

QoE-Aware Resource Allocation for Semantic Communication Networks

Lei Yan[†], Zhijin Qin[‡], Rui Zhang[†], Yongzhao Li[†], and Geoffrey Ye Li[§]

[†]State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an, China

[‡]School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK

[§]School of Electrical and Electronic Engineering, Imperial College London, London, UK

E-mail: lyang@stu.xidian.edu.cn, z.qin@qmul.ac.uk, {rz, yzhli}@xidian.edu.cn, geoffrey.li@imperial.ac.uk

Abstract—With the aim of accomplishing intelligence tasks, semantic communications transmit task-related information only, yielding significant performance gains over conventional communications. To guarantee user requirements for different types of tasks, we perform the semantic-aware resource allocation in a multi-cell multi-task network in this paper. Specifically, an approximate measure of semantic entropy is first developed to quantify the semantic information for different tasks, based on which a novel quality-of-experience (QoE) model is proposed. We formulate the QoE-aware semantic resource allocation in terms of the number of transmitted semantic symbols, channel assignment, and power allocation. To solve this problem, we first decouple it into two independent subproblems. The first one is to optimize the number of transmitted semantic symbols with given channel assignment and power allocation, which is solved by the exhaustive searching method. The second one is the channel assignment and power allocation subproblem, which is modeled as a many-to-one matching game and solved by the proposed low-complexity matching algorithm. Simulation results demonstrate the effectiveness and superiority of the proposed method on the overall QoE.

Index Terms—Quality of experience, resource allocation, semantic communications, semantic-aware networks.

I. INTRODUCTION

Different from conventional communications that focus on the engineering problem regardless of semantics and transmission tasks, semantic communications are task-oriented where only task-related information is extracted and transmitted [1]. To be specific, different types of tasks, including single-modal tasks, e.g., text transmission [2] and image retrieval [3], and multimodal tasks, e.g., visual question answering (VQA) [?], require different semantic transceivers to ensure the superior performance in transmission reliability and efficiency. Therefore, the conventional resource allocation model is not applicable and the resource allocation for such a semantic-aware network need to be revisited.

Some researchers have made several preliminary studies on this issue. In [5] and [6], the importance weights of the extracted semantic features are taken into account. Particularly, in addition to the conventional optimization variables, the partial semantic information to be transmitted was optimized in [5] for the text transmission task. The compression ratio of semantic features has been investigated in [6] for the image classification task. However, both works aimed to maximize the task performance, without evaluating the semantic com-

munication efficiency. Although Xia *et al.* [7] maximized the system throughput in message for text transmission, considering the dynamic background knowledge matching condition, they optimized the resource allocation at the message level rather than the semantic level.

To investigate the resource allocation at the semantic level, our prior work [8] has defined the semantic transmission rate (S-R) and the semantic spectral efficiency (S-SE), and formulated a problem to maximize the overall S-SE. Nevertheless, since the parameters related to the semantic information quantification are constants for a single task and are ignored to obtain the solutions, the work can only be used to allocate resources for the users with the same task. To cope with the coexistence of multiple tasks, we conduct further study on the semantic-aware resource allocation in this paper.

Particularly, since the performance of semantic communications depends on whether the user can complete the task, focusing more on the subjective user experience, semantic communications inherently have the advantage of enhancing quality-of-experience (QoE) [9], which considers both objective quality-of-service (QoS) and subjective user experience [10]. In this regard, we study the semantic-aware resource allocation based on the QoE metric in a multi-cell multi-task network. The main contributions of this paper are summarized as follows:

- A QoE-aware resource allocation problem is formulated for semantic communication networks. Specifically, based on the developed approximate semantic entropy, a novel QoE model is used to formulate the optimization problem in terms of the number of transmitted semantic symbols, channel assignment, and power allocation.
- The formulated problem is decoupled into two subproblems, which are solved by the exhaustive searching method and a low-complexity matching algorithm, respectively.
- Simulation results verify the superiority of the proposed QoE-aware semantic resource allocation method in terms of the overall QoE against the baselines.

