, this paper analyzed service parameters and used the revenue of services as the benefit evaluation function of the traffic distribution for concurrent multipath transfer. It built a traffic distribution model based on maximum revenue through mapping traffic distribution to bandwidth allocation. Based on the law of diminishing marginal utility, it proposed a traffic distribution method with marginal revenue priority for concurrent multipath transfer. The algorithm firstly obtained the traffic distribution results which had maximum revenue by a greedy strategy with marginal revenue priority, and then adjusted the traffic among the paths to ensure the minimum QoS. The case analysis and experimental results verify the effectiveness of the method. Compared with the traffic distribution method based on bandwidth ratio policy, the proposed method can achieve the optimal allocation of bandwidth resource and improve the operator's revenue while ensuring the service QoS.

Key words: concurrent multipath transfer; traffic distribution; maximum revenue; QoS

0 引言

随着三网融合的演进和用户带宽需求的日益增长, 多通道并行传输已经成了学者们关注和研究的热点^[1~3]。在多通道并行传输中, 业务数据通过多条通道传输到终端设备, 如何合理地分配业务流在多条通道中的流量, 保证业务的 QoS, 使得整个系统的效益最高, 是一个值得研究的问题。

目前相关的研究主要有基于贪婪、带宽比例和负载均衡策略的流量分配方法。文献[4]提出的 CMT-SCTP 以及文献[5]提出的多路径路由协议都是基于贪婪策略的流量分配, 其主要思想是在传输数据时优先选择第一个空闲的通道, 如果有多个

通道空闲, 则随机选择一个通道进行传输。文献[6]提出的 Westwood-SCTP 首先预测通道的带宽, 然后根据带宽比例进行通道间的流量分配。文献[7, 8]在研究多通道并行传输时, 都采用了基于负载均衡的策略进行通道间的流量分配。但是上述方法都不区分业务, 对不同的业务流采用相同的策略, 而且也不考虑带宽资源的收益, 而如何合理分配多业务流在不同通道间的流量使得系统的收益最大化, 是运营商最关心的问题。

针对上述问题, 本文从业务收益角度出发, 在分析业务参数模型基础上, 给出业务的收益函数模型, 然后提出一种基于收益最大化的多通道并行传输中的流量分配方法; 基于边际收益递减原理, 进行多业务流多通道并行传输的流量分配, 在保证业务质量的基础上实现运营商收益的最大化。最后通过仿

收稿日期: 2012-02-22; 修回日期: 2012-03-29 基金项目: 国家“863”计划资助项目(2011AA01A102); 国家科技支撑计划资助项目(2011BAH11B04)

作者简介: 郑艳伟(1985-), 男, 博士, 主要研究方向为网络新媒体技术、网络通信(zhengyw@dsp.ac.cn); 倪宏(1964-), 男, 主任研究员, 博导, 主要研究方向为网络通信、网络新媒体技术; 邓浩江(1971-), 男, 研究员, 硕士, 主要研究方向为数字音/视频处理、宽带多媒体通信; 刘磊(1980-), 男, 副研究员, 博士, 主要研究方向为新媒体技术、业务运营支撑。

真实实验与基于带宽比例策略的流量分配算法进行比较,验证了所提出的算法有效性。

1 模型建立与分析

1.1 业务参数模型

业务参数主要包括请求到达率、会话持续时间、拒绝率、延迟和丢包率等,本节依次对这些参数进行讨论。

1) 请求到达率 不同业务类型的请求到达率往往具有不同的分布。对于流媒体类业务,请求到达的时间间隔相互独立,请求到达率服从 Poisson 分布。对于数据类业务,数据包的到达时间间隔具有自相似性^[9],一般认为数据包的到达率服从 Pareto 分布。数据类业务一般不要求实时性,可利用资源的空闲间隔进行传输。本文主要针对实时性及业务质量要求较高的会话类和流媒体类业务,采用 Poisson 分布进行分析,令业务 S_i 的请求到达率为 λ_i 。

2) 会话持续时间 不同业务的会话持续时间在时间上相互独立,采用负指数分布来表示会话的持续时间^[10]。不同业务类型的请求持续时间并不相同,令业务 S_i 的平均会话持续时间为 T_i 。

3) 拒绝率 由于需要对每个接收的业务会话提供质量保证,而系统总带宽资源存在限制,所以当系统资源不足时,则无法继续接收请求,即每类业务存在最大可接收的并发会话数。当 S_i 的请求到达系统时,若正在服务的会话数小于最大并发会话数,则接收该请求,否则直接拒绝该请求,不进行排队。

