智能网联环境下面向语义通信的资源分配

陈九九[1]，郭彩丽[1,2]，冯春燕[1]，刘传宏[1]
（1. 北京邮电大学先进信息网络北京实验室，北京 100876；2. 北京邮电大学网络体系构建与融合北京市重点实验室，北京 100876）

摘　要：传统的资源分配方法难以满足智能网联环境下各种业务准确理解大量多媒体数据语义的需求。针对该挑战，以智能任务导向的车联网场景为例，首先，提出了两种面向语义通信的资源分配优化准则；然后，针对不同维度的资源，综述了面向语义通信的资源分配模型与算法；构建了面向语义通信的图像数据集，在车联网仿真场景下分析了所研究资源分配方法的性能优势；最后，给出了语义通信资源分配的未来挑战。

关键词：语义通信；资源分配；智能网联；优化准则；强化学习
中图分类号：TN929.5
文献标志码：A
doi: 10.11959/j.issn.2096−3750.2022.00279

Resource allocation for the semantic communication in the intelligent networked environment

CHEN Jiujiu[1], GUO Caili[1,2], FENG Chunyan[1], LIU Chuanhong[1]
1. Beijing Laboratory of Advanced Information Networks, Beijing University of Posts and Telecommunications, Beijing 100876, China
2. Beijing Key Laboratory of Network System Construction and Integration, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract: Traditional resource allocation methods are difficult to meet the needs of various services to accurately understand the semantics of a large amount of multimedia data in the intelligent networked environment. Facing with this challenge, taking intelligent task-oriented internet of vehicles scenarios as an example, two resource allocation optimization criteria for the semantic communication were firstly proposed. Then, according to different dimensions of resources, the models and algorithms of the resource allocation for the semantic communication were described. Then, a semantic communication-oriented image dataset was constructed, and the performance advantages of the proposed resource allocation methods in the simulation scenario of the internet of vehicles were analyzed. Finally, the future challenges of the resource allocation for the semantic communication were presented.

Key words: semantic communication, resources allocation, intelligent networked connection, optimization criteria, reinforcement learning

0　引言

随着 5G 通信技术、信息技术和汽车产业的发展，全球多个国家陆续将发展以自动驾驶为目标的智能网联汽车作为国家战略。以人工智能（AI, artificial intelligence）为代表的单车智能联合、以 V2X（vehicle to everything）为代表的汽车网联已成为自动驾驶的必由之路[1-2]。智能网联环境下，通信的信源和信宿不再是传统的收发信机，而是具有 AI 智能计算功能的智能体（如车辆、边缘计算服务器等）[3]。V2X 的通信也不再是传统通信而是智能体与智能体之间的通信。已有学者提出，在 5G 或

收稿日期：2022−03−01；**修回日期：**2022−06−15
通信作者：郭彩丽，guocaili@bupt.edu.cn
基金项目：北京市自然科学基金资助项目（No.4202049）；中央高校基本科研业务费专项资金资助项目（No.2021XD-A01-1）
Foundation Items: The Natural Science Foundation of Beijing (No.4202049), The Fundamental Research Funds for the Central Universities (No.2021XD-A01-1)

6G 网络驱动的智能场景中，智能体之间通信追求的目标不再是传统通信准确传输比特数据或者精确传递信号波形，而是准确理解传递的信息内容[4]，这里的“内容”即 Shannon 和 Weaver 定义的语义[5]，这种新的通信范式称为语义通信[6]。

在语义通信中，尤其是在智能网联环境下，随着车联网业务的快速发展，视频图像数据的爆发式增长和大量的视觉智能任务带来了极大的资源压力，资源的不合理配置会导致智能体难以获取准确的语义理解结果，因此，有必要研究更优的资源分配方法。

然而，当前传统通信技术以香农经典信息论为理论基础，期望使用最少的资源实现最大速率的数据比特的传输，同时保证误码率最低。常用的资源分配方法是基于服务质量（QoS, quality of service）[7]的资源分配方法。智能网联场景下车辆移动性和服务多样性等特性导致了不同的 QoS 需求，目前基于 QoS 的资源分配研究主要包括最大化网络效率[8-9]和资源利用率[10-11]两类。

相较于面向网络效率的 QoS 需求，面向用户体验质量（QoE, quality of experience）的需求更为主观[12]，其目的是满足人类用户的视觉体验等需求。智能网联环境下基于 QoE 的资源分配研究主要包括如何满足服务多样化需求[13-14]以及用户偏好[15-16]两类。

这些传统的资源分配方式不关注信息的语义，难以满足智能网联环境下自动驾驶等业务对语义理解的需求[17]。智能体之间的信息传输期望使用最少的资源代价获得语义理解准确率的最大化，因此有必要研究面向语义通信的资源分配新方法。基于上述分析，面向语义通信的资源分配目前面临的挑战如下。

1) 传统的资源分配优化准则主要面向网络效率和用户体验，不考虑从语义的角度对所传输的数据进行资源分配，难以满足智能任务对语义理解的需求。因此，如何构建面向语义通信的资源分配优化准则以满足智能化需求是一大挑战。

2) 在语义通信中，智能体完成智能任务时需要充分利用各维度的资源，包括如何利用通信资源完成传输、利用计算资源完成推理和决策、利用缓存资源完成模型下载和训练，这些资源之间的关系是耦合的，且智能网联环境又是动态变化的，因此如何设计多维度资源优化模型和算法，以解决语义通信中的资源紧耦合问题和场景非稳态问题，是另一大挑战。

3) 现有的数据集主要是面向智能任务的，没有考虑通信过程，难以直接用于语义通信的研究，且现有的 QoS 和 QoE 的指标不适用于面向语义通信的资源分配方法的性能评估。因此，如何验证和评估语义通信中资源分配算法的性能，也是一大挑战。

针对上述挑战，本文研究智能网联环境下面向语义通信的资源分配，主要贡献如下。

1) 提出了两种面向语义通信的资源分配优化准则，从理论角度，提出了最大化语义互信息的优化准则；从应用角度，提出了最大化语义理解准确率的优化准则，并对比分析了两种优化准则的优缺点和适用场景。

2) 针对不同维度的资源类型，研究了面向语义通信的 3 类资源分配模型和算法，包括：一维的通信资源分配算法，二维的通信和计算联合资源分配模型和算法，以及多维的通信、计算和缓存联合资源分配模型和算法，并且分析总结了不同资源分配算法、不同计算场景以及不同资源分配模型的差异。

3) 构建了面向语义通信的新数据集，该数据集可用于优化准则建模和算法性能验证，并且基于所构建数据集和部分现有数据集，在车联网仿真场景下，验证和分析了所研究的面向语义通信的资源分配方法的性能优势。

