

面向软件定义无线接入网的高效协作内容缓存算法

张 涛¹, 李 强¹, 张继良², 张蔡霞¹

(1. 华中科技大学电子信息与通信学院, 湖北武汉 430074; 2. 兰州大学信息科学与工程学院, 甘肃兰州 730000)

摘 要: 为了缓解海量的移动业务数据与容量受限的无线接入网回传链路之间的矛盾, 本文提出一种面向软件定义无线接入网(SD-RAN)的协作内容缓存网络架构. 在宏蜂窝基站(MBS)的控制管理下, 小蜂窝基站(SBS)可以在存储单元有序存储一些高流行度的内容. 针对 SBS 存储单元空间受限问题, 进一步提出 SD-RAN 网络架构下的协作内容缓存算法. 该算法中, 每个 SBS 缓存空间被分割成两部分: (1) 用于存储全网流行度最高的公共内容以保证各小蜂窝小区本地命中率. (2) 用于存储流行度较高的差异化的内容以促进 MBS 内 SBS 之间的协作. 在此基础上, 解析推导具有最优平均内容获取开销的分割参数闭合表达式. 仿真结果表明该算法在不同系统参数条件下能显著降低 SD-RAN 的平均内容获取开销.

关键词: 软件定义无线接入网; 内容流行度; 协作缓存; 内容获取开销

中图分类号: TN915.65 **文献标识码:** A **文章编号:** 0372-2112(2017)11-2649-07

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2017.11.011

An Efficient Cooperative Content Caching Algorithm for Software-Defined Radio Access Networks

ZHANG Tao¹, LI Qiang¹, ZHANG Ji-liang², ZHANG Cai-xia¹

(1. School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China;

2. School of Information Science and Engineering, Lanzhou University, Lanzhou, Gansu 730000, China)

Abstract: In order to alleviate the contradiction between the massive mobile data traffic and the limited capacity of wireless backhaul links, a cooperative content caching algorithm is proposed for Software-Defined Radio Access Networks (SD-RAN), where the Macro-cell Base Station (MBS) equipped with centralized control logic and programmable interfaces, coordinates the content caching performed at the Small-cell Base Stations (SBSs). Moreover, in order to strike a balance between the local hit ratio and miss ratio, a cooperative caching algorithm is proposed by dividing the cache space of each SBS into two portions, one for storing the most popular contents that are the same across SBSs, one for storing unique contents to facilitate inter-SBSs cooperation. The optimal partition factor at which the average content provisioning cost of the SD-RAN is minimized is analytically obtained in a closed form.

Key words: software-defined radio access networks; content popularity; cooperative caching; content provisioning cost

1 引言

随着移动互联网行业的蓬勃发展, 全球移动业务数据量在过去的十年间暴增了 4000 倍, 其中 2015 年较 2014 年增幅为 74%^[1]. 然而与此同时, 2015 年全球蜂窝移动网络的平均容量(如下行连接速度)却只增长了 20%^[1]. 按照这样的发展趋势, 未来通信网络的发展将无法再满足呈指数增长的海量移动数据业务需求.

另一方面, 以视频为主的多媒体内容贡献了约 55% 的移动业务负载, 而且这一比例还在不断上升中, 预计在 2020 年将达到 75%^[1]. 现有研究表明, 视频内容的流行度服从 Zipf 分布^[2], 即少部分流行度高的内容会被反复的观看、下载. 以 Youtube 为例, 流行度排名前 10% 的视频文件贡献了约 80% 的业务负载^[3]. 这将导致同样的内容被反复的经由“接入网-核心网-远端内容服务器”进行请求、调用及传输. 为了解决大量重复

传输所导致的资源利用率低、网络拥塞等问题,近年来研究者们提出了针对不同网络应用环境的“以存储换带宽”的内容缓存算法.在内容中心网络(Content-Centric Networks,CCN)中,一般的,每个中间节点,如网关、路由器等,都具有一定的存储能力用于缓存流行内容^[4].现有研究者针对内容放置及替换策略展开了广泛的研究^[4-14].总体而言,距离终端用户较近的网络节点应缓存流行度更高的内容^[4].在进行缓存内容替换时,可以根据内容流行度的动态变化采用最近最少使用(Least Recently Used,LRU),最少频繁使用(Least Frequently Used,LFU),以及最近最多使用(Most Recently Used,MRU)等策略^[5].通过这些手段,可以将流行度高的热点内容放置在距离用户更近的网络节点,提高资源利用率及用户体验^[4,6,7].

