

基于 NOMA 的车辆边缘计算网络优化策略

杜剑波¹, 薛哪哪¹, 孙艳¹, 姜静¹, 李树磊², 卢光跃¹

(1. 西安邮电大学陕西省信息通信网络及安全重点实验室, 陕西 西安 710071;

2. 天元瑞信通信技术股份有限公司, 陕西 西安 710121)

摘要: 目前, 车载网络正面临着为车辆提供无处不在的连接和大量计算密集型、时延敏感型智能服务的挑战。为了应对这些挑战, 非正交多址接入 (NOMA, non-orthogonal multiple access) 和移动边缘计算 (MEC, mobile edge computing) 被认为是两种较有前景的技术, 它们分别允许多个车辆共享相同的无线资源以及在车辆边缘使用强大的边缘计算资源。在基于 NOMA 的车辆边缘计算网络中, 在保证任务处理时延的情况下, 提出了任务卸载、用户分簇、计算资源分配和发射功率控制的联合优化问题, 使得任务处理成本最小化。由于所提出的问题难以解决, 因此将其分解为子问题, 并提出了低复杂度和易于实现的方法来解决。仿真结果表明, 与其他基准算法相比, 该算法在最小化成本方面表现良好。

关键词: 边缘计算; 非正交多址接入; 车联网

中图分类号: TN92

文献标识码: A

doi: 10.11959/j.issn.2096-3750.2021.00207

Optimization strategies in NOMA-based vehicle edge computing network

DU Jianbo¹, XUE Nana¹, SUN Yan¹, JIANG Jing¹, LI Shulei², LU Guangyue¹

1. Shaanxi Key Laboratory of Information Communication Network and Security, Xi'an University of Posts and Telecommunications, Xi'an 710071, China

2. Tianyuan Ruixin Communication Technology Co. Ltd, Xi'an 710121, China

Abstract: Nowadays, vehicular network is confronting the challenges to support ubiquitous connections and vast computation-intensive and delay-sensitive smart service for numerous vehicles. To address these issues, non-orthogonal multiple access (NOMA) and mobile edge computing (MEC) are considered as two promising technologies by letting multiple vehicles to share the same wireless resources, and the powerful edge computing resources were adopted at the edge of vehicular wireless access network respectively. A NOMA-based vehicular edge computing network was studied. Under the condition of guaranteeing task processing delay, the joint optimization problem of task offloading, user clustering, computing resource allocation and transmission power control was proposed to minimize the task processing cost. Since the proposed problem was difficult to solve, it was divided into sub-problems, and a low-complexity and easy-to-implement method was proposed to solve it. The simulation results show that compared with other benchmark algorithms, the proposed algorithm performs well in minimizing costs.

Key words: edge computing, non-orthogonal multiple access, vehicular network

收稿日期: 2020-12-31; 修回日期: 2021-02-03

通信作者: 卢光跃, tonylugy@163.com

基金项目: 陕西省自然科学基金资助项目 (No.2020JQ-844); 国家自然科学基金资助项目 (No.61901367, No.62001357, No.61871321, No.61901381, No.62071377); 陕西省宽带无线技术应用创新团队项目 (No.2017KCT-30-02); 国家科技重大专项资助项目 (No.2016ZX03001016)

Foundation Items: The Natural Science Foundation of Shaanxi Province (No.2020JQ-844), The National Natural Science Foundation of China (No.61901367, No.62001357, No.61871321, No.61901381, No.62071377), The Science and Technology Innovation Team of Shaanxi Province for Broadband Wireless and Application (No.2017KCT-30-02), The National Science and Technology Major Project (No.2016ZX03001016)

1 引言

近年来,随着车辆的普及以及车联网^[1]的发展,新型智能应用出现得越来越多,如智能交通、自动驾驶、虚拟现实和动态高精度地图导航等,给人们的生活带来了极大的便利^[2]。这些计算密集型应用通常需要大量的存储空间和计算资源,且对时延有极高的要求^[3]。对于车联网终端用户(以下简称车辆用户或者用户)来说,由于其处理能力有限,无法很好地支持这些智能应用。另外,越来越多的车辆涌入无线网络,与之矛盾的是频谱资源严重紧缺,传统的多址接入技术已无法满足无线系统的组网需求。

车辆边缘计算(VEC, vehicular edge computing)^[4]是将 MEC^[5-6]和车联网结合起来以降低任务处理时延和能耗等方面的问题。MEC 具备了为移动网络提供大量云计算的能力^[7],同时缓解了网络中时延的影响,其通过将计算和存储资源部署在网络边缘,为用户提供超低时延、高带宽的网络服务。其中,为了解决车载设备在资源存储、计算性能以及时延等方面的问题,任务卸载^[8-9]利用边缘计算服务器帮助解决该问题,从而实现高效操作^[10]。同时,卸载决策和资源分配^[11]问题也是 VEC 中的关键问题。由于车辆具有移动性,可能会导致任务卸载时所在的基站和任务回传时的基站不是同一个,并且每个基站所服务的范围会受基站无线通信的影响,因此,移动性也是 VEC 必须考虑的问题。