1 智能网联环境下面向语义通信的资源分配优化问题

1.1 面向语义通信的网络架构

智能网联环境下，一个典型的面向语义通信的网络架构如图 1 所示。考虑大量人工智能任务的实现需要丰富的计算资源，由于成本和能耗的限制，与服务器或云端相比，车辆自身的计算资源和计算能力十分有限，限制了对采集数据的高效快速理解与分析。车辆借助车联网将基于语义理解的计算任务进一步卸载到计算资源丰富的移动边缘计算服务器（通常部署在路边单元 RSU 侧）[18]或者远端云服务器进行协同计算[19]，以支持智能网联环境下多种不同业务的需求。如为了保障安全驾驶，针对时延敏感类的安全业务（如轨迹预测和目标检测等），通常采用边缘计算服务器和车辆协同计算（即

端边协同计算），对采集的数据（文本、图像、视频等）进行实时理解和决策；针对非实时需求的信息类和交通效率类业务（如路径优化和交通分析等），车辆和边缘计算服务器的数据可进一步汇聚到远端的云服务器上，进行更密集型计算（即边云协同计算），实现大数据分析挖掘以及算法模型的训练和升级。

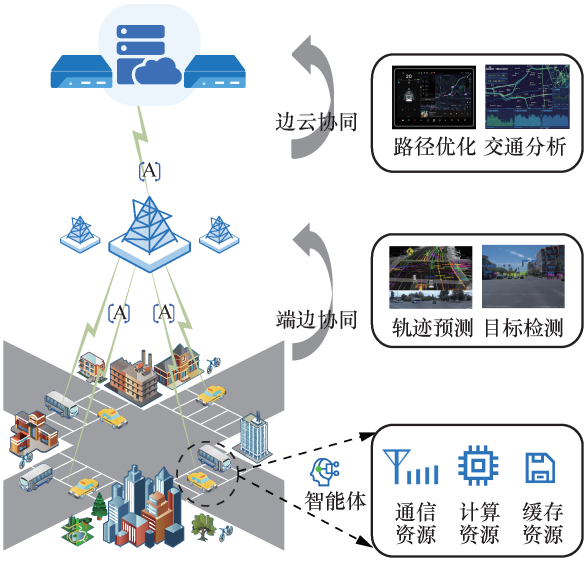


图 1　面向语义通信的网络架构

1.2　面向语义通信的资源分配优化准则

在香农信息论中，若不考虑传输形式，即在理想编码传输方案下，通常采用容量最大的优化准则求取通信资源分配的最优解；在实际资源分配的过程中，需要考虑具体系统需求和传输方案等，因此资源分配的优化准则通常选取传输速率最大化、能耗最小化或误码率最小化等 QoS 指标[7-9]。相应地，若不考虑具体的内容理解任务，在理想视频内容理解的智能算法模型下，资源分配的优化准则自然是使用最少的资源代价传输最准确的语义信息，即语义传输效率最高。基于上述分析，面向语义通信的资源分配优化准则的设计可以从以下两个角度考虑。

1) 从理论研究角度，智能网联环境对信息传输的理论基础和资源分配指导方式提出了新的需求，有必要将信息传输的理论基础扩展到语义信息论[20-21]。若不考虑传输形式以及具体的智能任务，面向语义通信的目的是使收发双方的语义互信息最大，即研究收发双方的语义互信息量的计算方法，在语义互信息计算公式或模型的指导下达到资源分配的优化上界，最大化资源利用率和最优化语义信息的传输。语义互信息的计算基于所推导的公式或模型，因此本文称该准则为模型驱动的资源分配优化准则，该资源优化准则可以形式化描述为

$$\begin{aligned}&\mathrm{P1}:\max_{\boldsymbol{\alpha}} I(X;Y)\\&\text{s.t.}\ \boldsymbol{\alpha}\leqslant D(\boldsymbol{\alpha})\end{aligned}\tag{1}$$

其中，$I(X;Y)$ 为收发双方的语义互信息，X 表示接收端采集数据（包括文本、图像、视频等），Y 表示接收端的语义理解结果，$\boldsymbol{\alpha}$ 表示资源优化变量，即资源参数，$D(\boldsymbol{\alpha})$ 表示可用资源的限制条件。

2) 从实际应用角度，若考虑具体智能任务，最大化资源利用率的目的是最大化智能任务的语义理解准确率，即资源分配的最优化目标为使用最少的资源代价获得最佳的语义理解结果。这些任务完成准确率的计算依赖于大量标记数据，本文称该准则为数据驱动的资源分配优化准则，该资源优化准则可以形式化描述为

$$\begin{aligned}&\mathrm{P2}:\max_{\boldsymbol{\beta}} F(\boldsymbol{\beta})\\&\text{s.t.}\ g(\boldsymbol{\beta})\leqslant \boldsymbol{\beta}_{\max}\end{aligned}\tag{2}$$

其中，$\boldsymbol{\beta}$ 表示资源优化变量，$F(\boldsymbol{\beta})$ 表示资源分配目标函数，表示语义理解准确率（即完成智能任务过程中智能体对语义信息的理解准确率）与资源参数之间的关系，$g(\boldsymbol{\beta})$ 表示资源限制条件，为约束函数，$\boldsymbol{\beta}_{\max}$ 为可用资源上限。在实际场景中，该优化模型需要根据具体场景和智能任务需求调整。

面向语义通信的资源优化准则见表 1。

1.2.1　模型驱动的资源分配优化准则

无语义的经典香农信息论中的信息表示为可调制的波形信号，对其表示和度量只需要考虑这些信息在传输过程中的统计概率，基于其统计概率利用信息熵计算信息量[22]。语义信息除了需要考虑信息传输过程的统计概率，还需要考虑任务计算过程中的语义理解概率，而这个概率通常为逻辑概率，

表 1　面向语义通信的资源优化准则

优化准则	优点	缺点	适用任务类型	适用场景
模型驱动	可解释性强，不需要额外计算开销	计算精度不够，模型中的概率难计算	一般语义理解任务	所有场景
数据驱动	计算精度高，方法可迁移性强	可解释性差，需要数据和算力支撑	特定语义理解任务	智能场景

目前仍未有完备的数学表示[23]。因此，设计理论指导的资源分配优化准则的难点在于：

- 语义互信息的形式化表达，即如何用数学式描述语义互信息和不同概率之间的关系；
- 语义理解概率的计算方法，语义信息不属于传统的二进制比特信息，语义理解概率也不属于统计型概率，其计算方法存在挑战。

为解决上述挑战，首先，基于端到端语义通信系统，借鉴已有语义信息论的思想[24]，描述信息统计概率与语义信息量之间的对数关系，以及语义理解概率与语义信息量之间的对数关系，给出语义互信息量的一般表示形式；其次，考虑不同的语义理解结果，利用全概率公式计算语义理解概率。

1.2.2 数据驱动的资源分配优化准则

在实际的面向语义通信的资源分配过程中，需要考虑具体的任务需求以及语义理解的智能算法带来的影响。面向特定任务时，理论指导的资源分配优化准则存在一定的局限，原因在于标签信息不足，条件概率以及语义理解概率难以计算。当智能体具有一定的算力和历史数据时，则可以通过经验学习到一个面向语义理解准确率的资源分配优化准则。然而，基于可获取的历史数据等通过学习的方式建立数据驱动的资源分配优化准则，存在以下挑战：

