

## 缓冲辅助中继网络中的缓冲资源分配方法

刘婷婷<sup>1</sup>, 陈开源<sup>1</sup>, 蒋诚智<sup>2</sup>, 余雨<sup>1</sup>, 蒋姝<sup>1</sup>

(1. 南京工程学院信息与通信工程学院, 南京, 211167; 2. 南京工程学院经济与管理学院, 南京, 211167)

**摘要:** 随着通信技术和智能终端的发展, 移动内容消费出现了明显的增长。人们随时随地享受无线网络带来的多媒体娱乐体验的同时, 也给蜂窝网络带来了巨大的压力, 尤其是热门视频的重复传输, 浪费了网络资源。预存储技术被认为是解决这类问题的有效方法。它是在非高峰时段, 将热门视频等预先存储在网络边缘, 以缓解高峰时段网络流量的压力。为了给用户提供更好的内容传输效率, 预存储技术中基于缓冲辅助的中继网络被提出。由于中继节点缓冲空间有限, 当其需要服务多个用户时, 有必要对中继节点的缓冲资源进行优化分配。基于此, 本文提出了一种中继节点的缓冲资源分配方法。首先, 根据中继节点数据到达模型和数据处理模型, 以及中继节点能忍受的缓冲区溢出概率, 得到中继节点可以贡献出的最大缓冲区比例; 接着, 根据用户的数据到达模型和数据服务模型, 给出用户缓冲区溢出概率与缓冲区比例关系的表达式; 然后, 建立优化问题, 以最小化网络中所有用户缓冲区溢出概率和为目标, 给出最优的缓冲区分配方法; 最后, 仿真结果证实了本文提出的方法比均匀分配法能获得更小的用户缓冲区溢出概率和。

**关键词:** 内容分发网络; 缓冲辅助中继网络; 资源分配

**中图分类号:** TN925; TN919.3; TN919.5      **文献标志码:** A

### Buffer Resource Allocation in Buffer-Aided Relay Networks

LIU Tingting<sup>1</sup>, CHEN Kaiyuan<sup>1</sup>, JIANG Chengzhi<sup>2</sup>, YU Yu<sup>1</sup>, JIANG Shu<sup>1</sup>

(1. School of Information and Communication Engineering, Nanjing Institute of Technology, Nanjing, 211167, China; 2. School of Economics and Management, Nanjing Institute of Technology, Nanjing, 211167, China)

**Abstract:** With the development of communication technology and the boost of intelligent terminals, mobile content consumption has been witnessed a significant increase. While people enjoy the multimedia entertainment experience brought by wireless network anytime and anywhere, it also brings great pressure to the cellular networks, especially the repeated transmissions of popular videos, which leads to the waste of network resources. The caching technology is considered to be an effective way to solve such problems. It stores popular videos in advance at the edge of the network in off-peak hours to relieve the pressure of network traffic during peak hours. To provide users with better content transmission efficiency, the buffer-aided relay network is proposed. Due to the limited buffer space of the relay node, it is necessary to study the allocation of buffer resources when the relay node needs to serve multiple users. To this end, this paper proposes a buffer resource allocation method for relay nodes. Firstly, according to the data arrival model and data processing model of the relay node, as well as the buffer overflow probability that the relay node

**基金项目:** 国家自然科学基金(61702258)资助项目。

**收稿日期:** 2020-07-20; **修订日期:** 2020-09-11

can tolerate, the maximum buffer proportion contributed by the relay node is provided. Then, according to the user's data arrival model and data service model, the expression between the user's buffer overflow probability and the buffer portion is given. Then, an optimization problem is established to minimize the sum of user buffer overflow probability in the network, and the optimal buffer allocation method is presented. Finally, simulation results show that the proposed method can obtain a smaller sum of user buffer overflow probability than the uniform allocation method.

**Key words:** content distribution network; buffer relay-aided network; resource allocation

## 引 言

近年来,由于智能终端的迅速发展,移动内容消费日趋增加,人们可以随时随地享受无线蜂窝网络带来的娱乐体验,但同时也给蜂窝网络带来了巨大的流量压力<sup>[1]</sup>。有证据表明,蜂窝网络中存在很多重复传输,比如热门视频,热点新闻短片等等。因此,有学者提出预先在网络边缘结点,比如移动手机终端,机顶盒,路由器等边缘存储设备中,存储热门视频的方式,缓解网络中,尤其是主干网中重复流量的问题<sup>[2-3]</sup>。在之前的研究中,也证实了可以通过预存储热门视频或文件,有效降低网络流量,提高网络传输效率,同时也增加用户的满意度。