NOMA^[12-13]是一个提高频谱利用率的有效技术。多址接入技术是无线通信网中提升系统容量^[14]的有效技术和核心技术。在 VEC 中,由于存在大量的车辆用户以及处理任务等问题,可能会导致频谱资源不足。因此,将 NOMA 与车辆边缘计算结合起来解决上述问题。传统车联网中,为了避免多址干扰,采用正交多址接入(OMA, orthogonal multiple access)方式,但其接收相对简单^[15-16]。然而,当今车联网要求低时延、高可靠等性能,OMA 技术已经无法满足需求。NOMA 技术是一种有效的解决方案,它允许不同终端用户共享相同的频谱、时间和空间等资源,并利用上述方式进行并行传输,通过在接收端进行串行干扰消除(SIC, successive interference cancellation)抵消干扰,从而实现正确解调^[17-18]。相较于 OMA 技术,NOMA 技术可以支持更多的用户,提高频谱利用率,缓解数据流量的拥塞,从而降低时延^[19]。

将 MEC 和 NOMA 引入车联网,可以大大提高系统任务处理性能,缩短任务处理时间并支持海量车辆大连接。目前,MEC 和车联网的结合以及 NOMA 和车联网的结合,都引起了广泛的关注。

对于车联网的特殊场景来说,车辆的能量供给充足,不需要考虑能耗问题,但其处理任务的时延是至关重要的。另外,如果执行任务卸载的话,需要向运营商支付相应的费用。因此,基于边缘计算的车联网主要考虑降低任务执行时延和经济开销两个因素。在文献[20]中,研究了基于边缘计算的车联网,通过优化任务的卸载时间、无线带宽分配和计算资源分配来最小化系统的总开销。文献[21]通过联合优化发送功率控制、上行卸载时间、任务卸载比例以及本地中央处理器(CPU, central processing unit)工作频率,使得车辆用户的计算开销最小化并将问题建模成一个非凸的问题,通过将其转换分解成两层子问题进行求解。文献[22]提出一个并行计算和传输方案,该方案将任务分成两部分,一部分卸载到 MEC 服务器执行,另一部分在本地执行,通过联合优化卸载比例、通信资源和无线资源分配,最小化车联网中虚拟现实(VR, virtual reality)应用的任务处理时延。

由于 NOMA 的主要作用是通过多个用户共用相同的无线资源来提高系统容量、频谱效率等。因此,把 NOMA 引入车联网也能起到相同的作用。文献[23]在长期演进(LTE, long term evolution)系统仿真平台上,将 NOMA 引入车辆与交通基础设施(V2I, vehicle-to-infrastructure)通信,并以吞吐量为性能指标,研究 NOMA 技术的不同用户配对方案和功率分配方案对系统性能的影响。文献[24]研究了基于 NOMA 的车联网中内容分发效率,在保证安全信息的服务质量(QoS, quality of service)的前提下,分析了 NOMA 传输的随机时延性能。

目前,在将 NOMA 和 MEC 引入车联网以提高系统性能时,大多数研究工作只关注将这两者之中的某一个与车联网结合。由于 NOMA 和 MEC 的关注点不同,因此将两者结合引入车联网,可以通过 MEC 减小任务处理时延和开销,再通过 NOMA 容纳更多的用户,获得更高的任务卸载传输速率,进一步提升任务卸载的性能。文献[25]从减小 MEC 服务器能耗的角度考虑,提出了一个基于 NOMA 的高能效任务卸载算法,在一条部署多个 MEC 服务器的马路上,通过优化每个 MEC 服务器要执行的

任务以及通过哪个路边通信单元把任务处理结果回传给用户，以最小化 MEC 服务器的能耗并同时保证任务的处理时延。然而，以车联网中用户的角度考虑，车辆卸载需要支付的费用、车辆任务处理时间都是很重要的因素。

本文的主要贡献概括如下。

1) 本文研究了在基于 NOMA 的 VEC 网络中，考虑在保证任务处理时延的前提下，使用户支付费用最小化。

2) 依据所提出的系统模型，本文考虑用户卸载任务到本地或 MEC、MEC 的资源分配、任务处理完成后需回传结果以及虚拟运营商需收取费用等现实情况，设计了一种联合优化卸载决策、用户分簇、计算资源分配以及发送功率控制的方案。由于所建模问题是混合整数非线性规划问题，因此，将其分解为子问题，并提出了复杂度较低且容易实现的方法进行求解。

3) 通过设置仿真实验将本文方案与随机卸载、随机分簇、随机计算资源分配方案进行对比，证明了本文方案在降低成本方面性能优越。

2 系统模型

考虑一个 VEC 网络，系统模型如图 1 所示。一条单车道公路由若干个接入点 (AP, access point) 覆盖且每个 AP 的覆盖范围为 D_g ，每个 AP 上部署了一个边缘计算服务器。仅考虑某一个特定的 AP，在其覆盖范围之内，随机分布了 N 个车辆用户，每个车辆的初始位置记为 l_n^0 (单位: m)，所有车辆以速率 v (单位: m/s) 匀速行驶。所有用户的集合记为 $\mathcal{N} = \{1, 2, \dots, N\}$ 。每个用户 n 都有一个计算密集型的智能应用，记为 $A_n = \{D_n, C_n, O_n, T_n^{\max}\}$ ，其中 D_n 表示任务的输入数据量，其单位为 bit; C_n 表示