- 智能体通信、计算等过程涉及的多个资源参数非独立，存在相互限制和转化的可能，需要人工干预；
- 大量参数以及参数间的复杂耦合关系，导致难以求解语义理解精度等目标函数 $F(\boldsymbol{\beta})$ 的精确数学表达式。

上述挑战实质上是多关联资源参数下以语义理解精度为目标函数的精确表达式的构建问题。可行的思路是：首先，生成实验数据，表征语义理解精度与参数间的关联性；其次，基于生成数据利用神经网络学习得到关系函数 $F(\boldsymbol{\beta})$ 模型；最后，评估模型精度，完成模型参数的调整优化。

2 面向语义通信的资源分配模型与算法

为了实现上述的资源分配优化准则，对语义最终理解起重要作用的语义信息理应被给予更可靠的资源保障。根据资源类别的不同，本文将面向语义通信的资源分配模型和算法分为以下 3 类：单一维度的通信资源分配；通信和计算二维资源的联合分配；通信、计算和缓存资源等多维度的资源联合分配。

2.1 面向语义通信的通信资源分配模型与算法

目前，已有车联网中关于通信资源分配的研究主要是面向网络效率或用户体验的，没有考虑语义信息，不适用于面向智能任务的语义通信系统，因此语义通信系统中的通信资源分配方法还存在很大的研究空间。考虑车联网的环境特性以及语义理解任务的需求复杂性，面向语义通信的通信资源分配研究还存在以下挑战：针对不同的任务需求如何构建通信资源分配模型；车联网动态环境下，如何设计高效的资源分配求解算法。

在智能网联场景下，无线传输会消耗通信资源，而整体通信资源有限，因此设计高效的通信资源分配模型及求解算法很重要[25]。首先，根据上述对资源分配优化准则的研究，针对特定任务，采用数据驱动的资源优化准则，在通信资源限制下，其优化准则为最大化语义理解准确率[26]。根据该资源分配的目标，则优化问题模型可以表示为

$$\begin{aligned}&\mathrm{P3}:\max_{B_m,P_m}\sum_{m=1}^{M}F_{\mathrm{mAP},m}(B_m,P_m)\\&\text{s.t.}\quad \mathrm{C1}:\sum_{m=1}^{M}B_m\leqslant B_{\max},\sum_{m=1}^{M}P_m\leqslant P_{\max}\\&\qquad \mathrm{C2}:F_{\mathrm{mAP},m}\geqslant F_{\min},\forall m\\&\qquad \mathrm{C3}:\text{其他}\end{aligned}\tag{3}$$

其中，M 为车辆总数，$F_{\mathrm{mAP},m}$ 表示语义理解准确率，B_m、P_m 分别表示分配的通信资源（如带宽和功率）。

优化问题 P3 中各约束条件的实际意义如下：约束条件 C1 表示通信资源的限制条件，如 $\sum_{m=1}^{M}B_m\leqslant B_{\max}$ 表示所有车辆所分配的带宽之和不大于总带宽资源 $B_{\max}$；约束条件 C2 表示语义理解准确率取值范围，在实际中，语义理解准确率应该大于或等于任务需求的阈值 $F_{\min}$；约束条件 C3 表示其他约束情况，如传输速率约束、时延约束、能耗约束等。

针对一般任务，采用模型驱动的资源优化准则，在通信资源限制下，其优化准则为最大化语义互信息量[27]。根据该资源分配的目标，优化问题模型可以表示为

$$\begin{aligned}&\mathrm{P4}:\max_{B_m,P_m}\sum_{m=1}^{M}\delta_m I_m\\&\text{s.t.}\quad \mathrm{C1}:\sum_{m=1}^{M}B_m\leqslant B_{\max},\sum_{m=1}^{M}P_m\leqslant P_{\max}\\&\qquad \mathrm{C2}:\text{其他}\end{aligned}\tag{4}$$

其中，I_m 表示第 m 个车辆与接收端的语义互信息量，δ_m 表示权重。

综上，针对一般任务的优化问题 P4 与特定任务的优化问题 P3 的区别在于指导函数的不同，P4 中约束条件不存在对语义互信息量的约束，这是因为面向一般任务，语义信息量与收发双方的语义理解情况相关，其数值是变化的。

上述构建的资源分配问题 P3 和 P4，其本质是面向语义通信的一维资源分配问题，属于小规模优化求解问题，当信道状态信息已知且时延容忍度较高时，为追求精确解，可利用凸优化或连续凸规划等方法来求解优化问题[28]。当时延约束严格或环境动态变化时，也可利用启发式算法或强化学习算法求解。启发式算法的优势在于收敛速度较快[29]，而以强化学习为代表的机器学习算法的优势在于无模型驱动，具有学习能力，能适用于高动态变化的场景[30]。不同资源分配算法对比见表 2，在实际情况中，可根据具体场景和优化问题约束选择适合的资源分配求解算法。

2.2　面向语义通信的通信和计算资源联合分配模型与算法

目前，以深度神经网络（DNN, deep neural network）为代表的深度学习算法是语义理解的核心技术[31]。传统方式将 DNN 的训练和推理任务直接部署在车辆或者将其加载至远端云服务器执行，但这两种方式性能较差（即端到端时延），难以实时地支持智能网联场景具有严格时延和超可靠要求的安全驾驶业务[32]。目前可行的解决方案有两种，一种是结合新兴的边缘计算技术，充分运用从云端下沉到网络边缘（如路边单元、蜂窝网络基站或 Wi-Fi 接入点等）端的计算能力，从而在具有适当计算能力的边缘计算设备（也称为边缘计算服务器）上实现边缘计算设备覆盖范围内低时延与高可靠的深度学习模型训练和推理[33]，车辆只负责采集数据并进行简单预处理。DNN 模型的训练和推理主要在边缘计算服务器上实现，这种称为计算卸载场景。另一种场景是为了充分利用车辆有限的计算能力，缓解服务器计算压力，服务器和车辆协同完成 DNN 模型的训练和推理[34]，这种称为协同计算场景。不同计算场景的对比分析见表 3。

无论是计算卸载还是协同计算场景下，DNN 训练和推理性能对车辆和服务器之间的可用传输带宽（即通信资源需求）以及 DNN 模型复杂度（即计算资源需求）高度敏感，因此，在高可靠、低时延的需求下，面向语义理解精度的资源分配优化的准则等价为优化 DNN 训练和推理的精度，有必要研究在这两种场景面向语义通信的以优化 DNN 训练和推理准确率为目标的通信资源和计算资源联合分配方法。

2.2.1　计算卸载场景中通信和计算资源联合分配模型与算法

为满足安全业务低时延和高准确率的需求，优化问题可转化为在确定任务处理时延需求、通信资源限制和计算资源限制的多约束条件下优化资源分配最大化语义理解准确率，即最大化 DNN 模型训练和推理的准确率。

