智能化高铁车载缓存

Intelligent High-Speed Railway Onboard Caching

崔新雨 /CUI Xinyu[1,2,3]，刘玲 /LIU Ling[1,2,3]，周一青 /ZHOU Yiqing[1,2,3]，潘振岗 /PAN Zhengang[4]

（1. 中国科学院计算技术研究所，中国 北京 100190；
2. 移动计算与新型终端北京市重点实验室，中国 北京 100190；
3. 中国科学院大学，中国 北京 100049；
4. 北京紫光展锐通信技术有限公司，中国 北京 100190）

(1. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China;
2. Beijing Key Laboratory of Mobile Computing and Pervasive Device, Beijing 100190, China;
3. University of Chinese Academy of Sciences, Beijing 100049, China;
4. Beijing Unisoc Technologies Co., Ltd., Beijing 100190, China)

摘要：通过对高铁车载缓存系统的概念、在工作流程的各个阶段遇到的问题以及相应的智能化解决方案进行详细介绍，指出高铁车载缓存可以有效地减少移动设备与路边基站直连通信的次数，提升乘客的内容服务体验。但是，高铁车载缓存系统在实际的应用中仍然存在诸多问题，具体包括高铁与路边基站的协同缓存、乘客请求数据缺失下的请求规律挖掘，以及通信、计算、缓存资源的联合管控。

关键词：高铁通信；车载缓存；智能

Abstract: An overview of the basic idea of the high-speed railway onboard caching system is presented. The main challenges at each stage of the workflow and the corresponding intelligent solutions are detailed. It is pointed out that the high-speed railway onboard caching can effectively reduce the number of direct communications between mobile devices and roadside base stations, improving the quality of service of passengers. However, there are still many problems to be solved in practice, including the collaborative caching of high-speed railway and roadside base stations, the mining of request features when passenger request data is missing, and the joint management and control of communication, computing, and caching resources.

Keywords: high-speed railway communications; onboard caching; intelligence

DOI：10.12142/ZTETJ.202104002
网络出版地址：https://kns.cnki.net/kcms/detail/34.1228.TN.20210722.1429.006.html

网络出版日期：2021-07-23
收稿日期：2021-06-20

近年来，高铁凭借其较高的行驶速度、舒适的乘车体验等优点，逐渐成为人们优选的远途出行方式。同时，随着移动通信技术的飞速发展[1-6]，在搭乘高铁时，乘客出于娱乐、办公等目的，可以通过手机等移动通信设备发起对视频等多媒体业务内容的请求[7]。不同于用户在静止或低速移动状态下通过移动通信设备与基站建立通信连接，高铁场景下移动通信设备与路边基站的直连通信面临着3个主要问题[8]。

（1）封闭式金属车厢引起的信号衰减。高铁车厢多采用金属全封闭式的结构，无线信号在穿过金属车厢时会受到较大的穿透损耗。

（2）高速移动引起的连接不稳定。由于路边基站的覆盖范围有限，高铁的高速移动将引起频繁的基站切换，从而影响移动设备与路边基站连接的稳定性。

（3）高速移动引起的多普勒频移。高铁与路边基站之间的相对运动会引起严重的多普勒频移，从而会影响无线信号的正确接收与解调。尤其是4G/5G采用的正交频分复用技术，对载波频率偏移十分敏感。

在高铁通信场景下，面向具有较大带宽、较低时延需求的多媒体类业务，移动通信设备与路边基站直连的方式通常难以满足业务的服务质量需求，乘客在观看视频时将面临频繁播放卡顿、加载慢等问题，观看服务体

基金项目：国家自然科学基金青年科学基金（61901452）；国家重点研发计划（2020YFB1806900）

验受到严重影响。

为了带给乘客更优的服务体验，可以在高铁上应用车载缓存技术[9-10]。通过在高铁上搭载缓存服务器，并在服务器中预先缓存乘客可能会请求的内容，使得乘客请求有一定的概率被车载缓存服务器就近响应，避免乘客移动通信设备与路边基站的直连通信。本文将结合现有研究，概述高铁车载缓存系统，分析高铁车载缓存系统面临的主要问题并总结相应的智能化解决方案，同时讨论高铁车载缓存系统未来面临的挑战，为车载缓存技术在高铁中的智能应用提供参考。