任务量，即处理完该任务所需的 CPU 圈数，其单位为 CPU cycles; O_n 表示任务执行完成后输出的数据量，其单位为 bit; T_n^{\max} 表示要成功执行完该任务可容忍的最大时延，其单位为 s。所有车辆的任务处理能力相同，记为 f^{loc} ，其单位为 CPU cycles/s。每个 AP 有 J 个子载波，每个子载波带宽为 B ，用于上传数据，所有子载波的集合记为 $\mathcal{J} = \{1, 2, \dots, J\}$ 。各个 AP 之间通过无线链路连接起来，速率记为 $R^{\text{AP-AP}}$ ；每个 AP 的下行速率为 R^{down} 。

每个用户的任务可以在本地处理，也可以卸载到 MEC 服务器处理。用 x_n 表示车辆 n 的任务 (简称任务 n) 的卸载决策， $x_n = 0$ 表示任务 n 在本地处理， $x_n = 1$ 表示任务 n 被卸载到 MEC 服务器处理。如果任务 n 在本地处理，其处理时间为

$$T_n^{\text{loc}} = \frac{C_n}{f^{\text{loc}}} \quad (1)$$

本文假定任务传输和处理不能同时进行，如果任务 n 在 MEC 服务器处理，任务的输入数据首先会被传送到 MEC 服务器，然后由 MEC 服务器处理。由于多个用户采用 NOMA 方式使用信道资源，即使用相同信道资源的用户被分到一个簇中，本文中每个簇使用一个子载波，故所有卸载的用户分成 J 个簇。假定每个簇中最多有 M 个用户，则一个簇中所有用户的编号集合可以记为 $\mathcal{M} = \{1, 2, \dots, M\}$ ，每个簇中的所有用户使用同一个子载波上传数据。因此，根据 NOMA 原理可知，不同簇之间没有干扰，但同一个簇中由于用户使用相同的资源，所以会相互干扰。因此，在接收端采用 SIC 技术来区分同一个簇中的不同用户。

对于每个簇中的用户，首先按照其信道增益进行降序排列，在解调时强用户的信号会被首先解

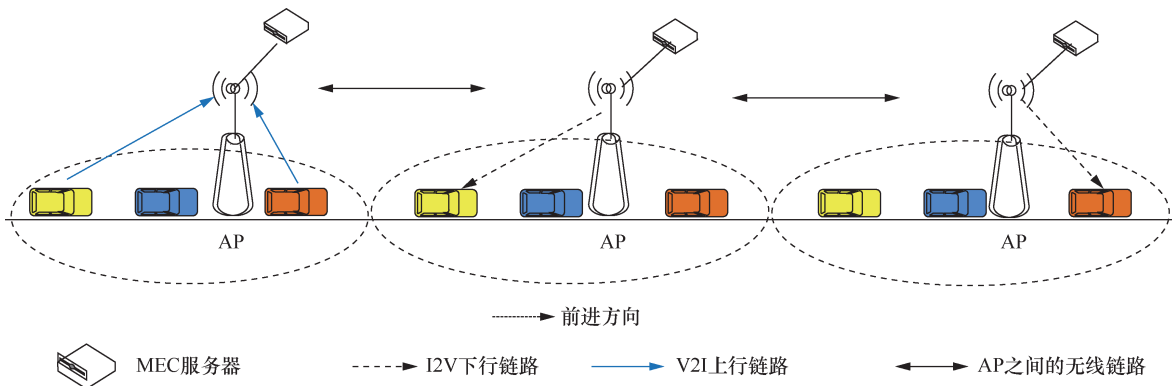


图 1 系统模型

调, 其余的弱用户会被当成干扰, 以此类推, 直到解调到最后一个用户。对于最后一个解调的用户, 此时信道中已没有其他用户, 因此不会受其他用户的干扰。 $\alpha_n^{j,m}, n \in \mathcal{N}, j \in \mathcal{J}, m \in \mathcal{M}$ 表示用户分簇, 其中 $\alpha_n^{j,m} = 1$ 表示用户 n 被分配到第 j 簇的第 m 个位置上, 反之 $\alpha_n^{j,m} = 0$ 。因此, 任务传输的速率为

$$R_n^{\text{up}} = \sum_{j \in \mathcal{J}} \sum_{m \in \mathcal{M}} \alpha_n^{j,m} B \text{lb} \left(1 + \frac{p_n h_n (d_n)^{-r}}{\sum_{\substack{l \in \mathcal{N}, \\ h_l < h_n}} p_l h_l (d_l)^{-r} + \sigma^2} \right) \quad (2)$$

其中, B 为子载波带宽, p_n 为终端 n 的发送功率, h_n 为信道增益, d_n 为终端 n 和 AP 之间的距离, σ^2 为高斯白噪声的功率, 且有

$$d_n = \sqrt{\left(\frac{D_g}{2} - l_n^0 \right)^2 + l_g^2} \quad (3)$$

因此, 在任务卸载时, 上行传输时延为

$$T_n^{\text{up}} = \frac{D_n}{R_n^{\text{up}}} \quad (4)$$

MEC 服务器中的计算资源以资源块的形式提供给用户, 每个 MEC 服务器中包含 U 个计算资源块, 每个计算资源块的处理能力相同, 记为 f^{unit} , 其单位为 CPU cycles/s。假设用户 u 分到 β_n 个计算资源块, 则任务 n 的处理时延为

$$T_n^c = \frac{C_n}{\beta_n f^{\text{unit}}} \quad (5)$$

在网络中, 有一个虚拟运营商管理所有 AP 及 MEC 服务器, 其为用户提供计算卸载服务并从用户处收取费用。用户若要进行计算卸载, 需要向虚拟运营商支付两部分费用, 分别为使用计算资源块的费用和使用无线资源费用。

