采用速率分割的无线网络能效语义通信

Zhaohui Yang、Mingzhe Chen、Zhaoyang Zhang 和 Chongwen Huang

arXiv:2301.01987v1 [cs.IT] 5 Jan 2023

摘要—本文研究采用速率分割的无线网络能效语义通信中的无线资源分配与语义信息提取问题。在所考虑的模型中，基站（BS）首先从大规模数据中提取语义信息，然后将小规模的语义信息传输给每个用户，用户则根据本地公共知识恢复原始数据。在基站端，采用概率图提取多层语义信息。在下行传输中，采用速率分割方案，通过私有消息传输小规模的私有语义信息，通过公共消息传输公共知识。由于无线资源有限，模型同时考虑计算能耗和传输能耗。该计算与通信联合问题被建模为一个优化问题，目标是在计算、时延和发射功率约束条件下，最小化网络的通信与计算总能耗。为解决该问题，本文提出一种交替算法，并在每一步求得语义信息提取率和计算频率的闭式解。数值结果验证了所提算法的有效性。

索引词——速率分割多址接入、语义通信、能效设计。

I. 引言

数字孪生、边缘学习和元宇宙等新兴应用的快速发展，要求无线网络支持高传输数据速率、超低时延和无缝连接 [1]–[4]。然而，由于频率和时间等无线资源有限，传统正交多址接入方案无法满足下一代无线通信网络对海量连接的需求 [5]。通过使用相同的时间或频率资源，多个用户可以通过非正交多址接入（NOMA） [6]–[11], 同时获得服务，其中用户可在功率域或码域进行划分。由于发射机采用叠加编码、接收机采用连续干扰消除（SIC）可以服务更多用户，因此 NOMA 的频谱效率通常高于传统的正交多址接入方案。

在下行非正交多址接入传输中，接收端会对所有接收到的强信号的干扰进行译码 [11]。因此，NOMA 译码的计算能力通常较高。为平衡目标信号与干扰信号之间的译码折衷，文献 [12]–[15] 提出了速率分割多址（RSMA）的概念。对于下行速率分拆多址传输，面向每个用户的传输消息会被分为公共部分和私有部分。所有用户都希望接收消息的公共部分，即公共消息；而只有部分用户希望接收并译码消息的特定私有部分，即私有消息。在用户端，首先译码公共消息，并将所有私有消息视为干扰；随后译码目标私有消息，此时仅将面向其他用户的私有消息视为干扰。通过动态控制私有消息和公共消息的分割方式，可以调整 RSMA 的计算复杂度，以满足特定的频谱效率要求。要在无线通信系统中实现 RSMA，仍面临诸多挑战，包括私有消息和公共消息的资源分配、译码顺序优化，以及非理想信道和硬件失配情况下的系统设计。

目前已有许多研究探讨无线通信系统中 RSMA 面临的问题。文献 [14] 指出了多输入多输出（MIMO）通信系统中 RSMA 面临的普遍挑战。为最大化所有用户的和速率，文献 [16] 提出了一种分布式速率分割技术。文献 [17] 对有限速率反馈的双接收机多输入单输出（MISO）通信系统进行了速率分析。文献 [18] 表明，与 NOMA 和空分多址接入（SDMA）相比，RSMA 在频谱效率和能量效率方面均可实现最佳性能[19]。具体而言，文献 [20] 研究了无人机辅助无线通信系统中 RSMA 和 NOMA 传输的能量效率优化。针对无线能量传输和信息传输，文献 [21] 研究了 RSMA 的线性预编码方法。对于信道状态信息不完美的情况，文献研究了基于部分信道

本研究部分得到了国家自然科学基金委员会项目（批准号：61725104 和 U20A20158）以及国家重点研发计划项目（批准号：2018 YFB1801104 和 2020YFB1807101）的资助。黄教授的研究工作得到了国家重点研发计划项目（批准号：2021YFA1000500）、国家自然科学基金委员会项目（批准号：62101492）、浙江省自然科学基金项目（批准号：LR22F010002）、国家自然科学基金优秀青年科学家基金项目（海外）、浙江大学教育基金会启真学者基金以及中央高校基本科研业务费项目（批准号：2021FZZX001-21）的资助。

Z. Yang 任职于中国浙江省杭州市的之江实验室，邮编 311121。Z. Yang、Z. Zhang 和 C. Huang 任职于中国浙江省杭州市浙江大学信息科学与电子工程学院（邮编 310027），以及浙江省信息处理、通信与网络重点实验室（IPCAN，位于中国浙江省杭州市，邮编 310007）。（电子邮箱：yang zhaohui@zju.edu.cn、ning ming@zju.edu.cn、chong- wenhuang@zju.edu.cn）

M. Chen 任职于美国佛罗里达州科勒尔盖布尔斯 33146 迈阿密大学电气与计算机工程系及数据科学与计算研究所（电子邮箱：mingzhe.chen@miami.edu）

状态信息的 RSMA 和速率最大化问题 [22], ; 同时，文献 [23] 研究了具有有界信道误差的下行 MISO 速率分割多址系统。

速率分割与新兴技术之间的相互作用已得到研究。借助可重构智能表面，文献 [24], 研究了可重构智能表面辅助 RSMA 的能效资源分配，其中对相移、速率分配和发射波束成形进行了联合调度。文献 [25]研究了用于速率分割无人机无线通信系统的基于学习的流量预测方法。文献 [26] 提出了一种神经网络，用于解决分层速率分割通信系统中的用户聚类问题。由于速率与功率分配之间存在耦合关系，RSMA 的资源分配通常会形成非凸问题，可利用深度强化学习等学习技术加以解决。研究人员设计了多种深度学习算法，以解决 RSMA 中各种复杂的资源分配问题，包括总功率最小化问题 [27], 、联合功率控制、波束成形设计与分割优化问题 [28], [29], 、信道状态信息知识受限时的功率分配问题[30], [31], 、发射功率、用户聚类和资源块分配联合优化问题 [32], , 以及可重构智能表面上的被动预编码与发射机上的主动预编码联合优化问题 [33]。在联邦学习框架下 [34], , 文献 [35] 采用 RSMA 进行上行模型传输，以最小化整个系统的总时延。文献 [36] 开发了一种基于模型的深度学习算法，用于解决 RSMA 的接收机设计问题。

近年来，语义通信受到了广泛关注 [37]–[46]。对于以香农容量为特征的无线通信系统，接收端需要恢复与传输信息完全相同的信息。然而，在虚拟现实、个性化医疗、自动驾驶和万物互联（IoE）等新兴无线应用中，无线通信系统需要利用海量数据满足多模态体验质量（QoE）需求，这使得传统的、以香农容量为特征的传输方式难以实现。尤其在人机交互场景中，人们可以通过语音和增强现实／虚拟现实指令同时控制多个 IoE设备，从而使小范围无线网络中的通信无处不在，这给传统的面向比特的通信带来了严峻挑战。通过文本、语音、图像以及增强现实／虚拟现实来支持实时人机交互和机器间交互，对未来无线通信至关重要。为了支持此类交互，最终接收到的重要信息主要取决于意图，而非对常识比特信息的依赖。这些应用采用先进的信号处理，推动面向任务的语义通信的发展 [3],[47], [48]。在语义通信中，发射机和接收机共享公共知识，可用于在发射机处提取少量信息，并在接收机处恢复原始信息 [49]。类似地，在下行速率分拆多址中，所有用户也都需要接收公共信息和私有信息。鉴于公共知识与公共消息之间存在内在相似性，可以利用 RSMA 提升下行语义通信的系统性能。据我们所知，目前尚无研究探讨语义通信与 RSMA 的融合。

本文的主要贡献包括：

- 本文研究了采用速率分割的无线网络节能语义通信中的无线资源分配与语义信息提取问题。在所考虑的模型中，基站首先从其大规模数据中提取语义信息，然后将小规模语义信息传输给每个用户，用户再基于本地公共知识恢复原始数据。

- 在下行传输中，采用速率分割方案：私有的小规模语义信息通过私有消息传输，而公共知识则通过公共消息传输。由于无线资源有限，必须同时考虑计算能耗与传输能耗。该计算与通信联合问题被建模为一个优化问题，目标是在时延约束下最小化网络的总能耗。

- 为解决该问题，本文提出了一种迭代算法，并在每一步中推导出语义信息提取率和计算频率的闭式解。数值结果表明，所提算法是有效的。

本文其余部分安排如下：第二节介绍系统模型和问题表述。第三节介绍算法设计。第四节分析仿真结果。第五节总结全文。

II. 系统模型与问题表述

如图 1 所示，考虑一个由一个多天线基站和 K 个单天线用户组成的下行语义无线通信（SWC）网络。基站配备 N 根天线，用户集合表示为 $\mathcal{K}$。每个用户 k都有大尺寸数据 $\mathcal{D}_k$ 需要接收。由于无线资源有限，基站需要从原始数据 $\mathcal{D}_k$中提取小尺寸语义信息。在所考虑的模型中，基站首先基于方向概率图提取语义信息，然后采用速率分拆技术传输语义信息。

A. 语义通信模型

在本部分中，我们利用方向概率图来刻画所传输信息的内在信息。在方向概率图中，每个顶点代表具有不同语义层级的语义实体。语义层级越高，

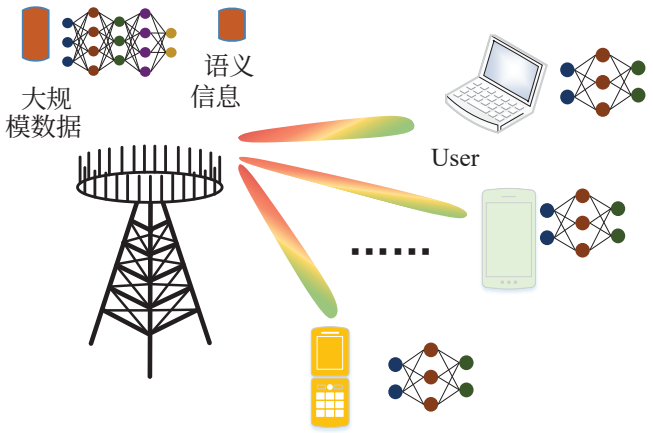

图1. 采用速率分割的所考虑SWC网络示意图。

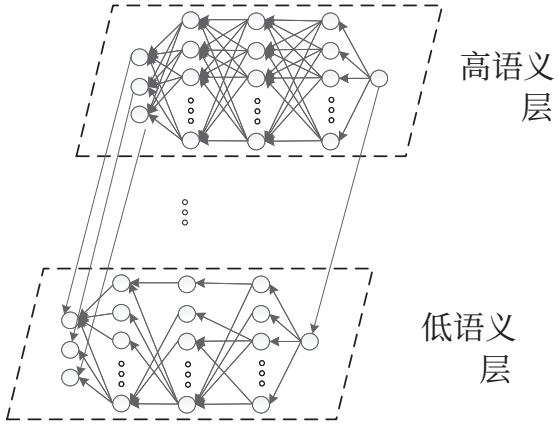

图2. 多级语义信息提取示例。

语义信息就越复杂。任意两个顶点之间的连线表示概率。

为构建方向概率图，我们使用深度神经网络对存储的数据集进行训练，主要包括三个步骤。第一步，从数据集中识别语义实体，其中语义实体指文本中的名称，包括人名、地名等。语义实体的名称具有高度开放性（类型多样、长度灵活、包含未登录词），蕴含丰富知识并凸显个性。识别语义实体可以采用三种常见方法，即规则方法、分类法和序列标注 [50] 。语义实体在方向概率图中表示为一个顶点。第二步，任意两个顶点之间的连边表示一个顶点与另一个顶点相连的概率。通过对数据集进行训练，可以利用卷积神经网络计算两个顶点之间的概率。第三步，进行语义信息融合。对于两个顶点，如果它们之间的连边概率高于预设阈值。因此，最终的方向概率图变为多层图，如图2所示。

