

车联网中基于移动边缘计算的内容感知分类卸载算法研究

赵海涛* 朱银阳 丁仪 朱洪波

(教育部泛在网络健康服务系统工程研究中心 南京 210003)

(江苏省无线通信重点实验室 南京 210003)

(南京邮电大学通信与信息工程学院 南京 210003)

摘要: 随着智能交通的快速发展, 车辆终端产生大量需要实时处理的数据消息, 而在有限资源上的竞争将会增加消息处理的时延, 且对终端设备造成很大的能量消耗。针对时延和能量损耗的均衡关系, 该文提出一种基于移动边缘计算(MEC)的内容感知分类卸载算法。首先根据层次分析法对安全消息进行优先级划分, 然后建立时延和能量损耗的最优任务卸载模型, 通过给时延和能量损耗赋予不同的权重系数构造关系模型, 并利用拉格朗日松弛法将非凸问题转化为凸问题, 从而结合次梯度投影法和贪婪算法得到问题的可行解。性能评估结果表明, 该算法在一定程度上改善了消息处理时延和能量损耗。

关键词: 移动边缘计算; 计算卸载; 消息优先级; 时延

中图分类号: TP399

文献标识码: A

文章编号: 1009-5896(2020)01-0020-08

DOI: [10.11999/JEIT190594](https://doi.org/10.11999/JEIT190594)

Research on Content-aware Classification Offloading Algorithm Based on Mobile Edge Calculation in the Internet of Vehicles

ZHAO Haitao ZHU Yinyang DING Yi ZHU Hongbo

(*Ministry of Education Ubiquitous Network Health Service System Engineering
Research Center, Nanjing 210003, China*)

(*Jiangsu Key Wireless Communication Laboratory, Nanjing 210003, China*)

(*College of Communication and Information Engineering, Nanjing University of
Posts and Telecommunications, Nanjing 210003, China*)

Abstract: With the rapid development of intelligent transportation, vehicle terminals generate a large number of data messages that need to be processed in real time. Competition on limited resources will increase the delay of message processing and energy consumption for terminal equipment. For the equilibrium relationship between delay and energy loss, this paper proposes a content-aware classification offloading algorithm based on Mobile Edge Computing (MEC). Firstly, the security message is prioritized according to the analytic hierarchy process, and then the optimal task unloading model of delay and energy loss is established. The relational model is established by assigning different weight coefficients to delay and energy loss. The Lagrangian relaxation method is used to transform the non-convex problem into a convex problem, which combines the sub-gradient projection method and the greedy algorithm to obtain the feasible solution. The performance evaluation results show that the algorithm improves the message processing delay and energy loss to some extent.

Key words: Mobile Edge Computing (MEC); Computational offloading; Message priority; Delay

收稿日期: 2019-08-06; 改回日期: 2019-11-14; 网络出版: 2019-11-28

*通信作者: 赵海涛 zhaoh@njupt.edu.cn

基金项目: 国家自然科学基金(61771252), 江苏省自然科学基金面上项目(BK20171444), 江苏省高校重点自然科学基金研究重大项目(18KJA510005), 江苏省“六大人才高峰”B类资助项目(DZXX-041), 江苏省科协青年科技人才托举工程资助培养项目, 江苏省研究生科研创新计划项目(KYCX19_0949)

Foundation Items: The National Natural Science Foundation of China (61771252), The Natural Science Foundation Project of Jiangsu Province (BK20171444), The University Natural Science Research Major Project of Jiangsu Province (18KJA510005), The "Six Talents High Peaks" Class B Funding Project of Jiangsu Province (DZXX-041), The Jiangsu Provincial Association for Science and Technology Talents Entrustment Project, Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX19_0949)

1 引言

随着智能交通的快速发展,车辆数量的不断增加导致了許多交通堵塞和交通事故等交通安全问题。车联网技术的主要目的是通过车辆间通信(Virtual to Virtual, V2V)和公共基础设施(Vehicle-to-Infrastructure, V2I)的通信来增强驾驶员安全性,保证车辆的道路行驶安全^[1]。目前车联网的发展重点已逐渐从核心网转向边缘网,在靠近车载终端的网络边缘分布式地部署各类网络资源,以尽可能地将各类消息任务从云平台迁移到边缘侧进行响应实现。但是,如今移动车载终端的业务环境更加复杂,消息的类型更加繁多,尤其是具有低时延、高可靠要求的安全类业务对计算资源的要求更加苛刻,例如刹车警告、违法交通法规和交通事故等,需要及时高效地处理^[2]。然而,移动车辆终端有限的资源计算能力使消息处理变得极为困难,一方面是有限资源的竞争会增加消息处理的时延,安全消息对网络带宽和传输速率有很高的要求,不能将消息的处理工作全部部署在云端,而需要考虑将消息任务部分卸载至边缘侧服务器,并以队列的形式保存消息。另一方面是对车载终端设备造成较大的能量消耗,当部分来自终端设备的工作任务部署到边缘服务器时,车载终端设备需要损耗一定的传输带宽和传输功率进行卸载任务,通过在边缘服务器中部署任务调度算法,以降低消息处理的时延,同时可以减少消息任务处理的能量损耗^[3]。这些问题给车辆终端处理消息带来了极大的负载,不能保证有效及时地处理这些安全消息,无法满足用户的服务需求。