1 高铁车载缓存系统

本节中，我们将从系统架构、用户请求规律、系统工作流程的角度对高铁车载缓存系统进行概述。

1.1 系统架构

高铁车载缓存系统架构由 4 部分组成，分别是高铁车载缓存服务器、高铁车载中继通信系统、路边接入网和云端计算中心，如图 1 所示[12-13]。高铁搭载了一定容量的缓存服务器，负责存储内容，调度用户请求，为乘客提供就近的内容接入服务。高铁车载中继通信系统由无线接入点和中继站组成，无线接入点分布在各节车厢，中继站安装在高铁顶部，无线接入点与中继站之间通过有线连接。乘客、车载缓存服务器分别通过无线、有线的方式接入车载中继通信系统，通过中继通信的方式与路边基站建立连接，以避免信号穿过车厢所产生的穿透损耗，获得更优的信号质量。此外，若乘客请求的内容被车载服务器所缓存，缓存内容将直接经车载中继通信系统发送给乘客。路边接入网由部署在铁路沿线的路边基站组成，为乘客提供移动网络接入服务。云端计算中心部署了内容源服务器，源服务器中存储了网络中的所有内容。云端计算中心经核心网与路边基站相连，当乘客请求的内容未被车载缓存服务器缓存时，将由源服务器为乘客提供所需内容。

1.2 用户请求规律

用户对内容的请求具有一定的统计规律，这些统计规律是高铁车载缓存系统工作时的重要参考依据。用户整体发起的内容请求通常服从 ZipF 定律，即少量内容吸引了大部分用户对其发起请求，而剩余的大量内容则很少受到用户们的关注[11]。这部分少量内容也被称为热点内容。研究中常用流行度表示内容受用户请求的频次，热点内容具有较高的流行度，非热点内容具有较低的流行度。在较短的时间内（例如一天），内容流行度并不会发生显著的变化。

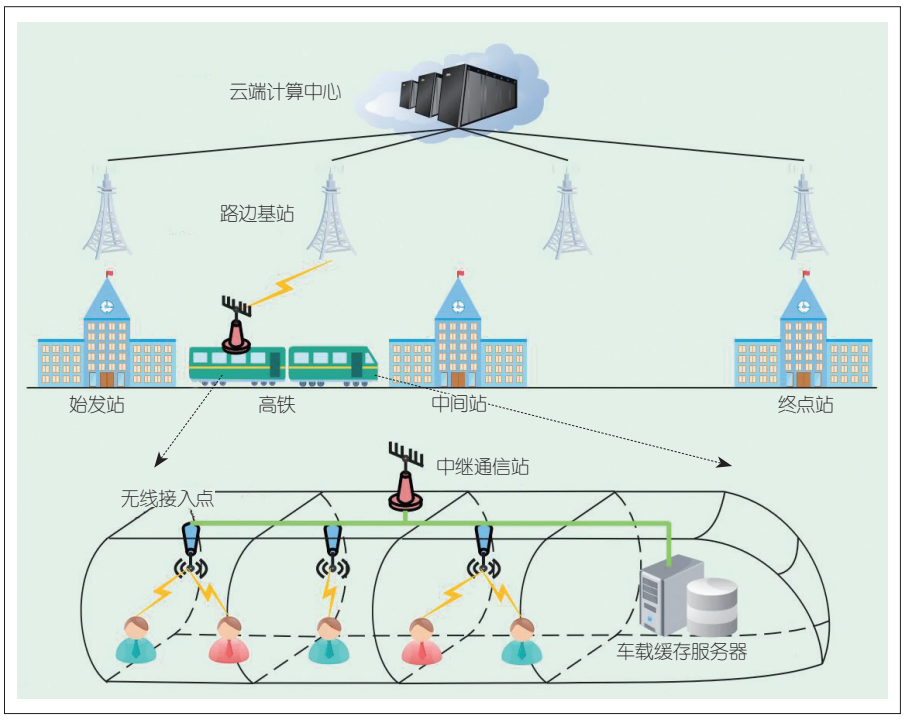
对于用户个体而言，常用喜好分布对用户过去一段时间的请求规律进行描述，喜好分布代表了用户对各个内容的喜好程度。喜好分布是极具个性化的，某一用户个体可能会对热点内容完全不感兴趣，不同用户的喜好分布也可能完全不同。由于喜好分布是根据用户以往的请求历史记录计算生成的，所以喜好分布仅能在一定程度上反应用户未来的请求倾向，用户未来的请求行为还会受到推荐系统、社交关系等多重因素的影响。

统计技术、智能化预测技术等可以量化内容的流行度和用户的喜好分布，为缓存内容的放置提供参考。在高铁场景下，得益于可提前获取乘客的班次信息，车载缓存系统可在列车启程前统计乘客的喜好分布，推算内容流行度，针对性地缓存乘客群体可能会请求的内容。