首先，分析通信资源和计算资源参数与目标检测精度之间的关系。由于通信资源与计算资源相互独立存在，但共同影响目标检测精度，可考虑将二者相乘，表示在仅分配通信资源优化语义理解率的

表 2　不同资源分配算法对比

算法类别	代表算法	优点	缺点	适用场景
精确算法	凸优化、动态规划等	能得到精确解，不依赖数据	复杂度高，不适合大规模问题	小规模求解，高可靠需求场景
启发式算法	遗传算法、粒子群算法等	通用性较强，收敛速度快	次优解，易陷入局部最优	局部开发，低时延需求场景
机器学习算法	强化学习等	无模型，与环境交互	对智能程度，计算能力要求较高	决策优化，高动态变化场景

表 3　不同计算场景的对比分析

计算场景	定义	优点	缺点
计算卸载	DNN 模型的训练和推理主要在边缘计算服务器上实现	适用于所有任务，实现简单	准确率较低，算力资源利用率较低
协同计算	服务器和车辆协同完成 DNN 模型的训练和推理	算力资源利用率高，准确率较高	实现复杂，不同任务间的兼容性差

基础上，如果此时增加或减少计算资源（如 GPU 算力）则会提升或削弱准确率[35]。基于以上理论分析，在一个可行的计算卸载场景下，面向语义通信的联合通信资源和计算资源分配的优化模型为

$$\begin{aligned}
&\mathrm{P5}: \max_{\boldsymbol{\alpha}_m, \boldsymbol{\beta}_m} \sum_{m=1}^{M} F_{2\mathrm{C}}\left(f_{\mathrm{mAP}}^{\mathrm{t}}(\boldsymbol{\alpha}_m) f_{\mathrm{mAP}}^{\mathrm{c}}(\boldsymbol{\beta}_m)\right) \\
&\text{s.t. C1: } \boldsymbol{\alpha}_m \leqslant D_t(\boldsymbol{\alpha}_m) \\
&\quad \text{C2: } \boldsymbol{\beta}_m \leqslant D_c(\boldsymbol{\beta}_m) \\
&\quad \text{C3: } T_m(\boldsymbol{\alpha}_m, \boldsymbol{\beta}_m) \leqslant \tau_m \\
&\quad \text{C4: } \left\{f_{\mathrm{mAP}}^{\mathrm{t}}, f_{\mathrm{mAP}}^{\mathrm{c}}\right\} \geqslant f_{\min} \\
&\quad \text{C5: 其他}
\end{aligned} \tag{5}$$

其中，$\boldsymbol{\alpha}_m$ 和 $\boldsymbol{\beta}_m$ 分别表示通信资源和计算资源变量，$f_{\mathrm{mAP}}^{\mathrm{t}}$ 和 $f_{\mathrm{mAP}}^{\mathrm{c}}$ 分别表示通信资源和计算资源分配的平均语义理解准确率，$F_{2\mathrm{C}}$ 表示两者的耦合关系。约束条件 C1 和 C2 分别表示通信资源和计算资源约束；C3 表示通信和计算过程造成的时延应在车辆智能任务的需求之内；C4 表示任务完成准确率应该大于或等于其最小需求 $f_{\min}$；C5 表示其他约束条件，如能耗约束等。

上述构建的资源分配问题 P5 的本质是面向语义通信的二维资源分配问题。智能网联计算卸载场景资源分配的实时性和环境的多变性，对资源分配求解算法的复杂度和稳定性提出了较高的要求。由于车辆具备一定的计算能力，一个可行的方法是考虑以车辆为智能体，以语义理解准确率作为奖励函数，根据从环境中获得的奖励优化资源分配问题。则该问题可转化为求解 M 个智能体的联合最优动作以获得最大化奖励函数，即多智能体 Q 学习问题。因此，可采用基于多智能体 Q 学习的联合资源分配算法，具体算法设计包括奖励函数设计、Q 值更新准则、动作选择准则等。

2.2.2 协同计算场景中通信和计算资源联合分配模型与算法

在协同计算场景中，车辆承担了本地模型训练或部分推理任务，不再需要上传完整的数据，而是将本地训练的 DNN 模型参数或推理模型中间参数上传至服务器，执行全局模型更新或推理模型后续计算。相较于计算卸载场景，协同计算场景对通信和计算资源的需求程度和利用方式不同，从而导致资源分配结果对视频语义理解精度和任务处理时延的影响不同。有必要针对这种场景研究面向语义通信的联合资源分配模型和算法。

协同计算的可行性得益于 DNN 模型分割技术[36]。DNN 是由多层神经网络相互叠加而成的，不同网络层的计算资源需求以及输出数据量（由模型每层输出的中间参数的大小决定，直接决定带宽需求量）都具有显著的差异性。为减小整个模型的计算时延，同时最小化通信带宽资源需求，需要寻找合适的模型切分点，尽量将计算量小的工作留在车辆，然后在通信量最少的地方进行切割，将中间结果传输至服务器执行复杂的计算任务，实现计算量和通信量之间的权衡。因此，在协同计算场景中，需要联合考虑 DNN 模型切分点、通信资源和计算资源等多种因素，研究面向语义通信的联合资源分配模型和算法。

假设系统中有 M 个车辆，每辆车视频理解类业务所需进行 DNN 推理的模型层数为 L_m。每层 DNN 模型都有拆分和不拆分两种选择，用 $D_{m,l}$ 表示 DNN 模型拆分情况，$D_{m,l}=1$ 表示第 m 辆车的 DNN 模型在第 l 层进行分割，$D_{m,l}=0$ 即为其他情况。

根据上述分析结果，基于 DNN 模型分割联合分配通信和计算资源，以最大化语义理解准确度为优化目标，一个可行的联合资源分配模型可以表示为

$$\begin{aligned}
&\mathrm{P6}: \max_{\boldsymbol{\alpha}_m, \boldsymbol{\beta}_m} \sum_{m=1}^{M} F_{2\mathrm{C}}\left(f_{\mathrm{mAP}}^{\mathrm{t}}(\boldsymbol{\alpha}_m) f_{\mathrm{mAP}}^{\mathrm{c}}(\boldsymbol{\beta}_m)\right) \\
&\text{s.t. C1: } \boldsymbol{\alpha}_m \leqslant D_{\mathrm{t}}(\boldsymbol{\alpha}_m) \\
&\quad \text{C2: } \boldsymbol{\beta}_m \leqslant D_{\mathrm{c}}(\boldsymbol{\beta}_m) \\
&\quad \text{C3: } T_m(\boldsymbol{\alpha}_m, \boldsymbol{\beta}_m) \leqslant \tau_m \\
&\quad \text{C4: } \left\{f_{\mathrm{mAP}}^{\mathrm{t}}, f_{\mathrm{mAP}}^{\mathrm{c}}\right\} \geqslant f_{\min} \\
&\quad \text{C5: } D_{m,l} \in \{0,1\} \\
&\quad \text{C6: } \sum_{l=1}^{L_m} D_{m,l} \leqslant 1 \\
&\quad \text{C7: 其他}
\end{aligned} \tag{6}$$