国内外已经有大量学者和研究机构对边缘计算进行了研究。文献^[4]针对目前车辆终端产生的数据量极大,拓扑结构快速变化以及流量不平衡的问题,介绍了一种基于边缘计算的分层架构,可以有效地分发大容量车载数据。文献^[5]考虑用户的安全隐私,提出了一种基于Lyapunov优化的隐私感知计算卸载方法,可以使用户的累积隐私量稳定在零附近,且总卸载频率与不考虑隐私的决策一致,有效保护了用户隐私,同时保持了较低的平均能耗。文献^[6]研究了在移动边缘计算(Mobile Edge Computing, MEC)中有限电源的物联网设备的能量损耗问题,提出了一种卸载决策算法,该算法分析计算策略中的权衡,提供了以高分辨率生成的视觉数据的低延迟处理。文献^[7]研究了车辆异构网络中基于MEC的任务卸载模式决策和资源分配问题。文献^[8]研究了移动边缘计算中的能效与性能可靠性问题,提出了一种能量最优化的卸载方案,其

在能耗和延迟性能方面优于局部计算和完全卸载方法。文献^[9]考虑系统总能耗,提出了卸载决策和资源分配的联合优化问题,可以在满足用户不同时延的要求下最小化系统能耗,有效地提升了系统性能。文献^[10]提出了一种排队网络模型,可以基于排队网络理论来分析用户之间的交互,并根据每种服务类型不同的优先级,降低平均延迟。文献^[11]解决了延迟敏感网络服务中的动态任务卸载和调度问题,在保证性能的同时降低了运行时间。文献^[12]研究了基于时分多址(Time Division Multiple Access, TDMA)和正交频分多址(Orthogonal Frequency Division Multiple Access, OFDMA)的多用户MECO系统的资源分配。文献^[13]考虑不同的优先级来优先处理边缘服务以实现最佳服务卸载,设计了一种自适应服务卸载方案,为MEC提供最大的收益和服务利用率。文献^[14]研究了移动设备在有限资源上的竞争将固有地增加用户所经历的延迟,提出了一种优先服务调度方案,以减少延迟敏感服务所经历的延迟。但是这些算法主要考虑了设备能耗问题,关于保证消息性能的研究较少,虽然节约能耗对MEC系统很重要,但对于安全消息,保证消息的可靠性和减少传输时延更加重要。

本文针对上述研究中存在的问题,提出一种基于移动边缘计算的内容感知分类卸载算法。通过引入移动边缘计算,使车辆终端产生的消息能够直接在边缘节点进行处理,并对安全消息进行优先级划分。在移动车辆端,研究了时延和能量最优的任务卸载策略,考虑时延和能量消耗之间的关系问题,在边缘器端部署了一种有效的优先级任务调度算法,在时延和能量损耗方面有一定的改善,保证优先级别最高的安全消息得到最及时地处理,从而改善用户的使用体验。

2 系统架构

本文选取车辆、路边单元与其部署的边缘服务器所构成的通信网络架构,在高速公路场景下,车辆的高速移动使得网络架构的拓扑快速变化,同时车辆会产生各种各样的安全消息^[15]。这会带来两个方面的问题,第1个问题是消息处理的时延问题,若车辆将消息上传到云处理则会造成非常大的时延问题,大大影响用户的使用体验,为了减小时延,本文将消息任务卸载到边缘服务器进行处理;第2个问题是消息处理的优先级问题,安全消息的类型不同,其处理优先级就不同,将消息统一卸载到边缘服务器就会带来优先级问题,因此本文将消息任务部分卸载至边缘侧服务器,以队列的形式保存

消息,并在边缘服务器中部署优先级任务调度算法,进一步降低消息的处理延迟,保证最紧急的消息获得最高的优先处理级别。本文提出了如图1所示的内容感知分类卸载的系统架构。

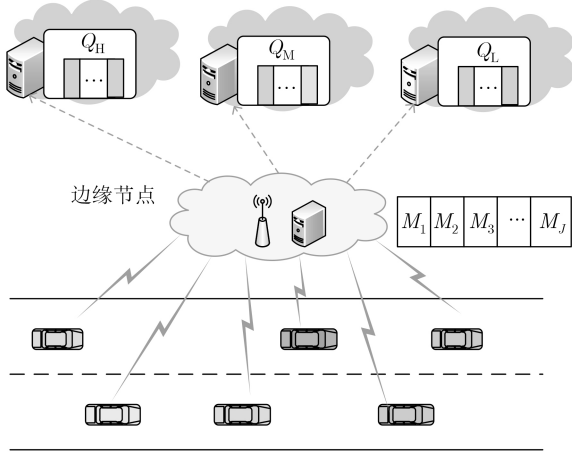


图1 系统架构

该系统架构由3层组成:移动车辆端、路边单元侧和边缘服务器层。在系统架构中,路边单元(RoadSide Unit, RSU)与车辆终端和边缘服务器进行通信,但由于移动车辆终端有限的计算能力,车辆终端将产生的安全消息以均匀分布的方式传输到附近的RSU中,在RSU侧部署具有强大计算能力的边缘服务器,其中消息任务将以队列的形式在服务器中保存,通过任务分类卸载将消息任务分配到对应的边缘服务器中。每个边缘服务器具有有限的通信带宽,为了提高边缘服务器可用带宽资源的利用率,每个边缘服务器具有动态计算消息队列,其负责保存不同类型的安全消息,消息队列主要取决于几个关键因素,例如消息任务到达率、最大延迟期限要求和车辆终端的计算能力可用性等,在将一组消息任务卸载到这些边缘服务器过程中,边缘服务器向RSU边缘节点通知消息队列的优先级。