用户若要使用无线资源, 需要从两方面收费, 即使用的带宽及使用时间。假定用户使用无线资源的价格为 ϕ_1 (单位: 元/(Hz·s)), 则用户使用无线资源需要支付的费用为

$$F_n^1 = \phi_1 \cdot B \cdot T_n^{\text{up}} \quad (6)$$

假定用户使用一个计算资源块的费用为 ϕ_2 (单位: 元), 则用户使用计算资源需要支付的费用为

$$F_n^2 = \phi_2 \beta_n \quad (7)$$

因此, 用户 n 任务卸载需要支付的总费用为 $F_n = F_n^1 + F_n^2$ 。

当用户任务处理完成之后, 需要将任务处理结果发送给用户。由于用户以速率 v 进行移动, 在任务处理完成之后, 用户可能在当前 AP 的覆盖范围内, 也可能已经驶离当前 AP 的覆盖范围。若是前者, 则将任务处理结果直接由下行链路发给用户即可; 否则, 将任务处理结果先转发到用户当前所在的 AP, 然后由下行链路发送给用户。如图 1 所示, 其中黄色小车的任务处理完成之后, 黄色小车处于第二个 AP 的覆盖范围之内, 任务处理结果需要传送到下一个 AP 才能转发给用户; 橙色小车任务处理完成后处于第三个 AP 的覆盖范围之内, 任务结果需要传送到第三个 AP 才能发给用户。用户任务处理结果在两个相邻 AP 之间传输的时间为

$$T_n^{\text{AP-AP}} = \frac{O_n}{R^{\text{AP-AP}}} \quad (8)$$

用户任务结果由 AP 下行回传给用户需要的时间为

$$T_n^{\text{down}} = \frac{O_n}{R^{\text{down}}} \quad (9)$$

因此, 用户将任务卸载到 MEC 服务器执行需要的总时间为

$$T_n^{\text{mec}} = T_n^{\text{up}} + T_n^c + \gamma T_n^{\text{AP-AP}} + T_n^{\text{down}} \quad (10)$$

其中, γ 为任务处理完成时, 用户驶离初始 AP 所经过的 AP 个数。用户的任务总处理时间为

$$T_n = (1 - x_n) T_n^{\text{loc}} + x_n T_n^{\text{mec}} \quad (11)$$

3 问题建模与求解

3.1 问题建模

本节通过联合优化任务卸载决策 $x = \{x_1, x_2, \dots, x_N\}$ 、用户分簇 $\alpha = \{\alpha_n^{j,m}\}$ 、计算资源块分配 $\beta = \{\beta_1, \beta_2, \dots, \beta_N\}$ 及发送功率控制 $p = \{p_1, p_2, \dots, p_N\}$, 在保证用户任务处理时延的前提下, 来最小化用户的总经济开销。所建模的优化问题为

$$\begin{aligned} P_1: & \min_{x, p, \beta, \alpha} \sum_{n \in \mathcal{N}} F_n \\ \text{s.t.} & \text{C1: } T_n \leq T_n^{\text{max}}, \forall n \in \mathcal{N} \\ & \text{C2: } 0 \leq p_n \leq p_n^{\text{max}}, \forall n \in \mathcal{N} \\ & \text{C3: } \beta_n \in \mathcal{N}, \forall n \in \mathcal{N} \\ & \text{C4: } \sum_{n \in \mathcal{N}} \beta_n \leq U \end{aligned}$$

$$\begin{aligned}
C5: & \alpha_n^{j,m} \in \{0,1\}, \forall n \in \mathcal{N}, \forall j \in \mathcal{J}, \forall m \in \mathcal{M} \\
C6: & \sum_{j \in \mathcal{J}} \sum_{m \in \mathcal{M}} \alpha_n^{j,m} = 1, \forall n \in \mathcal{N} \\
C7: & x_n \in \{0,1\}, \forall n \in \mathcal{N}
\end{aligned} \quad (12)$$

其中, C1 任务处理时延不能大于最大时延约束; C2 是用户发送功率约束; C3 是计算资源块取值的整数约束; C4 保证分给所有用户的计算资源块之和不能超过 MEC 服务器所拥有的总计算资源块的个数; C5 和 C6 保证每个用户只能分到一个簇中的一个位置; C7 表示用户的任务只能在本地或者 MEC 服务器上执行。由于问题 P_1 是一个混合整数非线性规划问题, 其求解很复杂, 因此, 提出低复杂度的启发式算法进行求解。

3.2 问题求解

3.2.1 卸载决策优化

由于本文的目标是在任务处理最大时延约束下, 最小化任务处理所需要的费用。因此, 可以通过卸载决策优化使费用最小化。对于所有本地可执行的任务, 即本地处理时间不超过本地处理最大时延, 令其卸载决策为 $x_n = 0$, 可以保证本地可行的任务不产生经济费用。对于所有本地不可执行的任务, 才将其卸载到 MEC 服务器处理, 即卸载决策为 $x_n = 1$ 。