1.3 系统工作流程

高铁车载缓存系统的工作流程分为 3 步，分别是内容放置、请求调度与内容分发、内容更新[13]。

（1）内容放置。高铁驶离始发站



▲图 1 高铁车载缓存系统架构

之前，车载缓存服务器可以获得搭乘本次高铁的乘客信息，并统计乘客们的喜好分布、乘客群体的内容流行度等能够反映乘客请求倾向的关键参考特征，然后根据这些关键参考特征，选择乘客可能会在乘车过程中请求的内容存入服务器。

（2）请求调度与内容分发。车载缓存服务器具备调度用户请求的功能。乘客在搭乘高铁的过程中发起对内容的请求后，车载缓存服务器将检测用户请求的内容是否被缓存，并根据检测结果对用户请求进行调度。如果车载缓存服务器缓存了乘客请求的内容，车载缓存服务器将通过中继通信系统向乘客分发缓存内容。如果乘客请求的内容未被车载服务器缓存，车载缓存服务器则将乘客请求转移至云端计算中心的内容源服务器处，由源服务器向用户分发请求内容。

（3）内容更新。高铁由始发站驶向终点站的过程中将途径多个中间站，乘客的流动使得缓存内容也需要进行相应的调整。高铁到达下一站点之前，车载缓存服务器将根据乘客在下一站点的流动状况对缓存内容进行更新，用上车乘客可能会请求的内容替换下车乘客可能会请求的内容。

2 高铁车载缓存面临的主要问题

高铁车载缓存系统在工作过程中会受到多方面的制约，影响缓存资源发挥效用。本节将分析高铁车载缓存面临的主要问题，包括内容放置阶段面临的缓存容量受限、请求调度阶段面临的请求随机性大、内容分发阶段面临的无线资源受限问题以及业务场景支持单一的问题。

2.1 内容放置阶段面临缓存容量受限的问题

云端计算中心的源内容服务器存储了动辄拍字节级的海量内容，但车载缓存服务器的容量通常在太字节级左右，只有较小比例的内容能被高铁车载缓存服务器缓存。通常，人们使用缓存命中率作为衡量缓存资源效用的指标。缓存命中率等于用户请求被缓存服务器满足的次数占用户请求总次数的比例，缓存命中率越高代表缓存资源越能得到充分利用。缓存容量受限增加了内容放置阶段的难度，选择哪些内容进行缓存直接决定了缓存命中率的高低。

2.2 请求调度阶段面临请求随机性大的问题

从提升乘客请求服务体验的角度出发，如果乘客的请求可以直接被车载缓存服务器满足，那么乘客可以获得低时延的请求服务体验，从而避免所提请求被调度至云端计算中心的情况。但是，乘客的请求行为在个人喜好、推荐系统、社交关系等多重因素的影响下较为随机，乘客根据个人喜好做出的选择，可能会受到推荐系统的影响而临时改变。此外，受社交关系的影响，乘客也可能会请求与个人喜好不匹配但受家人推送的内容。乘客请求的随机性，增加了请求被调度至云端计算中心的概率，降低了请求平均时延的性能。

2.3 内容分发阶段面临无线资源受限的问题

乘客的移动通信设备与高铁中继通信系统之间通过无线的方式建立连接。由于无线中继通信系统频谱资源有限，当较多的乘客发起请求，特别是请求超高清视频等大带宽业务内容时，容易引起网络拥塞，从而影响分发时延。即便在缓存命中率较高的情况下，如果网络拥塞问题不能解决，同样会影响到用户的服务体验。缓存资源与通信资源需要进行协同才能最大限度地发挥缓存资源的效用。

2.4 业务场景支持单一的问题

缓存服务器可以对多种业务提供缓存服务，如小文件加载、大文件下载、音视频点播等。小文件主要来自于各类门户网站的 html、js、jpg 等网页素材，使用缓存服务器对网页的小文件进行加速，可以减少连接的建立时间、首包时间等，优化网页的加载时间；大文件是指大小在 20 MB 以上的文件，例如应用安装包、应用更新包等，通过缓存服务器的加速，可以提升下载速度，减少下载总时间；音视频点播业务来自于各类音视频网站，缓存服务器通过对 MP4、Flash 视频（FLV）等主流的视频格式进行缓存，可以降低音视频的卡顿率，优化首播时间等。目前高铁缓存服务器仅对音视频点播业务进行了较好的支持，但仍需对小文件加载和大文件下载业务进行进一步优化，以提升乘客在浏览门户网站、下载大文件时的体验。