假设有 N 个移动车辆,每个车辆终端产生 j 个安全消息 $M = \{M_1, M_2, \dots, M_j\}$,其中 $j \in J, j = \{1, 2, \dots, J\}$,移动车辆产生的安全消息 M_j 可以通过一组参数建模,即 $M_j = \{b_j, C_j, T_j, P_j\}$,其中 b_j, C_j, T_j 和 P_j 分别表示消息数据大小、消息所需的CPU周期、截止期限要求和消息的优先级。

3 基于移动边缘计算的内容感知分类卸载方案

3.1 消息优先级划分

将车辆终端产生的安全消息卸载到边缘服务器时,由于消息的类型和对计算能力要求不一致等因

素,需要考虑对安全消息进行优先级划分。在划分安全消息优先级时,本文将利用层次分析法(Analytic Hierarchy Process, AHP)对安全消息进行优先级划分,层次分析法是一个多标准决策(Multiple Criteria Decision Making, MCDM)/多属性决策(Multi-Attribute Decision-Making, MADM)模型,AHP在各个领域都有很多应用,是一种适用于解决基于优先级调度的方法^[6]。

假设在 $0 \sim t$ 时间内,每辆车产生独立的安全消息 $M = \{M_1, M_2, \dots, M_j\}, j \in J, j = \{1, 2, \dots, J\}$,在确定消息的优先级时,本文主要考虑消息的数据大小、消息所需的CPU周期和截止期限要求这3个因素,其中截止期限要求相比消息数据大小、消息所需的CPU周期更重要,消息所需的CPU周期比消息数据大小重要。因此,在层次分析模型中,截止期限要求对于优先级的划分所占的权重最高。首先将同一层次的因素进行两两比较,构造层次分析矩阵 $\mathbf{A} = (a_{ij})_{n \times n}$,其中

$$a_{ij} = \begin{cases} \frac{1}{a_{ji}} = n, n = \{1, 2, \dots, 9\}, i \neq j \\ 1, & i = j \end{cases} \quad (1)$$

然后,本文计算比较矩阵权重向量时,使用最常用求和方法。权重所对应的矢量计算公式为

$$U_r^k = \frac{\sum_{j=1}^n a_{rj}}{\sum_{i=1}^n \sum_{j=1}^n a_{ij}} \quad (2)$$

其中, k 表示在决策过程中考虑的影响因素的个数, r 表示车辆产生的消息。则所有消息的权重所对应的矢量组成矩阵 Δ ,如式(3)

$$\Delta = \begin{bmatrix} u_1^1 & u_1^2 & u_1^3 \\ u_2^1 & u_2^2 & u_2^3 \\ \vdots & \vdots & \vdots \\ u_j^1 & u_j^2 & u_j^3 \end{bmatrix} \quad (3)$$

接着根据层次分析矩阵获得其权重所对应的特征值,用 Λ 表示,即 $\Lambda = [\lambda_1, \lambda_2, \lambda_3]^T$,其中

$$\lambda_k = \frac{1}{k} \sum_{i=1}^k \frac{\sum_{j=1}^k a_{rj}}{k}, k = 1, 2, 3 \quad (4)$$

最后得到每个消息的优先级向量, \mathbf{PV} 向量中的每个元素就代表消息的优先级值。它可以通过等式(5)计算为

$$\mathbf{PV} = \mathbf{\Delta} \times \mathbf{A} = \begin{bmatrix} u_1^1 & u_1^2 & u_1^3 \\ u_2^1 & u_2^2 & u_2^3 \\ \vdots & \vdots & \vdots \\ u_j^1 & u_j^2 & u_j^3 \end{bmatrix} \cdot [\lambda_1, \lambda_2, \lambda_3]^T = \begin{pmatrix} \sum_{k=1}^3 u_1^k \lambda_k \\ \sum_{k=1}^3 u_2^k \lambda_k \\ \vdots \\ \sum_{k=1}^3 u_j^k \lambda_k \end{pmatrix} \quad (5)$$

根据式(5)中 \mathbf{PV} 值可以获得每个消息任务的权重,也就是消息的优先级。

3.2 资源队列模型

每个边缘服务器具有有限的通信带宽,为了提高边缘服务器可用带宽资源的利用率,每个边缘服务器具有动态计算消息队列,其负责保存不同类型的安全消息,消息队列主要取决于消息任务到达率和边缘服务器的计算能力,在将一组消息任务卸载到边缘服务器过程中,边缘服务器向RSU边缘节点通知消息队列的其所对应的优先级。

在本文中,边缘节点采用 $M/M/s$ 的排队模型以处理安全消息, s 表示边缘服务器的数目。假设消息 j 的平均到达率为 λ_j 的泊松分布,边缘节点计算能力大小为 $\mu_i (i=1,2,\dots,s)$ 的指数分布,则 $M/M/s$ 系统的服务强度为

$$\rho_{ji} = \frac{\lambda_j}{i\mu_i} \quad (6)$$

为了保持系统稳定性, ρ_{ji} 必须小于等于1,相应的平衡分布为

$$\pi_k = \begin{cases} \frac{(i\rho_{jk})^k}{k!} \pi_0, & 0 \leq k \leq i-1 \\ \frac{\rho_{ji}^i}{i!} \pi_0, & k \geq i \end{cases} \quad (7)$$