在得到用户的卸载决策之后, 问题 P_1 中的约束 C7 随之确定, 然后用户被分成本地执行用户和计算卸载用户, 它们分别用集合 $\mathcal{N}_0 = \{1, 2, \dots, N_0\}$ 和 $\mathcal{N}_1 = \{1, 2, \dots, N_1\}$ 表示。此时将约束 C7 移除后优化问题变为

$$\begin{aligned}
P_2: & \min_{p, \beta, \alpha} \sum_{n \in \mathcal{N}_1} F_n \\
\text{s.t. } & C1: T_n \leq T_n^{\max}, \forall n \in \mathcal{N}_1 \\
& C2: 0 \leq p_n \leq p_n^{\max}, \forall n \in \mathcal{N}_1 \\
& C3: \beta_n \in \mathbb{N}, \forall n \in \mathcal{N}_1 \\
& C4: \sum_{n \in \mathcal{N}_1} \beta_n \leq U \\
& C5: \alpha_n^{j,m} \in \{0,1\}, \forall n \in \mathcal{N}_1, \forall j \in \mathcal{J}, \forall m \in \mathcal{M}
\end{aligned}$$

$$C6: \sum_{j \in \mathcal{J}} \sum_{m \in \mathcal{M}} \alpha_n^{j,m} = 1, \forall n \in \mathcal{N}_1 \quad (13)$$

3.2.2 用户分簇

在使用 SIC 解调时, 簇内用户信道差异越大, 系统的容量和性能越好。因此, 本文采取如下分簇策略。

首先, 将所有要执行任务卸载的用户按照信道增益降序排列, 即 $h_1 \geq h_2 \geq \dots \geq h_{N_1}$ 。每个 AP 有 J 个子载波, 将使用同一个子载波的用户分成一个簇, 共分成 J 个簇。

其次, 执行用户分簇。第 $\{j, j+J, j+2J, \dots\}$ 个用户分成一簇, 使用第 j 个子载波, 其中 $j \leq J$ 。用户分簇示意图如图 2 所示。图 2 中共有 10 个用户, 它们按照信道增益降序进行排列, 有 3 个子载波, 因此可以将用户分成 3 个簇, 其中第 1、4、7、10 号用户在簇 1, 第 2、5、8 号用户在簇 2, 第 3、6、9 号用户在簇 3。

当用户分簇完成后, 原问题退化为

$$\begin{aligned}
P_3: & \min_{p, \beta} \sum_{n \in \mathcal{N}_1} F_n \\
\text{s.t. } & C1: T_n^{\text{mec}} \leq T_n^{\max}, \forall n \in \mathcal{N}_1 \\
& C2: 0 \leq p_n \leq p_n^{\max}, \forall n \in \mathcal{N}_1 \\
& C3: \beta_n \in \mathbb{N}, \forall n \in \mathcal{N}_1 \\
& C4: \sum_{n \in \mathcal{N}_1} \beta_n \leq U
\end{aligned} \quad (14)$$

3.2.3 发送功率控制

由于各个簇使用相互正交的子载波, 各个簇在功率分配时互不干扰。因此, 本文只讨论某一个簇内的功率分配。本文的目的是在保证用户时延的情况下, 最小化任务处理的费用。另外, 由于车辆终端能量供给充足, 因此, 为了减少用户的费用, 最直观的方法是所有用户都使用最大功率发射, 即从功率方面保证速率最大化, 从而减少无线资源的使用费用。然而, 由于多个用户使用相同的资源, 为了保证解调时具有较好的性能, 提出如下低复杂度的功率控制策略。

1) 对于簇内最后一个用户 (即第 M 个用户), 也就是信道最差的用户, 由于它最后被解调, 解调

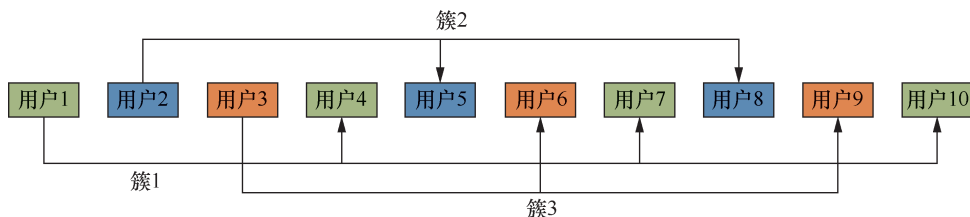


图2 用户分簇示意图

时无干扰，因此，可以采用最大的功率发送。

2) 当最后一个用户的发送功率确定后，其前一个用户（即用户 $M-1$ ）的传输速率如式(2)所示，就只取决于该用户的发送功率。为了保证该用户不会对其前面用户的解调造成太大的干扰，可以取满足最大时延需求的发送功率。

3) 同理，当用户 M 和 $M-1$ 的发送功率确定时，第 $M-2$ 个用户的传输速率就只取决于该用户的发送功率。因此，采用与 2) 中相同的方法，可以得到第 $M-2$ 个用户的功率，以此类推，直至得到第一个用户的发送功率。

3.2.4 计算资源块分配策略

在进行计算资源块分配时，由于用户使用计算资源块按照个数收费，因此，假定有足够多的计算资源块，给每个用户可按照需求分配足够的资源块。在此进行两轮计算资源块分配，具体如下。