其中，约束条件 C6、C7 表示 DNN 模型拆分情况，其他约束条件与优化问题 P5 相同。

上述构建的资源分配问题 P6 的本质是面向语义通信的整合整数非线性规划问题。在该优化模型中，整数变量是二进制变量，一个可行的方法是采用机器学习中的 BnB 算法，通过迭代搜索二叉树获得全局最优解[37]。具体来说，通过对整数约束线性松弛求解松弛子问题的最优解，树中的每个节点 n 都与所分解的非线性子问题关联。通过解决节点 n 处的相应非线性问题，获得其局部上限。检查所有

分支的解及目标函数值，若某分支的解是整数并且目标函数值大于或等于其他分支的目标值，则将其他分支剪去不再计算，若还存在非整数解使目标函数值大于整数解的目标函数值，需要继续分支，持续检查，直到得到最优解。

2.3　面向语义通信的多维资源联合分配模型与算法

传统的多维资源分配方法将各种资源分开考虑，且优化的目标通常是网络效率、用户偏好、业务类型或运营成本等，如基于能耗最小化的多维资源分配[38]，基于流行度、社会相似性和偏好的多维资源管理[39]，基于服务类型似然值的多维资源匹配[40]，以及基于运营商奖励函数最大化的多维资源分配等[41]。

另一方面，为完成语义理解任务，不仅需要考虑传输过程中的通信资源的分配，如带宽和功率，还需要考虑计算资源和缓存资源等，计算资源决定了传输到接收端的视频能否进行内容识别和分析等一系列智能计算，而缓存资源决定了数据和神经网络参数的获取时延，也影响最终的语义理解准确率和系统时延。因此，有必要研究面向语义通信的多维资源联合分配模型与算法。

在上述研究仅考虑通信和计算资源的基础上，可进一步联合考虑通信资源、计算资源和缓存资源构建语义理解准确率最大化的优化模型。通信资源、计算资源与缓存资源相互约束，且共同影响目标检测精度，一个可行的面向语义通信的多维资源联合分配优化模型为

$$
\begin{aligned}
\mathrm{P7}:\ & \max_{\boldsymbol{\alpha}_m,\boldsymbol{\beta}_m,\boldsymbol{\gamma}_m} \sum_{m=1}^{M} F_{3\mathrm{C}}\left(f_{\mathrm{mAP}}^{\mathrm{t}}(\boldsymbol{\alpha}_m) f_{\mathrm{mAP}}^{\mathrm{c}}(\boldsymbol{\beta}_m) f_{\mathrm{mAP}}^{\mathrm{r}}(\boldsymbol{\gamma}_m)\right) \\
\text{s.t. } & \mathrm{C1}: \boldsymbol{\alpha}_m \leqslant D_{\mathrm{t}}(\boldsymbol{\alpha}_m) \\
& \mathrm{C2}: \boldsymbol{\beta}_m \leqslant D_{\mathrm{c}}(\boldsymbol{\beta}_m) \\
& \mathrm{C3}: \boldsymbol{\gamma}_m \leqslant D_{\mathrm{r}}(\boldsymbol{\gamma}_m) \\
& \mathrm{C4}: \left\{f_{\mathrm{mAP}}^{\mathrm{t}}, f_{\mathrm{mAP}}^{\mathrm{c}}, f_{\mathrm{mAP}}^{\mathrm{r}}\right\} \geqslant f_{\min} \\
& \mathrm{C5}: F_{3\mathrm{C}}(\boldsymbol{\alpha}_m,\boldsymbol{\beta}_m,\boldsymbol{\gamma}_m) \leqslant D(\boldsymbol{\alpha}_m,\boldsymbol{\beta}_m,\boldsymbol{\gamma}_m) \\
& \mathrm{C6}: \text{其他}
\end{aligned} \tag{7}
$$

其中，$\boldsymbol{\alpha}_m$、$\boldsymbol{\beta}_m$ 和 $\boldsymbol{\gamma}_m$ 分别表示通信资源、计算资源和缓存资源变量，$f_{\mathrm{mAP}}^{\mathrm{t}}$、$f_{\mathrm{mAP}}^{\mathrm{c}}$ 和 $f_{\mathrm{mAP}}^{\mathrm{r}}$ 分别表示通信资源、计算资源和缓存资源分配下的平均语义理解准确率，$F_{3\mathrm{C}}$ 表示三者的耦合关系。约束条件 C1、C2 和 C3 分别表示通信资源、计算资源和缓存资源约束；C4 表示任务完成准确率应该大于或等于其最小需求 $f_{\min}$；C5 表示通信、计算和缓存资源的制约关系，在实际中，其他资源充足但某一资源紧缺都会导致不满足该条件；C6 表示其他约束条件，如时延或能耗约束等。

上述构建的联合资源分配问题，实际上是面向语义通信的具有复杂的非线性多约束的资源优化问题。由于环境的动态改变，多智能体之间的多维协同资源分配会随着环境的变化而改变，因此很难预先设计相应的算法实现最优的分配，或者随着时间推移最优的分配方法也会慢慢变差。通常需要根据环境的变化在线学习新的资源分配，才能提高智能体或整个系统的性能。一个可行的方法是利用基于多智能体深度强化学习的多维资源联合分配算法求解[42]。为了保证资源分配方法的鲁棒性以及约束条件下的语义理解准确率最高，考虑引入深度强化学习网络，在资源动态改变时可以利用现有的经验，且可使最终的分配方案具有较好的鲁棒性，实现智能网联环境下的多维资源联合分配。

基于上述分析，面向语义通信的不同维度的资源分配模型往往和场景需求相关，不同资源分配模型对比见表 4。

3　面向语义通信的数据集构建及资源分配性能分析

3.1　数据集构建

传统的计算机视觉数据集（如 Caltech[43] 和 Waymo[44] 交通视频数据集）可支持智能任务和深度学习中的网络训练和测试，而人工评分数据集则可以支持用户对视频或图片数据的质量、偏好等的评估。这两类数据集均没有考虑通信和计算的融合，无法支持语义通信的研究。因此，本文构建了面向语义通信的 SCO 数据集[45]，并用于本节的实验。

表 4　　**不同资源分配模型对比**

资源分配模型	目标函数	耦合关系	面向场景
通信资源分配	P3、P4	通信资源和语义理解准确率之间的关系	面向语义通信的无线传输
通信和计算资源联合分配	P5、P6	通信和计算相互制约，共同影响语义理解准确率	无线传输和智能计算融合场景
多维资源联合分配	P7	不同资源之间存在相互转化和制约，不同资源对语义理解准确率的影响不同	通信、计算和存储等一体化场景