然而,由于边缘结点存储空间有限,或者缺乏足够的激励,网络的边缘结点并不愿意贡献出足够的存储空间。因此,只有部分文件或部分热门视频,被存储在边缘结点中,当这些没有被存储/没有被完整存储的文件被请求时,需要向远端的服务器提出请求。由服务器通过网络传输未被存储的部分。因为服务器一般都位于较远的地区,文件传输过程中,会经历较长的时延,若文件到达网络边缘,比如基站,先被放入边缘结点的存储空间,再从存储空间读取,分发给请求用户,将进一步增加时延,容易导致用户满意度进一步降低。然而,在边缘结点中,除了存储空间,还有缓冲空间。缓冲空间是一种造价相对较高的短暂记忆高性能存储器。为了提高用户的满意程度,已有一部分学者开始研究如何使用缓冲辅助中继方法来提高从远端服务器传输文件的性能<sup>[4-5]</sup>。文献[4]提出了一种自适应链路选择的中继协议。为了避免数据丢失,在中继结点中配备了缓冲器,数据在被传输之前,在缓冲器中排队。笔者研究了时延约束和无时延约束两种条件下的链路选择方案。此外,还提出一种改进的链路选择协议,可以有效避免缓冲区溢出。文献[5]讨论了几种不同网络拓扑的缓冲辅助中继方案,包括单向单中继网络、单向多中继网络以及双向单中继网络。

上述论文主要是针对链路选择方案进行研究。然而,当到达数据流与离开数据流不平衡时,缓冲空间中容易产生数据积压,甚至数据溢出,也是影响缓冲辅助中继方法实现的重要因素之一。有必要对缓冲空间数据到达和离开与数据积压或缓冲区溢出关系进行研究。近年来,鞅理论被认为是分析数据积压/数据溢出的有效工具<sup>[6]</sup>。鞅理论适用范围较广,适于任何数据到达和离开过程,尤其对于有突发数据流的环境,能获得紧致的估计<sup>[7-10]</sup>。文献[7]运用鞅理论研究了在高速铁路网络中,数据的传输时延界,并将理论推导出的时延界与真实数据对比,证明了鞅理论获得的时延界具有紧致性;文献[8]运用鞅理论分析了车辆网中,车辆之间数据传输的时延界,实验结果也证明了鞅理论获得的时延界与真实数据模拟的时延界是一致的。随着研究的发展,还有学者将鞅理论运用于优化问题,将鞅理论获得的时延界作为优化条件或目标,增强网络性能。文献[9]将鞅理论推导出的时延界作为限制条件,优化网络的能效;文献[10]将鞅理论推导的时延违反概率作为目标函数,建立优化问题,获得计算任务最佳分配方案。

缓冲空间作为影响边缘结点或中继结点工作性能的重要部件,边缘/中继结点只愿意贡献出有

限的空间,用于缓冲辅助中继。需要在保证边缘/中继结点自有业务使用性能的前提下,对其贡献出的缓冲空间进行合理分配,以优化缓冲辅助中继的分发性能。因为边缘结点在缓冲辅助中继网络中有中继结点的功能,为了方便讨论,文中将不再区分边缘结点和中继结点,统一使用中继结点。本文将研究缓冲辅助网络中缓冲资源的最优分配方案。首先研究了在缓冲辅助中继网络中,为了保证中继结点自有业务使用性能,优化缓冲资源的方法;其次,使用鞅理论对中继结点缓冲空间内部的积压进行分析,确定缓冲区溢出概率与中继结点请求和输出参数的关系,并给出了能满足中继结点自有业务使用性能的缓冲区大小的表达式;接着,运用鞅理论,确定用户在中继结点处缓冲区溢出概率表达式,构建最小化所有用户缓冲区溢出概率和的优化问题,获得闭式解;最后,通过计算机仿真验证了中继结点贡献出的最大缓冲区比例与中继结点单位时间产生数据量的关系,用户分配到的最优缓冲区比例与用户距离中继结点的距离的关系,本文提出的最优缓冲区分配方法优于均匀分配方法。

## 1 系统模型

本文使用的系统模型如图1所示,该系统由远端服务器,回程链路,基站,中继结点和多个用户组成。在中继结点覆盖范围内有 $N$ 个用户,表示为 $U_i, i=1, \dots, N$ ,这些用户均有文件传输需求。该网络使用缓冲辅助中继方法提高文件传输性能,具体是指中继结点将缓冲空间分成不等的 $N$ 份,为 $N$ 个网络用户传输文件使用。