3 智能化高铁车载缓存关键技术

智能技术的发展为上述部分问题的解决提供了新的思路，本节分别针对高铁车载缓存系统在内容放置、请求调度、内容分发阶段面临的主要问题，总结相应的智能化解决方案，包括基于深度学习的缓存内容放置、基于内容推荐的用户请求调度和基于编码缓存的缓存内容分发。

3.1 基于深度学习的缓存内容放置

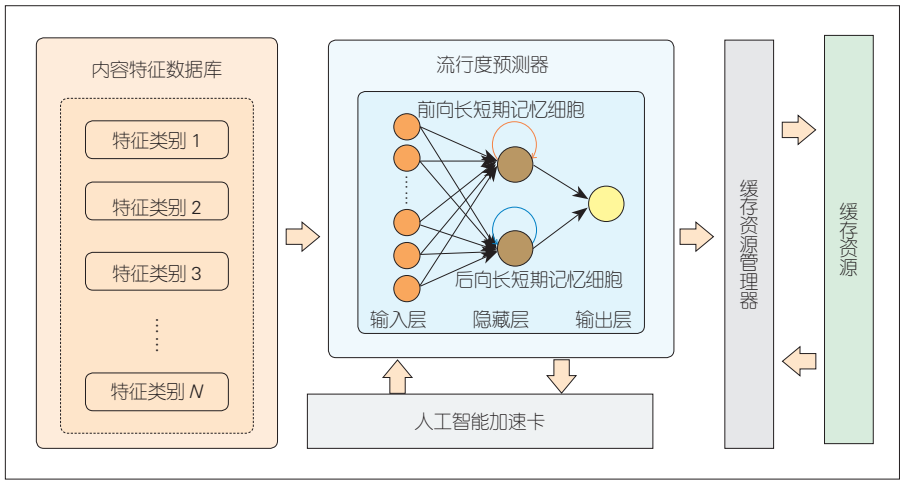
缓存容量的受限为缓存内容的放置带来了一定的挑战，选择哪些内容进行缓存将直接影响缓存命中率。为了在缓存容量受限的条件下尽可能地发挥缓存资源的效用，可以使用深度学习技术提高内容流行度的预测精度，

从而为缓存内容的放置决策提供更加可靠的参考信息[14]。

文献[14]在缓存服务器中引入了深度学习模块，如图2所示。缓存服务器除了具备基本的缓存资源、缓存资源管理器，还额外引入了与深度学习相关的内容特征数据库、流行度预测器、人工智能加速卡。缓存资源是内容存储的物理媒介，缓存资源管理器根据内容流行度的预测结果负责内容的放置、删除与更新。内容特征数据库负责采集乘客在过去一段时间的内容请求特征、乘客搭乘高铁出行时的内容请求特征、各内容的类别特征等，构建训练神经网络所需的数据集。流行度预测器内置流行度预测模型，预测模型可以从内容特征数据库中提取数据训练深度神经网络，并输出深度神经网络预测的内容流行度，供缓存资源管理器进行决策。人工智能加速卡负责加速深度神经网络的训练，使得流行度的预测精度可以尽快收敛。根据数据集规模的大小以及深度神经网络模型的复杂程度，高铁可以动态地配置缓存服务器中的人工智能加速卡资源，避免加速资源浪费或不足的问题。

缓存服务器内置的流行度预测模型采用双向长短期记忆网络，该网络由两个循环神经网络的变体组成，分别是双向循环神经网络和长短期记忆网络。循环神经网络是一种典型的深度学习模型，适合处理序列信息。例如，对内容流行度在过去一段时间的变化进行采样而得到的流行度序列信息。双向循环神经网络由两个循环神经网络组成，两个循环神经网络分别采用从前向后、从后向前的方式处理序列信息。在循环神经网络的基础上，长短期记忆网络对网络单元进行了改进，使用具备门系统、更多参数的细胞替代原有的网络单元，使得其在处理长信息序列方面更有优势。双向长短期记忆网络在双向循环神经网络的基础上，将隐藏层的每个单元替换为长短期记忆细胞。考虑到缓存服务器的计算能力和内容流行度预测精度需求，可以对双向长短期记忆网络的隐藏层层数进行一些动态的调整。

缓存服务器在进行缓存内容放置时，将依次经历数据预处理、基于深度学习的流行度预测、内容缓存3个过程。由于误操作等因素，个别用户产生的流行度序列信息存在异常，不能反映整体的流行度变化趋势，因此在训练神经网络之前需要对数据进行预处理，剔除异常的流行度序列信息。预处理后的数据被输入至双向长短期记忆网络，网络经过一定次数的迭代训练，待预测精度、交叉熵损失函数等反应训练效果的指标达到阈值后，将停止训练，输出对未来内容流行度的预测值。缓存资源管理器根据内容流行度的预测结果，选择流行度最高的内容存入缓存资源。