Notation: Bold-font variables represent matrices and vectors. Calligraphic-font variables represent sets. The superscript, \mathbf{H}^H , represents the conjugate transpose of \mathbf{H} . $\mathbb{E}(x)$ is the expectation of x . In addition, $x \sim U(a, b)$ means that x follows a uniform distribution over the interval $[a, b]$, and

$x \sim N(\mu, \sigma^2)$ means that x follows a normal distribution with mean μ and covariance σ^2 .

II. SYSTEM MODEL

In this article, we consider an uplink cellular network with B coordinated cells, where each cell has a base station (BS) with N_r receiving antennas and a set of single antenna users. Let \mathcal{U}^b denote the user set in the b -th cell. We assume that B BSs form a cooperation cluster, which can exchange channel state information (CSI) and scheduling information with a central controller through high-speed low-delay fiber backhaul links [11]. In addition, we consider the coexistence of different intelligence tasks in the network, including single-modal tasks and multimodal tasks. Task-specific semantic communication transmitters and receivers are equipped at users and BSs, respectively. The transmission model and semantic communication models are introduced below.

A. Transmission Model

In order to fully exploit spectrum resources, we assume that the B coordinated cells share the same channel set denoted by \mathcal{M} , and each user can occupy at most one channel with bandwidth W . In each cell, the channels are orthogonally allocated to users to eliminate the intra-cell interference. With the maximal ratio combining (MRC) detection at the BS, the receiving signal-to-interference-plus-noise ratio (SINR) of user $u, u \in \mathcal{U}^b$ at the b -th BS can be expressed as

$$\gamma_u^b = \sum_{m \in \mathcal{M}} \frac{\alpha_{u,m} p_u |\mathbf{W}_{u,m}^b \mathbf{H}_{u,m}^b|^2}{\|\mathbf{W}_{u,m}^b\|^2 \sigma^2 + I_{u,m}^b}, \quad (1)$$

where $\alpha_{u,m} \in \{0, 1\}$, $\alpha_{u,m} = 1$ if channel $m \in \mathcal{M}$ is allocated to user u and $\alpha_{u,m} = 0$ otherwise, p_u is the transmit power of user u , and $\mathbf{H}_{u,m}^b$ is the $N_r \times 1$ channel matrix from user u to the b -th BS over channel m . Accordingly, the MRC detection matrix can be obtained by $\mathbf{W}_{u,m}^b = (\mathbf{H}_{u,m}^b)^H$. In (1), σ^2 is the noise power spectral density and $I_{u,m}^b$ is the interference experienced by user u over channel m from the adjacent cells, that is

$$I_{u,m}^b = \sum_{b' \in \mathcal{B}, b' \neq b} \sum_{u' \in \mathcal{U}^{b'}} \alpha_{u',m} p_{u'} |\mathbf{W}_{u,m}^b \mathbf{H}_{u',m}^{b'}|^2. \quad (2)$$

B. Semantic Communication Models

We focus on two types of intelligence tasks in this paper, including a single-modal task and a bimodal task. However, the proposed algorithm can be extended to the case of multiple multimodal intelligence tasks easily. Assume that N_{Si} single-modal users and N_{Bi} bimodal user pairs are randomly deployed in the considered network, and the b -th cell has N_{Si}^b single-modal users and N_{Bi}^b bimodal user pairs. For ease of clarification, a single-modal user or a bimodal user pair is regarded as a user group. The index set of all user groups in the b -th cell is given by

$$\begin{aligned} \mathcal{Q}^b &= \{1, 2, \dots, q, \dots, |\mathcal{Q}^b|\} \\ &= \{1, 2, \dots, N_{\text{Bi}}^b, N_{\text{Bi}}^b + 1, \dots, N_{\text{Bi}}^b + N_{\text{Si}}^b\}, \end{aligned} \quad (3)$$

where $|\mathcal{Q}^b| = N_{\text{Bi}}^b + N_{\text{Si}}^b$. Accordingly, the q -th user group contains a bimodal user pair when $q \leq N_{\text{Bi}}^b$, and it contains a single-modal user when $q > N_{\text{Bi}}^b$. Besides, we denote the set of all user groups in the b -th cell by $\mathcal{G}^b = \{\mathcal{G}_q^b\}_{q \in \mathcal{Q}^b}$ where $\mathcal{G}_q^b \subseteq \mathcal{U}^b$ represents the user set of the q -th group.