根据上面的假设,会话类业务可建模为一个 M/M/C/C 排队模型。对于该模型,如果只有一种类型的业务,则可以采用 Erlang-B 分布^[11] 分析其行为特性,其业务拒绝率估算为

$$P_B = B(M, \rho) = \frac{\rho^M / M!}{\sum_{n=0}^M \rho^n / n!} \quad (1)$$

其中: $M = \lfloor W_{\max} / W_1 \rfloor$ 为系统能同时服务的最大并发会话数, $W_{\max}$ 为最大可用的带宽, W_1 为单个业务请求占用的带宽; $\rho = \lambda T_i$ 为业务的流量强度。

针对上述模型,如果存在多种类型的业务共享带宽资源,则可以采用文献[11,12]提出的多维 Markov 模型来分析求解各业务的拒绝率,但是多维情况下的求解比较复杂。文献[13]提出把多维空间转换为一维空间的 Kaufman-Roberts 迭代算法可以求解该问题。对此本文不再详述,感兴趣的读者可以阅读文献[11~13]。

4) 延迟 业务的延迟包括业务数据处理时间、链路传输时间及业务请求的排队缓冲时间。数据处理时间和链路传输时间一般很小,可以忽略不计。本文主要针对实时性强、业务质量要求高的业务,此类业务的特性如上面所述,当请求到来时,如正在服务的请求已达到最大并发会话数,则直接拒绝请求,不进入系统进行排队。因此本文在以下的分析中不考虑业务的排队延迟。

5) 丢包率 业务数据在传输过程中存在波动,当为业务分配的带宽较低时,业务传输波动对业务质量影响较大,需要较大的缓冲区;当为业务分配的带宽较高时,业务传输波动对业务质量的影响小,较小的缓冲区即能满足^[14]。当缓冲区不足时,业务数据包就会丢失。为了降低丢失率,需要为业务分配较大的带宽资源。一般认为,当为业务分配的带宽大于最低

需求带宽时,丢包率会低于最大允许的丢包率;小于最低需求带宽时,丢包率会高于最大允许的丢包率。本文采用业务的等效最小带宽来描述业务的最大丢包率约束。

1.2 收益函数

首先讨论单一通道情况下业务的收益函数。对于基于时长计费的业务 S_i , 每秒营业收入为该业务的并发会话数与每秒费用的乘积:

$$R_i = v_i E(N_i) \quad (2)$$

其中: $E(N_i)$ 表示 S_i 的并发会话数的数学期望, $E(N_i) = \rho_i (1 - P_{B_i})$ 。

对于基于计次计费的业务 S_j , 每秒营业收入为该业务每秒接收的请求到达数与每次费用的乘积:

$$R_j = v_j \lambda_j (1 - P_{B_j}) \quad (3)$$

假设业务 S_i 占用 W_i 的带宽资源所需要的成本为 $g(W_i)$, 则运营商开展 N 类业务的总净收益可表示为

$$R = \sum_{i=1}^N [R_i - g(W_i)] \quad (4)$$

在多业务流多通道并行传输情况下,需要将收益函数扩展到所有通道上的业务收益之和。假设有 M 条通道, N 类业务, 则收益函数演变为

$$R = \sum_{j=1}^M \sum_{i=1}^N [R_{ij} - g(W_{ij})] \quad (5)$$

如果 N 类业务中采用按次计费的有 m 个, 且所有业务对占用等量带宽资源具有相同的成本, 则运营商的总净收益表示为

$$R = \sum_{j=1}^M \left(\sum_{i=1}^m v_i \lambda_i (1 - P_{B_i}) + \sum_{i=m+1}^N v_i \rho_i (1 - P_{B_i}) \right) - g(W) \quad (6)$$

从上述分析可以看出, 业务的收益与业务价格、业务到达率、业务持续时间、单个业务请求占用的带宽以及为业务分配的带宽有关, 在固定前四者的情况下, 可以求出业务的收益与为业务分配的带宽之间的关系, 即业务的收益函数。

1.3 基本效用评估模型

在多业务流多通道并行传输中, 假设集合 $P(i)$ 为业务流 i 所用的通道, $S(j)$ 为使用通道 j 的所有业务流。当业务流 i 在通道 j 上获得的处理能力为 y_{ij} ($y_{ij} \geq 0$) 时, $\sum_{j \in P(i)} y_{ij}$ 为业务流 i 在各条通道上获得的处理能力总和。同时, y_{ij} 还应该满足 $\sum_{i \in S(j)} y_{ij} \leq C_j$, 其中 C_j 是通道 j 的最大处理能力, 即通道 j 为所有业务提供的处理能力总和不能超过通道 j 的最大处理能力。

假设业务流 i 获得的处理能力(速率)为 x_i 时, 其效用函数为 $U_i(x_i)$, 则整个系统的效益函数为各个业务流所获得的效用函数之和, 为 $\sum U_i(x_i)$ 。因此, 多业务流多通道并行传输的主要目标就是合理地将各通道的处理能力分配给各个业务流, 从而使得系统的效用函数达到最大值。此问题可抽象为如下最优化问题进行求解:

$$\begin{aligned} & \max \quad \sum_i U_i(x_i) \\ & \text{subject to: } \sum_{j \in P(i)} y_{ij} = x_i \\ & P0: \quad \sum_{i \in S(j)} y_{ij} \leq C_j \\ & \quad y_{ij} \geq 0 \end{aligned} \quad (7)$$