受此启发,为了缓解海量移动业务负载与容量受限的蜂窝网回传链路之间的矛盾,提高资源利用率,文献[8,9]将内容缓存技术应用于无线接入网中.文献[10]针对社交网络提出相邻的移动设备可通过协作方式缓存内容,并获得相应收益的算法.针对异构网络,文献[11,12]提出部署大量具有高存储容量、低回传速率的帮助者(如无线接入点),可以协助缓存用户感兴趣的内容从而提高网络效率,如时延.这一类具有多存储单元、多用户、多内容的分布式缓存优化问题通常是一个NP-难问题^[11].为了提高缓存内容的命中率,降低系统的内容获取开销,研究者们提出了不同的迭代算法,然而这些算法一般难以实现内容的最佳放置,且系统优化问题难以获得精确的解析解^[11,12].本文提出一种面向软件定义无线接入网(Software-Defined Radio Access Networks,SD-RAN)的协作内容缓存算法.在SD-RAN中,利用宏蜂窝基站(Macro-cell Base Station,MBS)的集中控制功能,优化小蜂窝基站(Small-cell Base Stations,SBSs)存储单元中的缓存内容^[13,14].该缓存策略的基本思想包括:(1)将每个SBS的缓存空间分割成两部分;(2)一部分用于缓存全网中流行度最高的公共内容以保证小蜂窝小区的本地命中率;(3)另一部分用于缓存差异化的内容以促进宏蜂窝内SBS间的协作传输,降低请求内容的错失率;(4)通过选择合适的分割参数,优化SD-RAN的平均内容获取开销.

2 系统模型

2.1 网络架构

如图1所示,在SD-RAN中,通过软件定义的方式,可以将控制平面和数据平面进行分离、解耦^[17].具体的,MBS具有逻辑集中的控制功能和应用编程接口,主要负责提供控制覆盖,如网络拓扑管理、路由、资源调度等.MBS经由相应的业务网关、分组数据网(Packet Data Net-

work,PDN)网关和核心网与远端内容服务器进行连接.为了提高频率复用,MBS的控制范围被进一步划分为 N 个互不重叠的小蜂窝小区,并由相应的SBS负责提供数据覆盖,交付用户请求的内容.每个SBS部署了一个容量为 s_c 的存储单元,它可以在MBS的统一控制和协调下进行协作内容缓存,以减少对相同内容的反复重传.

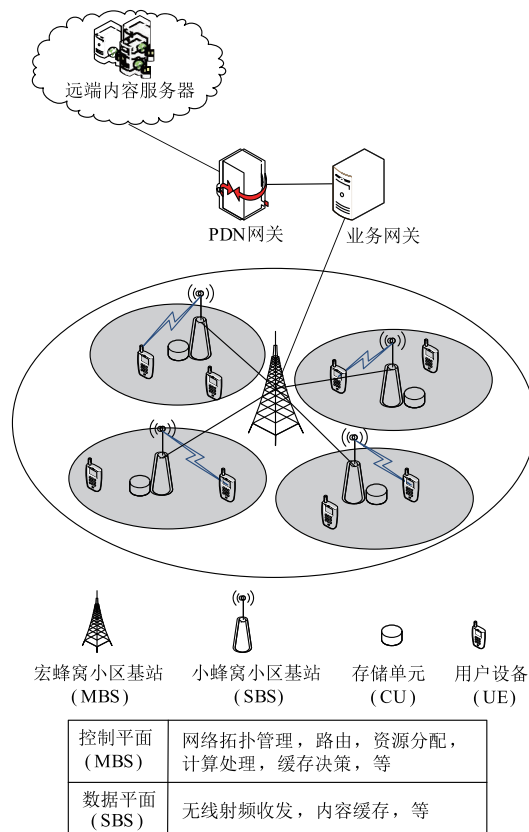


图1 具有内容缓存功能的软件定义无线接入网系统模型

2.2 内容流行度

流行度是用于描述一个内容平均被请求频率的指标.流行度越高,那么该内容就会被更频繁的请求、下载.考虑由 M 个不同内容构成的文件库,一般为多媒体(如视频、音乐等)内容文件^[2,11,18].为了便于对内容缓存问题的建模分析,假设每个内容大小均为1.在一定的时间内,这些内容的流行度变化较慢,例如,包含视频短片的新闻通常每2-3小时进行更新,新上线电影通常是每周进行更新,等.因此从无线接入网内容缓存及信息交付的时间维度来看,可以认为这些内容在一定的时间内服从静态的流行度分布^[18,19],并且系统可以确知这些内容流行度.随着大数据分析及内容预测技术的推进,用户内容流行度建模以及对内容流行度的估计和预测已变得更加准确、高效.对于这 M 个不同内容,可以按照其流行度降序排列为 $\{x_1, x_2, \dots, x_M\}$.现有研究表明,随机用户对文件库内容 x_m 发起请

求的概率服从 Zipf 分布^[2],即

$$p_m = \frac{1}{m^\alpha} \bigg/ \sum_{m=1}^M \frac{1}{m^\alpha} \quad (1)$$

其中 α 为 Zipf 指数,表征请求模型的偏斜程度: α 越大,用户请求越集中于排名靠前的流行内容;反之如果 α 越小,用户请求越分散. 根据不同的网络场景及用户需求, α 的取值变化范围一般为 $0.6 - 1.0$ ^[2,18].

