

DOI:10.13718/j.cnki.xsxb.2021.09.014

云环境下高效的车载多媒体任务调度架构^①

赵一瑾

云南交通职业技术学院 交通信息工程学院, 昆明 655000

摘要: 针对多媒体因资源成本、快速服务响应时间和体验质量严重影响车载通信性能等问题, 提出一种云环境下基于动态优先级的高效车载多媒体任务调度算法. 该算法利用基于动态优先级的车载多媒体作业队列(Job Queue, JQ)处理任何请求的优先级, 以确保向不同优先级的不同多媒体终端用户及时发送响应, 根据工作负载动态分配给每个计算服务器的计算资源, 以便根据多媒体用户体验质量(Quality of Experience, QoE)的要求来处理多媒体任务. 实验结果表明, 该算法在体验质量、服务响应时间和资源成本等方面均有明显改善.

关键词: 多媒体云计算; 智能交通系统; 车载自组网; 智能车辆; 体验质量

中图分类号: TP393

文献标志码: A

文章编号: 1000-5471(2021)09-0109-06

Efficient In-Vehicle Multimedia Task Scheduling Architecture in Cloud Environment

ZHAO Yijin

College of Traffic Information Engineering of Yunnan Jiaotong College, Kunming 655000, China

Abstract: To solve the program that the resource cost, fast service response time and experience quality of multimedia seriously affect the vehicle communication performance, an efficient in-vehicle multimedia task scheduling scheme based on dynamic priority in cloud environment has been proposed. In this scheme, dynamic priority based mobile multimedia job queue is used to process the priority of any request to ensure that the response is sent to different multimedia end users with different priorities in time, and the computing resources are dynamically allocated to each computing server according to the workload, so as to process the multimedia tasks according to the requirements of the quality of multimedia user experience. The experimental results show that the algorithm can significantly improve the quality of experience, service response time and resource cost.

Key words: multimedia cloud computing (MCC); intelligent transportation system; vehicle ad-hoc network; intelligent vehicle; quality of experience

① 收稿日期: 2021-03-15

基金项目: 云南省教育厅科学研究基金项目(2019J0521).

作者简介: 赵一瑾, 硕士, 副教授, 主要从事数字媒体及大数据可视化研究.

车载自组网(Vehicular Ad-hoc Networks, VANET)是传统的移动自组织网络在交通道路上的应用,是一种特殊的移动自组织网络^[1-2]. 在 VANET 中,车辆节点具有高度的移动性,在其运动过程中车辆之间使用路边基础设施进行通信以共享信息(用于安全驾驶的路线图、道路安全和交通拥堵信息)^[3-4]. 车辆产生大量的关键信息和多媒体数据,如监视视频、交通拥堵图像和其他娱乐数据. 由于有限的计算和存储能力,大量的多媒体相关数据无法在车载独立设备中进行处理,无法确保按时交付所需内容^[5-6],以维持最终用户所需的体验质量(Quality of Experience, QoE).

车载云计算(Vehicular Cloud Computing, VCC)是 VANET 倍受关注的网络之一,实现了在快速变化的自组织网络中进行通信和资源共享,为用户提供了低成本高效率的云服务^[7-8]. VCC 的目的是为配备低计算设备的车辆提供实时计算功能,优化行驶时间、避免事故和交通拥堵. 此外, VCC 提供技术上可行的无线传感器网络、移动云计算和智能交通系统的整合,提高道路安全并确保感知城市交通环境^[9-10]. 多媒体云计算(Multimedia Cloud Computing, MCC)关注如何为多媒体应用提供所需的服务质量^[11]. 车载数据的多媒体处理更加关键和具有挑战性,因为它要求处理及时、响应速度快、成本低.

为了方便车辆搜索和管理所需资源,已有车辆数据资源和资源分配方法相关的研究报道. 文献[12]提出一种车载云结构中资源分配和任务调度的多目标优化技术,它是一个由车辆云、路边云和集中云组成的 3 层体系结构. 采用遗传算法和自适应粒子群算法相结合的混合自适应粒子群优化算法进行优化,将所需资源分配给资源紧缺的应用和服务(即多媒体服务),以满足 3 层云体系结构中任何一种的服务质量要求(Quality of Service, QoS). 文献[13]提出通过车辆云感知体验质量的移动边缘缓存,满足日益增长的互联网内容需求并减少蜂窝基础设施的过载. 该算法在确保用户体验保证的前提下最大化流量卸载,设计 2 个可变的截止时间策略来考虑对用户体验的影响. 文献[14]提出一种基于优先级的动态信道分配方案,基于物联网的传输时间关键型应用. 该方案利用马尔可夫链模型对所有优先级类的次用户(Secondary User, SU)进行评估,并在此基础上为每个优先级类预留更多的信道以降低其呼叫阻塞概率,同时保持足够的信道利用率,并使低优先级 SU 的信道数目相似.