其中,

$$\pi_0 = \left[\sum_{k=0}^{i-1} \frac{(k\rho_{ji})^k}{k!} + \frac{i^i \rho_{ji}^i}{i!(1-\rho_{ji})} \right]^{-1} \quad (8)$$

消息任务 j 分配给对应的边缘服务器 i 的等待时间为

$$t_{ij} = \frac{\rho_{ji}}{\lambda_j(1-\rho_{ji})^2} \pi_i \quad (9)$$

3.3 卸载模型

在卸载过程中,终端设备能量损耗是本文考虑的关键问题之一^[17],但对于安全消息来说,由于其实时性要求较高,时延问题就显得更加重要了。因此,

在本文中,主要研究计算卸载时延和能量消耗之间的关系问题,希望在计算能力、传输带宽和任务等待时间要求的限制下,同时使能量消耗最小化。

$i = \{1, 2, \dots, I\}$ 分别表示存储不同种消息类型的边缘服务器,对于每个边缘服务器 i ,其通信带宽由 W_i 表示,最大计算速率由 V_i 表示。本文考虑在周期 t 时间内安全消息的卸载问题,在 $0 \sim t$ 时间内,安全消息 j 卸载到边缘服务器 i 时,边缘服务器 i 为消息 j 分配的通信带宽为 w_{ij} ,分配的计算速率由 v_{ij} 表示,在周期时间 $[1, t]$ 内,卸载到 i 的所有任务的总所需通信带宽和计算速率不应分别超过资源量 W_i 和 V_i 。

由于消息任务可能无法立即卸载,因此需要在边缘节点等待一段时间才能传输到相应的边缘服务器中。因此,系统卸载的总时延包括等待延迟、传输延迟和计算延迟,当在时段 t 将消息任务 j 卸载到边缘服务器 i 时,总延迟通过式(10)计算

$$d_{i,j,t} = t_{ij} + \frac{b_j}{r_{i,j}} + \frac{c_j}{v_{i,j}} \quad (10)$$

其中, t_{ij} 是边缘节点RSU进行调度之前消息任务 j 的等待时间, $b_j/r_{i,j}$ 是传输延迟, $c_j/v_{i,j}$ 是计算延迟,表示消息任务 j 与边缘服务器 i 之间的传输功率, c_j 表示消息任务的卸载大小。

在计算卸载过程中,移动车辆的能量消耗一般包含消息任务卸载传输消耗和计算处理消耗,本文将消息任务卸载到边缘服务器中进行处理,并定义一个函数 $h(x) = N_0 w_{ij} \left(2^{\frac{x}{w_{ij}}} - 1 \right)$,因此总的能量损耗可以表示为

$$p_{i,j} = \frac{b_j}{r_{i,j} S_{ij}^2} h(r_{i,j}) + \frac{c_j}{v_{i,j}} e_i \quad (11)$$

其中, S_{ij} 表示边缘服务器 i 与消息任务 j 之间的信道增益, e_i 表示边缘服务器 i 每CPU周期的功耗。

为了探讨计算卸载时延和能量消耗之间的关系问题,我们给卸载时延和能量消耗赋予不同的权重系数,主问题可以制定为

$$\begin{aligned} & \min \sum_{i \in I} \sum_{j \in J} \sum_{t \in T} (\theta_1 \cdot d_{i,j,t} + \theta_2 \cdot p_{i,j}) \cdot x_{i,j,t} \\ & \text{s.t.} \\ & C_1: \sum_{m \in [1,t]} \sum_{j \in J} w_{i,j,m} \cdot x_{i,j,m} \leq W_i, \forall i \in I, \forall t \in T \\ & C_2: \sum_{m \in [1,t]} \sum_{j \in J} v_{i,j,m} \cdot x_{i,j,m} \leq V_i, \forall i \in I, \forall t \in T \\ & C_3: \sum_{i \in I} \sum_{t \in T} x_{i,j,t} = 1, \forall j \in J \\ & C_4: x_{i,j,t} \in \{0, 1\}, \forall i \in I, \forall j \in J, \forall t \in T \\ & C_5: \frac{b_j}{r_{i,j} S_{ij}^2} h(r_{i,j}) + \frac{c_j}{v_{i,j}} e_i \leq C_j, \forall i \in I, \forall j \in J \end{aligned} \quad (12)$$

约束 C_1 要求从时间1到 t 内卸载到边缘服务器 i 的所有任务的占用通信带宽之和不能超过所有周期 t 的累积可用通信资源 W_i ; C_2 要求边缘服务器 i 从周期1到 t 的所有任务的所需计算速率之和不能超过所有周期 t 的累积可用计算资源 V_i ; C_3 指出每个任务 j 可以仅在一个周期 t 被分配给恰好一个边缘服务器 i ; C_4 是二进制变量约束, $x_{i,j,t}=1$ 表示消息任务 j 在 $0\sim t$ 时间内被卸载到相应的边缘服务器 i 中, 否则 $x_{i,j,t}=0$; C_5 表示任务消息的实际能量消耗必须小于预期的消息所需的CPU周期。