2.3 用户请求及内容获取过程

在 SD-RAN 网络架构下,MBS 控制每个 SBS 进行协作内容缓存,同时跟踪、记录每个 SBS 具体缓存了哪些内容. 在此基础上,当有用户发出请求时,获取内容的流程如图 2 所示. (1)小蜂窝小区的本地 SBS 接收到内容请求后,首先查询该内容是否被缓存在本地存储单元中:如果是,SBS 直接发送内容给用户;否则,SBS 将该请求上报至 MBS. (2)MBS 接收到内容请求后,查询该内容是否被缓存在宏蜂窝小区内其他 SBS 的存储单元中:如果是,可以在 MBS 的协调下建立 SBS 之间的协作传输^[8,13,14];否则,MBS 进一步将用户请求上报至远端内容服务器以获取相应的内容.

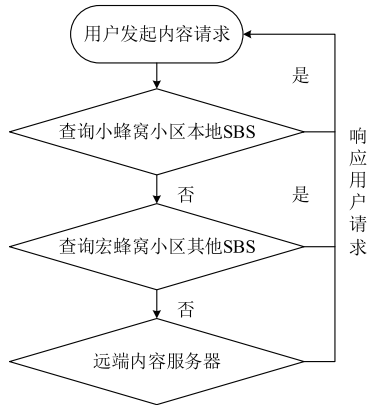


图2 用户内容请求及获取过程

3 缓存内容放置问题

由于 SBS 存储单元容量有限,不可能缓存文件库中的所有内容. 因此,需要对 SBS 的内容缓存策略展开研究. 为了建立准确的数学模型,首先定义一个内容放置变量 $U_m^n = \{0,1\}$, $n \in \{1, \dots, N\}$, $m \in \{1, \dots, M\}$, 用于表示内容 x_m 是否被缓存在第 n 个 SBS 的存储单元中. 对于宏蜂窝小区内的任一用户请求,定义 $p_m^{(n)}$ 为该请求发生在小蜂窝小区 n 中,且请求内容为 x_m 的概率. 因此内容 x_m 在宏蜂窝小区内的全局流行度可以表示为 $p_m = \sum_{n=1}^N p_m^{(n)}$, 由于均匀请求模型下各小区内用户请求频次相同,我们有 $p_m^{(n)} = \frac{1}{N} p_m$ ^[10]. 下面我们首先考虑两种

现有的内容缓存策略^[10,13,14].

3.1 贪婪缓存策略

为了优化小蜂窝小区的本地命中率,小蜂窝内的用户请求应尽可能在其本地小蜂窝小区内得到满足. 在这种情况下,目标函数为:

$$\begin{aligned} \text{Max}_{U_m^n} F_{\text{SBS}} &= \sum_{n=1}^N \sum_{m=1}^M U_m^n \cdot p_m^{(n)}, \\ \text{s. t.} \quad &\sum_{m=1}^M U_m^n \leq s_c, \quad \forall n \in \{1, \dots, N\} \end{aligned} \quad (2)$$

其中的约束条件 $\sum_{m=1}^M U_m^n \leq s_c$ 是为了确保每个 SBS 缓存的内容不超过其存储容量 s_c . 针对这一目标函数,最优的缓存策略为贪婪算法^[10,13,14],每个 SBS 尽可能存储相同的、最流行的内容,即:

$$U_m^n = \begin{cases} 1, & m \in \{1, \dots, s_c\}, \\ 0, & m > s_c, \end{cases}, \quad \forall n \in \{1, \dots, N\} \quad (3)$$

3.2 完全协作缓存策略

为了降低向远端内容服务器发送请求的概率,用户请求应尽可能的在宏蜂窝小区内部得到满足. 在这种情况下,目标函数为:

$$\text{Max}_{U_m^n} F_{\text{MBS}} = \sum_{n=1}^N \sum_{m=1}^M U_m^n \cdot p_m^{(n)} \quad (4a)$$

$$+ \sum_{n=1}^N \sum_{m=1}^M (V_m - U_m^n) \cdot p_m^{(n)} \quad (4b)$$

$$\text{s. t.} \quad \sum_{m=1}^M U_m^n \leq s_c, \quad \forall n \in \{1, \dots, N\}$$

其中变量 $V_m \in \{1,0\}$ 表示在宏蜂窝小区内 x_m 是否被至少一个 SBS 缓存. 式(4a)表示用户请求内容在其本地小蜂窝小区 SBS 的命中率,式(4b)表示用户请求内容在其他小蜂窝小区 SBS 的命中率.