▲图2 深度学习赋能的缓存服务器

在高铁车载缓存的场景下，可以在车载缓存服务器中引入基于深度学习的内容流行度预测模块。车载缓存服务器根据提前获知的列车乘客搭乘信息，获取乘客的历史请求记录，再通过深度学习技术对乘客群体的内容流行度进行精准预测，从而针对性地缓存乘客群体可能请求的内容。这样可以在缓存容量受限的条件下，提高缓存资源的效用。

3.2 基于内容推荐的用户请求调度

在请求调度阶段，乘客请求被车载缓存服务器满足得越多，平均请求时延就越低；但较大的乘客请求随机性使得大量的乘客请求被调度至云端计算中心，影响了车载缓存资源的效用，以及平均请求时延的性能。考虑到用户的请求易受推荐系统的影响，文献[15]和[16]提出通过向用户推荐缓存内容来降低其请求的随机性，从而提升用户请求缓存内容的概率。上述文献的不同之处在于：文献[15]提出了一种硬推荐机制，在用户发起请求前，向用户推荐缓存内容；而文献[16]提出了一种软推荐机制，待用户发起请求后，如果用户请求的不是缓存内容，再向用户推荐服务体验更佳的缓存内容，并询问用户是否改变自己的请求，具体如图3所示。

文献[15]首先量化建模了用户对内容的请求概率。在模型中，用户对内容的请求概率大小受用户喜好分布与推荐系统的双重影响，且不同用户受推荐系统影响的程度不一。对于推

▲图 3 缓存内容软推荐

荐给用户的内容，由于该内容获得了一定的曝光度，所以用户请求推荐内容的概率会获得一定幅度的提升；对于未推荐给用户的内容，由于用户被推荐内容所吸引，因此用户请求未推荐内容的概率会有一定幅度的下降。

通过上述模型，可以估算推荐系统向用户推荐缓存内容后，用户对缓存内容的请求概率。在用户发起请求前，选择推荐后请求概率最高的几个缓存内容并推荐给用户，以增加缓存内容的命中率。值得注意的是，推荐系统的初衷是帮助用户发现感兴趣的内容，如果缓存内容与用户喜好相差较多，则不会被推荐给用户。

在文献[16]提出的软推荐机制下，推荐系统在用户发起请求前向用户推荐内容，但不考虑推荐的内容是否被缓存。待用户发起请求后，缓存服务器检查用户请求的内容是否被缓存，如果用户请求的内容未被缓存，则再通过推荐系统进行二次推荐，向用户推荐符合用户喜好且缓存了的内容。在二次推荐时，推荐系统将对用户做出提示，请求二次推荐的内容可以获得更好的服务体验，用户可以根据需求自行决定是否改变原有的请求，请求二次推荐的内容。

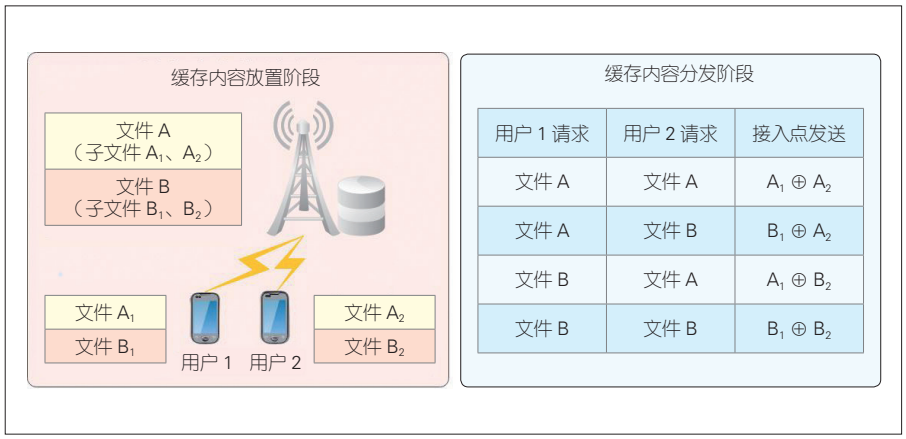
在高铁场景下，可以利用智能化的推荐系统，使用硬推荐或软推荐机制向乘客推荐车载缓存服务器中缓存的内容，以避免用户的请求被调度至云端计算中心，从而提高车载缓存服务器的服务效率。