θ_1 和 θ_2 是延迟和能量损耗的权重, 且 $\theta_1+\theta_2=1$, 可以通过改变目标函数中的权重系数, 研究能量损耗和延迟之间的关系。例如, 当为延迟设置较大的权重时, 该问题的主要目标是最小化总延迟, 所提出的算法倾向于使用更大的传输功率来分配任务。相反, 当对能量损耗设置较大的权重时, 该问题的主要目标就是最小化能量损耗, 则它将倾向于使用最小传输功率来分配任务。由于该问题含有两个决策变量 $d_{i,j,t}$ 和 $p_{i,j}$, 并且是两个变量的乘积, 因此该问题是非凸问题, 本文首先要通过拉格朗日松弛法将非凸问题转化为凸问题^[18]。其拉格朗日松弛函数如式(13)

$$\begin{aligned} L(x, \lambda) = & \sum_{i \in I} \sum_{j \in J} \sum_{t \in T} (\theta_1 \cdot d_{i,j,t} + \theta_2 \cdot p_{i,j}) \cdot x_{i,j,t} \\ & + \sum_{i \in I} \sum_{t \in T} \left(\lambda_{i,t}^0 \cdot \left(\sum_{m \in [1,t]} \sum_{j \in J} w_{i,j} \cdot x_{i,j,m} - W_i \right) \right. \\ & \left. + \lambda_{i,t}^1 \cdot \left(\sum_{m \in [1,t]} \sum_{j \in J} v_{i,j} \cdot x_{i,j,m} - V_i \right) \right) \\ & + \sum_{j \in J} \left(\lambda_j^2 \cdot \left(\sum_{i \in I} \sum_{t \in T} x_{i,j,t} - 1 \right) \right) \\ & + \sum_{i \in I} \sum_{j \in J} \left(\lambda_{i,j}^3 \cdot \left(\frac{b_j}{r_{i,j} S_{i,j}^2} h(r_{i,j}) + \frac{c_j}{v_{i,j}} e_i - C_j \right) \right) \end{aligned} \quad (13)$$

其中, $\lambda = \{\lambda^0, \lambda^1, \lambda^2, \lambda^3\}$, λ^0 与 $\lambda^1 \in R^{I \times T}$, $\lambda^2 \in R^J$, $\lambda^3 \in R^{I \times J}$, $\lambda \geq 0$ 。因此该问题就转换成求解如式(14)的最优化问题

$$g(\lambda) = \min_x L(x, \lambda) \quad (14)$$

容易知道函数 $g(\lambda)$ 是一个凸函数, 且 $g(\lambda) \leq L(x, p, \lambda)$, 所以上述问题进一步表述为其对偶问题

$$\begin{aligned} \max_{\lambda} g(\lambda) = & \max_{\lambda} [\min_x L(x, \lambda)], \\ \text{s.t. } & \lambda^0 \text{ 与 } \lambda^1 \in R^{I \times T} \\ & \lambda^2 \in R^J \\ & \lambda^3 \in R^{I \times J} \\ & \lambda \geq 0 \end{aligned} \quad (15)$$

$g(\lambda)$ 的次梯度投影为

$$\lambda_{i,t}^0(t+1) = \left[\lambda_{i,t}^0(t) - l_{i,t}^0 \cdot \left(\sum_{m \in [1,t]} \sum_{j \in J} w_{i,j} \cdot x_{i,j,m} - W_i \right) \right]^+ \quad (16)$$

$$\lambda_{i,t}^1(t+1) = \left[\lambda_{i,t}^1(t) - l_{i,t}^1 \cdot \left(\sum_{m \in [1,t]} \sum_{j \in J} v_{i,j} \cdot x_{i,j,m} - V_i \right) \right]^+ \quad (17)$$

$$\lambda_j^2(t+1) = \lambda_j^2(t) - l_j^2 \cdot \left(\sum_{i \in I} \sum_{t \in T} x_{i,j,m} - 1 \right) \quad (18)$$

$$\lambda_{i,j}^3(t+1) = \left[\lambda_{i,j}^3(t) - l_{i,j}^3 \cdot \left(\frac{b_j}{r_{i,j} S_{i,j}^2} h(r_{i,j}) + \frac{c_j}{v_{i,j}} e_i - C_j \right) \right]^+ \quad (19)$$

其中, $l_{i,t}^0$, $l_{i,t}^1$, $l_{i,t}^2$ 和 $l_{i,t}^3$ 分别是 λ^0 , λ^1 , λ^2 和 λ^3 所对应的步长。

在获得每个消息所对应的优先级后, 首先根据任务的优先级在消息队列中进行任务调度分配。消息卸载的边缘服务器具有高、中和低3个优先级别, 即 Q_H , Q_M 和 Q_L , 并且在每个边缘服务器中进一步按照消息优先级进行队列保存, 消息任务有3种类型, 即 $p_j = 1, 2, 3$, 其中3为高, 2为中, 1为低。因此, 所有具有严格延迟要求的消息应该被卸载到高优先级的边缘服务器队列中, 具有宽松延迟要求的消息将被卸载到中优先级边缘服务器的队列或低优先级边缘服务器的队列中, 主要是为了优先处理最紧急的安全消息。其具体算法流程如表1所示。

在完成任务队列调度后, 卸载的主问题采用次梯度投影和启发式贪婪算法得到可行解, 在次梯度投影中, 拉格朗日乘数的值以某些步长更新, 并且传输功率也根据KKT条件更新, 从计算复杂度考虑, 本文使用贪婪算法解决整数规划问题, 求得分配系数 x 和目标函数值。根据上述推导, 制定了消息任务卸载策略, 如表2所示。