针对这一目标函数,最优的缓存策略是完全协作算法^[10,13,14],即按照流行度高低每个 SBS 存储完全不同的内容,且每个内容只被缓存在一个 SBS 的存储单元中,从而使得宏蜂窝小区内缓存的内容样本数最大,即:

$$\sum_{m=1}^M U_m^n \leq 1, \quad \forall n \in \{1, \dots, N\} \quad (5)$$

通过对以上现有两种极端情况的简要分析,我们可以看到贪婪缓存策略中各 SBS 都缓存相同的最流行的内容,虽然能够优化各小蜂窝小区的本地内容命中率,但对于无法在本地 SBS 获得响应的内容就只能从远端内容服务器获取;而完全协作缓存策略则与之相反,其中各 SBS 缓存与其他 SBS 不同的流行内容,这样虽然能使宏蜂窝小区中缓存的内容样本数最大(降低从远端内容服务器中请求的概率),但是会损失小蜂窝小区的本地命中率,导致 SBS 之间频繁的协作传输并

造成不必要的内容获取开销. 由此可见, 这两种极端策略都存在局限或不足, 以这两种缓存策略作为基准情况, 下面我们对两种极端情况做出权衡提出一种高效的协作内容缓存算法.

4 高效协作内容缓存算法

4.1 协作内容缓存算法的目标与限制条件

为了便于分析, 我们分别定义 c_l, c_m, c_r 为从小蜂窝小区本地 SBS, 从宏蜂窝小区内其他 SBS, 以及从远端内容服务器获取单个内容的平均开销, 如带宽、时延、能耗等. 不失一般性, 我们有 $c_l < c_m < c_r$ [10].

为了优化 SD-RAN 中对任一用户请求的平均内容获取开销, 我们有目标函数:

$$\text{Min}_{U_m^n} \text{Cost}_{\text{avg}} = h_l \cdot c_l + h_m \cdot c_m + h_r \cdot c_r \quad (6)$$

$$\text{s. t.} \quad \sum_{m=1}^M U_m^n \leq s_c, \quad \forall n \in \{1, \dots, N\}$$

$$\text{其中,} \quad h_l = \sum_{n=1}^N \sum_{m=1}^M U_m^n \cdot p_m^{(n)} \quad (7)$$

$$h_m = \sum_{n=1}^N \sum_{m=1}^M (V_m - U_m^n) \cdot p_m^{(n)} \quad (8)$$

$$h_r = \sum_{n=1}^N \sum_{m=1}^M (1 - V_m) \cdot p_m^{(n)} \quad (9)$$

分别表示对于任一用户请求, 其在小蜂窝小区本地 SBS 的命中率, 在宏蜂窝小区内其他 SBS 的命中率, 以及在宏蜂窝小区内的错失率. 式(6)是一个常规的 0-1 多背包 NP 难问题 [11, 13]. 尤其针对大规模的文件库和多个小蜂窝小区的场景, 很难得到最优的缓存内容放置策略 U_m^n . 为了显著降低 SD-RAN 的内容获取开销, 下面我们将设计一种高效的协作内容缓存算法.

4.2 协作内容缓存算法设计

基于以上两种基准策略, 为了提高各小蜂窝小区的本地命中率, 每个 SBS 需要缓存相同的那部分最流行的内容. 但是由于缓存空间有限, 未被缓存的内容需要频繁的从远端内容服务器获取, 造成很大的开销. 另一方面, 通过建立 SBS 之间的协作缓存, 即每个 SBS 缓存完全不一样的内容, 可以最大化宏蜂窝小区内缓存的内容样本. 但是这样也会导致频繁的 SBS 间协作传输, 造成额外的开销. 因此, 为了优化 SD-RAN 的平均内容获取开销, 我们需要平衡小蜂窝小区命中率与宏蜂窝小区命中率. 鉴于此, 我们提出将每个 SBS 的缓存空间分割为公共部分及独有部分, 如图 3 所示. 另外, 通过算法 1 所示的分割内容缓存替换算法可实现这种缓存内容的最优放置, 该算法给出了在系统初始状态 (即缓存空间为空的状态) 下, 如何进行内容缓存和替换直至最终达到最优缓存空间分割的过程.

(1) 公共部分: 用于缓存流行度最高的那部分相同

内容, 以保证小蜂窝小区本地命中率.

(2) 独有部分: 用于存储各小蜂窝小区独有的内容, 以保证宏蜂窝小区内的内容命中率.

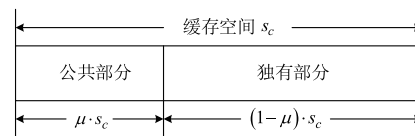


图3 缓存空间分割

算法 1 分割内容缓存替换算法

```

输入 SBSn 中用户请求内容 Onew
1: Omin1 = 公共部分流行度最低的内容
2: Omin2 = 独有部分流行度最低的内容
3: if Onew 从远端内容服务器获取
4:   if Onew 流行度 > Omin1 流行度
5:     Omin1 替换为 Onew
6:   else if Onew 流行度 > Omin2 流行度
7:     Omin2 替换为 Onew
8:   end
9: else if Onew 从 MBS 内其他 SBS 的公共部分获取
10:  if Onew 流行度 > Omin1 流行度
11:    Omin1 替换为 Onew
12:  end
13: end