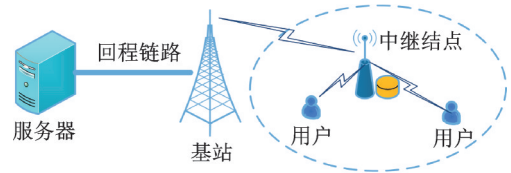


图1 系统模型

Fig.1 System model

### 1.1 中继结点能贡献出的缓冲区比例

在保证中继结点自有业务传输性能的前提下,其能够贡献出的缓冲区容量是有限的。下面根据中继结点对缓冲区溢出概率的要求,确定中继结点能贡献出的最大缓冲区的比例。

#### 1.1.1 中继结点的数据到达模型

中继结点 $R$ 由于自有业务传输性能要求,也要使用缓冲空间。所以中继结点在为周围用户提供缓冲辅助中继服务之前,根据数据到达模型和数据处理模型,需要预留出部分缓冲空间,供中继结点自有业务使用。中继结点在时间 $l$ 产生的数据量为 $r(l)$ 。 $r(l)$ 服从马尔科夫调制开关过程,具有两个状态 $\pi_R^0$ 和 $\pi_R^1$ 。状态 $\pi_R^0$ 表示没有数据到达,在这种状态下, $r(l)=0$ 。状态 $\pi_R^1$ 表示有数据到达,即 $r(l) \neq 0$ 。假设状态 $\pi_R^0$ 转移到状态 $\pi_R^1$ 的概率为 $p_r$ ,状态 $\pi_R^1$ 转移到状态 $\pi_R^0$ 的概率为 $q_r$ ,用转移矩阵表示为 $\mathbf{\Pi} = \begin{bmatrix} 1-p_r & p_r \\ q_r & 1-q_r \end{bmatrix}$ 。 $r(l)$ 在状态 $\pi_R^0$ 和 $\pi_R^1$ 的稳态分布可以分别表示为 $\frac{q_r}{p_r+q_r}$ 和 $\frac{p_r}{p_r+q_r}$ 。在一段时间 $[v, u]$ 内,中继结点累计到达数据量可表示为

$$R(v, u) = \sum_{l=v}^u r(l) \quad (1)$$

式中: $r(l)$ 表示中继结点在时间 $l$ 产生的数据量; $R(v, u)$ 表示中继结点在时间 $[v, u]$ 内累计到达的数据量。

#### 1.1.2 中继结点的数据处理模型

中继结点在时间 $k$ ,处理结点内部产生的数据速率可以表示为常数 $c$ 。在一段时间 $[v, u]$ 内,累计处理的数据量 $C(v, u)$ 可表示为

$$C(v, u) = \sum_{l=v}^u c \quad (2)$$

为使讨论有效,假设

$$E[r(l)] \leq c \leq \sup r(l) \quad \forall l \quad (3)$$

式(3)表示,中继节点的数据处理速率大于数据到达速率的期望,同时小于数据到达速率的峰值。这种情况会在中继节点内部产生数据积压,引起缓冲区溢出。

### 1.1.3 中继节点缓冲区溢出概率

中继节点内部积压  $Q_r$  定义为

$$Q_r = \sup \{ R(n) - C \} \quad (4)$$

式中:  $R(n)$  表示在时间  $[0, n]$  内累计到达的数据量;  $C$  表示在时间  $[0, n]$  内累计处理的数据量。从式(4)可以看出,缓冲中的积压值有随机的特性,很难给出确定的数值。因此,采用鞅理论,缓冲区溢出概率可表达为<sup>[11-12]</sup>

$$Pr(Q_r \geq (1 - \alpha)X) \leq \frac{E[h_r(r(0))]}{H_r} e^{-\theta_r^*(1-\alpha)X} \quad (5)$$

式中:  $Q_r$  表示中继节点自有业务产生的积压;  $\alpha$  表示用于缓冲辅助中继使用的缓冲区比例,  $X$  表示缓冲节点的总缓冲区容量,  $\alpha X$  表示中继节点贡献出的用于缓冲辅助中继的缓冲区大小,  $(1 - \alpha)X$  则表示中继节点留给自有业务使用的缓冲区大小,其中,  $0 \leq \alpha \leq 1$  表示中继节点贡献出的比例;根据鞅理论,  $E[h_r(r(0))]$  表示对  $h_r(r(0))$  求期望,  $h_r(r(0))$  表示  $r(0)$  的右特征向量,  $r(0)$  表示时间为 0 时请求的初始数据量。  $H_r = \min \{ h_r(r(l)) | r(l) - c > 0, l \geq 0 \}$ ,  $h_r(r(l))$  表示  $r(l)$  的右特征向量,  $\theta_r^* = \sup \{ \theta_r > 0 : K_r(\theta_r) \leq K_c(\theta_r) \}$ ,  $\theta_r$  是一辅助变量,  $K_r(\theta_r) = \frac{\ln sp(T_r^{\theta_r})}{\theta_r}$ , 其中,  $T_r^{\theta_r} \triangleq \mathbf{I} e^{\theta_r r(l)}$ ,  $sp(T_r^{\theta_r})$  表示  $T_r^{\theta_r}$  的谱半径,  $K_c(\theta_r) = -\frac{\ln E[e^{-\theta_r c}]}{\theta_r}$ 。