本文所构建的 SCO 数据集包含 100 张原始图像（来自 Pascal 数据集）和 5 000 张失真图像，为模拟实际通信过程，失真图像引入了 3 类失真（JPEG 压缩、高斯模糊、高斯噪声）。失真图像同时进行机器语义理解和人工体验评分。该数据集提供了基于完整通信过程的图像数据处理方法，能应用于以人类用户体验为目标和以语义理解准确率为目标的资源分配建模，面向语义通信和 QoE 资源分配的算法性能验证，以及面向分类和检测任务的网络训练和测试，还可以用于语义通信中基于 DNN 的编解码器的训练和测试。

3.2 仿真参数设置

为符合现实智能网联环境，以车联网场景为例，本小节基于 SUMO 交通仿真器、MATLAB R2019a、Pycharm 2019.1.1 平台完成了仿真系统的构建：

- 车联网仿真环境是基于 SUMO 生成的城市十字路口场景；
- 为符合车联网中真实的信道条件，路损模型使用 3GPP 标准中的 WINNER 模型，具体的系统仿真参数设置均基于 3GPP TR 36.885，系统仿真参数设置见表 5；
- 考虑实际车联网的时变非稳态特性，仿真中每次实验采用了 200 次独立的蒙特卡洛仿真的平均值，以消除异常数据带来的误差。

表 5 系统仿真参数设置

参数类别	参数数值
边缘服务器覆盖范围	0.5 km
车辆数目 M	3～18
车辆行驶速度 v_m	0～120 km/h
车辆发射功率	23 dBm
噪声功率谱密度	−174 dBm/Hz
中心载波频率	2 GHz

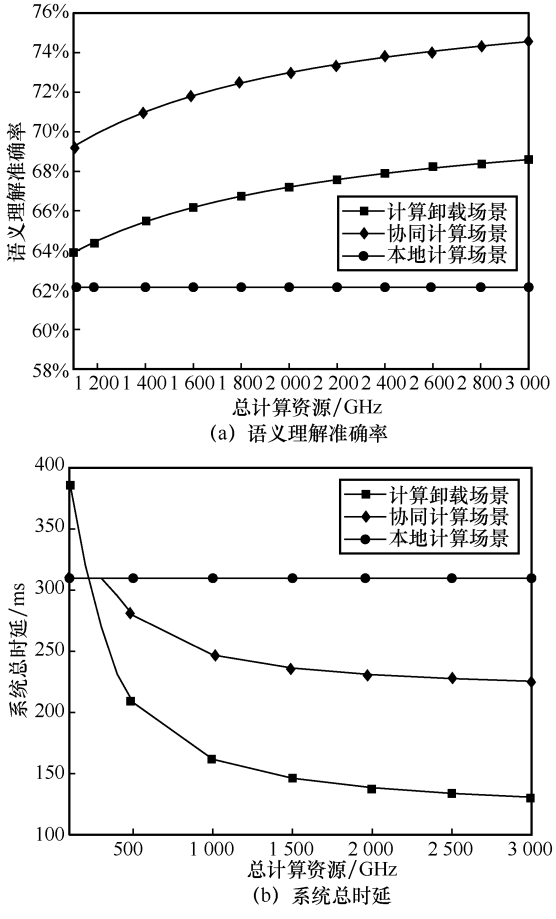
此外，为符合实际中车辆和服务器的运算能力，在实验中：GPU 型号为 Tesla M40，检测器使用 Faster-RCNN[46]，训练和测试环境为 Windows 10 + CUDA 8.0 + Tensorflow 2.1；视频数据的编译码器使用 H.265[47]；测试视频和图片来自本文所建立的数据集以及已有的数据集。实验部分均基于真实数据集，数据类型包括图片和视频，数据集类别包括：本文所构建的面向语义通信的 SCO 数据集；已有的 Caltech 交通视频数据集；已有的 Waymo 自动驾驶数据集。

3.3 性能分析

3.3.1 不同计算场景下资源分配方法对比分析

本部分的仿真实验对比了不同计算场景的性能差异，包括计算卸载场景和协同计算场景。此外，实验部分还引入了一种基础对比场景——本地计算场景，即所有任务在车辆本地完成，依赖于车辆有限的计算资源。

不同计算场景下系统性能曲线如图 2 所示，包括语义理解准确率和总时延性能。随着计算资源的增加，两种场景的语义理解准确率都逐渐提升。通过对总计算资源的高效利用，基于协同计算场景的联合资源分配方法的语义理解准确率始终高于其他两种资源分配方案。另外，随着计算资源的增加，两种场景的联合资源分配方法的系统时延逐渐降低，但基于协同计算的联合资源分配方法的时延始终大于基于计算卸载的联合资源分配方法，这是由于协同计算场景 CNN 模型的卷积层使输出数据量变大，产生了较大的传输时延。因此，实际情况需要根据不同任务的准确率和时延要求选择合适的计算场景和联合资源分配方法。



(a) 语义理解准确率

(b) 系统总时延

图 2 不同计算场景下系统性能曲线

3.3.2　面向语义通信的资源分配方法与传统方法对比分析

本部分的仿真实验对比了面向语义通信的资源分配方法与传统方法的性能差异，对比方案包括：面向语义通信的资源分配方案；基于 QoE 的资源分配方案；基于 QoS 的资源分配方案。仿真实验分析了车联网场景中常见的指标，以及接入车辆数目对资源分配方案的性能影响。

在不同资源分配方法下，语义通信资源分配方法与传统方法的性能曲线如图 3 所示。车辆接入数目的增多，造成了更激烈的资源竞争，整体性能都呈下降趋势。但是，随着车辆数目的增加，本文面向语义通信的资源分配方法具有更好的语义理解准确率，在资源紧缺条件下具有性能优势，能更好地服务于智能任务。有效的资源分配有助于在传输和计算过程中保留更多的语义信息，因此可以在接收端更好、更准确地完成理解语义任务，因此面向语义通信的资源分配方法更适合智能网联场景。

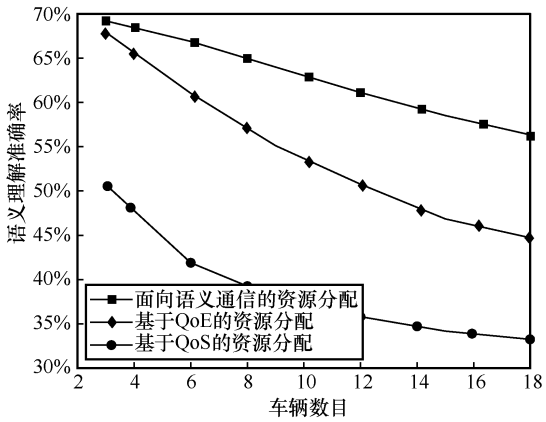


图 3　语义通信资源分配方法与传统方法的性能曲线

3.3.3　面向语义通信的不同维度资源分配方法对比分析

本部分的仿真实验对比了面向语义通信的不同维度资源分配方法的性能差异，包括：单维的通信资源分配方法；二维的通信和计算资源联合分配方法；多维的通信、计算和缓存联合资源分配方法。