```

如图 3 所示, 各个 SBS 的缓存空间被分割参数 μ 分为两部分, 其中 $0 \leq \mu \leq 1$. 因此, 在宏蜂窝小区中一共缓存了 $\Omega = \mu s_c + N(1 - \mu) s_c$ 个不同的内容. 通过这种方式, 在小蜂窝小区本地 SBS 公共部分的内容命中率可以表示为:

$$h_l^d = \sum_{n=1}^N \sum_{m=1}^{\mu s_c} p_m^{(n)} \quad (10)$$

另一方面, 我们假设独有的 $N(1 - \mu) s_c$ 个内容被均匀的缓存在各 SBS 的存储单元中. 那么小蜂窝小区本地 SBS 的独有部分的内容命中率可以表示:

$$h_l^u = \frac{\sum_{n=1}^N \sum_{m=\mu s_c+1}^{\Omega} p_m^{(n)}}{N} \quad (11)$$

类似的, 在宏蜂窝小区内其他 SBS 的独有部分的内容命中率为:

$$h_m^u = \frac{(N-1) \sum_{n=1}^N \sum_{m=\mu s_c+1}^{\Omega} p_m^{(n)}}{N} \quad (12)$$

根据式(10), 式(11), 可以得到 $h_l = h_l^d + h_l^u$, $h_m = h_m^u$, $h_r = 1 - h_l - h_m$. 因此, 为了优化 SD-RAN 中对任一用户请求的平均内容获取开销, 式(6)中的最优内容放置问题被转化为一个求解最优分割参数的问题, 即:

$$\begin{aligned}
\text{Min}_{\mu} \quad Cost_{avg} &= \left(\sum_{n=1}^N \sum_{m=1}^{\mu_s} p_m^n + \frac{\sum_{n=1}^N \sum_{m=\mu_s+1}^{\Omega} p_m^{(n)}}{N} \right) c_l \\
&+ \left(\frac{(N-1) \sum_{n=1}^N \sum_{m=\mu_s+1}^{\Omega} p_m^{(n)}}{N} \right) c_m \\
&+ \left(1 - \sum_{n=1}^N \sum_{m=1}^{\mu_s} p_m^{(n)} - \sum_{n=1}^N \sum_{m=\mu_s+1}^{\Omega} p_m^{(n)} \right) c_r \\
\text{s. t.} \quad &0 \leq \mu \leq 1 \quad (13)
\end{aligned}$$

为了便于分析,假设 $p_m^{(n)} = p_m^{(l)}$, $\forall n, l \in \{1, 2, \dots, N\}$, $n \neq l$, 即各小蜂窝小区中的用户在统计上有相同的请求频率和请求模式^[10]. 式(13)中的 $Cost_{avg}$ 可以进一步整理为:

$$\begin{aligned}
Cost_{avg} &= c_r - \sum_{m=1}^{\mu_s} p_m (c_r - c_l) \\
&- \sum_{m=\mu_s+1}^{\Omega} p_m \left(c_r - \frac{N-1}{N} c_m - \frac{c_l}{N} \right) \quad (14)
\end{aligned}$$

因此,式(13)的优化问题被转化为:

$$\max_{\mu} \left(\sum_{m=1}^{\mu_s} p_m (c_r - c_l) + \sum_{m=\mu_s+1}^{\Omega} p_m \left(c_r - \frac{N-1}{N} c_m - \frac{c_l}{N} \right) \right) \quad (15)$$

由于 $\sum_{m=1}^k m^{-\alpha} \approx \int_1^k m^{-\alpha} dm = \frac{k^{1-\alpha} - 1}{1-\alpha}$, 我们得到:

$$\sum_{m=1}^k p_m \approx \int_1^k \frac{m^{-\alpha}}{M} dm = \frac{k^{1-\alpha} - 1}{M^{1-\alpha} - 1} \quad (16)$$

将式(16)代入式(15), 然后对式(15)进行微分取导数并令其等于 0, 可以得到最优的分割参数为:

$$\mu^* \approx \frac{N}{\left(\frac{N \left(c_r - \frac{N-1}{N} c_m - \frac{c_l}{N} \right)}{c_m - c_l} \right)^{\frac{1}{\alpha}} + N - 1} \quad (17)$$

其中, 因为 $c_r > c_m$, 所以 $\left(\frac{N \left(c_r - \frac{N-1}{N} c_m - \frac{c_l}{N} \right)}{c_m - c_l} \right)^{\frac{1}{\alpha}} \geq 1$ 总是成立.

5 算法性能评估

下面我们将用数值仿真实验评估本文提出的协作内容缓存算法的性能. 不失一般性, 我们考虑由 $M = 5000$ 个不同内容构成的文件库, 这些内容服从 Zipf 模型静态流行度分布^[2, 18, 19]. $\beta = \frac{S_c}{M}$ 表示每个 SBS 存储单元容量与内容库的大小之比. 为了便于分析, 我们假设 $c_l = 1$, $c_r = 10$, 并且 $c_m = \tau c_r$, 那么 $0.1 < \tau < 1$. 如非单独说明, 我们假设 $N = 2$, $\beta = 0.3$, 用户的请求模式服从参数为 $\alpha = 0.7$ 的 Zipf

分布^[2, 18, 19]. 针对本文所优化的平均内容获取开销, 我们考虑任一用户发起的请求以及该被请求内容分别从本地 SBS 获取, 从宏蜂窝内其他 SBS 获取, 以及从远端内容服务器获取的所有可能情况下平均内容获取开销.