若给定中继节点能够容忍的缓冲区溢出概率  $\xi_r$ , 根据式(5)可得到中继节点贡献出的最大缓冲区比例为

$$\alpha \leq 1 - \frac{1}{\theta_r^* X} \ln \frac{E[h_r(r(0))]}{H_r \xi_r} \quad (6)$$

## 1.2 用户在中继节点处缓冲区溢出概率

基于1.1节的讨论,得到中继节点贡献给用户缓冲辅助中继的缓冲区大小为  $\alpha X$ 。假设用户  $U_i$  分配到的缓冲区大小为  $\alpha_i X$ ,  $0 \leq \alpha_i \leq \alpha$ , 且  $\sum_{i=1}^N \alpha_i \leq \alpha$ 。下面分析用户缓冲区溢出概率与缓冲区大小的关系。

### 1.2.1 用户的数据到达模型

假设用户  $U_i$  在时间  $l$  请求的数据量为  $a_i(l)$ 。  $a_i(l)$  服从马尔科夫调制开关过程, 具有两个状态  $\pi_{U_i}^0$  和  $\pi_{U_i}^1$ , 状态  $\pi_{U_i}^0$  表示没有数据到达, 即  $a_i(l) = 0$ , 状态  $\pi_{U_i}^1$  表示有数据到达, 且  $a_i(l) \neq 0$ 。用户  $U_i$  的状态转移矩阵表示为  $\mathbf{I}_i = \begin{bmatrix} 1-p_i & p_i \\ q_i & 1-q_i \end{bmatrix}$ , 两个状态的稳态分布可以表示为  $\frac{q_i}{p_i+q_i}$  和  $\frac{p_i}{p_i+q_i}$ 。在一段时间  $[v, u]$  内, 用户  $U_i$  累积到达中继节点的数据量  $A_i(v, u)$  可表示为

$$A_i(v, u) = \sum_{l=v}^u a_i(l) \quad (7)$$

式中  $a_i(l)$  表示用户  $U_i$  在时间  $l$  请求的数据量。

### 1.2.2 用户的数据服务模型

中继节点贡献给用户  $U_i$  的有限的缓冲容量为  $\alpha_i X$ , 用于缓冲从基站发送给用户的数据, 再利用近

距离优势,采用高速传输技术发送给用户,提高内容分发的效率。在时间 $l$ ,从中继结点到用户的服务速率可表示为

$$s_i = B \log_2 \left( 1 + \frac{P_i^c}{B \sigma_o^2} \right) \quad (8)$$

式中: $B$ 为分配给每个用户的带宽; $P_i^c$ 为用户的接收功率, $P_i^c = P_{tr} d_i^{-g}$ , $P_{tr}$ 为传输功率, $d_i$ 为中继结点到用户 $U_i$ 的距离, $g$ 为路径损耗参数; $\sigma_o^2$ 为噪声密度。在一段时间 $[v, u]$ 内,中继结点传输给用户的数据量 $S_i(v, u)$ 表示为

$$S_i(v, u) = \sum_{l=v}^u s_i(l) \quad (9)$$

式中 $s_i(l)$ 表示在时间 $l$ ,从中继结点到用户的服务数据量。为使讨论有效,假设

$$E[a_i(l)] \leq s_i \leq \sup a_i(l) \quad \forall l \quad (10)$$

式(10)表示,中继结点传输给用户 $U_i$ 的速率 $s_i$ 大于用户 $U_i$ 到达数据速率的期望 $E[a_i(l)]$ ,但小于用户 $U_i$ 请求数据到达的峰值 $\sup a_i(l)$ 。这种情况也会在中继结点内部产生积压,引起分配给用户 $U_i$ 的缓冲区溢出。

### 1.2.3 用户的缓冲区溢出概率

使用鞅理论,用户 $U_i$ 在中继结点内部的缓冲区溢出概率可写成

$$Pr(Q_i \geq \alpha_i X) \leq \frac{E[\mathbf{h}_i(a_i(0))]}{H_i} e^{-\theta_i^* \alpha_i X} \quad (11)$$