For the single-modal task, we take text transmission task as an example. Specifically, DeepSC [2] is adopted, where the sentence generated by the user is first mapped to semantic symbols that can be directly transmitted over the physical channel by the DeepSC transmitter, and then recovered by the DeepSC receiver at the BS. Moreover, semantic similarity is used to evaluate the performance, which is a function of the number of transmitted semantic symbols and SINR [8]. Thus, the semantic similarity of the user group $\mathcal{G}_q^b, q > N_{\text{Bi}}^b$ can be expressed as $\xi_q^b = f_{\text{Si}}(k_u, \gamma_u^b), u \in \mathcal{G}_q^b$ where k_u is the average number of transmitted semantic symbols of user u .

For the bimodal task, we take VQA task as an example and adopt the developed DeepSC-VQA model [?]. This task involves two users for text and image transmission, respectively. The two users first extract the semantic symbols from the text and image information through the DeepSC-VQA transmitter, respectively, and then send them to the BS. The received semantic symbols of text and image will be fused by the DeepSC-VQA receiver at the BS to predict the answer. As the two users jointly decide the task performance, the answer accuracy of this task could be modeled as a function with respect to the numbers of transmitted semantic symbols and the SINR of the two users, i.e., $\xi_q^b = f_{\text{Bi}}(k_{u_t}, k_{u_i}, \gamma_{u_t}^b, \gamma_{u_i}^b), u_t, u_i \in \mathcal{G}_q^b$, where u_t and u_i represent the user for text transmission and the one for image transmission, respectively.

III. PROBLEM FORMULATION

In this section, we first develop a method to obtain an approximate measure of semantic entropy, based on which a novel QoE model for semantic communications is proposed. Then, the semantic resource allocation is formulated as a QoE maximization problem.

A. Approximate Measure of Semantic Entropy

Information entropy measures the information based on the statistical characteristic of source symbols, while semantic entropy directly quantifies the semantic information of the source and plays a significant role in evaluating semantic communication efficiency and resource allocation. However, a common semantic information metric is still missing, and most existing metrics are non-calculable [1].

Since the semantic information carried by a source depends on the specific task, semantic entropy should be a measure with respect to the source and the task. Following Chattopadhyay *et al.* [12], we define the semantic entropy as following:

Definition 1. Given semantic source \mathcal{X} , semantic entropy is defined as the minimum number of semantic symbols about the data $X \in \mathcal{X}$ that are sufficient to predict task Y , i.e.,

$$\begin{aligned} H(X; Y) &\triangleq \min_E \mathbb{E}(|\text{Code}^E(X)|) \\ \text{s.t. } P(Y|\text{Code}^E(X)) &= P(Y|X), \end{aligned} \quad (4)$$

where $\text{Code}^E(x)$ denotes the semantic symbol vector extracted from X with the semantic encoder E , and $P(Y|X)$ is the conditional probability of Y given X .

From Definition 1, the semantic entropy of X given Y is actually defined as an expected value over the whole data set \mathcal{X} , i.e., the semantic entropy is a constant for the same task and dataset, which shares a similar philosophy as [8]. However, it is intractable to find an optimal E to derive the semantic entropy. To obtain a measure that is both meaningful and manipulable, we utilize a well-designed deep learning (DL) model as the semantic encoder to obtain the approximate semantic entropy for a task, which can be expressed as

$$\begin{aligned} \tilde{H}(X; Y) &\triangleq \min \mathbb{E}(|\text{Code}^{E_{\text{DL}}}(X)|) \\ \text{s.t. } P(Y|X) - P(Y|\text{Code}^{E_{\text{DL}}}(X)) &< \varepsilon, \end{aligned} \quad (5)$$

where the constraint indicates that the gap between $P(Y|X)$ and $P(Y|\text{Code}^{E_{\text{DL}}}(X))$ can not exceed ε .

According to the aforementioned method, the approximate semantic entropy of the considered tasks can be derived based the corresponding DL models. In particular, we first remove the channel models from DeepSC and DeepSC-VQA, then train them under different settings of the number of semantic symbols, and finally find the minimum number of semantic symbols that can guarantee a performance very close to the upper bound. Furthermore, we define the unit of the approximate semantic entropy as *sut* as in [8].