不同维度资源分配方法的性能曲线如图 4 所示。多维度的资源联合分配方法性能优于其他资源分配方法，这是由于对系统整体资源的优化分配，获得更好的智能任务性能，但同时也会带来更高的计算复杂度。单一维度的资源分配方法，如通信资源分配，整体性能虽然不如联合资源分配方法，但其在计算资源和缓存资源充足的条件下，更适用于动态无线传输环境，其较低的复杂度能有效提高系统效率。因此，在实际中需要根据具体的场景需求，选择合适的资源分配模型，并设计相应的求解算法。

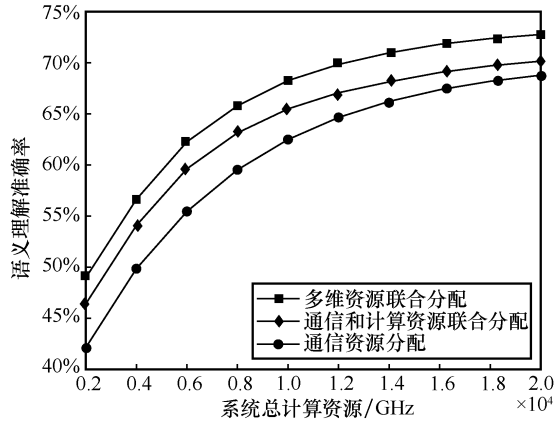


图 4　不同维度资源分配方法的性能曲线

4　未来挑战

在智能网联场景下，面向语义的资源分配方法相较传统资源分配方法具有性能优势。但是，面向语义通信的基础理论、总体结构等相关问题还不清楚，在这一领域需要进行更多的研究，主要挑战如下。

1) 语义信息的价值评估：针对接收端而言，不是所有的语义信息对完成智能任务都是有益的或必需的，如何评估语义的价值或重要程度是未来需要解决的问题。

2) 面向语义通信的虚拟资源管理与编排：在下一代网络中，除了通信、计算以及缓存资源，还存在网络切片等虚拟资源，如何对这些资源进行管理与编排也是智能网联场景需要解决的问题。

3) 面向语义通信系统需求差异化的资源分配：实际中存在用户需求差异以及业务差异等情况，如部分用户需要完成时延敏感型业务，另外的用户需要完成密集计算型业务，针对差异化的需求如何优化资源分配是语义通信中所面临的挑战。

5　结束语

本文探寻技术的交叉融合（即智能化和网联化高度融合）所驱使的通信变革新思路，突破传统的基于 QoS 或 QoE 的资源分配方式的局限，探索面向语义通信的新型资源分配理论模型，以及在该模型指导下的多维度资源分配新方法，实现智能网联环境下对语义信息的准确和高效理解，以满足车联网中以自动驾驶为代表的业务需求。本文所研究的

面向语义通信的资源分配方案旨在提升资源利用率，缓解智能网联环境中大量数据传输带来的资源压力；提升语义理解准确率，解决以语义理解任务为主的车联网自动驾驶中存在的交通安全、效率等问题。本文研究成果对未来数据海量化、应用需求多样化的无线移动通信网络高效传输和资源管理具有指导意义，尤其适用于动态变化的车联网场景。

参考文献：

[1] GYAWALI S, XU S J, QIAN Y, et al. Challenges and solutions for cellular based V2X communications[J]. IEEE Communications Surveys & Tutorials, 2021, 23(1): 222-255.

[2] 李克强. 中国智能网联汽车产业化过程中的挑战及发展对策[J]. 机器人产业, 2019(6): 54-57.
LI K Q. Challenges and development countermeasures in the industrialization of China's intelligent connected vehicles [J]. Robot Industry, 2019(6): 54-57.

[3] DENG S G, ZHAO H L, FANG W J, et al. Edge intelligence: the confluence of edge computing and artificial intelligence[J]. IEEE Internet of Things Journal, 2020, 7(8): 7457-7469.

[4] CALVANESE STRINATI E, BARBAROSSA S. 6G networks: beyond Shannon towards semantic and goal-oriented communications[J]. Computer Networks, 2021(190): 107930.

[5] SHANNON C E, WEAVER W, WIENER N. The mathematical theory of communication[J]. Physics Today, 1950, 3(9): 31-32.

[6] SHI G M, XIAO Y, LI Y Y, et al. From semantic communication to semantic-aware networking: model, architecture, and open problems[J]. IEEE Communications Magazine, 2021, 59(8): 44-50.

[7] LI G Y, BOUKHATEM L, WU J S. Adaptive quality-of-service-based routing for vehicular ad hoc networks with ant colony optimization[J]. IEEE Transactions on Vehicular Technology, 2017, 66(4): 3249-3264.

[8] KIM J W, KIM J W, JEON D K. A cooperative communication protocol for QoS provisioning in IEEE 802.11p/wave vehicular networks[J]. Sensors (Basel, Switzerland), 2018, 18(11): 3622.

[9] WU Z Y, LU Z H, HUNG P C K, et al. QaMeC: a QoS-driven IoVs application optimizing deployment scheme in multimedia edge clouds[J]. Future Generation Computer Systems, 2019(92): 17-28.

[10] SUN L, SHAN H G, HUANG A P, et al. Channel allocation for adaptive video streaming in vehicular networks[J]. IEEE Transactions on Vehicular Technology, 2017, 66(1): 734-747.

[11] ZHANG H X, MA Y B, YUAN D F, et al. Quality-of-service driven power and sub-carrier allocation policy for vehicular communication networks[J]. IEEE Journal on Selected Areas in Communications, 2011, 29(1): 197-206.

[12] LIANG H B, ZHANG X H, HONG X T, et al. Reinforcement learning enabled dynamic resource allocation in the Internet of vehicles[J]. IEEE Transactions on Industrial Informatics, 2021, 17(7): 4957-4967.

[13] TAO X M, DUAN Y P, XU M, et al. Learning QoE of mobile video transmission with deep neural network: a data-driven approach[J]. IEEE Journal on Selected Areas in Communications, 2019, 37(6): 1337-1348.

[14] BAIRAGI A K, ABEDIN S F, TRAN N H, et al. QoE-enabled unlicensed spectrum sharing in 5G: a game-theoretic approach[J]. IEEE Access, 2018, 6: 50538-50554.

[15] JALIL PIRAN M, TRAN N H, SUH D Y, et al. QoE-driven channel allocation and handoff management for seamless multimedia in cognitive 5G cellular networks[J]. IEEE Transactions on Vehicular Technology, 2017, 66(7): 6569-6585.

[16] ZHU H, CAO Y, WANG W, et al. QoE-aware resource allocation for adaptive device-to-device video streaming[J]. IEEE Network, 2015, 29(6): 6-12.