式中: $Q_i$ 表示用户 $U_i$ 产生的积压; $\alpha_i X$ 表示中继结点贡献给用户 $U_i$ 的用于缓冲辅助中继的缓冲区大小。根据鞅理论, $E[\mathbf{h}_i(a_i(0))]$ 表示对 $\mathbf{h}_i(a_i(0))$ 求期望, $\mathbf{h}_i(a_i(0))$ 表示 $a_i(0)$ 的右特征向量, $a_i(0)$ 表示时间为0时请求的初始数据量。 $H_i = \min\{\mathbf{h}_i(a_i(l)) | a_i(l) - s_i > 0, l \geq 0\}$ , $\mathbf{h}_i(a_i(l))$ 表示 $a_i(l)$ 的右特征向量, $\theta_i^* = \sup\{\theta_i > 0: K_a(\theta_i) \leq K_s(\theta_i)\}$ , $\theta_i$ 是一辅助变量, $K_a(\theta_i) = \frac{\ln sp(Tr_a^{\theta_i})}{\theta_i}$ ,其中,

$$Tr_a^{\theta_i} \triangleq \mathbf{\Pi}_i e^{\theta_i a_i(l)}, sp(Tr_a^{\theta_i}) \text{表示 } Tr_a^{\theta_i} \text{的谱半径}, K_s(\theta_i) = -\frac{\ln E[e^{-\theta_i s_i}]}{\theta_i}。$$

## 2 最优缓冲区资源分配方法

中继结点的缓冲区需要划分成不等的 $N$ 份,为了获得最小的缓冲区溢出概率和,建立如下优化问题

$$\min_{\alpha_i} \sum_{i=1}^N \frac{E[\mathbf{h}_i(a_i(0))]}{H_i} e^{-\theta_i^* \alpha_i X} \quad (12)$$

$$\text{s.t. } \frac{E[\mathbf{h}_r(r(0))]}{H_r} e^{-\theta_r^* (1-\alpha) X} \leq \xi_r \quad (13)$$

$$\sum_{i=1}^N \alpha_i \leq \alpha \quad (14)$$

$$\alpha_i \geq 0 \quad (15)$$

式(13)表示根据中继结点保留的缓存空间求出的缓冲区溢出概率小于等于中继结点能够容忍的缓冲区溢出概率 $\xi_r$ ;式(14)表示分给用户的缓冲区比例和不能超出中继结点分配给用户使用的缓冲区比例;式(15)约束了分配给用户的缓冲区比例值是非负值。对式(12)求一阶导数和二阶导数,可分别获得

$$\text{一阶导数 } -\theta_i^* X \frac{E[h_i(a_i(0))]}{H_i} e^{-\theta_i^* a_i X} \quad (16)$$

$$\text{二阶导数 } \theta_i^{*2} X^2 \frac{E[h_i(a_i(0))]}{H_i} e^{-\theta_i^* a_i X} \quad (17)$$

从式(16,17)可看出,目标函数的一阶导数小于零,二阶导数大于零,这是一个典型的凸函数。建立拉格朗日方程和 Karush-Kuhn-Tucher (KKT)条件

$$\mathcal{L} = \sum_{i=1}^N \left( \frac{E[h_i(a_i(0))]}{H_i} e^{-\theta_i^* a_i X} + \nu \alpha_i + \kappa_i \alpha_i \right) \quad (18)$$

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \alpha_i} = -\theta_i^* X \frac{E[h_i(a_i(0))]}{H_i} e^{-\theta_i^* a_i X} + \nu + \kappa_i = 0 \\ \kappa_i \geq 0 \\ \kappa_i \alpha_i = 0 \end{cases} \quad (19)$$

式中:参数  $\{\kappa_1, \kappa_2, \dots, \kappa_N\}$  和参数  $\nu$  为拉格朗日乘子。KKT 条件中的第3个条件,即  $\kappa_i \alpha_i = 0$ ,可以得到当  $\kappa_i \neq 0$  时,  $\alpha_i = 0$ ; 或者当  $\kappa_i = 0$  时

$$\alpha_i = \frac{1}{\theta_i^* X} \left[ \ln \theta_i^* X \frac{E[h_i(a_i(0))]}{H_i} - \ln \nu \right] \quad (20)$$