B. QoE Model for Semantic Communications

We formulate the QoE model based on two objective metrics, *semantic accuracy* and *semantic rate*. The former, the accuracy of message transmission, corresponds to the semantic similarity in the considered single-modal task and the answer accuracy in the considered bimodal task. The latter is defined as the amount of semantic information emitted to the transmission medium per second, measured in *suts/s*. The semantic rate of the single-modal user $u \in \mathcal{G}_q^b, q > N_{\text{Bi}}^b$ is given as

$$\varphi_u = \frac{\tilde{H}_{\text{Si}}}{k_u/W}, \quad (6)$$

where \tilde{H}_{Si} represents the DeepSC [2] based approximate semantic entropy. The semantic rate of bimodal users u_t and u_i , $u_t, u_i \in \mathcal{G}_q^b, q \leq N_{\text{Bi}}^b$ are expressed as

$$\varphi_{u_t} = \frac{\tilde{H}_{\text{Bi},t}}{k_{u_t}/W}, \quad \text{and } \varphi_{u_i} = \frac{\tilde{H}_{\text{Bi},i}}{k_{u_i}/W}, \quad (7)$$

respectively, where $\tilde{H}_{\text{Bi},t}$ and $\tilde{H}_{\text{Bi},i}$ represent the DeepSC-VQA [?] based approximate semantic entropy for text transmission user and image transmission user, respectively.

Note that the semantic rate is significantly different from the S-R that is defined as the amount of *successfully* delivered semantic information per second in [8]. With Γ_u denoting the S-R, we have $\Gamma_u = \varphi_u \xi_q$, i.e., the difference between them lies in that whether the semantic accuracy is considered. The S-R couples semantic accuracy and semantic rate. However,

from the perspective of users, accuracy and efficiency of message transmission are different, and users may have their own preferences on them depending on the applications. For example, some users prefer high accuracy but are delay-tolerated, while others may desire higher semantic rate but do not need a very high accuracy. In order to reflect the QoE requirements of users more properly, we consider the two parameters to formulate the QoE model as

$$\begin{aligned} QoE_q^b &= \sum_{u \in \mathcal{G}_q^b} w_u G_u^R + (1 - w_u) G_u^A \\ &= \sum_{u \in \mathcal{G}_q^b} \frac{w_u}{1 + e^{\beta_u(\varphi_u^{\text{req}} - \varphi_u)}} + \frac{(1 - w_u)}{1 + e^{\lambda_u(\xi_u^{\text{req}} - \xi_q^b)}}, \end{aligned} \quad (8)$$

where QoE_q^b is the QoE of the q -th user group in the b -th cell, w_u and $(1 - w_u)$ are the weights of semantic rate and semantic accuracy at user u respectively, G_u^R and G_u^A are the scores of semantic rate and semantic accuracy at user u respectively, and β_u and λ_u represent the growth rates of G_u^R and G_u^A respectively. Additionally, φ_u^{req} and ξ_u^{req} represent the minimum semantic rate and semantic accuracy to acquire the 50% of the scores, respectively. Due to the simplicity and generality, we choose the logistic function to model the correlation between QoE and QoS metrics while other QoE function types are applicable as well for the proposed method, such as MOS-based function or exponential function. Here, QoE_q^b , G_u^R , and G_u^A are between 0 and 1.

C. Problem Formulation

In this part, we formulate the semantic-aware resource allocation with the goal of maximizing the overall QoE of all users in terms of channel assignment, power allocation, and the number of transmitted semantic symbols. The optimization problem can be expressed as follows:

$$(\text{P0}) \quad \max_{\{\{\alpha_{u,m}\}, \{p_u\}, \{k_u\}\}} \sum_{b \in \mathcal{B}} \sum_{q \in \mathcal{Q}^b} QoE_q^b \quad (9)$$

$$\text{s.t. } C_1 : \alpha_{u,m} \in \{0, 1\}, \quad \forall u \in \mathcal{U}^b, \forall m \in \mathcal{M}, \forall b \in \mathcal{B}, \quad (9a)$$

$$C_2 : \sum_{u \in \mathcal{U}^b} \alpha_{u,m} \leq 1, \quad \forall m \in \mathcal{M}, \forall b \in \mathcal{B}, \quad (9b)$$