1.4 增强效用评估模型

为了能使上面的效用模型和运营商收益结合起来, 选用收益函数作为评估模型的效用函数, 则上面的最优化问题就转换为基于收益最大化的多通道并行传输流量分配问题, 即在满足

各业务质量和总带宽资源限制的条件下,通过合理分配各业务流在各条通道中的流量,实现资源在业务之间的优化配置,从而使运营商获得最大化的资源收益。此时模型的目标函数演化

$$\sum_{j=1}^M \sum_{i=1}^N R_{ij} - g(W)$$

由 1.3 节的分析可知,PO 问题中的 y_{ij} 是业务 i 在通道 j 中获得的速率,而业务获得的速率是由为业务分配的带宽决定的,故可将业务 i 在通道 j 中获得的处理能力 y_{ij} 转换为业务 i 在通道 j 中所分配的带宽 W_{ij} 。

流量分配的目标是将各业务流的流量(即到达率 λ_i)合理地分配到各条通道中,令 λ_{ij} 代表业务 i 在通道 j 中分配的流量,则 $\sum_{j=1}^M \lambda_{ij} = \lambda_i$ 。

在满足业务质量的前提下,业务 i 分配到通道 j 中的流量与业务 i 在通道 j 中所分配的带宽有关,通道 j 为业务 i 分配的带宽越大,业务 i 分配到通道 j 中的流量也越高,可以用比例式来近似表示两者的关系,即

$$\lambda_{ij}/\lambda_i = W_{ij}/W_j \quad (8)$$

根据上面的分析,建立增强型的效用评估模型为

$$\begin{aligned} & \max \left(\sum_{j=1}^M \sum_{i=1}^N R_{ij} \right) - g(W) \\ & \text{subject to: } P_{B_i} \leq P_{B_i}^m \\ & \sum_{i=1}^N W_{ij} \leq W_j^{\max} \\ P1: & \sum_{j=1}^M W_{ij} \geq W_i^{\text{base}} \\ & \sum_{j=1}^M \lambda_{ij} = \lambda_i \\ & \lambda_{ij}/\lambda_i = W_{ij}/W_j \end{aligned} \quad (9)$$

通过求解上述评估模型,就可以获得各业务流在各条通道中获得的处理能力 W_{ij} ,即可得各业务流在各条通道上分配的流量 λ_{ij} 。

2 算法描述

根据 1.4 节分析可知,多业务流多通道并行传输中基于收益最大化的流量分配问题是一个非线性规划问题,其非线性约束条件非常复杂,难以直接求解。本章首先分析业务带宽边际收益的概念,在保证为各类业务分配最低需求带宽的基础上,采用边际收益最大优先的思想设计多通道并行传输中的流量分配算法。

2.1 业务的边际收益

在描述算法以前,先给出业务带宽的边际收益的定义。借鉴经济学中边际收益的概念^[15],带宽资源的边际收益定义为每增加单位带宽资源所增加的收益值。在离散函数值情况下,业务 i 的边际收益可以表示为

$$q_i = \frac{\Delta R}{\Delta W} = \frac{R(W+w) - R(W)}{w} \quad (10)$$

其中: w 为带宽的最小分配粒度。

根据经济学的边际收益递减原理^[15],即在技术水平不变的条件下,增加某种生产要素的投入,当该生产要素投入数量增加到一定程度以后,增加一单位该要素所带来的产量增加量是递减的。

基于边际收益递减原理,文献[14]提出了如下定理:在带宽资源边际收益递减的前提下,当带宽资源完全分配且业务质量为最低要求时,所有获得资源分配的业务的边际收益相等是

目标收益函数取最大值的充分必要条件。

虽然文献[14]给出了该定理的充分条件和必要条件的证明,但是实际上作者在证明过程中忽略了业务质量这个约束。因为如果考虑业务的最低质量要求,就可能存在这样的情况:业务 1 的边际收益最低,但为了满足业务 1 的最低质量要求,必须为之分配一定的带宽资源。而这与定理中提到的边际收益相等矛盾,按照定理的要求,应该为该业务分配的带宽资源为 0。因此不能只根据边际收益相等来确定业务带宽资源的分配。

2.2 边际收益最大优先的流量分配算法

基于上述业务边际收益的定义,本文提出的多通道并行传输中的流量分配算法分为两个阶段,分别是:

a) 基于边际收益最大的原则将各通道的带宽按照最小粒度依次分配给各个业务。在分配时,首先选择边际收益最大的业务和通道进行分配,然后重新计算已分配业务的边际收益,再重新选择边际收益最大的业务和通道进行分配,如此循环,直至所有通道的带宽都分配完毕为止。