3.3 基于编码缓存的缓存内容分发

对高铁实际网络数据的分析表明，其每分钟的活跃乘客数量保持在 100 人左右，约占 8 编组列车满员人数的 1/5，每分钟的活跃乘客数量峰值可达 200 人左右 [17]。由于无线中继通信系统的频谱资源有限，当较多的乘客发起请求时，系统容易引起网络拥塞。文献 [18] 提出了一种基于用户与接入点协同的编码缓存机制，通过在流量负载低峰期向用户本地的存储资源中缓存特定的子文件，接入点可以在流量负载高峰期通过异或编码多播的方式进行分发。相比于非编码缓存，编码缓存可以大幅减少高峰期的流量负载，降低对有限频谱资源的要求。

为了阐述编码缓存的机理，图 4 给出了一个编码缓存的简单例子。其中，接入点（如基站）缓存了两个大小为 F 的文件 A、B，分别被分割为两个大小相同的子文件 A_1、A_2、B_1、B_2；用户 1、2 具有相同容量的本地缓存资源，均可以缓存两个子文件，接入点与用户之间通过无损无线链路建立连接。在流量负载低峰期（如深夜）对子文件进行缓存放置，如图 4 中左图所示，用户 1 缓存子文件 A_1、B_1，用户 2 缓存子文件 A_2、B_2。当用户在流量负载高峰期提出文件请求时，例如用户 1 请求文件 A，用户 2 请求文件 B，系统将进行子文件的分发。如图 4 中右图所示，由于用户 1 缺少子文件 A_2，用户 2 缺少子文件 B_1，那么接入点可以将子文件 A_2 和 B_1 进行异或操作形成编码子文件 $A_2 \oplus B_1$，然后通过多播的方式发送给用户 1、2。用户 1 将本地缓存的子文件 B1 与编码子文件 $A_2 \oplus B_1$ 再次进行异或操作，即可得到所需的子文件 A_2；用户 2 将本地缓存的子文件 A_2 与编码子文件 $A_2 \oplus B_1$ 再次进行异或操作，即可得到所需的子文件 B_1。由于异或操作并不改变文件大小，所以上述分发过程中产生的数据量为 F/2。而在非编码缓存机制下，接入点需要分别向用户 1 发送 A_2，向用户 2 发送 B_1，分发过程中产生的数据量为 F。可见，编码缓存较非编码缓存可以降低分发时的流量负载。

文献 [18] 从理论的角度证明了编码缓存机制的优越性。在用户数量为 K、文件总量为 N、用户本地缓存容量为 M 的情况下，编码缓存机制下的流量负载是非编码缓存机制下流量负载的 $1/(1+KM/N)$ 倍，理论最优流量负载下界最多是编码缓存机制下的流量负载的 1/12。相比于非编码缓存机制够能大幅降低流量负载，采用编码缓存机制能够有效降低对频谱资源的需求。



用户 1 请求	用户 2 请求	接入点发送
文件 A	文件 A	$A_1 \oplus A_2$
文件 A	文件 B	$B_1 \oplus A_2$
文件 B	文件 A	$A_1 \oplus B_2$
文件 B	文件 B	$B_1 \oplus B_2$

▲图 4 编码缓存简例

在高铁场景下，可以在乘客乘车前按照一定的规则，在乘客的移动通信设备上缓存子文件。乘客在乘车的过程中，通过异或编码多播的方式获得缺失的子文件，以降低中继通信系统在高峰期的流量负载压力，避免网络拥塞。

4 未来挑战

高铁车载缓存相关的研究虽然已取得了一定的进展，在实际中也获得了一定的应用，但仍然存在许多需要解决的问题。本节对其中的 3 个主要问题进行讨论，包括高铁与路边基站的协同缓存，乘客请求数据缺失下的请求规律挖掘，以及通信、计算、缓存资源的联合管控。

（1）高铁与路边基站的协同缓存

缓存服务器可以部署在高铁中，也可以部署在铁路沿线的路边基站中。当高铁车载缓存服务器无法满足乘客请求时，可以将乘客请求调度至路边基站的缓存服务器中。如果路边基站的缓存服务器无法满足用户请求，再将用户请求调度至云端计算中心。通过这种两级缓存的方式，可以增加乘客就近获得请求内容的概率，降低乘客平均请求时延。但是，由于高铁是高速移动的，且在各个路边基站覆盖范围内的停留时间较为短暂，这导致路边基站需要在较短的时间内完成缓存内容的分发。因此，高铁与路边基站如何协同缓存内容的放置，在有限的时间内完成缓存内容的顺利分发是未来需要解决的一个难点问题。