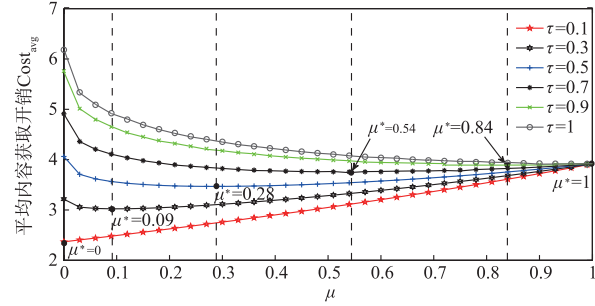
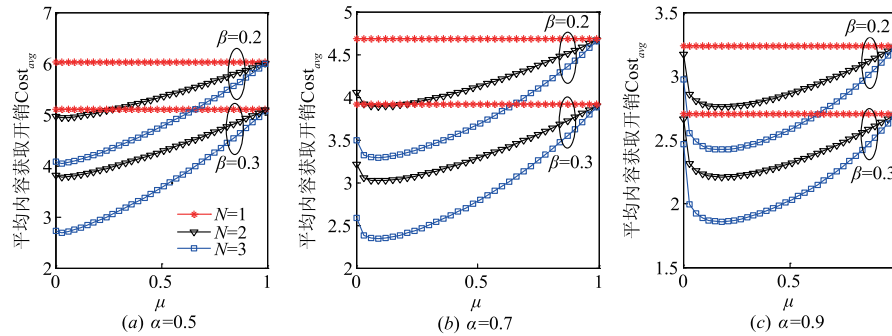


图4 分割参数 μ 对 SD-RAN 平均内容获取开销的影响

如图 4 所示, 我们首先研究分割参数 μ 的取值对 SD-RAN 开销性能的影响. 其中 $\mu = 0$ 对应完全协作缓存策略, 而 $\mu = 1$ 对应贪婪缓存策略. 针对不同的 τ 值 (即 SBS-SBS 间协作传输开销 c_m), 总是存在一个合适的分割参数使得 SD-RAN 的平均内容获取开销最小. 具体的, 针对 $\tau = 0.1$ ($c_m = c_l$) 的极端情况, 此时将内容缓存在本地 SBS 与将内容缓存在其他的 SBS 之间没有差别. 因此最优的缓存策略是完全协作算法, 即 $\mu^* = 0$. 随着 τ 的增加, 如 $\tau = 0.3, 0.5, 0.7, 0.9$, SBS-SBS 间协作传输的开销增加, 因此各 SBS 倾向于在公共部分缓存更多的流行内容, 这会导致更大的最优分割参数, 如 $\mu^* = 0.09, 0.28, 0.54, 0.84$. 另一方面, 针对 $\tau = 1$ ($c_m = c_r$) 的极端情况, 此时 SBS-SBS 间的协作传输与从远端内容服务器传输内容具有相同的开销. 这将完全抑制 SBS-SBS 间的协作, 因此最优的缓存策略是贪婪算法, 即 $\mu^* = 1$. 如图 4 所示, 对于任一固定的 τ 值, 选取最优分割参数 μ^* 可以将系统的平均内容获取开销降低 15% - 40%. 例如当 $\tau = 0.5$ 时, 最优分割参数 $\mu^* = 0.28$ 对应的平均内容获取开销 $Cost = 3.469$ 比最差情况 $\mu = 0$ 对应的平均内容获取开销 $Cost = 4.063$ 降低了约 15%.

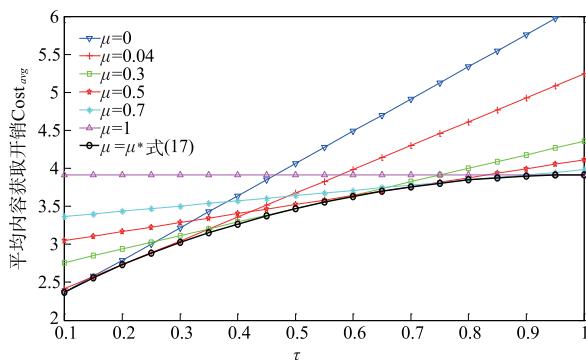
接下来我们讨论系统参数 α, β, N 对 SD-RAN 内容获取开销的影响. 当 $\alpha = 0.5$ 时, 由图 5(a) 可见, 随着 β 的增加, 如 $\beta = 0.2, 0.3$, 由于每个 SBS 具有更大的缓存空间能够存储更多的流行内容, 因此在其他参数保持不变的情况下, 能够有效降低 SD-RAN 总体的内容获取开销. 另一方面, 我们考虑宏蜂窝小区分别被划分为 $N = 1, 2, 3$ 个小蜂窝小区的情况. 由图 5(a) 可见, 当 $N = 1$ 时, 由于不存在 SBS-SBS 间的协作传输, 因此内容获取开销与分割参数 μ 的取值无关. 随着小蜂窝小区个数的增加, 如 $N = 2, 3$, 由于宏蜂窝内能够缓存更多的内容, 同时可以建立 SBS-SBS 间的协作传输, 因此相比 $N = 1$ 的情况, 可以获得更低的平均内容获取开销. 此外, 与图