$$C_3 : \sum_{m \in \mathcal{M}} \alpha_{u,m} \leq 1, \quad \forall u \in \mathcal{U}^b, \forall b \in \mathcal{B}, \quad (9c)$$

$$C_4 : \sum_{u \in \mathcal{G}_q^b, q \leq N_{\text{Bi}}^b} \alpha_{u,m} \in \{0, 2\}, \quad \forall m \in \mathcal{M}, \forall b \in \mathcal{B}, \quad (9d)$$

$$C_5 : k_u \in \mathcal{K}_u, \quad \forall u \in \mathcal{U}^b, \forall b \in \mathcal{B}, \quad (9e)$$

$$C_6 : 0 \leq p_u \leq P_{\text{max}}, \quad \forall u \in \mathcal{U}^b, \forall b \in \mathcal{B}, \quad (9f)$$

$$C_7 : G_u^R, G_u^A \geq G_{\text{th}}, \quad \forall u \in \mathcal{U}^b, \forall b \in \mathcal{B}, \quad (9g)$$

where C_1 constrains the range of $\alpha_{u,m}$, C_2 ensures the orthogonal channels for each user in a cell, C_3 restricts each user to occupy at most one channel, C_4 ensures that a bimodal user pair will be allocated no channel or two channels as only

one channel assigned to one of them will lead to a failure, C_5 specifies the range of the number of transmitted semantic symbols for each user, and $\mathcal{K}_u = \mathcal{K}_{Si}, \mathcal{K}_{Bi,t}$, and $\mathcal{K}_{Bi,i}$ when u is the user with the single-modal task, the user with the bimodal task for text transmission, and the user with the bimodal task for image transmission, respectively, C_6 constrains the range of transmit power, and C_7 limits the minimum required scores of semantic rate and semantic accuracy.

IV. A MATCHING THEORY BASED SOLUTION

In this section, we first decouple (P0) into two independent subproblems, and then solve them respectively to obtain a suboptimal solution.

A. Problem Decoupling

By observing (P0), the optimization variables, $\{\alpha_{u,m}\}$ and $\{p_u\}$, jointly decide the SINR γ_u^b . Additionally, the objective function can be expressed as a function with respect to γ_u^b and k_u . Therefore, given γ_u^b , the optimal $\{k_u\}$ for each \mathcal{G}_q^b can be obtained by solving the following optimization problem:

$$\begin{aligned} (\mathbf{P1}) \quad & \max_{\{k_u\}, u \in \mathcal{G}_q^b} QoE_q^b \\ \text{s.t.} \quad & C_5 \text{ and } C_7. \end{aligned} \quad (10)$$

Then the channel assignment and power allocation problem can be written as

$$\begin{aligned} (\mathbf{P2}) \quad & \max_{\{\alpha_{u,m}\}, \{p_u\}} \sum_{b \in \mathcal{B}} \sum_{q \in \mathcal{Q}^b} QoE_q^b \\ \text{s.t.} \quad & C_1, C_2, C_3, C_4, \text{ and } C_6. \end{aligned} \quad (11)$$

Therefore, for each option of $\{\{\alpha_{u,m}\}, \{p_u\}\}$, the SINR of each user in the network can be calculated using equation (1). Then the optimal k_u and the corresponding maximum QoE_q^b can be obtained by solving (P1). Since ξ_q^b can only be obtained by the look-up table method, the exhaustive searching method is utilized to solve (P1) here. Note that the complexity is acceptable due to the limitation of $|\mathcal{K}_u|$. Finally, with the obtained maximum QoE_q^b , we can solve (P2) to get the solutions, which will be detailed in the next subsection.

B. Channel Assignment and Power Allocation Subproblem

To cope with the tight coupling among users in multiple cells and those in a bimodal user pair, we construct a matching game to model (P2) and propose a low-complexity matching algorithm to obtain the stable matching in this part.

Problem (P2) is a three-sided many-to-one matching game among users, channels, and power levels, where each user can select at most one channel and one power level while each channel or power level can serve multiple users. For the convenience of algorithm implementation, we put all combinations of channels and power levels together to form a resource set $\mathcal{T} = \{(m, p), \forall m \in \mathcal{M}, \forall p \in \mathcal{P}\}$ where \mathcal{P} denotes the power set. Then the three-sided many-to-one matching can be converted to a two-sided matching. In addition, since the QoE of each user depends not only on the opposite partners to be matched, but also on the users of other cells sharing the same

channel. More specifically, this is a matching problem with externalities. The idea of swap matching [13] can be used for reference to obtain the solution.