第一轮，给每个用户分配若干计算资源块，保证满足最低时延即可。

第二轮，根据移动性，对第一轮分配结果进行修正。由于在上传及处理过程中，用户一直在移动。因此，当任务处理完成后，需要将结果传给用户，这时用户可能经过了若干个 AP，处理结果需要通过多个 AP 中继才能到达当前用户所处的 AP 覆盖范围之内。因此，在任务处理完成后，首先根据任务上传和处理时间计算出用户的当前位置，当前位置确定后，可以计算出上传、处理、下传的总时延。如果计算出某个用户的总时延超出了任务处理最大时延约束，则给该用户继续分配计算资源块，直到满足时延约束为止。

4 仿真结果分析

本文默认参数为：用户数为 $N=40$ ，子载波个数为 $K=4$ ，用户行驶速率为 $v=10$ km/h，每个 MEC 服务器之间的有线传输速率为 $R^{\text{wired}}=10$ Mbit/s，每个 MEC 服务器的下行速率为 $R^{\text{down}}=5$ Mbit/s，用户任务输入数据为 $D_n=0.5\sim 1$ MB，用户卸载的任务量为 $C_n=100\sim 400$ CPU cycles，任务执行最大时延为 $T_n^{\text{max}}=0.05$ s，用户本地处理能力为 $f^{\text{loc}}=100$ CPU cycles/s，无线子载波带宽为 $B=180\times 10^2$ Hz，MEC 服务器所拥有的计算资源块总个数为 $U=150$ 个，无线资源单价为 $\phi_1=1$ 元/(Hz·s)，MEC 服务器计算资源块单价为 $\phi_2=12$ 元，MEC 服

务器每个计算资源块的处理能力为 $f^{\text{mit}}=10^9$ CPU cycles/s，用户初始位置随机生成，AP 高度为 $l_g=20$ m，小区覆盖直径为 $D_g=500$ m。

当车辆的行驶速度不同时，速度与总成本比较如图 3 所示，随着车辆行驶速度的提升，当 MEC 服务器将接收的任务处理完成时，车辆可能会行驶经过多个接入点，导致计算结果回传过程中使用的无线资源费用增加，进一步增加了系统的总成本。从图 3 中可以看出，每条线的趋势和本文的分析情况一致。通过与随机卸载机制、随机计算资源分配机制和随机分簇机制 3 种情况进行对比，可以看出本文所提算法的性能是最好的，能够最大限度降低系统成本。

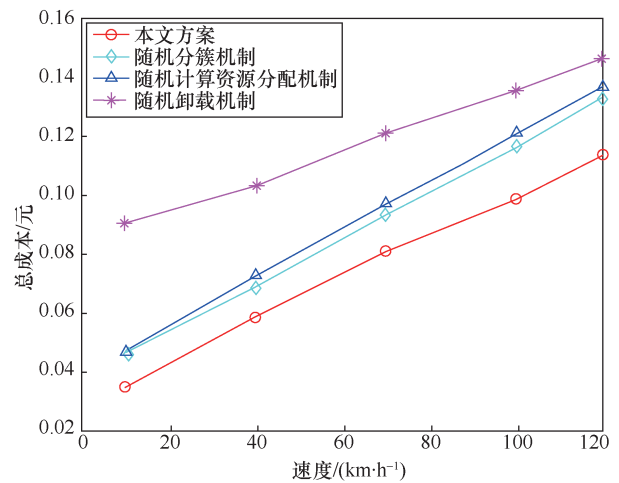


图 3 速度与总成本比较

任务处理最大时延对系统总成本的影响如图 4 所示。根据计算资源分配算法可以得到每个车辆用户分配的计算资源块个数，当任务处理最大时延增大时，车辆用户分配到的计算资源的个数会减少，那么使用计算资源的费用就会降低，进一步地使系统的总成本降低。由于随机分簇机制的分簇是随机的，而本文提出的算法根据要卸载的车辆用户的信道增益的降序排列后进行分簇，两种分簇方式的不同导致随机分簇的总成本比本文提出的分簇算法的总成本大。在随机计算资源分配机制中，分配给用户的计算资源块个数是随机的，而本文中的计算资源分配是在判断其是否满足最大任务处理时延的情况下进行分配的，随机分配的计算资源个数可能比本文提出的算法分配的计算资源个数多，所以随机计算资源分配机制的成本大于本文方案的成本。

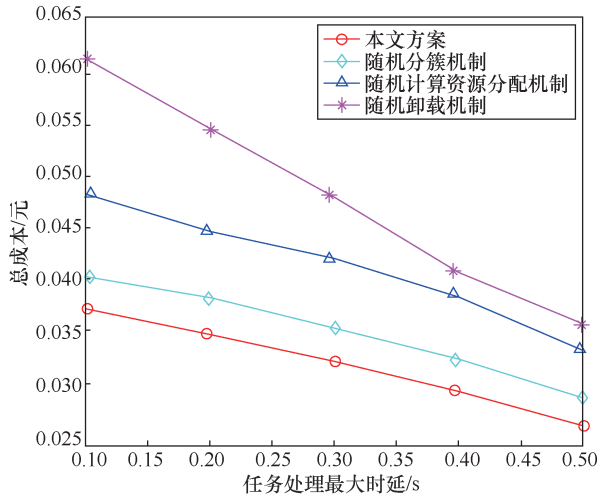


图 4 任务处理最大时延对系统总成本的影响

无线资源的单价对系统总成本的影响如图 5 所示。在本文提出的系统模型中，系统总成本为车辆用户使用计算资源成本和无线资源成本之和，当使用无线资源的单价增加时，由于其他条件不变，所以系统的总成本会增加，单位价格越高，总成本也会越高。在随机卸载机制中，由于其卸载方式与本文所提算法中的卸载方式不同，所以可能使需要卸载的车辆用户数增加，从而导致使用的无线资源增加，使得系统的总成本大于本文所提算法。通过对比可以看出，本文所提算法的成本最低。