由于上述算法是基于贪心策略的思想,因此当带宽分配完毕时可能有业务未达到其最低质量要求,需要对已分配的带宽资源在业务间进行调整,即算法的第二阶段。

b) 将所有业务分为两个集合,即已达标和未达标的业务集合。从已达标的业务集合中选取边际收益最小的业务,从未达标的业务集合中选取边际收益最大的业务,减少已达标业务占用的带宽,同时增加未达标业务占用的带宽,如此循环,直至所有业务均至少达到最低质量要求为止。

算法的阶段 a) 是根据边际收益递减原理进行的带宽资源分配,所以当算法结束时,系统整体收益达到最大,且已达标集合中业务的边际收益都大于未达标集合中业务的边际收益;而算法的阶段 b) 则是分别从已达标和未达标的业务集合中选取边际收益最小和最大的业务进行带宽的修改,每次修改都尽量使得系统收益的降低程度最小。因此,经过上述两个阶段后,算法能够在保持业务质量最低需求的基础上,实现带宽资源收益的最大化。

假设共有 N 类业务和 M 条通道,第 i 类业务的最低需求带宽为 W_i^{base} ,第 j 条通道的最大可分配带宽为 $W_j^{\max}$, V_{ij}^k 是第 i 类业务在第 j 条通道上带宽为 kw 时的收益,则第 i 类业务在第 j 条通道上的收益函数可用二维数组的形式表示为

$$f_{ij}(W, V) = [(w, V_{ij}^1), (2w, V_{ij}^2), \dots, (kw, V_{ij}^k), \dots, (T_j w, V_{ij}^{T_j})] \quad (11)$$

其中: T_j 由第 j 条通道带宽的上限确定为 $T_j = W_j/w$ 。

基于边际收益最大优先的思想,算法的阶段 a) 的具体步骤为:

(a) 为各类业务分配初始带宽 $W_{ij} = 0$;

(b) 计算各类业务在各条通道中的第一个边际收益,组成二维边际收益矩阵 $CQ = [Q_1, Q_2, \dots, Q_M]$, 其中, $Q_j = [V_{1,j}^1 - V_{1,j}^0, V_{2,j}^1 - V_{2,j}^0, \dots, V_{N,j}^1 - V_{N,j}^0]$, 为各类业务在第 j 条通道上的第一个边际收益组成的集合;

(c) 选择矩阵 CQ 中最大的元素,假设为 $V_{ij}^2 - V_{ij}^1$, 则在第 j 条通道中为第 i 类业务增加分配带宽 w , 即 $W_{ij} = W_{ij} + w$, 同时第 j 条通道的可分配带宽减去 w , 即 $W_j^{\text{rest}} = W_j^{\text{rest}} - w$;

(d) 重新计算第 i 类业务在第 j 条通道中的边际收益, 设为 $V_{ij}^{t+1} - V_{ij}^t$, 用其代替矩阵 CQ 中的第 j 行第 i 列的元素;

(e)判断第 j 条通道的剩余带宽是否为 0,如果是,则从矩阵 CQ 删除第 j 行元素;

(f)判断矩阵 CQ 是否为空,如果是,则转步骤(g),否则转步骤(c)继续执行;

(g)阶段 a)分配结束,为第 i 类业务分配带宽 $\sum_{j=1}^M W_{ij}$,其收益可从 $\sum_{j=1}^M f_{ij}(W, V)$ 中求出。

在算法阶段 a)的分配结束后,可能存在业务未达到其最低质量要求,因此需要对已分配的带宽资源进行调整,即算法的阶段 b),其具体步骤为:

(a)判断第 i 类业务在所有通道中已分配的带宽 $\sum_{j=1}^M W_{ij}$ 是否达到最低需求带宽 W_i^{base} ,如果是,则将业务 i 加入到已达标业务集合 CS_1 中,否则,将业务 i 加入到未达标业务集合 CS_2 中;

(b)分别对已达标业务和未达标业务重新计算其在各条通道中的第一个边际收益,组成边际收益矩阵 $CQ_1 = [Q_1, Q_2, \dots, Q_M]$ 和 $CQ_2 = [Q'_1, Q'_2, \dots, Q'_M]$,其中 $Q_j = [V_{1,j}^2 - V_{1,j}^1, V_{2,j}^2 - V_{2,j}^1, \dots, V_{k,j}^2 - V_{k,j}^1]$ 和 $Q'_j = [V_{k+1,j}^2 - V_{k+1,j}^1, V_{k+2,j}^2 - V_{k+2,j}^1, \dots, V_{N,j}^2 - V_{N,j}^1]$ 分别为已达标业务和未达标业务在第 j 条通道上的第一个边际收益组成的集合;