4 仿真结果与分析

本节对所提出的基于移动边缘计算的消息内容感知分类卸载方案进行了仿真以评估其性能, 为了评估移动边缘计算场景中的不同任务分配算法, 本

表1 任务队列调度算法

(1)	输入消息的数据大小、消息所需的CPU周期、截止时间要求和消息的优先级 b_j , C_j , T_j 和 P_j ;
(2)	for 边缘服务器中的每个安全消息 M_j ;
(3)	if $p_j=3$, 则
(4)	将消息 M_j 放置在 Q_H 队列中;
(5)	构建层次分析矩阵 $A=(a_{ij})_{n \times n}$;
(6)	计算影响因素所对应的权重向量 U_r^k ;
(7)	根据层次分析矩阵获得其权重所对应的特征值 $\lambda = [\lambda_1, \lambda_2, \lambda_3]^T$;
(8)	通过 $PV = \Delta \times A$ 得到每个消息的优先级向量, 即消息的优先级值;
(9)	根据 PV 值的大小在 Q_H 队列中按顺序排列;
(10)	else if $p_j=2$, 则
(11)	将消息 M_j 放置在 Q_M 队列中;
(12)	重复步骤(4)–步骤(7);
(13)	根据 PV 值的大小在 Q_M 队列中按顺序排列;
(14)	else if $p_j=1$, 则
(15)	将消息 M_j 放置在 Q_L 队列中;
(16)	重复步骤(4)–步骤(7);
(17)	根据 PV 值的大小在 Q_L 队列中按顺序排列;
(18)	End if;
(19)	End for;
(20)	End

表2 消息任务卸载策略

(1)	输入: 任务集 M , 边缘计算服务器集 I , 分配的通信带宽为 w_{ij} , 分配的计算速率由 v_{ij} ;
(2)	输出: 分配系数 x 和目标函数值 z^* ;
(3)	for $i \in I$ 和 $j \in M$;
(4)	初始化拉格朗日乘数 $\lambda^0, \lambda^1, \lambda^2, \lambda^3$, 并根据式(11)求得传输功率 $p_{i,j}$;
(5)	计算 $W_{i,j}$ 和 $V_{i,j}$, 设 $z^*=0$;
(6)	if $W_{i,j} < W_i$ 和 $V_{i,j} < V_i$;
(7)	$x=1$;
(8)	else
(9)	$x=0$;
(10)	End if;
(11)	利用 x 更新目标函数式(15);
(12)	根据 $g(\lambda)$ 的次梯度投影更新拉格朗日乘数, 并利用KKT条件更新传输功率 $p_{i,j}$;
(13)	End for;
(14)	End.

文使用CloudSim作为模拟器^[19]。在移动边缘计算场景中, 每个RSU的通信范围为1000 m, 所有车辆统一位于此范围内, 车辆的速度设定为40 km/h, 同一时间内车辆产生的安全消息数不超过30个, 每个消息的大小为400~10000 Byte, 延迟期限为

100~2000 ms, 消息的初始优先级为 $p_j = \{1, 2, 3\}$, 消息所需的CPU大小为 $1 \times 10^6 \sim 6 \times 10^6$ cycle/s, 安全消息以均匀分布的方式到达RSU。RSU与边缘服务器之间的距离为100~1000 m, 传输路径损耗为3 dB, 边缘服务器的通信带宽为1~3 MHz, 计算速率为 $1 \sim 15 \times 10^8$ cycle/s。本文所提的算法是消息分类感知卸载和优先级调度, 下面讨论本文算法与各种任务分配策略的性能比较, 主要包括常规任务调度算法、文献[5]的能量最优卸载算法与文献[7]的DTOS算法。

图2和图3分别显示了不同算法下消息处理的总时延与能量损耗随着安全消息数目变化的关系。从图2中可以看出, 本文算法与其他3个算法相比实现了更低的处理时延, 与常规任务调度算法相比少95 ms, 主要是由于常规调度算法没有采取卸载策略, 总体上消息处理时延较大, 且随着安全消息数目的增多, 算法的复杂性也逐渐变大, 因此消息处理时延会出现显著的增加; 而能量最优卸载算法以时延为代价, 换取了能量损耗最优化, 从图3可以看出, 能量最优卸载算法与其他算法相比能量损耗