式中  $\nu$  是一变量。可通过调节  $\nu$  的值,来满足限制条件  $\sum_{i=1}^N \alpha_i \leq \alpha$ 。因此,可以获得优化问题的解为

$$\alpha_i^* = \frac{1}{\theta_i^* X} \left( \ln \frac{E[h_i(a_i(0))]}{H_i} \theta_i^* X - \frac{\sum_{i=1}^N \frac{1}{\theta_i^* X} \ln \frac{E[h_i(a_i(0))]}{H_i} \theta_i^* X + \frac{1}{\theta_r^* X} \ln \frac{E[h_r(r(0))]}{H_r \xi_r}}{\sum_{i=1}^N \frac{1}{\theta_i^* X}} - 1 \right)^+ \quad (21)$$

式中符号  $(\cdot)^+$  表示:若括号内的数据比零小,则取零值;若括号内的数据比零大,则取该数值。

综上,本文提出的算法步骤如图2所示。首先,根据鞅理论确定中继节点缓冲区溢出概率表达式;然后,根据中继节点能够容忍的缓冲区溢出概率值,求出中继节点能贡献出的最大缓冲区比例;接着,运用鞅理论,确定用户在中继节点内部的缓冲区溢出概率表达式;接着,建立最小化用户缓冲区溢出概率和的优化问题,受限于中继节点能贡献出的最大缓冲区比例和每个用户分配到的缓冲区比例必须为大于等于零的数;最后,运用拉格朗日方程和KTT条件,求出缓冲区最优分配比例。

### 3 仿真结果

本文提出的缓冲资源分配方法的数值仿真参数如表1所示。

图3给出了中继节点贡献出的最大缓冲区比例随着中继节点单位时间产生的数据量  $r$  的关系图。从图3可以看出,随着中继节点单位时间产生数据量  $r$  不断增大,中继节点贡献出的最大缓冲区比例逐渐降低。这是因为随着中继节点单位时间产生的数据量增加,需要的缓冲空间也在逐渐增大。中继节点需要留一定的缓冲空间,以满足中继节点缓冲区溢出概率值。因此,剩余可贡献的缓冲空间就逐渐减少。此外,图中还对比了在不在

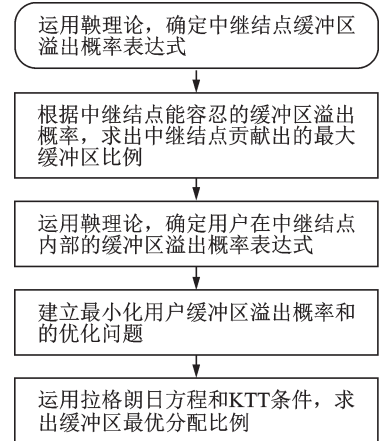


图2 本文提出的算法步骤

Fig.2 Steps of the proposed algorithm

同中继节点服务速率的情况下,中继节点贡献出的最大缓冲区比例的区别。当中继节点服务速率为 2.4 Mbit/s 的情况下,中继节点贡献出的最大缓冲区比例比中继节点服务速率为 2.65 Mbit/s 的情况下小,并且在中继节点单位时间产生的数据量大于 4.5 Mbit/s 的时候,差距逐渐加大。这是由于中继节点服务速率减少,中继节点缓冲区积压将变多,为了满足同样的缓冲区溢出概率要求,中继节点需要保留更多的缓冲区空间,所以贡献出的缓冲区比例变小。

图 4 给出了 6 个用户,距离中继节点距离为 20 m,用户的数据到达率分别为 4、4.2、4.4、4.6、4.8 和 5 Mbit/s。序号 1—6 分别代表用户 1 到用户 6。随着用户到达速率增加,而中继节点给用户的传输速率是一样的,那么用户在中继节点内部的积压值也会增加,为了获得较低的缓冲区溢出概率值,需要给到达速率较大的用户分配较大的缓冲区比例。从图 4 可看出,用户 6,即到达速率为 5 Mbit/s 的用户,获得了最大的缓冲区比例,而到达速率最低的用户 1,则被分配了最小的缓冲区比例。此外,图中还对比了中继节点数据到达速率不同时,用户分配的最优缓冲区比例的区别。可以看出,中继节点数据到达率较高时,即 4.8 Mbit/s 时,分配给用户的最优缓冲区比例小于中继节点数据到达率为 3 Mbit/s 时的最优缓冲区比例。这是因为当中继节点数据到达速率增大时,中继节点能贡献出的最大缓冲区比例降低,导致分配给每个用户的最优缓冲区比例也降低。