(c)选择矩阵 CQ_1 中最小的元素,假设为 $V_{ij}^2 - V_{ij}^1$,将第 i 类业务在第 j 条通道中已分配的带宽减去 w ,即 $W_{ij} = W_{ij} - w$;

(d)选择矩阵 CQ_2 中第 j 行最大的元素,假设为 $V_{kj}^2 - V_{kj}^1$,在第 j 条通道中为第 k 类业务增加分配带宽 w ,即 $W_{kj} = W_{kj} + w$;

(e)重新计算第 i 类业务和第 k 类业务在第 j 条通道中的边际收益,设分别为 $V_{ij}^{t+1} - V_{ij}^t$ 和 $V_{kj}^{t+1} - V_{kj}^t$,分别用其代替矩阵 CQ_1 中第 j 行第 i 列的元素和 CQ_2 中第 j 行第 k 列的元素;

(f)判断第 i 类业务在所有通道上已分配的带宽是否降到了其最低带宽需求,如果是,则从矩阵 CQ_1 删除第 i 列元素;

(g)判断第 k 类业务在所有通道上已分配的带宽是否满足其最低带宽需求,如果是,则从矩阵 CQ_2 删除第 k 列元素;

(h)判断矩阵 CQ_2 是否为空,如果是,则转步骤(i),否则转步骤 c)继续执行;

(i)分配结束,为第 i 类业务分配带宽 $\sum_{j=1}^M W_{ij}$,其收益从 $\sum_{j=1}^M f_{ij}(W, V)$ 中求出。

本算法充分利用了边际收益的思想,首先算法的阶段 a)基于边际收益最大优先的贪心策略将所有通道的带宽依次分配给各业务,然后算法的阶段 b)在业务之间对已分配的带宽进行调整,以保证各业务的最低带宽需求。因此,在各通道总带宽资源分配完成并调整完毕时,算法能够在保证业务最低质量要求的基础上,获取最大化的收益。

2.3 算法复杂度分析

算法的复杂度体现在算法的 a)b)两个阶段。从上面的描述可以看出,无论是阶段 a)还是阶段 b),在给出各类业务收益函数的基础上,计算复杂度主要集中在求取边际收益矩阵 CQ 中最大元素或最小元素的过程,其时间复杂度为 MN ,而两个阶段分别需要计算 $\sum_{j=1}^M W_j/w$ 和 $\sum_{i=1}^N W_i^{\text{base}}/w$ 次,因此算法的总时间复杂度为 $O(MN(\sum_{j=1}^M W_j + \sum_{i=1}^N W_i^{\text{base}})/w)$,为多项式时间计算复杂度。

本文提出的算法将多业务流在多通道中并行传输时的流量分配问题转换成了多业务共享多通道的带宽资源分配问题,

在综合分析业务的流量特征和参数模型的基础上进行多通道带宽资源的分配,并不是业务流的实时调度,因此算法对业务流在多通道中传输的时间性能影响不大。

3 实验仿真分析

本章以视频通话、视频点播 1 (VOD1)、视频点播 2 (VOD2)为例,对提出的基于收益最大化的流量分配算法进行实验和仿真分析,仿真在 MATLAB 环境下进行。

假设上述三种业务通过两条通道进行传输,通道 1 可提供的带宽为 800 Mbps,通道 2 可提供的带宽为 200 Mbps,带宽分配的最小粒度为 5 Mbps。实验中所需要的业务参数有业务到达率、业务流量强度、业务的价格、单个业务请求所占用的带宽等,参数取值如表 1 所示。由于通道特性不同,同一种类型的业务在不同的通道中传输所产生的价格收益也不同,因此假设每一种业务在两条通道中都有不同的价格。

表 1 业务仿真参数

类型	到达率	流量强度	价格		占用带宽/M	最大拒绝率
			通道 1	通道 2		
视频通话	2.0	10	0.5	1	5	0.05
VOD1	0.5	30	20	15	10	0.05
VOD2	0.4	24	40	30	15	0.05

3.1 边际收益计算结果

根据上述仿真参数和 1.1 节的相关公式可以得到三种业务的拒绝率函数,如图 1 所示。可以看出,随着带宽的增加,业务的拒绝率迅速下降,而又以视频通话业务的拒绝率下降最快,这是因为视频通话业务所需要的带宽最小。根据各业务的最大拒绝率则可以得到三种业务的最低需求带宽分别为 75、360 和 450 Mbps。

根据各业务在两条通道中的价格和 1.2 节的相关公式可以得到三种业务的收益函数,如图 2、3 所示。有了收益函数,就可以计算出各业务在两条通道中的边际收益函数,如图 4、5 所示。从图中可以看出,随着带宽的增加,业务的收益函数先是急速上升,然后逐渐趋向于平稳,对应的边际收益函数也是先增加,再急速递减,然后趋向于 0。这是由于在开始阶段,业务的边际收益比较大,因此业务的收益也就快速增加,而随着带宽的增加,业务的边际收益逐渐减小,也就造成业务的收益逐渐趋向于平稳。