为实现小尺寸语义通信，提取过程包括两部分，如图 3 所示。在第一部分中，利用方向概率图提取语义信息，输出记为 $\mathcal{G}(\mathcal{D}_k)$。为高效传输信息，在第二部分中，用户 k 从 $\mathcal{G}(\mathcal{D}_k)$ 中选择一个子集 $\mathcal{S}_k$ ，用于数据传输。

在用户侧，每个用户利用共享的方向概率图恢复原始数据，恢复数据记为 $\mathcal{R}(\mathcal{S}_k)$。恢复数据的语义准确率

$$u_k(\mathcal{D}_k,\mathcal{S}_k)=\frac{\sum_{i=1}^{|\mathcal{R}(\mathcal{S}_k)|}\min\{\sigma(\mathcal{R}(\mathcal{S}_k),s'_{ki}),\sigma(\mathcal{D}_k,s'_{ki})\}}{\sum_{i=1}^{|\mathcal{R}(\mathcal{S}_k)|}\sigma(\mathcal{R}(\mathcal{S}_k),s'_{ki})}, \tag{1}$$

其中， $|\mathcal{R}(\mathcal{S}_k)|$ 是 $\mathcal{R}(\mathcal{S}_k)$ 中的比特数， s'_{ki} 表示 $\mathcal{R}(\mathcal{S}_k)$ 中文本或视频帧里的第 i 个词，并且

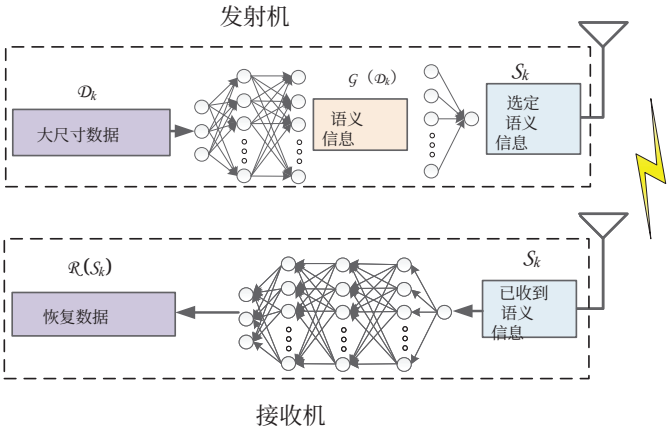

图 3. SWC模型示意图。

$\sigma(\mathcal{R}(\mathcal{S}_k)\ s'_{ki})$ 是 $\mathcal{R}(\mathcal{S}_k)$ 中 s'_{ki} 的出现次数。

B. RSMA模型

在 RSMA 中，每个用户的目标消息可以分成两部分，即公共部分和私有部分 [51]。所有用户的公共部分汇集并合并为一条公共消息。通过所有用户共享同一码本，公共消息被编码为公共消息 s_0，所有用户都需要对其进行解码。每个用户 k 的私有部分被编码为私有流 s_k，该流面向特定用户 k。因此，基站的发射信号 $\boldsymbol{x}$ 可表示为:

$$\boldsymbol{x}=\sqrt{p_0}\boldsymbol{w}_0 s_0+\sum_{k=1}^{K}\sqrt{p_k}\boldsymbol{w}_k s_k, \tag{2}$$

其中， $\boldsymbol{w}_0$ 是公共消息 s_0 的发射波束成形， $\boldsymbol{w}_k$ 是面向用户 k 的私有消息 s_k 的发射波束成形， p_0 是公共消息 s_0 的发射功率， p_k 是私有消息 s_k 的发射功率。

对于用户 k，接收到的消息可表示为：

$$\boldsymbol{h}_{\boldsymbol{k}}^{H}\boldsymbol{x}+n_k=\boldsymbol{h}_k^H\sqrt{p_0}\boldsymbol{w}_0 s_0+\sum_{j=1}^{K}\sqrt{p_k}\boldsymbol{h}_k^H\boldsymbol{w}_k s_j+n_k, \quad (3)$$

其中，$\boldsymbol{h}_k$ 表示用户 k 与基站之间的信道。为解码公共消息 s_0，用户 k 的速率可表示为：

$$c_k=B\log_2\left(1+\frac{p_0|\boldsymbol{h}_k^H\boldsymbol{w}_0|^2}{\sum_{j=1}^{K}p_j|\boldsymbol{h}_k^H\boldsymbol{w}_j|^2+\sigma^2}\right). \quad (4)$$

其中，B 为基站的带宽。请注意，所有用户都需要解码同一条公共消息。为确保所有用户都能成功解码该公共消息，其速率可设为 [18]

$$c_0=\min_{k\in\mathcal{K}}c_k. \quad (5)$$

在我们所考虑的采用速率分割的 SWC 中，所有用户共享公共知识。因此，语义通信所需的公共知识可以编码在公共消息中。此外，公共消息还包含分配给不同用户的部分，即公共消息中分配给用户 k 的速率记为a_k。因此，公共消息的速率约束可表示为

$$a_0+\sum_{k=1}^{K}a_k\le c_k,\quad \forall k\in\mathcal{K}, \quad (6)$$

其中，a_0 为分配给所有用户都需要接收的更新公共知识的速率。在 SWC 中，a_0 表示传输更新的方向概率图信息的速率。

对于每个用户，首先解码公共消息，然后可以将公共消息消除，以解码私有消息。因此，用户 k 解码私有消息的速率可计算为

$$r_k=B\log_2\left(1+\frac{p_k|\boldsymbol{h}_k^H\boldsymbol{w}_k|^2}{\sum_{j=1,j\neq k}^{K}p_j|\boldsymbol{h}_k^H\boldsymbol{w}_j|^2+\sigma^2}\right). \quad (7)$$

C. 传输与计算模型

对于每个用户 k，从数据 $\mathcal{D}_k$ 中提取语义信息所需的计算时间为

$$t_{1k}=\frac{y_{1k}(\mathcal{D}_k,\mathcal{S}_k)}{f_k}, \quad (8)$$

其中，$y_{1k}(\mathcal{D}_k,\mathcal{S}_k)$ 是从 $\mathcal{D}_k$ 中计算出 $\mathcal{S}_k$ 所需的 CPU 周期数，f_k 是用户 k 的计算能力。本地计算能耗可表示为：

$$E_{1k}=\kappa y_{1k}(\mathcal{D}_k,\mathcal{S}_k)f_k^2, \quad (9)$$

其中，κ 是用于衡量有效开关电容的常数系数。

在私有速率 (7) 和分配的公共速率 a_k 下，传输 $\mathcal{S}_k$ 的下行传输时间为

$$t_{2k1}=\frac{Z(\mathcal{S}_k)}{r_k+a_k}, \quad (10)$$

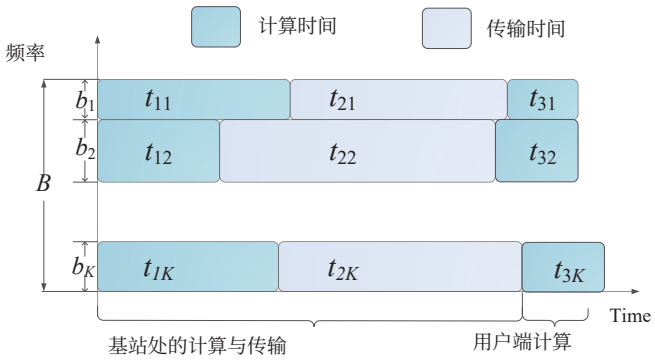

图4. 计算与通信时间示意图。

其中，$Z(\mathcal{S}_k)$ 是集合 $\mathcal{S}_k$ 的数据大小。为传输知识库的更新信息，即方向概率图的更新信息，所有用户的传输时间可表示为

$$t_0=\frac{K_0}{a_0}, \quad (11)$$

其中，K_0 是方向概率图更新信息的大小。结合式（10）和式（11），用户 k 的下行传输时间为

$$t_{2k}=\max\{t_{2k1},t_0\}. \quad (12)$$

发送 $\mathcal{S}_k$ 的传输能耗为

$$E_{2k}=t_{2k1}p_k, \quad (13)$$

广播方向概率图更新信息的传输能耗为

$$E_{20}=t_0p_0. \quad (14)$$

对于用户 k，要恢复原始数据，用户需要计算语义信息 $\mathcal{S}_k$。用户 k的计算时间

$$t_{3k}=\frac{y_{2k}(\mathcal{D}_k,\mathcal{S}_k)}{g_k}, \quad (15)$$

其中，y_{2k}（$\mathcal{D}_k$，$\mathcal{S}_k$）是从 $\mathcal{S}_k$ 恢复 $\mathcal{D}_k$ 所需的计算周期数，g_k 是用户 k的计算能力。用户 k 的总完成时间包括基站的计算时间、下行传输时间以及用户 k的计算时间，如图4所示。用户 k 同时进行计算和计算的总完成时间为

$$\begin{aligned}t_k&=t_{1k}+t_{2k}+t_{3k}\\&=\frac{y_{1k}(\mathcal{D}_k,\mathcal{S}_k)}{f_k}+\max\left\{\frac{Z(\mathcal{S}_k)}{r_k+a_k},\frac{K_0}{a_0}\right\}+\frac{y_{2k}(\mathcal{D}_k,\mathcal{S}_k)}{g_k}.\end{aligned} \quad (16)$$