（2）乘客请求数据缺失下的请求规律挖掘

缓存内容的放置取决于乘客的请求规律，而乘客请求规律的挖掘需要乘客请求历史数据的支持。但是，大部分的请求历史数据受到隐私协议的保护，因此车载缓存服务器所能获得的历史请求数据是不完整的。同时，部分乘客的请求历史数据较少，即便隐私授权后可以得到全部的请求历史数据，也不足以支持对乘客请求规律进行挖掘。因此，还需进一步探索智能化技术在高铁车载缓存中的应用，在乘客请求历史数据集较小情况下，提高请求规律挖掘精度。

（3）通信、计算、缓存资源的联合管控

高铁车载缓存服务器可以为传统的多媒体类业务提供良好的支持，但对虚拟现实、增强现实等新兴的计算密集型业务而言，仅有缓存服务器的支持是不够的，还需要在高铁上搭载移动边缘计算服务器进行协助。不同的业务对通信、计算、存储资源的需求是不同的。为了满足不同业务的服务质量需求，人们需要设计高效的资源分配机制，从而对高铁车载通信、计算、缓存资源进行合理调度。

5 结束语

受高铁全封闭式金属车厢引起的穿透信号损耗、高铁高速移动引起的通信连接不稳定和多普勒频移等因素的影响，乘客移动设备与路边基站直连的通信方式往往难以保障乘客的通信服务体验。高铁车载缓存技术通过在高铁上搭载缓存服务器，并在缓存服务器中预先存储乘客未来可能请求的内容，使得乘客未来的请求有一定几率被缓存服务器就近满足，移动通信设备无须与路边基站进行直连通信。

本文首先从系统架构、用户请求特性、系统工作流程等方面对高铁车载缓存系统进行了概述，然后分析了缓存放置阶段面临的缓存容量受限问题、请求调度阶段面临的请求随机性较大问题、内容分发阶段面临的无线资源受限问题以及业务场景支持单一的问题。针对上述的问题，我们分别总结了相应的智能化解决方案，通过使用深度学习技术对内容流行度进行精准预测，为缓存内容的放置提供更可靠的参考依据；考虑乘客的请求易受推荐系统的影响，通过推荐系统向乘客推荐缓存内容，以降低乘客请求的随机性；通过人车协同，在乘客移动设备中预先缓存精心设计的子文件，使得缓存内容可以通过编码多播的方式分发给乘客，降低对无线通信资源的需求。高铁车载缓存技术的发展仍面临诸多挑战，例如高铁与路边基站的协同缓存，乘客请求数据缺失下的请求规律挖掘，通信、计算、缓存资源的联合管控等。上述挑战的解决将有利于高铁车载缓存资源进一步发挥效用，为乘客带来更好的通信服务体验。

致谢

特别感谢中国科学院计算技术研究所曹梦华、邢旺同学在文章撰写过程中提供的大力支持。

参考文献

[1] ZHOU Y Q, LIU L, WANG L, et al. Service aware 6G: an intelligent and open network based on convergence of communication, computing and caching [J]. Digital communication networks, 2020, 6(3): 253–260. DOI: 10.1016/j.dcan.2020.05.003

[2] LIU L, ZHOU Y Q, YUAN J H, et al. Economically optimal MS association for multimedia content delivery in cache-enabled heterogeneous cloud radio access networks [J]. IEEE journal on selected areas in communications, 2019, 37(7): 1584–1593. DOI: 10.1109/JSAC.2019.2916280

[3] LIU L, ZHOU Y Q, GARCIA V, et al. Load aware joint CoMP clustering and inter-cell resource scheduling in heterogeneous ultra dense cellular networks [J]. IEEE transactions on vehicular technology, 2018, 67(3): 2741–2755. DOI: 10.1109/TVT.2017.2773640

[4] GARCIA V, ZHOU Y Q, SHI J L. Coordinated multipoint transmission in dense cellular networks with user-centric adaptive clustering [J]. IEEE transactions on wireless communication, 2014, 13(8): 4297–4308. DOI: 10.1109/TWC.2014.2316500

[5] ZHOU Y Q, LIU H, PAN Z G, et al. Two-stage cooperative multicast transmission with optimized power consumption and guaranteed coverage [J]. IEEE journal on selected areas in communications, 2014, 32(2): 274–284. DOI: 10.1109/JSAC.2014.141208

[6] ZHOU Y Q, WANG J Z, SAWAHASHI M. Downlink transmission of broadband OFCDM systems—Part I: hybrid detection [J]. IEEE transactions on communication, 2005, 53(4): 718–729. DOI: 10.1109/tcomm.2004.841961