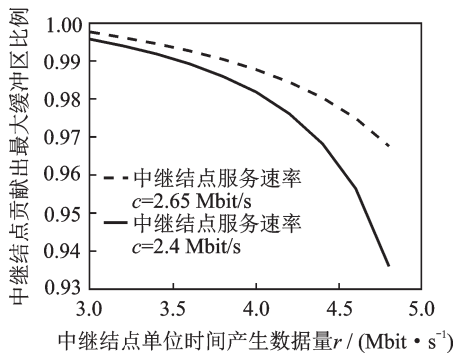


图 3 中继节点贡献出的最大缓冲区比例

Fig.3 Maximum buffer ratio contributed by relay node

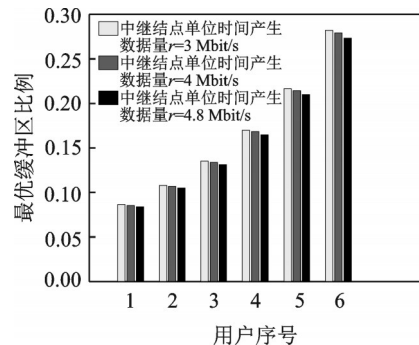


图 4 用户分配到的最优缓冲区比例

Fig.4 Optimal buffer ratio allocated by users

图 5 给出了用户 1 分配到的最优缓冲区比例随着中继节点单位时间产生的数据量  $r$  之间的关系。由图 3 的讨论可知,中继节点贡献出的缓冲区大小随着中继节点单位时间产生的数据量逐步降低,因此分配给用户的缓冲区大小也逐步降低。由图 5 可看出,用户 1 分配到的最优缓冲区比例随着  $r$  变大而变小。此外,图中还对比了中继节点服务速率不同时,用户 1 分配的最优缓冲区比例关系。在中继节点服务速率为 2.65 Mbit/s 时,用户 1 分配到的缓冲区比例大于服务速率为 2.4 Mbit/s 时分配到的缓冲区比例,这与图 3 的趋势一致。这是由于中继节点服务速率降低,中继节点能贡献的最大缓冲区比例下降,分配给用户 1 的缓冲区比例也下降。

图 6 给出了本文提出的最优分配法与均匀分配法的对比。本文提出的算法是以最小化缓冲区溢出

表 1 系统仿真参数

Table 1 System simulation parameters

参数	数值
$p_r$	0.4
$q_r$	0.5
$N$	6
$c/(\text{Mbit} \cdot \text{s}^{-1})$	2.4—2.65
$r/(\text{Mbit} \cdot \text{s}^{-1})$	3—4.8
$\xi_r$	0.05
$B/\text{kHz}$	83
$P_{tr}/\text{dBm}$	23
$a_i/(\text{Mbit} \cdot \text{s}^{-1})$	[4, 5]
$d_i/\text{m}$	[10, 30]
$g$	4
$\sigma_o^2/\text{dBm}$	-174
$X/M$	600

概率和为目标设计的,这里使用简单的均分缓冲区方法作为对比方法。从图5可以看出,使用本文提出的最优分配法,缓冲区溢出概率和在 $10^{-8}$ 量级,而均匀分配法则是在 $10^{-5}$ 量级,两者之间相差1000倍。此外,还可以看出,随着中继节点单位时间产生的数据量 $r$ 增多,缓冲区溢出概率和增加了,这是因为 $r$ 增多,中继节点贡献给用户使用的缓冲区变小了,用户在中继节点内部的溢出概率相应增加。均分缓冲区的方法只是简单的将缓冲区根据用户的个数进行均分,并未考虑用户由于距离中继节点远近,带来的传输速率差别,使得距离中继节点远的用户,在中继节点内部会产生积压,极易产生缓冲区溢出。而本文提出的算法是以最小化缓冲区溢出概率和为目标的,因此,最终给出的分配方案是考虑了用户距离中继节点的远近,距离中继节点较远的用户分配了较多的缓冲区,以减缓由于传输速率低带来的缓冲区积压问题,进一步降低了缓冲区溢出的概率。