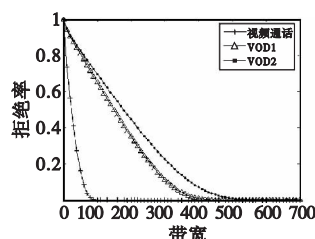


图1 三种业务的拒绝率函数

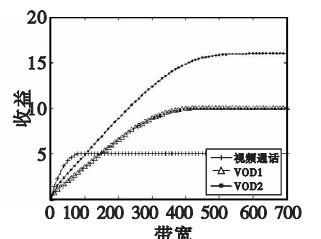


图2 三种业务在通道1上的收益函数

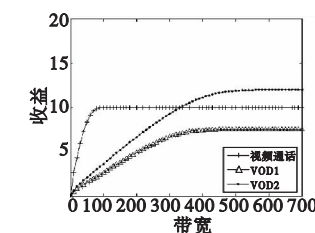


图3 三种业务在通道2上的收益函数

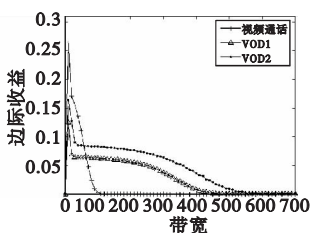


图4 三种业务在通道1上的边际收益函数

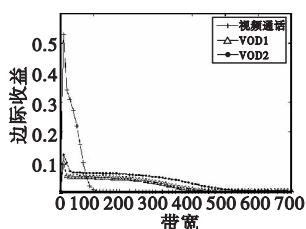


图5 三种业务在通道2上的边际收益函数

3.2 算法仿真结果

根据上述的边际收益函数,执行2.2节提出的算法,就可以得到三种业务在两条通道中的带宽分配结果,再根据1.4节的式(8)就可以获得三种业务在两条通道中的流量分配结果。

算法的阶段a)按照边际收益最大优先的原则为各类业务分配带宽资源,直至所有通道的带宽都分配完毕,仿真结果如表2所示。从表2中可以看出,为VOD1业务分配的带宽小于其最低需求带宽($320 + 10 < 360$),因此算法还需要执行阶段b)的带宽调整步骤。

算法的阶段b)在业务之间调整带宽的分配,保证所有业务均能满足其最低带宽要求,仿真结果如表3所示。从表3中可以看出,所有业务获得的带宽都能满足其最低带宽要求($65 + 60 > 75$, $340 + 20 \geq 360$, $395 + 120 > 450$)。

表2 算法阶段a)的带宽分配结果

业务类型	通道类型	
	通道1 /Mbps	通道2 /Mbps
视频通话	70.0	70.0
VOD1	320.0	10.0
VOD2	410.0	120.0

表3 算法阶段b)的带宽分配结果

业务类型	通道类型	
	通道1 /Mbps	通道2 /Mbps
视频通话	65.0	60.0
VOD1	340.0	20.0
VOD2	395.0	120.0

当所有通道的带宽资源完全分配且调整完毕时,各业务的收益、对应的拒绝率及在两条通道中分配的流量如表4所示。可以看出,当算法执行完毕时,各业务的拒绝率都能满足其最大拒绝率要求。

表4 带宽分配完毕时各业务的收益

业务类型	收益	拒绝率	通道1的 流量强度	通道2的 流量强度
视频通话	7.399 8	$2.927e-5$	5.2	4.8
VOD1	9.438 2	0.043	28.333	1.667
VOD2	14.901 6	0.011	18.408	5.592

3.3 算法性能比较

本节将基于收益最大化的流量分配算法与基于带宽比例的流量分配算法进行比较。

仿真环境采用3.1节的业务参数和通道带宽参数,基于收益最大化的流量分配算法采用3.2节的流量分配结果,基于带宽比例的流量分配算法采用文献[6]的方法,各业务在不同通道中流量按带宽比例进行分配,各业务在同一通道中传输时采用基于完全共享的带宽分配机制。

两种算法的比较结果如表5、6所示,其中表5是各业务在不同算法下的收益,表6是各业务在不同算法下的拒绝率。

表5 各业务在不同算法下的收益

业务类型	算法类型	
	基于带宽比例	基于收益最大化
视频通话	5.963 2	7.399 8
VOD1	9.359 4	9.438 2
VOD2	14.79 2	14.901 6

表6 各业务在不同算法下的拒绝率

业务类型	算法类型	
	基于带宽比例	基于收益最大化
视频通话	0.006 1	$2.927e-5$
VOD1	0.014 8	0.042 9
VOD2	0.026 8	0.011 0

从表5中可以看出,基于带宽比例的流量分配算法的总收益为30.1148,而基于收益最大化的流量分配算法的总收益为31.7396,比基于带宽比例的流量分配算法的总收益提高了5.4%。从表6中可以看出,两种流量分配方法的拒绝率都较低,均能够满足业务的最低需求。