图5 系统参数 α, β, N 对SD-RAN平均内容获取开销的影响

4类似,对 $N=2,3$,总是存在一个最优的分割参数 $0 < \mu < 1$ 使得SD-RAN的平均内容获取开销最小。

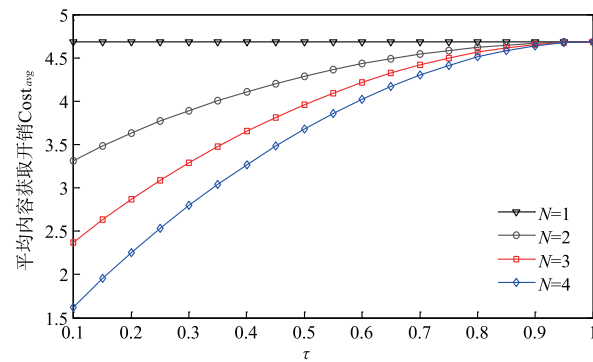
图5(b,c)分别显示了 $\alpha=0.7, 0.9$ 时的内容获取开销性能。从图中可以看到在其他参数保持不变的情况下(如 $\beta=0.2, N=2$),随着 α 的增加,流行度分布会更加集中于排名靠前的流行内容。这将会增加小蜂窝小区本地以及宏蜂窝小区内内容命中率,并减少从远端内容服务器获取内容的概率,因此在同样的网络拓扑及存储容量情况下能有效降低系统的内容获取开销。

接下来我们验证本文协作内容缓存算法的最优内容获取开销性能。在不同的 τ 值下,图6给出了不同分割参数能够实现的开销性能,如 $\mu=0, 0.04, 0.3, 0.5, 0.7, 1$ 。与图4类似,当 $\tau=0.1$ 时,由于 $c_m=c_r$,因此 $\mu=0$ 的完全协作算法可以取得最优的开销性能。然而随着 τ 的增加,SBS-SBS间的协作传输开销增大,此时 $\mu=0$ 的完全协作算法将导致很高的内容获取开销。另一方面,当 $\tau=1$ 时,由于 $c_m=c_r$,因此 $\mu=1$ 的贪婪算法可以取得最优的开销性能。然而随着 τ 的减小,SBS-SBS间的协作传输开销降低,此时贪婪算法将导致很高的内容获取开销。类似的,当 μ 分别取值 $0.04, 0.3, 0.5, 0.7$ 时,虽然能够在某些 τ 值的情况下显著降低内容获取开销,但是在另外一些 τ 值的情况下,却会导致更高的开销。而通过本文提出的协作内容缓存算法并选择式(17)中的最优分割参数 μ^* ,由图6可见,在不同的 τ 值情况下SD-RAN总能实现最优的平均内容获取开销。此外,在 $\tau=0.1, 1$ 的极端情况下,

图6 SD-RAN平均内容获取开销与 τ 的关系

可以看到采用最优分割参数 μ^* 最高能够将系统的平均内容获取开销降低40%左右。

图7给出了当 $\beta=0.2, \mu=\mu^*$ 时平均内容获取开销与参数 τ 以及小蜂窝小区数目 N 的关系。有如下两个方面:(1)随着 τ 的增加,由于SBS-SBS间协作传输开销增大,因此在其他参数不变的前提下,平均内容获取开销呈现增加的趋势。(2)随着 N 的增加,由于宏蜂窝内能够缓存更多的内容,同时可以建立SBS-SBS间的协作传输,因此可以实现更低的平均内容获取开销。

图7 最优分割参数 μ^* 情况下平均内容获取开销与 τ, N 的关系

6 结束语

本文提出一种面向SD-RAN的高效协作内容缓存算法。通过选择合适的分割参数,该算法可以实现小蜂窝小区本地命中率与宏蜂窝小区命中率之间的折中,从而显著降低SD-RAN系统的内容获取开销。在此基础上,解析推导并验证了最优分割参数的闭合表达式。数值结果表明,相比于现有的贪婪策略及完全协作策略,我们提出的最优分割参数能够将SD-RAN的平均内容获取开销降低40%左右。此外,进一步探讨了不同网络环境参数下协作内容缓存算法的性能。结果表明,在不同网络环境下,本文提出的协作内容缓存算法均能够有效降低用户获取内容的平均开销。