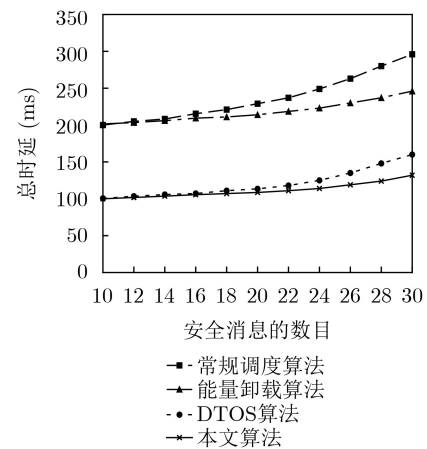


图2 时延与安全消息数目的关系

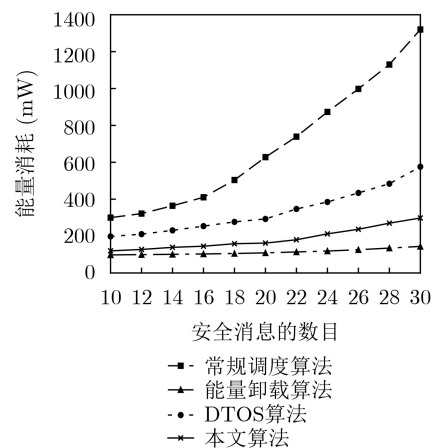


图3 能量损耗与安全消息数目的关系

最低, 但是其处理时延依然较大; DTOS算法进行动态任务卸载和调度, 在一定程度上可以降低时延和能量损耗, 但未考虑任务的多样性和任务的优先级, 不能及时地将时延最敏感的任务进行优先处理, 因此处理时延和能量损耗会随着任务种类和数目的增多而增大。本文所提出的算法综合考虑了消息处理时延和能量损耗两个因素, 并制定了合理的安全消息优先级判定策略, 在边缘服务器中部署任务优先级调度算法, 每个边缘服务器分别处理不同优先级别的安全消息, 本文算法使用比最小传输功率稍微大一点的功率进行传输消息, 因此其计算延迟有所降低, 与DTOS算法相比, 处理时延减少了15.3%, 但能量损耗会比能量最优卸载算法增加。

图4显示了在 θ_1 的不同设置下, 延迟和能量损耗之间的关系变化, 图5表示了在各个算法下不同优先级消息的平均时延, 实验中, 本文选取消息数目为20个。从图4可以看出, θ_1 取值越大表示所提算法对延迟度量的要求越高, 虚线表示了在不同的 θ_1 值下延迟的变化趋势, 目标值是平均时延和能量损耗的加权和, 即对目标函数支配的占比, 随着 θ_1 从0.001增加到0.999, 延迟度量逐渐支配目标函数的值。因此, 能量损耗将由于目标函数中的权重 θ_2 的减小而增加。从图5可以看出, 每个消息的优

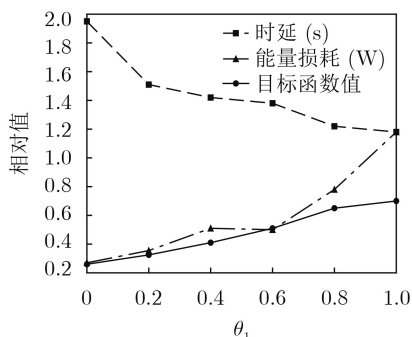


图4 时延和能量损耗的关系

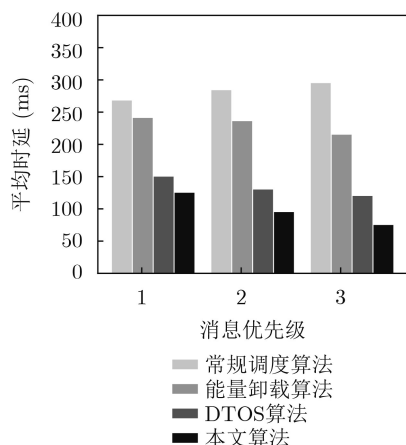


图5 平均时延和消息优先级的关系

优先级不同, 其对消息处理时延的期望不同, 随着消息优先级别的提高, 本文所提出的改进算法在平均时延上优势越明显, 主要是因为本文算法制定了合理的安全消息优先级判定策略, 这样可以保证优先级别高的消息进行最优先地处理。

5 结束语

本文提出一种基于移动边缘计算的消息内容感知分类卸载方案, 制定了时延和能量最优的任务卸载策略, 并将其性能与现有的计算卸载和调度方案进行了比较, 结果表明, 本文所提基于MEC的消息内容感知分类卸载方案改善了消息处理时延和能量损耗, 且得到了时延与能量消耗之间的关系。本文提供了一种适用于其他物联网边缘系统的方法, 未来可能会在移动设备端使用计算卸载算法, 并在算法中设计动态优先调度策略。

参考文献

- [1] LIN Kai, LI Chensi, FORTINO G, *et al.* Vehicle route selection based on game evolution in social internet of vehicles[J]. *IEEE Internet of Things Journal*, 2018, 5(4): 2423–2430. doi: 10.1109/JIOT.2018.2844215.
- [2] MINH Q T, KAMIOKA E, and YAMADA S. CFC-ITS: Context-aware fog computing for intelligent transportation systems[J]. *IT Professional*, 2018, 20(6): 35–45. doi: 10.1109/MITP.2018.2876978.
- [3] TANG Shanjiang, LEE B S, and HE Bingsheng. Fair resource allocation for data-intensive computing in the cloud[J]. *IEEE Transactions on Services Computing*, 2018, 11(1): 20–33. doi: 10.1109/TSC.2016.2531698.
- [4] LUO Guiyang, YUAN Quan, ZHOU Haiibo, *et al.* Cooperative vehicular content distribution in edge computing assisted 5G-VANET[J]. *China Communications*, 2018, 15(7): 1–17. doi: 10.1109/CC.2018.8424578.
- [5] 赵星, 彭建华. 基于Lyapunov优化的隐私感知计算卸载方法[J]. 电子与信息学报, 2020. doi: 10.11999/JEIT190170. ZHAO Xing and PENG Jianhua. A privacy-aware computation offloading method based on Lyapunov optimization[J]. *Journal of Electronics & Information Technology*, 2020. doi: 10.11999/JEIT190170.
- [6] TRINH H, CALYAM P, CHEMODANOV D, *et al.* Energy-aware mobile edge computing and routing for low-latency visual data processing[J]. *IEEE Transactions on Multimedia*, 2018, 20(10): 2562–2577. doi: 10.1109/TMM.2018.2865661.
- [7] 张海波, 栾秋季, 朱江, 等. 车辆异构网中基于移动边缘计算的任务卸载与资源分配[J]. 物联网学报, 2018, 2(3): 36–43. doi: 10.11959/j.issn.2096-3750.2018.00062. ZHANG Haiibo, LUAN Qiuji, ZHU Jiang, *et al.* Task