用户 k 的能耗为

$$E_{3k}=\kappa y_{2k}(\mathcal{D}_k,\mathcal{S}_k)g_k^2. \quad (17)$$

根据上述考虑的模型，系统的总通信与计算能耗为

$$\begin{aligned} E =& \sum_{k=1}^{K}(E_{1k}+E_{2k}+E_{3k})+E_0 \\ =& \sum_{k=1}^{K}\left(\kappa y_{1k}(\mathcal{D}_k,\mathcal{S}_k)f_k^2+\frac{Z(\mathcal{S}_k)p_k}{r_k+a_k}+\kappa y_{2k}(\mathcal{D}_k,\mathcal{S}_k)g_k^2\right) \\ &+\frac{K_0p_0}{a_0}. \end{aligned} \tag{18}$$

我们的目标是在考虑完成时间、传输信息准确率、计算能力、速率分配和功率分配约束的情况下，最小化整个系统的总能耗。数学上，所构建的总能耗最小化问题可表示为：

$$\min_{\mathcal{S},\boldsymbol{f},\boldsymbol{g},\boldsymbol{p},\boldsymbol{a},\boldsymbol{w}} E, \tag{19}$$

$$\text{s.t.}\quad \frac{y_{1k}(\mathcal{D}_k,\mathcal{S}_k)}{f_k}+\max\left\{\frac{Z(\mathcal{S}_k)}{r_k+a_k},\frac{K_0}{a_0}\right\} + \frac{y_{2k}(\mathcal{D}_k,\mathcal{S}_k)}{g_k}\le T,\quad \forall k\in\mathcal{K}, \tag{19a}$$

$$u_k(\mathcal{D}_k,\mathcal{S}_k)\ge A_k,\quad \forall k\in\mathcal{K}, \tag{19b}$$

$$\mathcal{S}_k\subseteq\mathcal{G}(\mathcal{D}_k)\quad \forall k\in\mathcal{K}, \tag{19c}$$

$$a_0+\sum_{k=1}^{K}a_k\le c_k,\quad \forall k\in\mathcal{K}, \tag{19d}$$

$$\sum_{k=1}^{K}f_k\le F^{\max} \tag{19e}$$

$$\sum_{k=0}^{K}p_0\le P^{\max} \tag{19f}$$

$$a_k,f_k,p_k\ge 0,\quad \forall k, \tag{19g}$$

$$\|\boldsymbol{w}_k\|=1,\quad \forall k\in\mathcal{K}\cup\{0\}, \tag{19h}$$

$$0\le g_k\le g_k^{\max},\quad \forall k\in\mathcal{K}, \tag{19i}$$

其中，$\mathcal{S}=\{\mathcal{S}_1,\cdots,\mathcal{S}_K\}$，$\boldsymbol{f}=[f_0,f_1,\cdots,f_K]^T$，$\boldsymbol{g}=[g_1,\cdots,g_K]^T$，$\boldsymbol{p}=[p_1,\cdots,p_K]^T$，$\boldsymbol{a}=[a_0,\cdots,p_K]^T$，$\boldsymbol{w}=[\boldsymbol{w}_0;\ \boldsymbol{w}_1;\ \cdots;\ \boldsymbol{w}_K]$，$T$为系统的最大通信时延，$A_k$ 为用户 k的最低语义准确率，$F^{\max}$ 为基站的最大计算能力，$P^{\max}$ 为基站的传输功率，$g_k^{\max}$为用户k的最大本地计算能力。由于目标函数和约束条件 (19a)-(19c) 均为非凸，该问题通常难以求解。为解决这一问题，我们提出了一种结合交替法与连续凸近似（SCA）方法的迭代算法。

III. 算法设计

本节提出一种交替算法，通过优化三个子问题来迭代求解问题 (19)，即语义信息提取子问题、计算容量子问题，以及联合功率控制、速率分配和波束成形设计子问题。

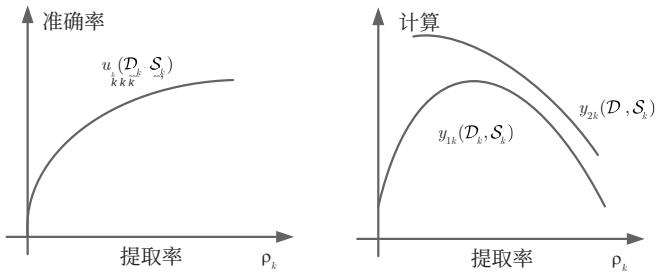

图5. 准确率和计算函数随提取率变化的示意图。

A. 语义信息提取

给定计算能力、功率控制、速率分配和波束成形设计后，问题 (19) 可简化为

$$\min_{\mathcal{S}}\sum_{k=1}^{K}\left(\kappa y_{1k}(\mathcal{D}_k,\mathcal{S}_k)f_k^2+\frac{Z(\mathcal{S}_k)p_k}{r_k+a_k}+\kappa y_{2k}(\mathcal{D}_k,\mathcal{S}_k)g_k^2\right) +\frac{K_0p_0}{a_0} \tag{20}$$

$$\text{s.t.}\quad \frac{y_{1k}(\mathcal{D}_k,\mathcal{S}_k)}{f_k}+\max\left\{\frac{Z(\mathcal{S}_k)}{r_k+a_k},\frac{K_0}{a_0}\right\} + \frac{y_{2k}(\mathcal{D}_k,\mathcal{S}_k)}{g_k}\le T,\quad \forall k\in\mathcal{K}, \tag{20a}$$

$$u_k(\mathcal{D}_k,\mathcal{S}_k)\ge A_k,\quad \forall k\in\mathcal{K}, \tag{20b}$$

$$\mathcal{S}_k\subseteq\mathcal{G}(\mathcal{D}_k)\quad \forall k\in\mathcal{K}. \tag{20c}$$

由于存在两个普遍的难点，问题（20）难以求解。第一个难点在于变量 $\mathcal{S}_k$ 的离散取值空间，这会导致离散优化问题，而求得最优解的复杂度通常极高。第二个难点在于准确度函数 $u_k(\mathcal{D}_k,\mathcal{S}_k)$ 以及计算函数 $f_{1k}(\mathcal{D}_k,\mathcal{S}_k)$ 和 $f_{2k}(\mathcal{D}_k,\mathcal{S}_k)$ 的表达式是隐式的。

为解决第一个难点，我们引入新变量——提取率 ρ_k，其定义如下：

$$\rho_k=\frac{Z(\mathcal{S}_k)}{Z(\mathcal{G}(\mathcal{D}_k))}. \tag{21}$$

显然，ρ_k 的取值范围为 (0,1]。下文中，我们用变量替代 $\mathcal{S}_k$，以便深入了解提取率。

为解决第二个难点，我们首先分析准确度函数和计算函数的变化趋势。对于准确度函数，随着提取率的增加，准确率总是会提高，因为可以利用更多信息来恢复原始数据，如图5所示。因此，最低准确度约束 (20c) 可等价地表示为

$$\rho_k\ge\Gamma_k, \tag{22}$$

其中，Γ_k 是满足$u_k(\mathcal{D}_k,\Gamma_k)=A_k$的最低提取率。对于计算函数 $y_{1k}(\mathcal{D}_k,\rho_k)$，它表示计算

提取率为 ρ_k 的信息所需的CPU周期数。 $\mathcal{D}_k$, $y_{1k}(\mathcal{D}_k, \mathcal{S}_k)$ 包括两部分。第一部分是计算方向概率图，这可以建模为仅与 $\mathcal{D}_k$的大小相关的函数，即 $y_{3k}(\mathcal{D}_k)$。第二部分是从方向概率图中选出提取率为 ρ_k 的信息。由于 $\rho_k = 0$ 或 $\rho_k = 1$，选择方案很简单，因此所需的计算周期最低。因此，第二部分的计算量随提取率ρ_k先增大后减小。例如，计算函数 $y_{1k}(\mathcal{D}_k, \rho_k)$可表示为

$$y_{1k}(\mathcal{D}_k, \rho_k) = y_{3k}(\mathcal{D}_k) + C_{k1}(\rho_k - C_{k2})^{C_{k3}}, \tag{23}$$

其中， $C_{k1} > 0$、 $C_{k2} \in$ (0,1)，且 $C_{k3} > 0$ 为常数参数，这些参数可通过仿真获得。对于计算函数 $y_{2k}(\mathcal{D}_k, \rho_k)$，随着 ρ_k 增加，计算周期数会减少，因为更多的语义信息有助于恢复原始信息。例如，计算函数 $y_{2k}(\mathcal{D}_k, \rho_k)$可表示为

$$y_{2k}(\mathcal{D}_k, \rho_k) = C_{k4}\rho_k^{-C_{k5}}, \tag{24}$$

其中， $C_{k4} > 0$ 和 $C_{k5} > 0$ 是通过仿真获得的常数参数。

通过上述变量代换（21）以及表达式（22）–（24），问题（20）可重新表述为:

$$\begin{aligned}
\min_{\boldsymbol{\rho}} \ & \sum_{k=1}^{K} \Big(\kappa f_k^2 (y_{3k}(\mathcal{D}_k) + C_{k1}(\rho_k - C_{k2})^{C_{k3}}) \\
& + \frac{Z(\mathcal{G}(\mathcal{D}_k)) p_k \rho_k}{r_k + a_k} + \kappa C_{k4} \rho_k^{-C_{k5}} g_k^2 \Big) \\
& + \frac{K_0 p_0}{a_0} \qquad (25) \\
\text{s.t.} \ & \frac{y_{3k}(\mathcal{D}_k) + C_{k1}(\rho_k - C_{k2})^{C_{k3}}}{f_k} \\
& + \max\left\{ \frac{Z(\mathcal{G}(\mathcal{D}_k))\rho_k}{r_k + a_k}, \frac{K_0}{a_0} \right\} + \frac{C_{k4}\rho_k^{-C_{k5}}}{g_k} \le T, \\
& \forall k \in \mathcal{K}, \qquad (25a) \\
& \Gamma_k \le \rho_k \le 1, \quad \forall k \in \mathcal{K}, \qquad (25b)
\end{aligned}$$

其中， $\boldsymbol{\rho} = [\rho_1, \cdots, \rho_K]^T$。由于目标函数和可行集均为凸，问题（25）是一个凸问题。因此，我们可以采用对偶方法求得 Karush-Kuhn-Tucker（KKT）点。为计算问题（25）的解，我们可以得到以下定理。