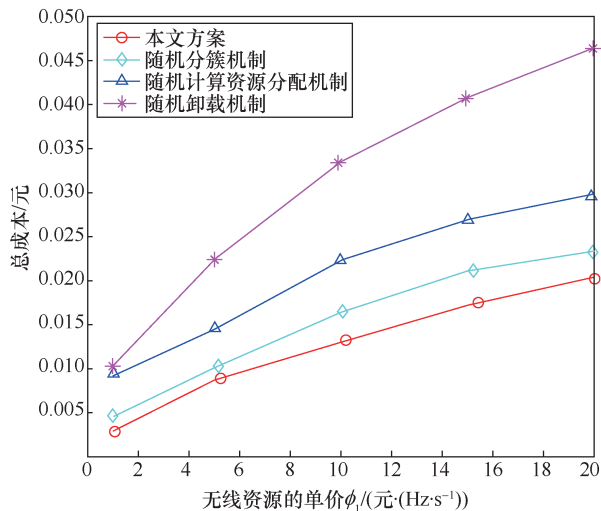


图 5 无线资源的单价对系统总成本的影响

计算资源的单价对系统总成本的影响如图 6 所示。由系统模型可知，系统的总成本会随着使用计算资源的单价的增加而增加。在随机计算资源分配机制中，由于其分配计算资源的方式与本文提出的方式不同，所以可能导致给车辆用户分配的计算资

源块的个数增加，从而导致系统的总成本比本文所提算法的成本高。通过与其他 3 种机制进行对比可以看出，本文提出的算法最有效。

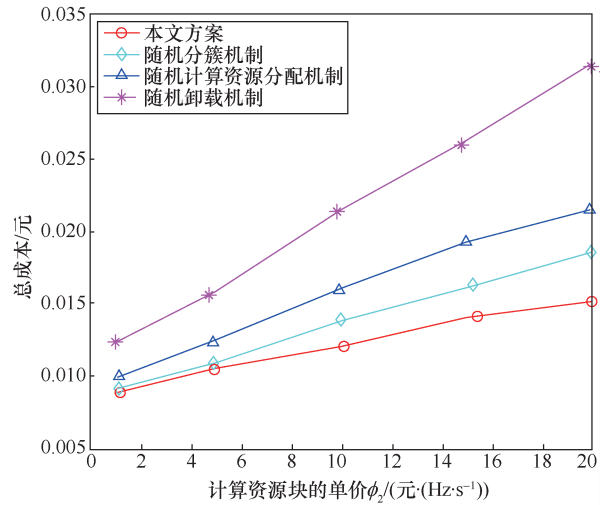


图 6 计算资源的单价对系统总成本的影响

5 结束语

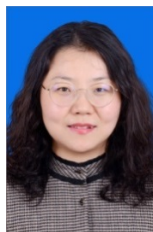
本文研究了基于 NOMA 的 VEC 网络，以任务执行期限为前提，对卸载决策、用户分簇、计算资源分配以及发送功率控制进行联合优化，以最大限度地降低任务处理成本。为了有效解决该问题，将优化问题分解为子问题，先进行卸载决策和用户分簇的优化，基于上述优化的结果再优化发送功率，最后进行计算资源分配。通过两轮计算资源分配，在保证用户任务处理时延的前提下，有效降低了计算资源的使用成本，并解决了车辆移动性的问题。仿真结果表明，本文提出的联合优化方案在保证最大时延的前提下能有效降低任务处理的成本。

参考文献:

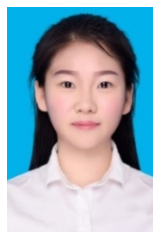
- [1] CHEN C, WANG C, QIU T, et al. Caching in vehicular named data networking: architecture, schemes and future directions[J]. IEEE Communications Surveys & Tutorials, 2020, 22(4): 2378-2407.
- [2] LAI C Z, LU R X, ZHENG D, et al. Security and privacy challenges in 5G-enabled vehicular networks[J]. IEEE Network, 2020, 34(2): 37-45.
- [3] CHEN C, HU J N, QIU T, et al. CVCG: cooperative V2V-aided transmission scheme based on coalitional game for popular content distribution in vehicular Ad-hoc networks[J]. IEEE Transactions on Mobile Computing, 2019, 18(12): 2811-2828.
- [4] LIU L, CHEN C, PEI Q Q, et al. Vehicular edge computing and networking: a survey[J]. Mobile Networks and Applications, 2020: 1-24.
- [5] CHEN X, SHI Q, YANG L, et al. Thrifty edge: resource-efficient edge