上述文献对独立的机载设备如此大量的数据没有足够的处理能力,无法为车载多媒体应用和服务提供快速、高效的资源调度,以维持最终用户所需的 QoE. 本文提出一种基于动态优先级的车载多媒体云计算高效资源分配与计算方案,以较低的成本处理基于多媒体任务优先级的时延敏感和多媒体计算. 该方案将多媒体任务分为 4 个子任务,并分配给相应的专用计算集群进行处理. 为了确保向具有不同优先级的不同多媒体终端用户及时发送响应,该方案利用基于动态优先级的车载多媒体作业队列(Job Queue, JQ)处理任何请求的优先级或紧急性. 为了根据多媒体用户 QoE 的要求来处理多媒体任务,将工作负载动态分配给每个计算服务器的计算资源. 实验结果表明,该算法在体验质量、服务响应时间和资源成本等方面均有明显提升.

1 模型构建

车载多媒体云计算(Vehicle Multimedia Cloud Computing, VMCC)用于处理车载多媒体图像或视频分析等相关任务. VMCC 体系结构分为 4 个主要部分:请求单元(Request Unit, RU)、负载管理器(Load Manager, LM)、计算集群单元(Computing Cluster Unit, CCU)以及车辆到云网关. 车辆多媒体通过无线保真(Wireless Fidelity, WiFi)、全球微波接入互操作性(World Interoperability for Microwave Access, WiMAX)、长期演进(Long-Term Evolution, LTE)等无线接口向 RU 提交多媒体计算请求,请求在请求队列中被接收,并将其转发给 LM 进行进一步处理,LM 分析传入请求的性质并将其分配给特定的 CCU. CCU 处理接收到请求,并根据请求类型将它们发送到下一个 CCU 或传输单元(Transmission Unit, TU).

最后，TU 广播/单播处理后的请求转发到云网关中。该方案利用作业队列 (Job Queue, JQ) 来处理任何处理请求的优先级或紧急性。在本文提出的方案中，每个 CCU 和 TU 都包含一个作业队列来存储所有的多媒体请求。VMCC 中的 LM 模块是一个智能化的模块，它不仅可以分析任务调度所需的媒体任务性质，而且还可以估计出任意时刻的总负载 a_i 。这些负载估计信息将被进一步用于优化每个 DCC 的资源分配。

CCU 分为 4 个子计算集群：动态转换集群 (Dynamic Conversion Cluster, DCC)、动态抽取集群 (Dynamic Extraction Cluster, DEC)、动态匹配集群 (Dynamic Matching Cluster, DMC) 以及动态重构集群 (Dynamic Reconstruction Cluster, DRC)，如图 1 所示。每个 DCC 包含一个优先级队

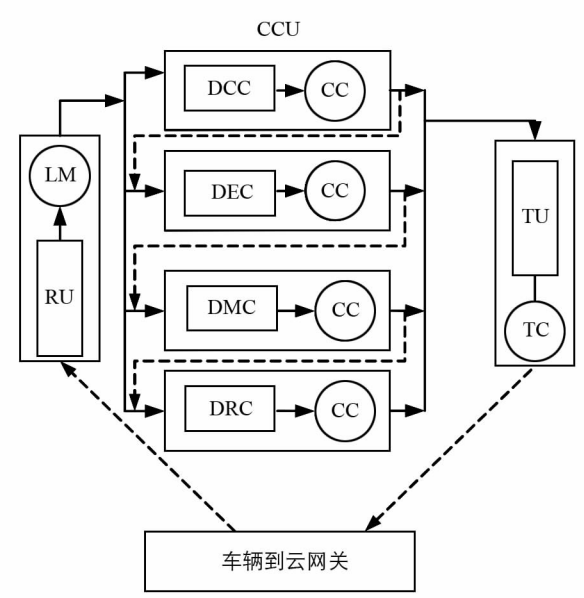


图 1 车载多媒体云计算架构

列，LM 根据预先定义的优先级策略 (值) 提交车载多媒体任务进行计算。本文假设每个车载多媒体任务的优先级是根据任务类型和性质来计算和分配的。例如，最不重要和最不关键的任务被赋予最低的优先级，最重要和最关键的任务被赋予最高的优先级。因此，由 CCU 处理的媒体任务被放置在 TU 的作业队列中。TU 可能会应用一些转换，将处理后的作业转发到云网关，然后再通过路边基础设施进一步转发到请求的车辆。