4 结束语

随着三网融合的演进和用户带宽需求的增长,多通道并行传输越发重要,而如何合理地分配业务流在多条通道中的流量,保证业务的QoS,是多通道并行传输中的关键问题。本文在相关研究的基础上,分析了业务参数模型,引入业务的收益作为多通道并行传输中流量分配的效益评价函数,建立了基于收益最大化的多通道并行传输流量分配模型。该模型为非线性规划的数学模型,且约束条件比较复杂,求解困难。针对此问题,本文基于边际收益递减原理,提出一种基于边际收益最大优先的流量分配算法,在保证业务最低质量要求的基础上,实现了系统收益的最大化。仿真实验中对各业务的收益函数和边际收益函数进行了分析,给出了算法的带宽分配结果,与基于带宽比例策略的流量分配算法相比,所提出的算法能够在保证业务质量的前提下实现带宽资源的优化配置,使运营商获得更大的收益。

参考文献:

- [1] YUAN Yao, ZHANG Zi-di, LI Juan, *et al.* Extension of SCTP for concurrent multipath transfer with parallel subflows [C]//Proc of IEEE Wireless Communications and Networking Conference. 2010:1-6.
- [2] BUDZISZ L, FERRUS R, CASADEVALL F, *et al.* On concurrent multipath transfer in SCTP-based handover scenarios [C]//Proc of IEEE International Conference on Communications. Piscataway: IEEE Press, 2009: 1-6.
- [3] KIM K H, JEONG K M, KANG C H, *et al.* A transmission control SCTP for real-time multimedia streaming [J]. *Computer Networks*, 2010, 54(9): 1418-1425.
- [4] IYENGAR J R, AMER P D, STEWART R. Concurrent multipath transfer using SCTP multihoming over independent end-to-end paths [J]. *IEEE/ACM Trans on Networking*, 2006, 14(5): 951-964.
- [5] YANG Wen-jing, YANG Xin-yu, LIU Guo-zheng, *et al.* An interference avoidance multipath routing protocol based on greedy forwarding in MANETs [C]//Proc of IEEE International Conference on Wireless Communications, Networking and Information Security. 2010: 483-487.
- [6] CASETTI C, GAIOTTO W. Westwood SCTP-load balancing over multipaths using bandwidth-aware source scheduling [C]//Proc of the 60th IEEE Vehicular Technology Conference. 2004: 3025-3029.
- [7] JOE I W, YAN Si-jia. SCTP throughput improvement with best load sharing based on multihoming [C]//Proc of the 5th International Joint Conference on INC, IMS and IDC. Washington DC: IEEE Computer Society, 2009: 138-142.

方案 6(0,2,6,4,5,7,3,1,8,10,14,12,13,15,11,9)。

对星座图进行旋转并不改变其判决区域,通过排列组合可知此六种星座图仅为有的最小判决区域为 $\pi/2$ 的星型 16QAM 星座图。

3 仿真结果与分析

为了验证本文提出的星座映射设计,构造了星形 16QAM 调制下的 LDPC-BICM-ID 系统,图 5 给出了在 AWGN 信道下,最小判决区域为 $\pi/2$ 的六种映射方式和图 3 所示的映射方式误码率曲线比较图,仿真参数的区别仅在于映射方式。

仿真参数设置如下:采用码长为 1024 bit,行重为 6、列重为 3、码率为 1/2 的规则 LDPC 码,迭代解映射和译码 8 次,每次迭代中 LDPC 译码迭代 10 次,译码输出作为先验信息重新解映射再译码,比特交织器采用随机交织,LDPC 译码算法采用最基本的置信传播(belief propagation, BP)算法^[11],解映射算法采用文献[12]中提出的算法。

由图 5 可见,方案 1~6 误码曲线基本重合,误码性能差距在仿真点数足够大时可以忽略不计,说明几种映射方式是等价的;但是相比于图 3 的星座映射方式,系统误码性能有较大的提升,当误码率达到 10^{-6} 时,前者比后者性能提高了 2.2 dB,本文研究的设计准则得证。而且 Rayleigh 衰落信道条件下此准则同样适用,图 6 给出误码性能曲线,与图 5 仿真参数设置区别仅在于信道。