- computing for intelligent IoT applications[J]. IEEE Network, 2018, 32(1): 61-65.
- [6] CHEN X, LI W Z, LU S L, et al. Efficient resource allocation for on-demand mobile-edge cloud computing[J]. IEEE Transactions on Vehicular Technology, 2018, 67(9): 8769-8780.
- [7] DU J B, YU F R, LU G Y, et al. MEC-assisted immersive VR video streaming over terahertz wireless networks: a deep reinforcement learning approach[J]. IEEE Internet of Things Journal, 2020, 7(10): 9517-9529.
- [8] LIU L, FENG J, PEI Q Q, et al. Blockchain-enabled secure data sharing scheme in mobile-edge computing: an asynchronous advantage actor-critic learning approach[J]. IEEE Internet of Things Journal, 2021, 8(4): 2342-2353.
- [9] FENG J, RICHARD Y F, PEI Q Q, et al. Cooperative computation offloading and resource allocation for blockchain-enabled mobile-edge computing: a deep reinforcement learning approach[J]. IEEE Internet of Things Journal, 2020, 7(7): 6214-6228.
- [10] FENG J, YU F R, PEI Q Q, et al. Joint optimization of radio and computational resources allocation in blockchain-enabled mobile edge computing systems[J]. IEEE Transactions on Wireless Communications, 2020, 19(6): 4321-4334.
- [11] DU J B, YU F R, CHU X L, et al. Computation offloading and resource allocation in vehicular networks based on dual-side cost minimization[J]. IEEE Transactions on Vehicular Technology, 2019, 68(2): 1079-1092.
- [12] LI X W, LI J J, LIU Y W, et al. Residual transceiver hardware impairments on cooperative NOMA networks[J]. IEEE Transactions on Wireless Communications, 2020, 19(1): 680-695.
- [13] LI X W, ZHAO M L, LIU Y W, et al. Secrecy analysis of ambient backscatter NOMA systems under I/Q imbalance[J]. IEEE Transactions on Vehicular Technology, 2020, 69(10): 12286-12290.
- [14] DING Z G, FAN P Z, POOR H V. Impact of non-orthogonal multiple access on the offloading of mobile edge computing[J]. IEEE Transactions on Communications, 2019, 67(1): 375-390.
- [15] ZHAI D S, LI H, TANG X, et al. Height optimization and resource allocation for NOMA enhanced UAV-aided relay networks[J]. IEEE Transactions on Communications, 2020(99): 1.
- [16] ZHAI D S, ZHANG R N, CAI L, et al. Energy-efficient user scheduling and power allocation for NOMA-based wireless networks with massive IoT devices[J]. IEEE Internet of Things Journal, 2018, 5(3): 1857-1868.
- [17] LIU B H, LIU C X, PENG M G. Resource allocation for energy-efficient MEC in NOMA-enabled massive IoT networks[J]. IEEE Journal on Selected Areas in Communications, 2020(99): 1.
- [18] FANG F, XU Y Q, DING Z G, et al. Optimal resource allocation for delay minimization in NOMA-MEC networks[J]. IEEE Transactions on Communications, 2020, 68(12): 7867-7881.
- [19] FANG F, XU Y Q, PHAM Q V, et al. Energy-efficient design of IRS-NOMA networks[J]. IEEE Transactions on Vehicular Technology, 2020, 69(11): 14088-14092.
- [20] YANG C, LIU Y, CHEN X, et al. Efficient mobility-aware task offloading for vehicular edge computing networks[J]. IEEE Access, 2019(7): 26652-26664.
- [21] WANG J, FENG D Q, ZHANG S L, et al. Computation offloading for mobile edge computing enabled vehicular networks[J]. IEEE Access, 2019(7): 62624-62632.
- [22] ZHOU J, WU F, ZHANG K, et al. Joint optimization of offloading and resource allocation in vehicular networks with mobile edge computing[C]//Proceedings of 2018 10th International Conference on Wireless Communications and Signal Processing (WCSP). Piscataway: IEEE Press, 2018: 1-6.
- [23] 张璐. NOMA技术与车联网V2I通信结合系统性能仿真研究[J]. 信息通信, 2019, 32(2): 285-289.
- ZHANG L. Performance simulation of the combination of NOMA and V2I communication[J]. Information & Communications, 2019, 32(2): 285-289.
- [24] FANG S S, CHEN H Y, KHAN Z, et al. On the content delivery efficiency of NOMA assisted vehicular communication networks with delay constraints[J]. IEEE Wireless Communications Letters, 2020, 9(6): 847-850.
- [25] PEIRAN D, ZHAOLONG N, RONG M, et al. NOMA-based energy-efficient task scheduling in vehicular edge computing networks: a self-imitation learning-based approach[J]. China Communications, 2020, 17(11): 1-11.

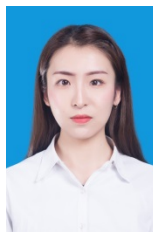
[作者简介]



杜剑波（1982-），女，博士，西安邮电大学讲师，主要研究方向为边缘计算、非正交多址、人工智能等。



薛哪哪（1999-），女，西安邮电大学通信与信息工程学院硕士生，主要研究方向为边缘计算、非正交多址等。



孙艳（1996-），女，西安邮电大学通信与信息工程学院硕士生，主要研究方向为边缘计算、非正交多址等。

姜静（1974-），女，博士，西安邮电大学教授、博士生导师，主要研究方向为无线通信、大规模 MIMO 等。

李树磊（1981-），男，天元瑞信通信技术股份有限公司总工程师，主要研究方向为移动网络通信、区块链、边缘计算等。

卢光跃（1971-），男，博士，西安邮电大学教授、博士生导师，主要研究方向为认知无线电、携能通信和边缘计算等。