定理1. 问题 (25) 的最优解为

$$\rho_k^* = \begin{cases} \rho_{k1}^*(\lambda_{1k1}) & \text{if } \rho_{k1}^*(\lambda_{11}) \ge \frac{K_0(a_k + r_k)}{a_0 Z(\mathcal{G}(\mathcal{D}_k))} \\ \rho_{k2}^*(\lambda_{1k2}) & \text{if } \rho_{k2}^*(\lambda_{12}) < \frac{K_0(a_k + r_k)}{a_0 Z(\mathcal{G}(\mathcal{D}_k))} \end{cases}, \tag{26}$$

其中， $\rho_{k1}^*(\lambda_{1k})$ 和 $\rho_{k2}^*(\lambda_{1k})$ 分别是式 (30) 和 (32) 中 $\frac{\partial \mathcal{L}_1(\boldsymbol{\rho}, \boldsymbol{\lambda}_{1k})}{\partial \rho_k} = 0$ 的解， λ_{1k1} 和 λ_{1k2} 分别满足

$$\begin{aligned}
& \frac{y_{3k}(\mathcal{D}_k) + C_{k1}(\rho_{k1}^*(\lambda_{1k1})|_{\Gamma_k}^1 - C_{k2})^{C_{k3}}}{f_k} + \frac{Z(\mathcal{G}(\mathcal{D}_k))}{r_k + a_k} \\
& + \frac{C_{k4}(\rho_{k1}^*(\lambda_{1k1})|_{\Gamma_k}^1)^{-C_{k5}}}{g_k} = T, \qquad (27)
\end{aligned}$$

$$\begin{aligned}
& \frac{y_{3k}(\mathcal{D}_k) + C_{k1}(\rho_{k2}^*(\lambda_{1k2})|_{\Gamma_k}^1 - C_{k2})^{C_{k3}}}{f_k} + \frac{K_0}{a_0} \\
& + \frac{C_{k4}(\rho_{k2}^*(\lambda_{1k2})|_{\Gamma_k}^1)^{-C_{k5}}}{g_k} = T, \qquad (28)
\end{aligned}$$

其中 $a|_b^c = \min\{\max\{a, b\}, c\}$。

证明. 令 $\lambda_{11}, \cdots, \lambda_{1K} > 0$ 为与约束 (25a) 相关联的拉格朗日乘子变量，可得问题 (25) 的拉格朗日函数为

$$\begin{aligned}
\mathcal{L}_1(\boldsymbol{\rho}, \boldsymbol{\lambda}_1) = & \sum_{k=1}^{K} \Big(\kappa f_k^2 (y_{3k}(\mathcal{D}_k) + C_{k1}(\rho_k - C_{k2})^{C_{k3}}) \\
& + \frac{Z(\mathcal{G}(\mathcal{D}_k)) p_k \rho_k}{r_k + a_k} + \kappa C_{k4} \rho_k^{-C_{k5}} g_k^2 \Big) + \frac{K_0 p_0}{a_0} \\
& + \sum_{k=1}^{K} \lambda_{1k} \Big(\frac{y_{3k}(\mathcal{D}_k) + C_{k1}(\rho_k - C_{k2})^{C_{k3}}}{f_k} \\
& + \max\left\{ \frac{Z(\mathcal{G}(\mathcal{D}_k))\rho_k}{r_k + a_k}, \frac{K_0}{a_0} \right\} \\
& + \frac{C_{k4}\rho_k^{-C_{k5}}}{g_k} - T \Big), \qquad (29)
\end{aligned}$$

其中 $\boldsymbol{\lambda}_1 = [\lambda_{11}, \cdots, \lambda_{1K}]^T$。其一阶导数为 (29)

$$\begin{aligned}
\frac{\partial \mathcal{L}_1(\boldsymbol{\rho}, \boldsymbol{\lambda}_1)}{\partial \rho_k} = & \kappa f_k^2 C_{k1} C_{k3} (\rho_k - C_{k2})^{C_{k3}-1} \\
& + \frac{Z(\mathcal{G}(\mathcal{D}_k)) p_k}{r_k + a_k} - \kappa C_{k4} C_{k5} \rho_k^{-C_{k5}-1} g_k^2 \\
& + \lambda_{1k} \Big(\frac{C_{k1} C_{k3} (\rho_k - C_{k2})^{C_{k3}-1}}{f_k} + \frac{Z(\mathcal{G}(\mathcal{D}_k))}{r_k + a_k} \\
& - \frac{C_{k4} C_{k5} \rho_k^{-C_{k5}-1}}{g_k} \Big) \qquad (30)
\end{aligned}$$

for

$$\rho_k \ge \frac{K_0(a_k + r_k)}{a_0 Z(\mathcal{G}(\mathcal{D}_k))}, \tag{31}$$

and

$$\begin{aligned}
\frac{\partial \mathcal{L}_1(\boldsymbol{\rho}, \boldsymbol{\lambda}_1)}{\partial \rho_k} = & \kappa f_k^2 C_{k1} C_{k3} (\rho_k - C_{k2})^{C_{k3}-1} \\
& + \frac{Z(\mathcal{G}(\mathcal{D}_k)) p_k}{r_k + a_k} - \kappa C_{k4} C_{k5} \rho_k^{-C_{k5}-1} g_k^2 \\
& + \lambda_{1k} \Big(\frac{C_{k1} C_{k3} (\rho_k - C_{k2})^{C_{k3}-1}}{f_k} - \frac{C_{k4} C_{k5} \rho_k^{-C_{k5}-1}}{g_k} \Big) \qquad (32)
\end{aligned}$$

for

$$\rho_k < \frac{K_0(a_k + r_k)}{a_0 Z(\mathcal{G}(\mathcal{D}_k))}, \tag{33}$$

将方程 (30) 和 (32) 关于 $\frac{\partial \mathcal{L}_1(\boldsymbol{\rho}, \boldsymbol{\lambda}_1)}{\partial \rho_k} = 0$ 的解分别记为 $\rho_{k1}^*(\lambda_1)$ 和 $\rho_{k2}^*(\lambda_{1k})$。注意到 (30) 和 (32) 的左侧关于 ρ_k 单调递增，因此可通过二分法求得解 $\rho_{k1}^*(\lambda_{1k})$ 和 $\rho_{k2}^*(\lambda_{1k})$。考虑约束条件 (25b)、(31) 和 (33)，拉格朗日乘子应满足 KKT 条件，即该问题的最优解可表示为 (26)。

□

B. 最优计算能力

给定语义信息提取、功率控制、速率分配和波束成形设计后，问题 (19) 可简化为

$$\min_{\boldsymbol{f},\boldsymbol{g}} \sum_{k=1}^{K}\left(\kappa y_{1k}(\mathcal{D}_k,\mathcal{S}_k)f_k^2+\frac{Z(\mathcal{S}_k)p_k}{r_k+a_k}+\kappa y_{2k}(\mathcal{D}_k,\mathcal{S}_k)g_k^2\right)+\frac{K_0p_0}{a_0} \tag{34}$$

$$\text{s.t.}\quad \frac{y_{1k}(\mathcal{D}_k,\mathcal{S}_k)}{f_k}+\max\left\{\frac{Z(\mathcal{S}_k)}{r_k+a_k},\frac{K_0}{a_0}\right\}+\frac{y_{2k}(\mathcal{D}_k,\mathcal{S}_k)}{g_k}\le T,\quad \forall k\in\mathcal{K}, \tag{34a}$$

$$\sum_{k=1}^{K} f_k\le F^{\max} \tag{34b}$$

$$f_k\ge 0,\quad \forall k, \tag{34c}$$

$$0\le g_k\le g_k^{\max},\quad \forall k\in\mathcal{K}. \tag{34d}$$

问题（34）的拉格朗日函数可表示为

$$\begin{aligned}\mathcal{L}_2(\boldsymbol{f},\boldsymbol{g},\boldsymbol{\lambda}_2,\lambda_3)=&\sum_{k=1}^{K}\left(\kappa y_{1k}(\mathcal{D}_k,\mathcal{S}_k)f_k^2+\frac{Z(\mathcal{S}_k)p_k}{r_k+a_k}\right.\\&\left.+\kappa y_{2k}(\mathcal{D}_k,\mathcal{S}_k)g_k^2\right)+\frac{K_0p_0}{a_0}\\&+\sum_{k=1}^{K}\lambda_{2k}\Big(\frac{y_{1k}(\mathcal{D}_k,\mathcal{S}_k)}{f_k}+\max\left\{\frac{Z(\mathcal{S}_k)}{r_k+a_k},\frac{K_0}{a_0}\right\}\\&+\frac{y_{2k}(\mathcal{D}_k,\mathcal{S}_k)}{g_k}-T\Big)+\lambda_3\left(\sum_{k=1}^{K}f_k-F^{\max}\right),\end{aligned} \tag{35}$$

其中，$\boldsymbol{\lambda}_2=[\lambda_{21},\cdots,\lambda_{2K}]^T$是与约束（34a）相关联的拉格朗日乘子，$\lambda_3>0$ 是与约束（34b）相关联的拉格朗日乘子。（35）式的一阶导数为

$$\begin{aligned}&\frac{\partial\mathcal{L}_2(\boldsymbol{f},\boldsymbol{g},\boldsymbol{\lambda}_2,\lambda_3)}{\partial f_k}\\=&2\kappa y_{1k}(\mathcal{D}_k,\mathcal{S}_k)g_k-\frac{\lambda_{2k}y_{1k}(\mathcal{D}_k,\mathcal{S}_k)}{g_k^2}+\lambda_3\end{aligned} \tag{36}$$

$$\begin{aligned}&\frac{\partial\mathcal{L}_2(\boldsymbol{f},\boldsymbol{g},\boldsymbol{\lambda}_2,\lambda_3)}{\partial g_k}\\=&2\kappa y_{2k}(\mathcal{D}_k,\mathcal{S}_k)f_k-\frac{\lambda_{2k}y_{2k}(\mathcal{D}_k,\mathcal{S}_k)}{f_k^2}\end{aligned} \tag{37}$$