[17] CHEN X, HWANG J N, DE MENG, et al. A quality-of-content-based joint source and channel coding for human detections in a mobile surveillance cloud[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 27(1): 19-31.

[18] ZHU L, YU F R, WANG Y G, et al. Big data analytics in intelligent transportation systems: a survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(1): 383-398.

[19] ZHOU Z, CHEN X, LI E, et al. Edge intelligence: paving the last Mile of artificial intelligence with edge computing[J]. Proceedings of the IEEE, 2019, 107(8): 1738-1762.

[20] FLORIDI L. Outline of a theory of strongly semantic information[J]. Minds and Machines, 2004, 14(2): 197-221.

[21] BAO J, BASU P, DEAN M K, et al. Towards a theory of semantic communication[C]//Proceedings of 2011 IEEE Network Science Workshop. Piscataway: IEEE Press, 2011: 110-117.

[22] SHANNON C E. Communication in the presence of noise[J]. Proceedings of the IRE, 1949, 37(1): 10-21.

[23] 徐文伟, 张弓, 白铂, 等. 后香农时代 ICT 领域的十大挑战问题[J]. 中国科学: 数学, 2021, 51(7): 1095-1138.
XU W W, ZHANG G, BAI B, et al. Ten key ICT challenges in the post-Shannon era[J]. Scientia Sinica (Mathematica), 2021, 51(7): 1095-1138.

[24] ZHONG Y X. A theory of semantic information[J]. China Communications, 2017, 14(1): 1-17.

[25] ZHU M Y, FENG C Y, CHEN J J, et al. Video semantics based resource allocation algorithm for spectrum multiplexing scenarios in vehicular networks[C]//Proceedings of 2021 IEEE/CIC International Conference on Communications in China (ICCC Workshops). Piscataway: IEEE Press, 2021: 31-36.

[26] 陈九九, 冯春燕, 郭彩丽, 等. 车联网中视频语义驱动的资源分配算法[J]. 通信学报, 2021, 42(7): 1-11.
CHEN J J, FENG C Y, GUO C L, et al. Video semantics-driven resource allocation algorithm in Internet of vehicles[J]. Journal on Communications, 2021, 42(7): 1-11.

[27] CHEN J J, GUO C L, FENG C Y, et al. Content driven and reinforcement learning based resource allocation scheme in vehicular network[C]//Proceedings of ICC 2021 - IEEE International Conference on Communications. Piscataway: IEEE Press, 2021: 1-6.

[28] BOYD S, VANDENBERGHE L. Convex optimization[M]. Cambridge: Cambridge University Press, 2004.

[29] ZANDAVI S M, CHUNG V Y Y, ANAISSI A. Stochastic dual simplex algorithm: a novel heuristic optimization algorithm[J]. IEEE Transactions on Cybernetics, 2021, 51(5): 2725-2734.

[30] HUSSAIN F, HASSAN S A, HUSSAIN R, et al. Machine learning for resource management in cellular and IoT networks: potentials, current solutions, and open challenges[J]. IEEE Communications Surveys & Tutorials, 2020, 22(2): 1251-1275.

[31] XIE H Q, QIN Z J, LI G Y, et al. Deep learning enabled semantic communication systems[J]. IEEE Transactions on Signal Processing, 2021, 69: 2663-2675.

[32] SHARMA S, KAUSHIK B. A survey on Internet of vehicles: applications, security issues & solutions[J]. Vehicular Communications, 2019(20): 100182.

[33] ZHANG K, MAO Y M, LENG S P, et al. Predictive offloading in cloud-driven vehicles: using mobile-edge computing for a promising network paradigm[J]. IEEE Vehicular Technology Magazine, 2017(99): 1.

[34] HAN X, TIAN D X, SHENG Z G, et al. Reliability-aware joint optimization for cooperative vehicular communication and computing[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(8): 5437-5446.

[35] CAO X W, WANG F, XU J, et al. Joint computation and communication cooperation for energy-efficient mobile edge computing[J]. IEEE Internet of Things Journal, 2019, 6(3): 4188-4200.

[36] LI E, ZENG L K, ZHOU Z, et al. Edge AI: on-demand accelerating deep neural network inference via edge computing[J]. IEEE Transactions on Wireless Communications, 2020, 19(1): 447-457.

[37] LEE M Y, YU G D, LI G Y. Learning to branch: accelerating resource allocation in wireless networks[J]. IEEE Transactions on Vehicular Technology, 2020, 69(1): 958-970.

[38] LUO S Q, CHEN X, ZHOU Z, et al. Fog-enabled joint computation, communication and caching resource sharing for energy-efficient IoT data stream processing[J]. IEEE Transactions on Vehicular Technology, 2021, 70(4): 3715-3730.

[39] XU L M, YANG Z X, WU H Q, et al. Socially driven joint optimization of communication, caching, and computing resources in vehicular networks[J]. IEEE Transactions on Wireless Communications, 2022, 21(1): 461-476.

[40] WANG G, WANG L, CHUAN J B, et al. LRA-3C: learning based resource allocation for communication-computing-caching systems[C]//Proceedings of 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data. Piscataway: IEEE Press, 2019: 828-833.

[41] 贺颖. 基于深度强化学习的无线网络多维资源分配技术研究[D]. 大连: 大连理工大学, 2018.
HE Y. Research on multi-dimensional resource allocation of wireless networks based on deep reinforcement learning[D]. Dalian: Dalian University of Technology, 2018.

[42] LI Z, GUO C L. Multi-agent deep reinforcement learning based spectrum allocation for D2D underlay communications[J]. IEEE Transactions on Vehicular Technology, 2020, 69(2): 1828-1840.

[43] DOLLÁR P, WOJEK C, SCHIELE B, et al. Pedestrian detection: anevaluation of the state of the art[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(4): 743-761.

[44] SUN P, KRETZSCHMAR H, DOTIWALLA X, et al. Scalability in perception for autonomous driving: waymo open dataset[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2020: 2443-2451.

[45] CHEN J, GUO C, WEI S, et al. SCO-Dataset[EB]. 2022.

[46] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[47] SULLIVAN G J, OHM J R, HAN W J, et al. Overview of the high efficiency video coding (HEVC) standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(12): 1649-1668.

[作者简介]

陈九九（1994– ），男，北京邮电大学博士生，主要研究方向为车联网资源分配、语义通信、强化学习算法等。

郭彩丽（1977– ），女，博士，北京邮电大学教授、博士生导师，主要研究方向为语义通信、无线移动通信技术、认知无线电、信号检测与估值、车联网、可见光通信、视觉智能计算、社交跨媒体数据挖掘与分析等。

冯春燕（1963– ），女，博士，北京邮电大学教授、博士生导师，主要研究方向为无线通信信息传输与处理、宽带通信网络理论与技术、社交网络分析和信息检索、电信大数据分析与挖掘等。

刘传宏（1998– ），男，北京邮电大学博士生，主要研究方向为深度学习、语义通信、资源分配等。