Considering two cases may exist in each cell, i.e., $|\mathcal{U}^b| \leq |\mathcal{M}|$ and $|\mathcal{U}^b| > |\mathcal{M}|$, the markets of different sides will be focused on to maximize the overall QoE. In particular, the proposed algorithm aims to keep the utility of users increasing by the swap operation to achieve the stable matching for the first case, while focuses on the utility of channels for the second case.

Case 1: $|\mathcal{U}^b| \leq |\mathcal{M}|$. Firstly, we add $|\mathcal{M}| - |\mathcal{U}^b|$ virtual single-modal users to enable every channel to match with a user, making the swap operation easier. Then, the user group set in the b -th cell can be updated as $\mathcal{G}_A^b = \mathcal{G}^b \cup \mathcal{U}_0^b$, where \mathcal{U}_0^b is the established virtual single-modal user set with $|\mathcal{U}_0^b| = |\mathcal{M}| - |\mathcal{U}^b|$. Correspondingly, we denote the updated index set of \mathcal{G}_A^b as \mathcal{Q}_A^b .

As the users belonging to a user group jointly decide the task performance, a bimodal user pair (*rather than an individual bimodal user*) or a single-modal user is regarded as a selfish and rational player to make decisions, i.e., there are totally $|\mathcal{G}_A^b|$ players in the b -th cell and each of them is denoted by $\mathcal{G}_q^b \in \mathcal{G}_A^b$. Clearly, the resource sets of the players with the single-modal task and those with the bimodal task are different, which are $\mathcal{T}_{Si} = \mathcal{T}$ and $\mathcal{T}_{Bi} = \{(m, m', p, p'), \forall m, m' \in \mathcal{M}, m \neq m', \forall p, p' \in \mathcal{P}\}$, respectively. Thus, by denoting \mathcal{T}_q^b as the resource set of the player \mathcal{G}_q^b , we have $\mathcal{T}_q^b = \mathcal{T}_{Bi}$ when $q \leq N_{Bi}^b$ and $\mathcal{T}_q^b = \mathcal{T}_{Si}$, otherwise.

Then, we can define a matching Φ as a function from set $\mathcal{G}_A \cup \mathcal{T}_A$ mapping into set $\mathcal{G}_A \cup \mathcal{T}_A$, where $\mathcal{G}_A = \bigcup_{b \in \mathcal{B}} \mathcal{G}_A^b$ and $\mathcal{T}_A = \mathcal{T}_{Si} \cup \mathcal{T}_{Bi}$. For a matching pair $(\mathcal{G}_q^b, t), t \in \mathcal{T}_A$, we have $\Phi(\mathcal{G}_q^b) = t$ and $\Phi(t) = \mathcal{G}_q^b$. The utility function of player \mathcal{G}_q^b under matching Φ is defined as

$$U_q^b(\Phi) = \begin{cases} QoE_q^b, & \text{if } q \in \mathcal{Q}^b; \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

Based on this, whether the considered matching is stable can be judged from the following definition.

Definition 2. A matching Φ is stable if and only if, for each player $\mathcal{G}_q^b \in \mathcal{G}_A^b, \forall b \in \mathcal{B}$ with $\Phi(\mathcal{G}_q^b) = t$, there does not exist a blocking matching Φ_q^B such that, $\exists t' \in \mathcal{T}_q^b$:

- 1) $\forall i \in \mathcal{G}_{(t)} \cup \mathcal{G}_{(t')}, U_i^b(\Phi_q^{(t,t')}) \geq U_i^b(\Phi_q^B)$ and
- 2) $\exists i \in \mathcal{G}_{(t)} \cup \mathcal{G}_{(t')}, U_i^b(\Phi_q^{(t,t')}) > U_i^b(\Phi_q^B)$,

where $\mathcal{G}_{(t)}$ is formed by all players who have at least one same channel as those included in t , and $\Phi_q^{(t,t')}$ represents the swap matching, where the player \mathcal{G}_q^b swap t with t' , the players who are in both $\mathcal{G}_{(t')}$ and the b -th cell swap their channels accordingly, and the remaining players keep unchanged.