令 $\frac{\partial\mathcal{L}_2(\boldsymbol{f},\boldsymbol{g},\boldsymbol{\lambda}_2,\lambda_3)}{\partial f_k}=0$ 和 $\frac{\partial\mathcal{L}_2(\boldsymbol{f},\boldsymbol{g},\boldsymbol{\lambda}_2,\lambda_3)}{\partial g_k}=0$，可得

$$2\kappa y_{1k}(\mathcal{D}_k,\mathcal{S}_k)f_k^3+\lambda_3f_k^2-\lambda_{2k}y_{1k}(\mathcal{D}_k,\mathcal{S}_k)=0, \tag{38}$$

$$g_k=\left(\frac{\lambda_{2k}y_{2k}(\mathcal{D}_k,\mathcal{S}_k)}{2\kappa y_{2k}(\mathcal{D}_k,\mathcal{S}_k)}\right)^{\frac{1}{3}}. \tag{39}$$

通过求解式（38）中的三次函数，可得到 f_k 的数值。得到计算能力 f_k 和 g_k 的数值后，可通过以下方式更新拉格朗日乘子：梯度法。在第 t 次迭代中，λ_{2k} 和 λ_3 的数值按以下方式更新：

$$\begin{aligned}\lambda_{2k}(t)=\Bigg[&\lambda_{2k}(t-1)-\upsilon(t)\Big(\frac{y_{1k}(\mathcal{D}_k,\mathcal{S}_k)}{f_k}+\\&\max\left\{\frac{Z(\mathcal{S}_k)}{r_k+a_k},\frac{K_0}{a_0}\right\}+\frac{y_{2k}(\mathcal{D}_k,\mathcal{S}_k)}{g_k}-T\Big)\Bigg]^+,\end{aligned} \tag{40}$$

and

$$\lambda_3(t)=\left[\lambda_3(t-1)-\upsilon(t)\left(\sum_{k=1}^{K}f_k-F^{\max}\right)\right]^+, \tag{41}$$

其中，$[a]^+=\max a、0$ 和 $\upsilon(t)>0$ 表示动态步长。通过迭代更新（f_k，g_k）和（λ_{2k}，λ_3），整个过程可得到问题（34）的全局最优解。

C. 联合功率控制、速率分配与波束成形设计

给定语义信息提取和计算能力后，问题 (19) 可简化为

$$\min_{\boldsymbol{p},\boldsymbol{a},\boldsymbol{w}} \sum_{k=1}^{K}\left(\kappa y_{1k}(\mathcal{D}_k,\mathcal{S}_k)f_k^2+\frac{Z(\mathcal{S}_k)p_k}{r_k+a_k}+\kappa y_{2k}(\mathcal{D}_k,\mathcal{S}_k)g_k^2\right)+\frac{K_0p_0}{a_0}, \tag{42}$$

$$\text{s.t.}\quad \frac{y_{1k}(\mathcal{D}_k,\mathcal{S}_k)}{f_k}+\max\left\{\frac{Z(\mathcal{S}_k)}{r_k+a_k},\frac{K_0}{a_0}\right\}+\frac{y_{2k}(\mathcal{D}_k,\mathcal{S}_k)}{g_k}\le T,\quad \forall k\in\mathcal{K}, \tag{42a}$$

$$a_0+\sum_{k=1}^{K}a_k\le c_k,\quad \forall k\in\mathcal{K}, \tag{42b}$$

$$\sum_{k=0}^{K}p_0\le P^{\max} \tag{42c}$$

$$a_0,a_k,p_k\ge 0,\quad \forall k, \tag{42d}$$

$$\|\boldsymbol{w}_k\|=1,\quad \forall k\in\mathcal{K}\cup\{0\}, \tag{42e}$$

问题 (42) 是非凸的，这是由于目标函数以及约束条件 (42a)、(42b) 和 (42e) 均为非凸。为处理目标函数的非凸性，我们引入新变量 r_k 并使用变量 p_k^2 替代功率 p_k。因此，

问题 (42) 可等价地转化为

$$\min_{\boldsymbol{p},\boldsymbol{a},\boldsymbol{r},\boldsymbol{w}} \sum_{k=1}^{K}\left(\kappa y_{1k}(\mathcal{D}_k,\mathcal{S}_k)f_k^2+\frac{Z(\mathcal{S}_k)p_k^2}{r_k+a_k}+\kappa y_{2k}(\mathcal{D}_k,\mathcal{S}_k)g_k^2\right)+\frac{K_0p_0^2}{a_0}, \tag{43}$$

$$\text{s.t.}\quad \frac{y_{1k}(\mathcal{D}_k,\mathcal{S}_k)}{f_k}+\max\left\{\frac{Z(\mathcal{S}_k)}{r_k+a_k},\frac{K_0}{a_0}\right\}+\frac{y_{2k}(\mathcal{D}_k,\mathcal{S}_k)}{g_k}\le T,\quad \forall k\in\mathcal{K}, \tag{43a}$$

$$a_0+\sum_{k=1}^{K}a_k\le B\log_2\left(1+\frac{p_0^2|\boldsymbol{h}_k^H\boldsymbol{w}_0|^2}{\sum_{j=1}^{K}p_j^2|\boldsymbol{h}_k^H\boldsymbol{w}_j|^2+\sigma^2}\right),\quad \forall k\in\mathcal{K}, \tag{43b}$$

$$r_k\le B\log_2\left(1+\frac{p_k^2|\boldsymbol{h}_k^H\boldsymbol{w}_k|^2}{\sum_{j=1,j\neq k}^{K}p_j^2|\boldsymbol{h}_k^H\boldsymbol{w}_j|^2+\sigma^2}\right),\quad \forall k\in\mathcal{K}, \tag{43c}$$

$$a_0,a_k,p_k\ge 0,\quad \forall k, \tag{43d}$$

$$\|\boldsymbol{w}_k\|\le 1,\quad \forall k\in\mathcal{K}\cup\{0\}, \tag{43e}$$

其中 $\boldsymbol{r}=[r_0,r_1,\cdots,r_K]^T$，目标函数是凸的，且将约束条件 (43e) 替换为不等式不会损失一般性。在问题 (43) 中，我们只需处理约束条件 (43b) 和 (43c) 的非凸性。通过引入松弛变量 γ_k 和 η_k，问题 (43) 可改写为:

$$\min_{\boldsymbol{p},\boldsymbol{a},\boldsymbol{r},\boldsymbol{w},\boldsymbol{\gamma},\boldsymbol{\eta}} \sum_{k=1}^{K}\left(\kappa y_{1k}(\mathcal{D}_k,\mathcal{S}_k)f_k^2+\frac{Z(\mathcal{S}_k)p_k^2}{r_k+a_k}+\kappa y_{2k}(\mathcal{D}_k,\mathcal{S}_k)g_k^2\right)+\frac{K_0p_0^2}{a_0}, \tag{44}$$

$$\text{s.t.}\quad \frac{y_{1k}(\mathcal{D}_k,\mathcal{S}_k)}{f_k}+\max\left\{\frac{Z(\mathcal{S}_k)}{r_k+a_k},\frac{K_0}{a_0}\right\}+\frac{y_{2k}(\mathcal{D}_k,\mathcal{S}_k)}{g_k}\le T,\quad \forall k\in\mathcal{K}, \tag{44a}$$

$$a_0+\sum_{k=1}^{K}a_k\le B\log_2\left(1+\eta_k\right),\quad \forall k\in\mathcal{K}, \tag{44b}$$

$$r_k\le B\log_2\left(1+\gamma_k\right),\quad \forall k\in\mathcal{K}, \tag{44c}$$

$$a_0,a_k,p_k\ge 0,\quad \forall k, \tag{44d}$$

$$\|\boldsymbol{w}_k\|\le 1,\quad \forall k\in\mathcal{K}\cup\{0\}, \tag{44e}$$

$$\frac{p_k^2|\boldsymbol{h}_k^H\boldsymbol{w}_k|^2}{\sum_{j=1,j\neq k}^{K}p_j^2|\boldsymbol{h}_k^H\boldsymbol{w}_j|^2+\sigma^2}\ge\gamma_k,\quad \forall k\in\mathcal{K}, \tag{44f}$$

$$\frac{p_0^2|\boldsymbol{h}_k^H\boldsymbol{w}_0|^2}{\sum_{j=1}^{K}p_j^2|\boldsymbol{h}_k^H\boldsymbol{w}_j|^2+\sigma^2}\ge\eta_k,\quad \forall k\in\mathcal{K}, \tag{44g}$$

其中 $\boldsymbol{\gamma}=[\gamma_1,\cdots,\gamma_K]^T$和 $\boldsymbol{\eta}=[\eta_1,\cdots,\eta_K]^T$。在问题 (44) 中，目标函数已转化为凸函数。由于约束条件 (44f) 和 (44g) 是非凸的，因此问题 (44) 是非凸问题。接下来，我们采用逐次凸近似法处理这两个非凸约束条件。

对于约束条件 (44f)，可将其等价地转化为

$$p_k^2|\boldsymbol{h}_k^H\boldsymbol{w}_k|^2\ge\gamma_k\alpha_k, \tag{45}$$

$$\sum_{j=1,j\neq k}^{K}p_j^2|\boldsymbol{h}_k^H\boldsymbol{w}_j|^2+\sigma^2\le\alpha_k, \tag{46}$$

其中，α_k 是一个非负松弛变量。在 (45) 中，通过改变波束成形 $\boldsymbol{w}_k$的相位，我们总能使项 $\boldsymbol{h}_k^H\boldsymbol{w}_k$ 取实数值。因此，约束 (45) 可重写为

$$\mathcal{R}(\boldsymbol{h}_k^H\boldsymbol{w}_k)\ge\frac{\sqrt{\gamma_k\alpha_k}}{p_k}, \tag{47}$$

此时左侧为凸函数。利用一阶泰勒级数替换 (47) 的右侧，可将约束 (47) 近似为

$$\mathcal{R}(\boldsymbol{h}_k^H\boldsymbol{w}_k)\ge\frac{\sqrt{\gamma_k^{(n)}\alpha_k^{(n)}}}{p_k^{(n)}}+\frac{\sqrt{\gamma_k^{(n)}}}{2p_k^{(n)}\sqrt{\alpha_k^{(n)}}}(\alpha_k-\alpha_k^{(n)})+\frac{\sqrt{\alpha_k^{(n)}}}{2p_k^{(n)}\sqrt{\gamma_k^{(n)}}}(\gamma_k-\gamma_k^{(n)})-\frac{\sqrt{\gamma_k^{(n)}\alpha_k^{(n)}}}{(p_k^{(n)})^2}(p_k-p_k^{(n)}), \tag{48}$$