本文将整个多媒体处理结构分为 4 个阶段：转换、提取、匹配和重建。所有这 4 个子阶段都是由 4 个不同的、独立的、专用计算集群而不是单个计算集群来执行的。在本文提出的方案中，多媒体处理的每个子阶段可以单独地执行或根据用户需求与其他子阶段相关联地执行。为了最小化计算成本，根据需要或负载信息动态地分配给每个计算集群的计算资源。在本文所提出的 VMCC 架构中，每个专用计算集群 (Computing Cluster, CC) 包含一个用于存储作业的优先级队列，以保持车载多媒体任务的优先级。为了获得更高的计算性能，这 4 个 CC 都以协作的方式执行车载多媒体任务。

2 基于动态优先级的车载多媒体云计算任务调度

本文提出的算法主要由车载多媒体作业队列模型和基于动态优先级的车载多媒体任务调度两部分构成。该算法利用基于动态优先级的车载多媒体 JQ 处理任何处理请求的优先级，根据工作负载动态分配给每个计算服务器的计算资源，并根据多媒体用户 QoE 的要求来处理多媒体任务。

2.1 基于优先级的多媒体作业队列模型

车载多媒体计算中基于优先级的作业队列模型有 3 种不同的类型：调度队列、计算或作业队列以及传输队列。本文假设多媒体请求的优先级类，因此每个多媒体请求一次对应一个优先级类。较小的优先级值代表较高的优先级。例如，优先级为 1 的多媒体请求类具有比优先级为 2 的多媒体请求类更高的优先级。

任何优先级类别的多媒体请求都可以通过 4 个参数进行分类：用来指定指令总数的平均任务大小 TS 、优先级为 i 级的多媒体请求的平均到达率 AR_i 用来表示每秒的请求数量、平均结果大小 RS_i (以位为单位)、服务响应时间的上限 BS_i (以秒为单位)。在本文提出的方案中，每个多媒体任务进一步分为 4 个子任务。每个多媒体请求的优先级根据其重要性或关键性分配。

2.1.1 计算或作业队列

计算或作业队列存储车辆媒体任务, 并进行云处理. 在本文提出的方案中, 每个优先级为 i 的多媒体任务 F_i 被分为 4 个子任务或由 F_i^k 表示的子阶段, 其中 $k = 1, 2, 3, 4$, 每个子任务代表一个唯一的计算阶段, 优先级最高的多媒体任务首先以优先级方式进行计算.

2.1.2 调度队列

调度队列是指所有车辆提交需要媒体云处理的媒体任务. 调度服务器(Scheduling Server, SS) 根据分配或分配的优先级类, 以调度速率 SR 调度多媒体请求. 每个计算群集包含 N 个优先级队列, 如图 1 所示. 类似地, N 个传输队列存储已处理的数据或结果, 以传输速率 TR_i 转发给预期的最终用户. 利用平均到达率为 AR_i 的泊松分布, 对优先级为 i 类的多媒体请求进行建模.

2.1.3 传输队列

通过调度和计算队列处理的多媒体任务最终以传输速率 TR_i 发送到传输队列, 其中 N 个优先级队列存储 RS_i 大小的处理结果, 然后将其转发给预定的优先级等级为 i 的最终用户.

2.2 用户体验质量(QoE)

用户 QoE 以端到端或 RT(Response Time, RT) 时间来衡量, 即从车辆媒体请求启动阶段到车辆收到所需响应的时延, QoE 与 RT 值成正比. 如果特定车辆及时收到所需的响应, 那么它将体验到更好的 QoE. 如果特定车辆没有及时收到要求的响应, 则其 QoE 较低.

2.3 基于动态优先级的车载多媒体任务调度

基于动态优先级的车载多媒体云计算任务调度算法的主要目标是: 高效处理、准时交付及最小化计算成本. 在本文提出的动态资源分配方案中, 云有 4 个专用的媒体集群, 安装的虚拟机负责处理多媒体车辆任务. 根据 LM 估计的车载多媒体请求负载动态地分配给每个 CC 的计算资源. 如果车辆请求多于分配的资源, 则将更多的计算资源分配给 CC; 如果估计的负载小于已经分配的计算资源, 则计算资源被移除. 因此, 基于车辆负载估计, 计算资源会定期更新, 以最小化计算成本为移动车辆提供更好的 QoE.