[7] 艾渤，马国玉，钟章队．智能高铁中的 5G 技术及应用 [J]. 中兴通讯技术，2019, 25(6): 42–47. DOI: 10.12142/ZTETJ.201906007

[8] GAO M L, AI B, NIU Y, et al. Edge caching and content delivery with minimized delay for both high-speed train and local users [C]// 2019 IEEE Global Communications Conference. Waikoloa, HI, USA: IEEE, 2019. DOI: 10.1109/GLOBECOM38437.2019.9013389

[9] KANAI K, MUTO T, KATTO J, et al. Proactive content caching for mobile video utilizing transportation systems and evaluation through field experiments [J]. IEEE journal on selected areas in communications, 2016, 34(8): 2102–2114. DOI: 10.1109/JSAC.2016.2577238

[10] ZHOU Y Q, TIAN L, LIU L, et al. Fog computing enabled future mobile communication networks: a convergence of communication and computing [J]. IEEE communications magazine, 2019, 57(5): 20–27. DOI: 10.1109/MCOM.2019.1800235

[11] DUTTA L K, XIONG J, GUI L, et al. On hit rate improving and energy consumption minimizing in cache-based convergent overlay network on high-speed train [C]// 2019 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting. Jeju, Korea: IEEE, 2019. DOI: 10.1109/BMSB47279.2019.8971853

[12] WU Y, FANG X M, YAN L. Performance analysis of on-board content caching and retrieval for high-speed railways [C]//2019 IEEE 2nd 5G World Forum. Dresden, Germany: IEEE, 2019. DOI: 10.1109/5GWF.2019.8911642

[13] LI B, XIONG J, LIU B, et al. Cache-based popular services pushing on high-speed train by using converged broadcasting and cellular networks [J]. IEEE transactions on broadcasting, 2019, 65(3): 577–588. DOI: 10.1109/TBC.2018.2863102

[14] JIANG Y X, FENG H J, ZHENG F C, et al. Deep learning-based edge caching in fog radio access networks [J]. IEEE transactions on wireless communications, 2020, 19(12): 8442–8454. DOI: 10.1109/TWC.2020.3022907

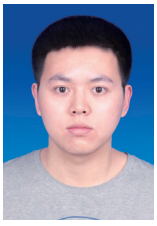
[15] CHATZIELEFTHERIOU L E, KARALIOPOULOS M, KOUTSOPOULOS I. Jointly optimizing content caching and recommendations in small cell networks [J]. IEEE transactions on mobile computing, 2019, 18(1): 125–138. DOI: 10.1109/TMC.2018.2831690

[16] SERMPEZIS P, GIANNAKAS T, SPYROPOULOS T, et al. Soft cache hits: improving P14erformance through recommendation and delivery of related content [J]. IEEE journal on selected areas in communications, 2018, 36(6): 1300–1313. DOI: 10.1109/JSAC.2018.2844983

[17] 王忠峰，王富章，孙华龙．高铁动车组 WiFi 运营服务系统服务质量的测量与分析 [J]. 电子技术应用，2018, 44(5): 77–81. DOI: 10.16157/j.issn.0258-7998.175166

[18] MADDAH-ALI M A, NIESEN U. Fundamental limits of caching [J]. IEEE transactions on information theory, 2014, 60(5): 2856–2867. DOI: 10.1109/TIT.2014.2306938

作 者 简 介



崔新雨，中国科学院计算技术研究所在读博士生；主要研究方向为通信与计算融合、移动边缘缓存等。

刘玲，中国科学院计算技术研究所助理研究员；主要研究方向为超密集网络中的干扰管控与资源管理、通信与计算融合等；曾获 IEEE ICC 2018 最佳论文奖；已发表论文 20 余篇。

周一青，中国科学院大学教授，中国科学院计算技术研究所“百人计划”研究员、博士生导师，无线通信技术研究中心副主任，移动计算与新型终端北京市重点实验室副主任；主要研究方向为移动通信、通信与计算融合等；曾获得 WCSP2019、IEEE ICC2018、ISCIT2016、IEEE PIMRC2015、ICCS2014、WCNC013 最佳论文奖，并获《China Communications》《IEEE Transactions on Vehicular Technology》最佳编辑奖等；已发表论文 150 余篇。

潘振岗，北京紫光展锐通信技术有限公司中央研究院先进通信技术实验室主任；主要研究方向为无线信号处理、多天线系统、信道编解码、跨层优化。