其中，上标 (n) 表示变量在第 n 次迭代中的数值。此外，(46) 可重新表述为

$$\sum_{j=1,j\neq k}^{K}\frac{1}{4}((p_j^2+|\boldsymbol{h}_k^H\boldsymbol{w}_j|^2)^2-(p_j^2-|\boldsymbol{h}_k^H\boldsymbol{w}_j|^2)^2)=\sum_{j=1,j\neq k}^{K}p_j^2|\boldsymbol{h}_k^H\boldsymbol{w}_j|^2+\sigma^2\le\alpha_k. \tag{49}$$

用其一阶泰勒近似替换 (49) 的左侧，可得

$$\sum_{j=1,j\neq k}^{K}\frac{1}{4}\Big[((p_j^{(n)})^2+|\boldsymbol{h}_k^H\boldsymbol{w}_j^{(n)}|^2)^2+4((p_j^{(n)})^2+|\boldsymbol{h}_k^H\boldsymbol{w}_j^{(n)}|^2)p_j^{(n)}(p_j-p_j^{(n)})-(p_j^2-|\boldsymbol{h}_k^H\boldsymbol{w}_j|^2)^2+4((p_j^{(n)})^2+|\boldsymbol{h}_k^H\boldsymbol{w}_j^{(n)}|^2)(\mathcal{R}(\boldsymbol{h}_k^H\boldsymbol{w}_j^{(n)}\boldsymbol{h}_k^H\boldsymbol{w}_j)-|\boldsymbol{h}_k^H\boldsymbol{w}_j^{(n)}|^2)\Big]+\sigma^2\le\alpha_k. \tag{50}$$

类似地，我们可以引入松弛变量 β_k，并将约束 (44g) 重写为:

$$\frac{1}{4}((p_0^2+|\boldsymbol{h}_k^H\boldsymbol{w}_0|^2)^2-(p_0^2-|\boldsymbol{h}_k^H\boldsymbol{w}_0|^2)^2)=p_0^2|\boldsymbol{h}_k^H\boldsymbol{w}_0|^2\ge\beta_k\eta_k=\frac{1}{4}((\beta_k+\eta_k)^2-(\beta_k-\eta_k)^2), \tag{51}$$

$$\sum_{j=1}^{K}p_j^2|\boldsymbol{h}_k^H\boldsymbol{w}_j|^2+\sigma^2\le\beta_k. \tag{52}$$

注意，无法通过改变 $\boldsymbol{w}_0$ 的相位，使所有 k 对应的 $\boldsymbol{h}_k^H\boldsymbol{w}_0$ 都成为实数值。为处理式 (51) 的非凸性，我们对式 (51) 两侧使用一阶泰勒近似，这与 [52] 中的方法不同。

考虑两侧的一阶泰勒近似后，式 (51) 可转换为

$$\begin{aligned}&((p_0^{(n)})^2+|\boldsymbol{h}_k^H\boldsymbol{w}_0^{(n)}|^2)^2+4((p_0^{(n)})^2\\&+|\boldsymbol{h}_k^H\boldsymbol{w}_0^{(n)}|^2)p_0^{(n)}(p_0-p_0^{(n)})-(p_0^2-|\boldsymbol{h}_k^H\boldsymbol{w}_0|^2)^2\\&+4((p_0^{(n)})^2+|\boldsymbol{h}_k^H\boldsymbol{w}_0^{(n)}|^2)(\mathcal{R}(\boldsymbol{h}_k^H\boldsymbol{w}_0^{(n)}\boldsymbol{h}_k^H\boldsymbol{w}_0)-|\boldsymbol{h}_k^H\boldsymbol{w}_0^{(n)}|^2)\\&\geq(\beta_k+\eta_k)^2-(\beta_k^{(n)}-\eta_k^{(n)})(\beta_k-\eta_k)+(\beta_k^{(n)}-\eta_k^{(n)})^2,\end{aligned}\tag{53}$$

对于约束条件 (52)，我们可以采用类似的方法处理 (46) 的非凸性。因此，(52) 可以近似为

$$\begin{aligned}&\sum_{j=1}^{K}\frac{1}{4}\Big[((p_j^{(n)})^2+|\boldsymbol{h}_k^H\boldsymbol{w}_j^{(n)}|^2)^2+4((p_j^{(n)})^2\\&+|\boldsymbol{h}_k^H\boldsymbol{w}_j^{(n)}|^2)p_j^{(n)}(p_j-p_j^{(n)})-(p_j^2-|\boldsymbol{h}_k^H\boldsymbol{w}_j|^2)^2\\&+4((p_j^{(n)})^2+|\boldsymbol{h}_k^H\boldsymbol{w}_j^{(n)}|^2)(\mathcal{R}(\boldsymbol{h}_k^H\boldsymbol{w}_j^{(n)}\boldsymbol{h}_k^H\boldsymbol{w}_j)-|\boldsymbol{h}_k^H\boldsymbol{w}_j^{(n)}|^2)\Big]\\&+\sigma^2\leq\beta_k.\end{aligned}\tag{54}$$

通过上述近似，我们可以用相应的凸近似项近似非凸约束条件 (44f) 和 (44g)。因此，原问题 (44) 可以近似为以下凸问题:

$$\begin{aligned}\min_{\boldsymbol{p},\boldsymbol{a},\boldsymbol{r},\boldsymbol{w},\boldsymbol{\gamma},\boldsymbol{\alpha},\boldsymbol{\eta},\boldsymbol{\beta}}\ &\sum_{k=1}^{K}\left(\kappa y_{1k}(\mathcal{D}_k,\mathcal{S}_k)f_k^2+\frac{Z(\mathcal{S}_k)p_k^2}{r_k+a_k}\right.\\&\left.+\kappa y_{2k}(\mathcal{D}_k,\mathcal{S}_k)g_k^2\right)+\frac{K_0p_0^2}{a_0},\end{aligned}\tag{55}$$

$$\text{s.t.}\quad (44a)-(44e),(48),(50),(53),(54),\tag{55a}$$

$$\alpha_k\geq0,\beta_k\geq0,\quad\forall k\in\mathcal{K},\tag{55b}$$

其中 $\boldsymbol{\alpha}=[\alpha,\cdots,\alpha]^T$ 和 $\boldsymbol{\beta}=[\beta_1,\cdots,\beta_K]^T$。凸问题 (55) 可以通过现有的凸优化工具箱求解。

D. 算法分析

算法 1 展示了采用 RSMA 的 SWC 的整体联合通信与计算资源分配。根据算法 1，在每次迭代中求解问题 (19) 的复杂度取决于求解三个子问题。对于语义信息提取子问题，利用定理1中的 (26) 计算最优解，其复杂度为$\mathcal{O}(K\log_2(1/\epsilon_1))$；其中，采用准确率为$\epsilon_1$的二分法求解 (27) 和 (28) 的复杂度为 $\mathcal{O}(\log_2(1/\epsilon_1))$。对于计算容量子问题，其复杂度为 $\mathcal{O}(N_1K)$，其中 N_1 表示使用对偶方法求解计算容量子问题的迭代次数。对于联合功率控制、速率分配与波束成形设计子问题，其复杂度取决于求解近似凸问题 (55)。求解问题 (55) 的复杂度为 $\mathcal{O}(M_1^2M_2)$ [53], 其中，$M_1=(N+7)K+N+2$ 是变量总数，$M_2=13K+1$ 是约束条件总数。求解联合功率控制、速率分配与波束成形设计子问题的总复杂度为 $\mathcal{O}(N_2N^2K^3)$，其中N_2 是逐次凸近似法的迭代次数。作为

算法 1 联合通信与计算资源

采用 RSMA 的 SWC 速率分配1：初始化 $\mathcal{S}^{(0)}$、$\boldsymbol{f}^{(0)}$、$\boldsymbol{g}^{(0)}$、$\boldsymbol{p}^{(0)}$、$\boldsymbol{a}^{(0)}$、$\boldsymbol{w}^{(0)}$。设置迭代次数 $n=1$。2：重复 3：在给定 $\boldsymbol{f}^{(n-1)}$、$\boldsymbol{g}^{(n-1)}$、$\boldsymbol{p}^{(n-1)}$、$\boldsymbol{a}^{(n-1)}$、$\boldsymbol{w}^{(n-1)}$ 的情况下，求解语义信息提取子问题并获得解 $\mathcal{S}^{(n)}$。4：在给定 $\mathcal{S}^{(n)}$、$\boldsymbol{p}^{(n-1)}$、$\boldsymbol{a}^{(n-1)}$、$\boldsymbol{w}^{(n-1)}$ 的情况下，求解计算容量子问题并获得解 $\boldsymbol{f}^{(n)}$、$\boldsymbol{g}^{(n)}$。5：在给定 $\mathcal{S}^{(n)}$、$\boldsymbol{f}^{(n)}$、$\boldsymbol{g}^{(n)}$ 的情况下，求解联合功率控制、速率分配和波束成形设计子问题，其解为 $\boldsymbol{p}^{(n)}$、$\boldsymbol{a}^{(n)}$、$\boldsymbol{w}^{(n)}$。6：设置 $n=n+1$。7：直到目标值 (19) 收敛。

表 I 主要系统参数

参数	数值
基站带宽 B	20 兆赫
噪声功率的功率谱密度	-174 分贝毫瓦/赫兹
最大发射功率 $P^{\max}$	30 分贝毫瓦
有效开关电容 κ	10^{-28}
用户数量 K	5

因此，所提出的算法 1 的总复杂度为$\mathcal{O}(N_3K\log_2(1/\epsilon_1)+N_1N_3K+N_2N_3N^2K^3)$，其中 N_3 是算法 1 的外层迭代次数。

IV. 仿真结果

仿真中，考虑区域内有 $K=5$ 个用户。对于每个用户与基站之间的路径损耗模型，我们设定 $128.1+37.6\log_{10}d$(d 以 km 为单位) [54]，阴影衰落的标准差为 4 dB [52]。此外，系统总带宽为 $B=20$ 兆赫，噪声功率的功率谱密度为 -174 分贝毫瓦/赫兹。除非另有说明，我们设定最大发射功率为 $P^{\max}=30$ 分贝毫瓦，本地计算中的有效开关电容为 $\kappa=10^{-28}$，最大本地计算能力为 $g_1^{\max}=\cdots=g_K^{\max}=2$ GHz。对于所考虑的语义信息任务，我们采用与 [55] 中相同的参数。主要系统参数汇总于表 I。

本文提出的采用 RSMA 的 SWC 联合通信与计算资源分配方案记为'RSMA'。为比较所提出方案的结果，我们考虑传统的正交多址接入、频分多址 (FDMA) [56], 以及 NOMA 的总能耗最小化问题 [57], , 后者记为 'NOMA'。为更好地展示多天线方案的性能，我们考虑文献 [18]中的 SDMA 系统。