基于优先级的任务调度和处理过程: ①对来自不同车辆用户的实时和多媒体任务进行初始分配和分析, 所有接收到的多媒体请求根据其优先级值以优先级非抢占的形式进行分析和排序. 在本文中, 优先级值越低, 表示要处理的任务优先级越高. ②为 4 个计算集群中的每一个集群都分配了用于作业处理的资源, 计算资源分配给每个 CC 是基于对在时间 a_i 中接收到的初始工作负载的分析. 因此, 计算资源量将根据接收到的工作量而变化, 有效地利用计算资源, 满足多媒体任务延迟的要求. LM 将每个多媒体任务分配给其相应的 CC, 并将该任务放入其作业队列进行进一步处理. CC 处理多媒体任务, 根据任务处理的性质将其放入下一计算单元或传输单元的任务队列中. 若任务被完全处理, 则将其发送到传输单元的任务队列传输给预期的车辆用户; 若任务处理尚未完成, 则将其发送到下一个 CC 的任务队列. ③处理后的多媒体任务被同时转发给其预期的车辆用户, 避免任何进一步的排队延迟, 满足多媒体任务的延迟截止时间, 以获得较好的用户 QoE.

3 实验结果与分析

为了对实验进行性能评估, 在配置为 Intel(R) Core(TM)i5-3320 CPU @2.60 GHz 和 4 GB RAM 的 Windows 10 Professional 64 位操作系统上, 使用 CloudSim 模拟器对本文所提出的方案进行性能分析, CloudSim 是一个基于 Java 的库模型, 广泛用于云的模拟. CloudSim 模拟的有效性通过基于云的内容分发服务进行, 该服务根据服务的计算请求设置相应的应用复杂度, 保证了仿真的有效性. 本实验设计将云建模为 3 个主要组件: LU, CCU, TU. 为了在多媒体图像上执行仿真, 修改 Cloudlet 类的特性并根据要求对

其进行建模. 车辆媒体任务的到达遵循 t 时间间隔的泊松分布.

图 2 给出了在各种仿真设置下, 采用静态资源分配方案的性能评估结果.

由图 2 可以看出, 当云数量较少时, 静态资源分配方案的性能会更好, 但是随着云数量增加, 响应时间也会增加, 从而降低了终端设备上的 QoE. 本文提出的算法中资源会根据负载信息定期动态更新, 因此即使在更多的任务数下也能取得更好的性能, 并且在接收车辆处提供了更好和有保证的 QoE.

图 3 给出了在各种仿真设置下, 采用基线单集群方案的性能评估结果. 由图 3 可以看出, 本文所提出的算法优于基线单集群的计算方案. 该算法提供了最短的计算时间, 从而支持车辆的保证用户 QoE. 在基线单集群的计算方案中, RT 随着云的增加而增加. 这是因为只有一个集群负责执行所有 4 个与映像相关的任务. 在本文提出的算法中, 图像处理任务分为 4 个子任务, 子任务被进一步分配给 4 个专用计算集群. 因此, 在响应时间方面, 该算法的性能优于基线单集群计算方案.