Case 2: $|\mathcal{U}^b| > |\mathcal{M}|$. Evidently, if we follow the solution for **Case 1**, no channel will be tentatively matched with $|\mathcal{U}^b| - |\mathcal{M}|$ of all users even if the overall QoE will increase. To deal with this problem, we focus on the market of channels to achieve the stable matching. However, when a channel is matched with

a user with the bimodal task, its utility depends not only on which user it matches but on which channel the other user in the same group matches. Hence, we propose the user groups to perform the swap operation but the channels to make the swap decision based on their utility changes.

Similar to **Case 1**, we first add $|\mathcal{U}^b| - |\mathcal{M}|$ virtual channels to enable the swap operation at each user. Then, the channel set in the b -th cell can be updated as $\mathcal{M}_A^b = \mathcal{M} \cup \mathcal{M}_0^b$, where \mathcal{M}_0^b is the established virtual channel set with $|\mathcal{M}_0^b| = |\mathcal{U}^b| - |\mathcal{M}|$. Then, the resource sets can be reformed in the similar manner but based on \mathcal{M}_A^b . For simplicity, we use the same notation \mathcal{T}_q^b to denote the resource set of each player $\mathcal{G}_q^b \in \mathcal{G}^b$. Furthermore, the utility function of channel $m \in \mathcal{M}_A^b$ in the b -th cell under a matching Φ is defined as

$$U_m^b(\Phi) = \begin{cases} 0, & \text{if } m \in \mathcal{M}_0^b; \\ U_{\mu(m)}^b(\Phi)/2, & \text{if } m \in \mathcal{M}^b, \mu(m) \leq N_{\text{Bi}}^b; \\ U_{\mu(m)}^b(\Phi), & \text{if } m \in \mathcal{M}^b, \mu(m) > N_{\text{Bi}}^b. \end{cases} \quad (13)$$

where $\mu(m) \in \mathcal{Q}^b$ represents the index of the player that matches channel m . When the player is a bimodal user pair, $U_m^b(\Phi)$ is set as the half of the utility of the player, and when the player is a single-modal user, $U_m^b(\Phi)$ is equal to the utility of the player. Then, we give the following definition.

Definition 3. A matching Φ is stable if and only if, for each channel $m \in \mathcal{M}_A^b$ with $\mu(m) \in \mathcal{Q}^b$, there does not exist a blocking matching Φ_m^B such that, $\exists t' \in \mathcal{T}_{\mu(m)}^b$:

- 1) $\forall i \in \mathcal{M}_{(\mathcal{G})}, U_i^b(\Phi_{\mu(m)}^{(t,t')}) \geq U_i^b(\Phi_m^B)$ and
- 2) $\exists i \in \mathcal{M}_{(\mathcal{G})}, U_i^b(\Phi_{\mu(m)}^{(t,t')}) > U_i^b(\Phi_m^B)$,

where $\mathcal{M}_{(\mathcal{G})}$ is formed by all channels of all players in \mathcal{G} , $\mathcal{G} = \mathcal{G}_{(t)} \cup \mathcal{G}_{(t')}$.

Based on the two definitions, we develop an efficient matching algorithm to obtain the stable matching, as shown in Algorithm 1. The proposed algorithm starts by a random initial matching. Then the users will search their resource set to find the blocking matching and thus update the current matching. Once no blocking matching is found, the stable matching Φ will be the output.

Remark 1: The worst-case complexity of Algorithm 1 is as $\mathcal{O}(V \times |\mathcal{B}| \times ((|\mathcal{G}_A^b| - N_{\text{Bi}}^b) \times |\mathcal{M}_A^b| \times |\mathcal{P}| + N_{\text{Bi}}^b \times \binom{2}{|\mathcal{M}_A^b|} \times |\mathcal{P}|^2))$ where V is the number of iterations. However, the complexity of exhaustive searching method here is as $\mathcal{O}((\frac{\max(|\mathcal{U}^b|, |\mathcal{M}|)!}{\min(|\mathcal{U}^b|, |\mathcal{M}|)!} \times |\mathcal{P}|^{|\mathcal{U}^b|})^{|\mathcal{B}|})$.