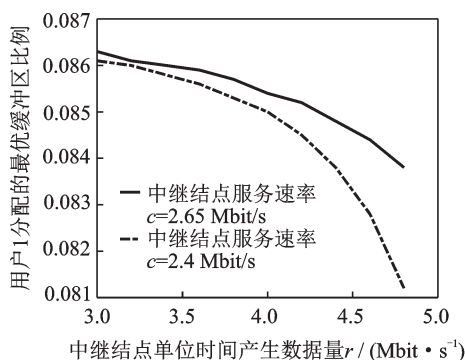


图5 用户1分配的最优缓冲区比例

Fig.5 Optimal buffer ratio allocated by user 1

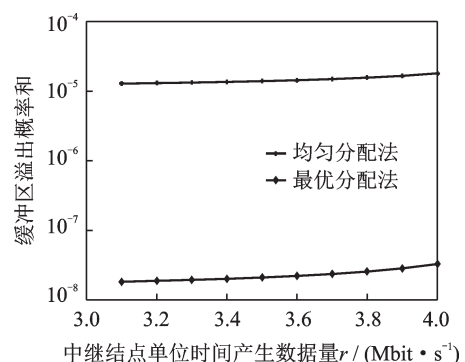


图6 最优分配方法与均匀分配方法的对比

Fig.6 Comparison between optimal allocation method and uniform allocation method

## 4 结束语

本文针对缓冲辅助中继网络中,中继节点缓冲区资源的分配进行研究。首先根据中继节点的数据产生和服务模型,以及中继节点能容忍的缓冲区溢出概率值,给出了中继节点能贡献出的最大缓冲区比例。接着根据用户的数据产生和服务模型,给出了用户的缓冲区溢出概率值。构建最小化所有用户缓冲区溢出概率值之和的优化问题,获得分配给各个用户的最优缓冲区分配方法。从仿真结果可以看出,最优缓冲区分配方法与中继节点单位时间产生的数据量有关系,也与用户距离中继节点的距离有关,并且本文提出的最优缓冲区分配方法优于均匀分配方法。

### 参考文献:

- [1] WANG X, CHEN M, TALEB T, et al. Cache in the air: Exploiting content caching and delivery techniques for 5G systems [J]. IEEE Communications Magazine, 2014, 52(2): 131-139.
- [2] LIU D, CHEN B, YANG C, et al. Caching at the wireless edge: Design aspects, challenges, and future directions[J]. IEEE Communications Magazine, 2016, 54(9): 22-28.
- [3] ZHOU S, GONG J, ZHOU Z, et al. GreenDelivery: Proactive content caching and push with energy-harvesting-based small cells[J]. IEEE Communications Magazine, 2015, 53(4): 142-149.
- [4] ZLATANOV N, SCHÖBER R, POPOVSKI P. Buffer-aided relaying with adaptive link selection[J]. IEEE Journal on Selected Areas in Communications, 2013, 31(8): 1530-1542.
- [5] ZLATANOV N, IKHLEF A, ISLAM T, et al. Buffer-aided cooperative communications: Opportunities and challenges[J]. IEEE Communications Magazine, 2014, 52(4): 146-153.



- [6] POLOCZEK F, CIUCU F. Service-martingales: Theory and applications to the delay analysis of random access protocols[C]// Proceedings of IEEE Conference on Computer Communications (INFOCOM). Hong Kong, China: IEEE, 2015: 945-953.
- [7] HU Y, LI H, CHANG Z, et al. Scheduling strategy for multimedia heterogeneous high-speed train networks[J]. IEEE Transactions on Vehicular Technology, 2017, 66(4): 3265-3279.
- [8] HU Y, LI H, CHANG Z, et al. End-to-end backlog and delay bound analysis for multi-hop vehicular ad hoc networks[J]. IEEE Transactions on Wireless Communications, 2017, 16(10): 6808-6821.
- [9] ZHAO L, CHI X, ZHU Y. Martingales-based energy-efficient D-ALOHA algorithms for MTC networks with delay-insensitive/URLLC terminals co-existence[J]. IEEE Internet of Things Journal, 2018, 5(2): 1285-1298.
- [10] LIU T, LI J, SHU F, et al. Quality-of-service driven resource allocation based on martingale theory[C]//Proceedings of IEEE Global Communications Conference: Communication QoS, Reliability and Modeling (Globecom CQRM). Abu Dhabi, United Arab Emirates: IEEE, 2018: 1-6.
- [11] LIU T, SUN L, CHEN R, et al. Martingale theory-based optimal task allocation in heterogeneous vehicular networks[J]. IEEE Access, 2019, 7: 122354-122366.
- [12] LIU T, CHANG Z, LI J, et al. Optimal buffer resource allocation in wireless caching networks[C]//Proceedings of 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). Cannes, France: IEEE, 2019: 1-5.

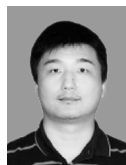
#### 作者简介:



刘婷婷(1982-),通信作者,女,副教授,研究方向:边缘计算、资源分配、博弈论、信息经济学, E-mail: liutt@njit.edu.cn。



陈开源(2001-),男,本科生,研究方向:中继网络、资源分配。



蒋诚智(1982-),男,高级工程师,研究方向:分布式机器学习、基于agent的机器学习。



余雨(1990-),男,讲师,研究方向:物理信道建模、分布式优化、网络资源调度与优化。



蒋姝(1988-),女,讲师,研究方向:通信系统仿真、电子信息工程。

(编辑:王静)