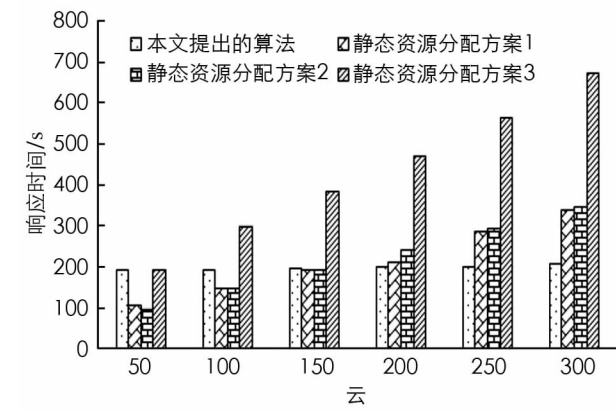


图 2 本文提出的算法与静态资源分配方案

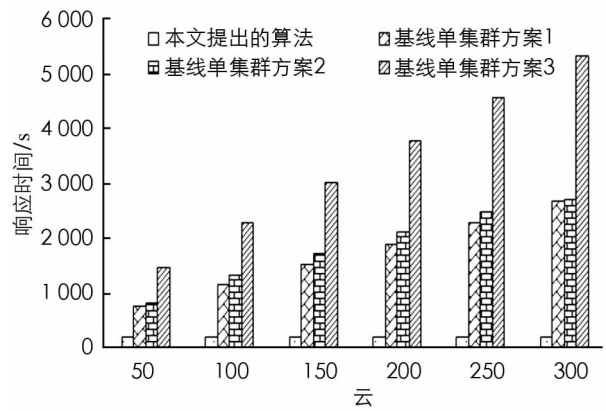


图 3 本文提出的算法与基线单集群方案

4 结 语

为了解决快速响应、保证体验质量和最小化计算资源成本等挑战, 本文提出一种云计算环境下基于动态优先级的高效车载多媒体任务调度算法. 为了确保向具有不同优先级的不同多媒体终端用户及时发送响应, 该算法利用基于动态优先级的车载多媒体作业队列 (Job Queue, JQ) 处理任何处理请求的优先级或紧急性. 为了根据多媒体用户 QoE 的要求来处理多媒体任务, 将工作负载动态分配给每个计算服务器的计算资源. 利用 Cloudsim 模拟器, 结合静态资源分配方案和基于单集群的基线计算方案, 对本文提出算法的性能进行了评估. 实验结果表明, 该算法在体验质量、服务响应时间和资源成本等方面均有明显改善. 未来的工作是针对动态场景中车辆计算资源共享问题, 研究提升网络任务卸载能力的算法, 满足用户的服务要求.

参考文献:

- [1] ZHANG D, GE H, ZHANG T, et al. New Multi-Hop Clustering Algorithm for Vehicular Ad Hoc Networks [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(4): 1517-1530.
- [2] BOUSSOUFA-LAHLAH S, SEMCHEDINE F, BOUALLOUCHE-MEDJKOUNE L. Geographic Routing Protocols for Vehicular Ad Hoc NETWORKS (VANETs): aSurvey [J]. Vehicular Communications, 2018, 11: 20-31.
- [3] QIU T, WANG X, CHEN C, et al. TMED: aSpider-Web-Like Transmission Mechanism for Emergency Data in Vehicular Ad Hoc Networks [J]. IEEE Transactions on Vehicular Technology, 2018, 67(9): 8682-8694.
- [4] SUNY X, GUOX Y, ZHOUS, et al. Learning-Based Task Offloading for Vehicular Cloud Computing Systems [C]//

- 2018 IEEE International Conference on Communications (ICC). Kansas City: IEEE, 2018.
- [5] ASHOK A, STEENKISTE P, BAI F. Vehicular Cloud Computing through Dynamic Computation Offloading [J]. Computer Communications, 2018, 120: 125-137.
- [6] LIH W, LUR X, MISIC J, et al. Security and Privacy of Connected Vehicular Cloud Computing [J]. IEEE Network, 2018, 32(3): 4-6.
- [7] 郭剑岚, 陈俞强. 基于 Stackelberg 博弈的车载云计算任务卸载优化 [J]. 计算机应用研究, 2019, 36(12): 3752-3755, 3784.
- [8] KIM T, MIN H, CHOI E, et al. Optimal Job Partitioning and Allocation for Vehicular Cloud Computing [J]. Future Generation Computer Systems, 2020, 108: 82-96.
- [9] BOUKERCHE A, DE GRANDE R E. Vehicular Cloud Computing: Architectures, Applications, and Mobility [J]. Computer Networks, 2018, 135: 171-189.
- [10] ASHRAF M U, ARIF S, BASIT A. Provisioning Quality of Service for Multimedia Applications in Cloud Computing [J]. International Journal of Information Technology and Computer Science, 2018, 10(5): 40-47.
- [11] BHUSHAN K, GUPTA B B. Network Flow Analysis for Detection and Mitigation of Fraudulent Resource Consumption (FRC) Attacks in Multimedia Cloud Computing [J]. Multimedia Tools and Applications, 2019, 78(4): 4267-4298.
- [12] MIDYA S, ROY A, MAJUMDER K, et al. Multi-Objective Optimization Technique for Resource Allocation and Task Scheduling in Vehicular Cloud Architecture: aHybrid Adaptive Nature Inspired Approach [J]. Journal of Network and Computer Applications, 2018, 103: 58-84.
- [13] VIGNERI L, SPYROPOULOS T, BARAKAT C. Quality of Experience-Aware Mobile Edge Caching through a Vehicular Cloud [J]. IEEE Transactions on Mobile Computing, 2020, 19(9): 2174-2188.
- [14] ALI A, FENG L, BASHIR A K, et al. Quality of Service Provisioning for Heterogeneous Services in Cognitive Radio-Enabled Internet of Things [J]. IEEE Transactions on Network Science and Engineering, 2020, 7(1): 328-342.

责任编辑 夏 娟