图 6 展示了总通信能耗和计算能耗如何随每个用户的最大发射功率变化。根据该图，EXH-RSMA 方案代表穷举搜索方法，该方法可以

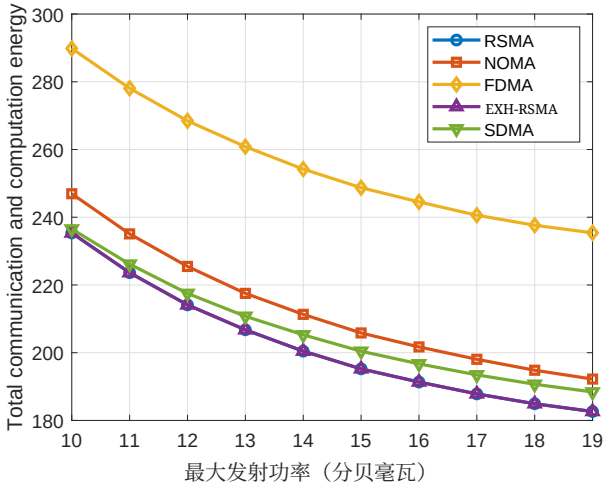

图 6. 总通信与计算能耗与最大发射功率的关系。

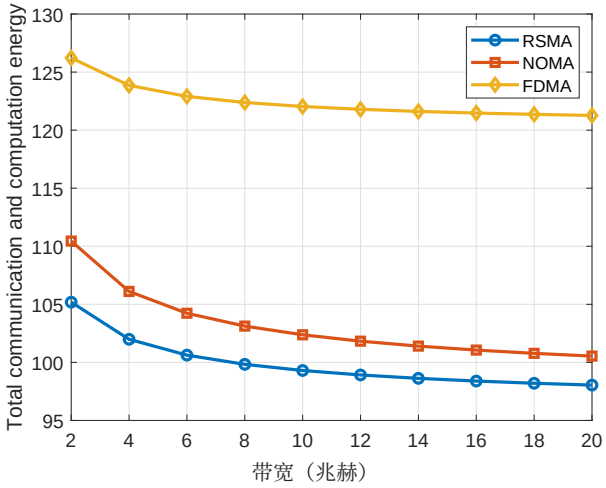

图 7. 系统的总通信与计算能耗随带宽的变化。

通过运行所提出的算法并使用 1000 个初始解，得到接近全局最优的解。从图中可以看出，基站的最大发射功率越大，总能耗越低。这是因为较大的发射功率可以缩短传输时间，从而为计算留出更多时间，并降低总能耗。可以观察到，所提出的 RSMA 优于 FDMA 和 NOMA，因为 RSMA 的频谱效率高于 FDMA 和 NOMA。 与 SDMA 相比，RSMA 仍能实现更低的能耗，尤其是在最大发射功率较高时。原因在于，SDMA 更倾向于为信道增益较高的用户提供服务，而信道增益较低的用户往往需要较长的传输时间，并且任务计算需要较高的计算功率，因此其总能耗高于 RSMA。 还可以发现，所提出的 RSMA 性能接近 EXH-RSMA，这表明所提出方案的有效性。

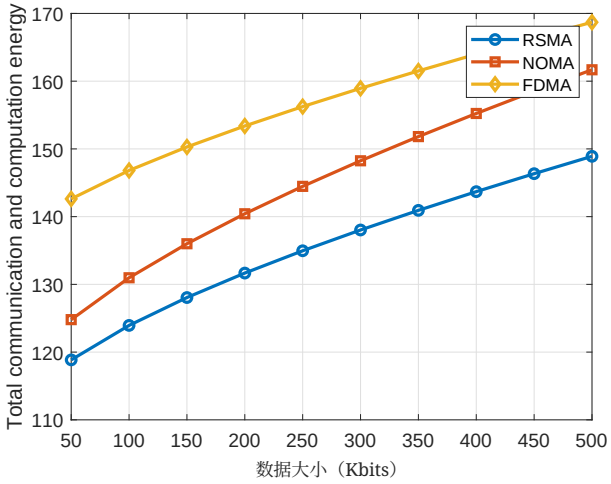

图 8. 每个用户的总通信与计算能耗随传输数据大小的变化。

图 7 展示了系统总能耗随带宽的变化。由图可见，对于所有方案，系统带宽增加时，总通信与计算能耗会降低。这是因为较大的带宽会缩短用户与基站之间的传输时间，从而留出更长的计算时间，进而降低本地计算能耗。

图 8 展示了总通信与计算能耗随每个用户传输数据大小的变化趋势。可以观察到，对于所有方案，数据大小增加时总能耗也会增加。这是因为需要传输更多信息，从而增加发射功率和计算功率。可以发现，所提出的 RSMA 的总能耗增长速度低于 NOMA 和 FDMA，这体现了 RSMA 的鲁棒性。

为展示计算能力对系统性能的影响，图 9 给出了总通信与计算能耗随每个用户的最大计算能力变化的情况。根据该图可知，总能耗先迅速下降，随后趋于某个固定值。其原因在于，在计算能力较低的区间内，提高最大计算能力可以大幅缩短计算时间，从而有更多时间用于传输，进而降低发射功率和总能耗。在计算能力较高的区间内，每个用户都已选择其最优计算能力，提高最大计算能力不会影响计算能力分配结果，因此能耗趋于稳定。

V. 结论

本文研究了采用速率分割的无线网络能效语义通信中的无线资源分配与语义信息提取问题。在所考虑的模型中，基站首先从其大规模

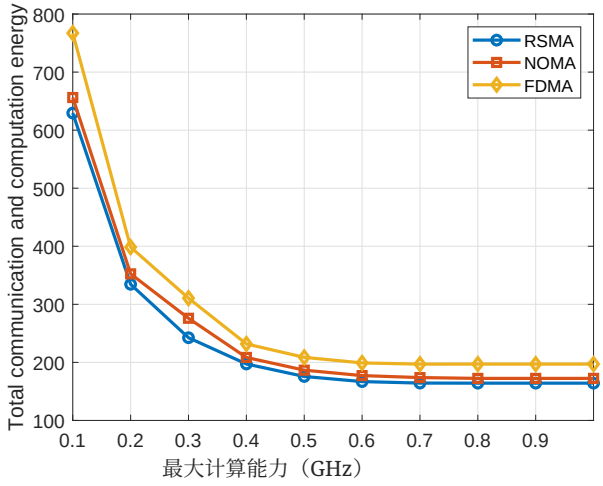

图 9. 总通信与计算能耗与每个用户的最大计算能力之间的关系。

数据中提取语义信息，然后将小规模的语义信息传输给各用户，用户根据本地公共知识恢复原始数据。在下行传输中采用速率分割方案，其中，小规模的私有语义信息通过私有消息传输，而公共知识通过公共消息传输。由于无线资源有限，必须同时考虑计算能耗和传输能耗。本文将这一计算与通信联合问题建模为一个优化问题，目标是在满足任务完成和语义准确率约束的条件下，最小化网络总能耗。本文提出了一种迭代算法来求解该问题，并在每一步推导出语义信息提取率和计算频率的最优解。数值结果表明了所提算法的有效性。

参考文献

[1] Y. Mao、O. Dizdar、B. Clerckx、R. Schober、P. Popovski 和 H. V. Poor，《速率分裂多址：基础、综述与未来研究趋势》，IEEE 通信调查与教程，2022年，已接受，即将发表。[2] B. Clerckx、Y. Mao、E. A. Jorswieck、J. Yuan、D. J. Love、E. Erkip 和 D. Niyato，《速率分裂多址入门：教程、误区与常见问题》，arXiv 预印本 arXiv:2209.00491，2022年。[3] W. Xu、Z. Yang、D. W. K. Ng、M. Levorato、Y. C. Eldar 等，《面向 B5G网络的分布式信号处理边缘学习：语义通信、边缘计算与无线感知》，arXiv 预印本 arXiv:2206.00422，2022年。[4] M. Chen、D. Gunduz、K. Huang、W. Saad、M. Bennis、A. V. Feljan 和 H. V. Poor，《无线网络中的分布式学习：近期进展与未来挑战》，IEEE 通信领域精选期刊，第39卷，第12期，第3579–3605页，2021年12月。[5] W. Saad、M. Bennis 和 M. Chen，《6G无线系统愿景：应用、趋势、技术与开放研究问题》，IEEE Network，第34卷，第3期，第134–142页，2020年5月/6月。[6] Y. Liu、Z. Qin、M. Elkashlan、Z. Ding、A. Nallanathan 和 L. Hanzo，《面向5G及后续网络的非正交多址》，IEEE Proceedings，第105卷，第12期，第2347–2381页，2017年12月。[7] M. Vaezi、Z. Ding 和 H. V. Poor，《5G及后5G无线网络的多址接入技术》，Springer，2018。