- offloading and resource allocation in vehicle heterogeneous networks with MEC[J]. *Chinese Journal on Internet of Things*, 2018, 2(3): 36–43. doi: [10.11959/j.issn.2096-3750.2018.00062](https://doi.org/10.11959/j.issn.2096-3750.2018.00062).
- [8] TAO Xiaoyi, OTA K, DONG Mianxiong, *et al.* Performance guaranteed computation offloading for mobile-edge cloud computing[J]. *IEEE Wireless Communications Letters*, 2017, 6(6): 774–777. doi: [10.1109/LWC.2017.2740927](https://doi.org/10.1109/LWC.2017.2740927).
- [9] 张海波, 李虎, 陈善学, 等. 超密集网络中基于移动边缘计算的任务卸载和资源优化[J]. *电子与信息学报*, 2019, 41(5): 1194–1201. doi: [10.11999/JEIT180592](https://doi.org/10.11999/JEIT180592).
ZHANG Haibo, LI Hu, CHEN Shanxue, *et al.* Computing offloading and resource optimization in ultra-dense networks with mobile edge computation[J]. *Journal of Electronics & Information Technology*, 2019, 41(5): 1194–1201. doi: [10.11999/JEIT180592](https://doi.org/10.11999/JEIT180592).
- [10] ALAMEDDINE H A, SHARAFEDDINE S, SEBBAH S, *et al.* Dynamic task offloading and scheduling for low-latency IoT services in multi-access edge computing[J]. *IEEE Journal on Selected Areas in Communications*, 2019, 37(3): 668–682. doi: [10.1109/JSAC.2019.2894306](https://doi.org/10.1109/JSAC.2019.2894306).
- [11] GUO Shuaishuai, WU Dalei, ZHANG Haixia, *et al.* Resource modeling and scheduling for mobile edge computing: A service provider's perspective[J]. *IEEE Access*, 2018, 6: 35611–35623. doi: [10.1109/ACCESS.2018.2851392](https://doi.org/10.1109/ACCESS.2018.2851392).
- [12] YOU Changsheng, HUANG Kaibin, CHAE H, *et al.* Energy-efficient resource allocation for mobile-edge computation offloading[J]. *IEEE Transactions on Wireless Communications*, 2017, 16(3): 1397–1411. doi: [10.1109/TWC.2016.2633522](https://doi.org/10.1109/TWC.2016.2633522).
- [13] SAMANTA A and CHANG Zheng. Adaptive service offloading for revenue maximization in mobile edge computing with delay-constraint[J]. *IEEE Internet of Things Journal*, 2019, 6(2): 3864–3872. doi: [10.1109/JIOT.2019.2892398](https://doi.org/10.1109/JIOT.2019.2892398).
- [14] ALNOMAN A and ANPALAGAN A. A dynamic priority service provision scheme for delay-sensitive applications in fog computing[C]. The 29th Biennial Symposium on Communications, Toronto, Canada, 2018: 1–5. doi: [10.1109/BSC.2018.8494691](https://doi.org/10.1109/BSC.2018.8494691).
- [15] KIM G, JEON Y, and KIM J. A secure message service using the secure domain of a mobile security solution[C]. The 2014 International Conference on Information and Communication Technology Convergence, Busan, South Korea, 2014: 619–620. doi: [10.1109/ICTC.2014.6983230](https://doi.org/10.1109/ICTC.2014.6983230).
- [16] LI Junjie. Analyzing key factors in Taiwanese teachers teaching in China with analytic hierarchy process[J]. *Journal of Interdisciplinary Mathematics*, 2018, 21(2): 307–316. doi: [10.1080/09720502.2017.1420561](https://doi.org/10.1080/09720502.2017.1420561).
- [17] YUAN Jie and LI Xiaoyong. A reliable and lightweight trust computing mechanism for iot edge devices based on multi-source feedback information fusion[J]. *IEEE Access*, 2018, 6: 23626–23638. doi: [10.1109/ACCESS.2018.2831898](https://doi.org/10.1109/ACCESS.2018.2831898).
- [18] LING Ping, KONG Xiangrui, FANG Chen, *et al.* Novel distributed state estimation method for the AC-DC hybrid microgrid based on the Lagrangian relaxation method[J]. *The Journal of Engineering*, 2019, 2019(18): 4932–4936. doi: [10.1049/joe.2018.9329](https://doi.org/10.1049/joe.2018.9329).
- [19] RANI E and KAUR H. Study on fundamental usage of CloudSim simulator and algorithms of resource allocation in cloud computing[C]. The 8th International Conference on Computing, Communication and Networking Technologies, Delhi, India, 2017: 1–7. doi: [10.1109/ICCCNT.2017.8203998](https://doi.org/10.1109/ICCCNT.2017.8203998).
- 赵海涛：男，1983年生，博士，副教授，研究方向为物联网与移动边缘计算。
- 朱银阳：男，1993年生，硕士，研究方向为移动边缘计算与资源优化。
- 丁仪：女，1995年生，硕士，研究方向为物联网路由优化和边缘计算。
- 朱洪波：男，1956年生，博士，教授，研究方向为移动通信与宽带无线技术、无线通信与电磁兼容。