参考文献

- [1] Cisco visual networking index; Global mobile data traffic

- forecast update,2015-2020[R]. Cisco White Paper,2016.
- [2] Breslau L, Cao P, Fan L, et al. Web caching and Zipf-like distributions: evidence and implications[A]. INFOCOM 99 Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings [C]. New York, USA:IEEE,1999. 126 – 134.
- [3] Cha M, Kwak H, Rodriguez P, et al. I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system[A]. Proc ACM Internet Measurement Conference[C]. San Diego, USA:ACM,2007. 1 – 14.
- [4] Li J, Liu B, Wu H. Energy-efficient in-network caching for content-centric networking [J]. IEEE Commun Letters, 2013, 17(4):797 – 800.
- [5] Podlipnig S, Boszormenyi L. A survey of web cache replacement strategies[J]. ACM Computing Surveys, 2003, 35(4):374 – 398.
- [6] Ahleghagh H, Dey S. Video caching in radio access network: impact on delay and capacity[A]. Wireless Communications and Networking Conference[C]. Shanghai, China:IEEE WCNC,2013. 2276 – 2281.
- [7] Li X, Wang X, Xiao S, et al. Delay performance analysis of cooperative cell caching in future mobile networks [A]. 2015 IEEE International Conference on Communications [C]. London, UK:IEEE,2015. 5652 – 5657.
- [8] Pantisano F, et al. In-Network caching and content placement in cooperative small cell networks[A]. 2014 1st International Conference on 5G for Ubiquitous Connectivity [C]. Akaslompolo, Finland:IEEE,2014. 128 – 133.
- [9] Yang CC, Yao Y, Chen ZY, et al. Analysis on cache-enabled wireless heterogeneous networks [J]. IEEE Trans Wireless Commun. ,2016, 15(1):131 – 145.
- [10] Taghizadeh M, Micinski K, Ofria C, et al. Distributed cooperative caching in social wireless networks[J]. IEEE Trans. Mobile Computing, 2013, 12(6):1037 – 1053.
- [11] Shanmugam K, et al. FemtoCaching: Wireless content delivery through distributed caching helpers[J]. IEEE Trans Inf Theory, 2011, 59(12):8402 – 8413.
- [12] Song J, Song H, Wan C. Optimal caching placement of caching system with helpers[A]. 2015 IEEE International Conference on Communications[C]. London, UK:IEEE, 2015. 1825 – 1830.
- [13] Wang X, Chen M, Taleb T, et al. Cache in the air: Exploiting content caching and delivery techniques for 5G systems[J]. IEEE Commun Mag, 2014, 52(2):131 – 139.
- [14] Wang X, Li X, Leung VCM, et al. A Framework of cooperative cell caching for the future mobile networks[A]. Hawaii International Conference on System Sciences[C]. Kauai, USA:2015. 5404 – 5413.
- [15] Ge X H, Yang B, Ye J, et al. Spatial spectrum and energy efficiency of random cellular networks [J]. IEEE Trans Commun. ,2015, 63(3):1019 – 1030.
- [16] Ge X H, Tu S, Han T, Q Li, et al. Energy efficiency of small cell backhaul networks based on Gauss-Markov mobile models[J]. IET Networks, 2015, 4(2):158 – 167.
- [17] Mohamed A, Onireti O, Imran M A, et al. Control-data separation architecture for cellular radio access networks: a survey and outlook [J]. IEEE Commun, Surveys Tuts, 2016, 18(1):446 – 465.
- [18] Shi L, Gu Z M, Wei L, et al. An applicative study of Zipf's law on web cache[J]. International Journal of Information Technology, 2006, 12(2):49 – 58.
- [19] Golrezaei N, Shanmugam K, Dimakis A G, et al. FemtoCaching: Wireless video content delivery through distributed caching helpers [A]. 2012 Proceedings IEEE INFOCOM[C]. Orlando, USA:IEEE,2012. 1107 – 1115.

作者简介



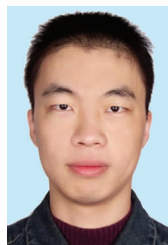
张 涛 男,1987 年生,硕士生,研究方向为下一代无线通信系统、内容缓存、软件定义网络、D2D 通信。

E-mail:tao_n_zhang@hust.edu.cn



李 强(通信作者) 男,1984 年生,华中科技大学副教授,硕士生导师,主要研究方向涉及下一代移动通信、无线协作通信、软件定义网络、认知无线电、网络编码、信息论等。

E-mail:qli_patrick@hust.edu.cn



张继良 男,1985 年出生,兰州大学副教授,目前研究的方向集中于高能效无线通信技术,多天线技术,无线定位技术和无线电波传播特性建模等。

E-mail:zhangjiliang@lzu.edu.cn



张蔡霞 女,1992 年生,硕士生,研究方向为下一代无线通信系统、软件定义网络、内容缓存。

E-mail:cxzhang_rosy@hust.edu.cn