[8] L. Dai, B. Wang, Y. Yuan, S. Han, C. l. I, 和 Z. Wang，“面向5G的非正交多址接入：解决方案、挑战、机遇与未来研究趋势”，IEEE Co mmun. Mag.，第53卷，第9期，第74–81页，2015年9月。[9] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, 和 V. K. Bhargava，“面向5G网络的非正交多址接入综述：研究挑战与未来趋势”，IEEE J. Sel. Areas Commun.，第35卷，第10期，第2181–2195页，2017年10月。[10] K. Wang, Y. Liu, Z. Ding, 和 A. Nallanathan，“非正交多址网络中的用户关联”，载于 IEEE 国际通信会议论文集，美国密苏里州堪萨斯城，2018年5月，第1–6页。[11] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, 和 K. Higuchi，“用于蜂窝未来无线接入的非正交多址接入（NOMA）”，载于 IEEE 车辆技术会议论文集，德国德累斯顿，2013年6月，第1–5页。[12] B. Rimoldi 和 R. Urbanke，“高斯多址信道的速率分割方法”，IEEE Trans. Inf. Theory，第42卷，第2期，第364–375页，1996年3月。[13] T. Han 和 K. Kobayashi，“干扰信道的一个新的可达速率区域”，IEEE Trans. Inf. Theory，第27卷，第1期，第49–60页，1981年1月。[14] B. Clerckx, H. Joudeh, C. Hao, M. Dai, 和 B. Rassouli，“用于MIMO无线网络的速率分割：LTE演进的一种有前景的物理层策略”，IEEE Commun. Mag.，第54卷，第5期，第98–105页，2016年5月。[15] H. Liu, T. A. Tsiftsis, K. J. Kim, K. S. Kwak, 和 H. V. Poor，“通过增强公平性和中断性能的速率分割实现上行链路NOMA”，IEEE Trans. Wireless Commun.，第19卷，第7期，第4657–4670页，2020年。[16] J. Cao 和 E. M. Yeh，“通过分布式速率分裂实现渐近最优的多址通信”，IEEE Trans. Inf. Theory，第53卷，第1期，第304–319页，2007年1月。[17] C. Hao, Y. Wu, 和 B. Clerckx，“有限速率反馈下双接收机单输入多输出广播信道的速率分析：一种速率分裂方法”，IEEE《通信汇刊》，第63卷，第9期，第3232–3246页，2015年9月。[18] Y. Mao, B. Clerckx, 和 V. O. K. Li，“下行链路通信系统的速率分割多址：衔接、推广并优于SDMA和NOMA”，EURASIP J. Wireless Commun. Network.，第2018卷，第1期，第1–54页，2018年5月。[19] X. Li, Y. Fan, R. Yao, P. Wang, N. Qi, N. I. Miridakis, 和 T. A. Tsiftsis，“认知卫星地面网络中基于速率分割多址的安全性分析”，IEEE Trans. Veh. Technol.，2022年。[20] A. Rahmati, Y. Yapıcı, N. Rupasinghe, I. Guvenc, H. Dai, 和 A. Bhuyany，“蜂窝连接毫米波无人机网络中RSMA和NOMA的能量效率”，arXiv预印本 arXiv:1902.04721，2019年2月。[21] Y. Mao, B. Clerckx, 和 V. O. K. Li，“用于多用户多天线无线信息与能量传输的速率分割”，arXiv预印本 arXiv:1902.07851，2019年。[22] H. Joudeh 和 B. Clerckx，“部分信道状态信息在发射端条件下线性预编码下行多用户单输入多输出系统的和速率最大化：一种速率分割方法”，IEEE《通信汇刊》，第64卷，第11期，第4847–4861页，2016年11月。[23] ——，“下行多用户单输入多输出系统中的鲁棒传输：一种速率分割方法”，IEEE Tran. Signal Process.，第64卷，第23期，第6227–6242页，2016年12月。[24] Z. Yang, J. Shi, Z. Li, M. Chen, W. Xu, 和 M. Shikh-Bahaei，“采用可重构智能表面的高能量效率速率分割多址（RSMA）”，载于 IEEE 国际通信会议研讨会（ICC Workshops）论文集。IEEE，2020年，第1–6页。[25] L. Lu, Y. Hu, Y. Zhang, G. Jia, J. Nie, 和 M. Shikh-Bahaei，“用于基于速率分割多址的无人机预测性部署的机器学习”，载于 IEEE 全球通信大会研讨会（GC Wkshps. IEEE，2020年，第1–6页。[26] R. Pereira, A. A. Deshpande, C. J. Vaca-Rubio, X. Mestre, A. Zanella, D. Gregoratti, E. de Carvalho, 和 P. Popovski，“使用机器学习进行速率分割的用户聚类”，arXiv预印本 arXiv:2205 11373，2022年。。[27] M. R. Camana, C. E. Garcia, 和 I. Koo，“基于速率分割多址的底层单输入多输出无线信息与功率传输系统中的深度学习辅助功率最小化”，IEEE Access，第10卷，第62137–62156页，2022年。[28] J. Liu, C.-H. R. Lin, Y.-C. Hu, 和 P. K. Donta，“结合深度强化学习与博弈论的无线信息与功率传输赋能物联网网络联合波束成形、功率分配与分裂控制”，Sensors，第22卷，第6期，第2328页，2022年。[29] Z. Yang, M. Chen, W. Saad, W. Xu, 和 M. Shikh-Bahaei，“上行链路速率分割多址（RSMA）通信的和速率最大化”，IEEE Trans. Mobile Computing，2020年。

[30] N. Q. Hieu, D. T. Hoang, D. Niyato, 和 D. I. Kim，“通过深度强化学习优化速率分裂通信的功率分配”，IEEE Wireless Commun. Lett.，第10卷，2823，2021年12，2820 no. pp. – 。[31] J. Huang, Y. Yang, L. Yin, D. He, 和 Q. Yan，“基于深度强化学习的6G低地球轨道卫星通信系统速率分割多址功率分配”，IEEE Wireless Commun. Lett.，第11卷，第10期，第2185–2189页，2022年。 [32] M. Z. Hassan, M. J. Hossain, J. Cheng, 和 V. C. Leung，“采用速率分割多址和基于强化学习的用户聚类实现雾无线接入网的吞吐量-功率联合优化”，IEEE Trans. Veh. Technol.，第70卷，第8期，第8019–8036页，2021年。 [33] M. Wu, Z. Gao, Y. Huang, Z. Xiao, D. W. K. Ng, 和 Z. Zhang，“面向可重构智能表面辅助太赫兹大规模MIMO的深度学习速率分割多址”，arXiv预印本 arXiv:2209.08456，2022年。 [34] M. Chen, Z. Yang, W. Saad, C. Yin, H. V. Poor, 和 S. Cui，“面向无线网络联邦学习的学习与通信联合框架”，IEEE Trans. Wireless Commun.，第20卷，第1期，第269–283页，2021年1月。 [35] S.-H. Park 和 H. Lee，“采用速率分割传输的雾无线接入网辅助联邦学习完成时间最小化”，IEEE Trans. Veh. Technol.，2022年。 [36] R. C. Loli, O. Dizdar, B. Clerckx, 和 C. Ling，“面向速率分割多址的基于模型的深度学习接收机设计”，arXiv预印本 arXiv:2205.00849，2022年。 [37] Z. Weng 和 Z. Qin，“用于语音传输的语义通信系统”，IEEE J. Sel. Areas Commun.，第39卷，第8期，第2434–2444页，2021年。 [38] Y. Wang, M. Chen, T. Luo, W. Saad, D. Niyato, H. V. Poor, 和 S. Cui，“语义通信的性能优化：一种基于注意力机制的强化学习方法”，IEEE J. Sel. Areas Commun.，第40卷，第9期，第2598–2613页，2022年7月。 [39] H. Xie, Z. Qin, G. Y. Li, 和 B.-H. Juang，“深度学习赋能的语义通信系统”，IEEE Tran. Signal Process.，第69卷，第2663–2675页，2021年。 [40] M. Chen, Y. Wang, 和 H. V. Poor，“面向能量收集网络中无线语义通信的性能优化”，载于 IEEE国际声学、语音与信号处理会议（ICASSP）论文集。IEEE，2022年，第8647–8651页。 [41] J. Kang, H. Du, Z. Li, Z. Xiong, S. Ma, D. Niyato, 和 Y. Li，“面向任务的语义通信中的个性化显著性：图像传输与性能分析”，arXiv预印本 arXiv:2209 12274，2022年。。[42]Z. Yang, M. Chen, Z. Zhang, 和 C. Huang，“安全语义通信：基础与挑战”，arXiv预印本 arXiv:2301.01421，2023年。 [43] R. Kaewpuang, M. Xu, W. Y. B. Lim, D. Niyato, H. Yu, J. Kang, 和 X. S. Shen，“面向语义通信的量子密钥分发（QKD）网络协作资源管理”，arXiv预印本 arXiv:2209.11957，2022年。 [44] S. Xie, Y. Wu, S. Ma, M. Ding, Y. Shi, 和 M. Tang，“采用数字调制的面向任务通信的鲁棒信息瓶颈”，arXiv预印本 arXiv:2209.10382，2022年。 [45] S. Wang, J. Dai, Z. Liang, K. Niu, Z. Si, C. Dong, X. Qin, 和 P. Zhang，“无线深度视频语义传输”，arXiv预印本 arXiv:2205.13129，2022年。 [46] H. Zou, C. Zhang, S. Lasaulce, L. Saludjian, 和 V. Poor，“面向目标的量化：分析、设计及其在资源分配中的应用”，arXiv预印本 arXiv:2209.15347，2022年。 [47] D. Gunduz 等，“超越比特传输：上下文、语义与面向任务的通信”，arXiv预印本 arXiv:2207.09353，2022年。 [48] Z. Qin, X. Tao, J. Lu, 和 G. Y. Li，“语义通信：原理与挑战”，arXiv预印本 arXiv:2201.01389，2021年。 [49] H. Tong, Z. Yang, S. Wang, Y. Hu, W. Saad, 和 C. Yin，“基于联邦学习的无线网络语音语义通信”，载于 IEEE全球通信大会（GLOBECOM）论文集。IEEE，2021年，第1–6页。 [50] L. Deng 和 Y. Liu，《自然语言处理中的深度学习》。Springer，2018年。[51] B. Clerckx, Y. Mao, R. Schober, 和 H. V. Poor，“速率分割统一SDMA、正交多址、NOMA和MISO广播信道中的多播：简单的双用户速率分析”，IEEE Wireless Communications Letters，第9卷，第3期，第349–353页，2019年。 [52] Z. Yang, M. Chen, W. Saad, 和 M. Shikh-Bahaei，“速率分割多址（RSMA）的速率分配与功率控制优化”，IEEE《通信汇刊》，第69卷，第9期，第5988–6002页，2021年6月。

[53] M. S. Lobo、L. Vandenberghe、S. Boyd 和 H. Lebret，“二阶锥规划的应用”，《线性代数及其应用》，第284卷，第1-3期，第193–228页，1998年。[54] M. S. Ali、E. Hossain、A. Al-Dweik 和 D. I. Kim，“多小区网络中 CoMP-NOMA 的下行功率分配”，IEEE《通信汇刊》，第66卷，第9期，第3982–3998页，2018年9月。[55] W. Ammar、D. Groeneveld、C. Bhagavatula、I. Beltagy、M. Crawford、D. Downey、J. Dunkelberger、A. Elgohary、S. Feldman、V. Ha 等，“Semantic Scholar 文献图谱的构建”，arXiv 预印本 arXiv:1805.02262，2018。[56] K. Seong、M. Mohseni 和 J. M. Cioffi，“OFDMA下行系统的最优资源分配”，载于 IEEE 信息理论国际研讨会论文集，美国华盛顿州西雅图，2006年，第1394–1398页。[57] Z. Yang、C. Pan、J. Hou 和 M. Shikh-Bahaei，“NOMA 移动边缘计算网络的高效资源分配：完成时间与能量最小化”，IEEE《通信汇刊》，第67卷，第11期，第7771–7784页，2019年。