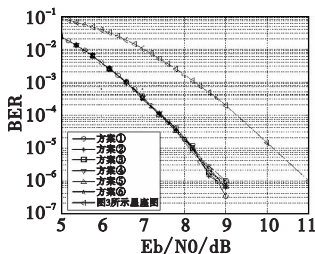


图5 AWGN信道不同映射方式误码特性比较

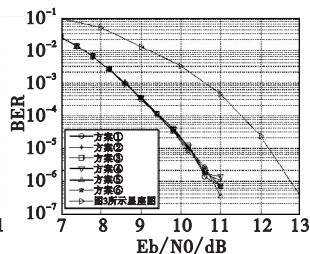


图6 瑞利信道不同映射方式误码特性比较

4 结束语

LDPC 编码的 BICM-ID 系统方案已经被应用于广播通信系统,易于实现的优势使得星型 16QAM 调制方式也受到广泛关注。本研究通过分析星座映射对系统性能的影响,提出一种适合于星型 16QAM 调制 LDPC 编码的 BICM-ID 系统的星座映射新规则,结合本文提出的映射设计方法,可有效地设计出性能优化的映射方式,理论分析及计算机仿真结果都表明了该映

射规则下误码性能的优越性,为实际应用中映射方式的选择提供了依据。

参考文献:

- [1] CAIRE G, TARICCO G, BIGLIERI E. Bit-interleaved coded modulation[J]. *IEEE Trans on Information Theory*, 1998, 44(3): 927-946.
- [2] LI Xiao-dong, CHINDAPOL A, RITCEY J A. Bit-interleaved coded modulation with iterative decoding and 8 PSK signaling[J]. *IEEE Trans on Communication*, 2002, 50(8): 1250-1257.
- [3] WANG Ying, XIE Lei, CHEN Hui-fang, et al. Improved decoding algorithm of bit-interleaved coded modulation for LDPC code[J]. *IEEE Trans on Broadcasting*, 2010, 56(1): 103-109.
- [4] WEI Ruey-yi. Differential encoding by a look-up table for quadrature-amplitude modulation[J]. *IEEE Trans on Communication*, 2011, 59(1): 84-94.
- [5] HOFFMANN S, WORDEHOFF C, Al-Bermani A, et al. Hardware-efficient phase estimation for digital coherent transmission with star constellation QAM[J]. *IEEE Photonics Society*, 2010, 2(2): 174-180.
- [6] 王军选, 尧文元, 廖汉程. 多径衰落下基于多码检测的多天线 CDMA 信道容量分析[J]. *北京邮电大学学报*, 2006, 29(3): 99-102.
- [7] TRAN N H, NGUYEN H H. Signal mappings of 8-gary constellations for bit interleaved coded modulation with iterative decoding[J]. *IEEE Trans on Broadcasting*, 2006, 52(1): 92-99.
- [8] TAN Jun, GORDON L S. Analysis and design of symbol mappers for iteratively decoded BICM[J]. *IEEE Trans on Wireless Communication*, 2005, 4(2): 662-672.
- [9] DON T, MATTHEW C V. Constellation labeling maps for low error floors[J]. *IEEE Trans on Wireless Communication*, 2008, 7(12): 5401-5407.
- [10] LIANG Dan-dan, NG S X, HANZO L. Soft-decision star-QAM aided BICM-ID[J]. *IEEE Signal Processing Letters*, 2011, 18(3): 169-172.
- [11] OHTSUKI T. LDPC codes in communications and broadcasting[J]. *IEICE Trans on Communication*, 2007, E90-B(3): 440-453.
- [12] 谢求亮, 彭克武, 潘长勇, 等. 比特交织 LDPC 编码调制系统中的迭代解映射和译码算法[J]. *清华大学学报: 自然科学版*, 2009, 49(8): 1201-1204.
- [11] GRIFFITH D, SRIRAM K, KRIVULINA L, et al. Resource planning and bandwidth allocation in hybrid fiber-coax residential networks[C]//Proc of the 1st International Conference on Broadband Network. Washington DC: IEEE Computer Society, 2004: 263-268.
- [12] 徐站. 下一代广播电视网开放业务平台中网络边缘资源控制技术[D]. 合肥: 中国科学技术大学, 2011.
- [13] KAUFMAN J S. BLOCKING in a shared resource environment[J]. *IEEE Trans on Communications*, 1981, 29(10): 1474-1480.
- [14] 董元元. 开放业务平台精细化运营的支撑技术研究[D]. 北京: 中国科学院声学研究所, 2011.
- [15] 百度百科. 边际效用递减法则[EB/OL]. (2011-05-13) [2012-02-22]. <http://baike.baidu.com/view/1508604.htm>.

(上接第 4264 页)

- [8] DONG Yu, WANG Ding-ding, PISSINOU N, et al. Multipath load balancing in transport layer[C]//Proc of the 3rd EuroNGI Conference on Next Generation Internet Networks. 2007: 135-142.
- [9] GORDON J. Pareto process as a model of self-similar packet traffic[C]//Proc of IEEE Global Telecommunications Conference. Piscataway: IEEE Press, 1995: 2232-2236.
- [10] VILAS M, PANEDA X G, GARCIA R, et al. User behavior analysis of a video-on-demand service with a wide variety of subjects and lengths[C]//Proc of the 31st EUROMICRO Conference on Software Engineering and Advanced Applications. Washington DC: IEEE Computer Society, 2005: 330-337.