Remark 2: Since the utility of user groups or channels is bounded by 1 and will increase monotonically by the swap operation, Algorithm 1 will terminate to a final matching after a finite number of iterations.

V. SIMULATION RESULTS

In this section, we present simulation results to demonstrate the performance of the proposed QoE-aware resource allocation method in a multi-cell multi-task network.

We consider three cells in the simulation, where users are randomly deployed. The radius of each cell is set as 500 m.

Algorithm 1: Proposed Matching Algorithm for Channel Assignment and Power Allocation

Input: $\mathcal{B}, \mathcal{U}^b, \mathcal{M}_A^b, \mathcal{G}_A^b, \mathcal{T}_q^b, \forall q \in \mathcal{Q}_A^b, \forall b \in \mathcal{B}$.

```

1 Initialization: For each cell  $b \in \mathcal{B}$ , initial the matching
   channels and power levels of all users with a
   permutation of all channels and the minimum power
   level, respectively. Denote the current matching as  $\Phi$ .
2 repeat
3   for all  $b \in \mathcal{B}$  do
4     for all  $\mathcal{G}_q^b \in \mathcal{G}_A^b$  do
5        $t = \Phi(\mathcal{G}_q^b)$ ;
6       for all  $t' \in \mathcal{T}_q^b, t' \neq t$  do
7         if  $|\mathcal{U}^b| \leq |\mathcal{M}|$  then
8           If  $\Phi$  is a blocking matching
             according to Definition 2, update
             the matching as  $\Phi = \Phi_q^{(t,t')}$ ;
             otherwise, keep the current
             matching state.
9         else
10          If  $\Phi$  is a blocking matching
            according to Definition 3, update
            the matching as  $\Phi = \Phi_{\mu(m)}^{(t,t')}$  where
             $m$  is the channel in  $t$ ; otherwise,
            keep the current matching state.
11 until No blocking matching is found;
Output: the stable matching  $\Phi$ .

```

Each BS is equipped with two receiving antennas and each user is with single antenna. For the channel model, both large-scale fading, including pathloss and shadowing, and small-scale Rayleigh fading are considered. The adopted pathloss model is $128.1 + 37.6 \lg[d(\text{km})]$ dB and the shadowing factor is set as 6 dB. The bandwidth of each channel is $W = 180$ kHz. The noise power spectral density is $\sigma^2 = -174$ dBm/Hz. In the simulation, we consider the available power levels as $\mathcal{P} = \{-10, -5, 0, 5, 10, 15, 20\}$ dBm. Meanwhile, each user generates its QoE-related parameters as $w_u \sim U(0, 1)$, $\xi_u^{\text{req}} \sim U(0.8, 0.9)$, and $\lambda_u \sim N(55, 2.5^2)$. In addition, for text transmission users, we set $\varphi_u^{\text{req}} \sim U(50, 70)$ in ksuts/s and $\beta_u \sim N(0.2, 0.05^2)$. For image transmission users, we set $\varphi_u^{\text{req}} \sim U(80, 100)$ in ksuts/s and $\beta_u \sim N(0.1, 0.02^2)$. The threshold of the scores is set as $G_{\text{th}} = 0.5$. Further, we run the DeepSC model and DeepSC-VQA model to obtain the mapping relations $\xi_q^b = f_{\text{Si}}(k_u, \gamma_u^b)$ and $\xi_q^b = f_{\text{Bi}}(k_{ut}, k_{ui}, \gamma_{ut}^b, \gamma_{ui}^b)$ based on the settings as $\mathcal{K}_{\text{Si}} = \{1, 2, \dots, 20\}$, $\mathcal{K}_{\text{Bi},t} = \{2, 4, 6, 8, 10\}$, and $\mathcal{K}_{\text{Bi},i} = \{394, 788, 1576, 2364, 3152\}$, respectively.

We first verify the superiority of the developed QoE based formulation. Fig. 1 compares the QoE maximization and S-R maximization methods [8]. Here, the upper bound is obtained by assuming that the maximum number of users in each cell, i.e., $\max(|\mathcal{U}^b|, |\mathcal{M}|)$, can be served and their QoE can reach 1. Thus the upper bound is irrelevant to G_{th} , which is compared to demonstrate the effectiveness of the proposed method. As

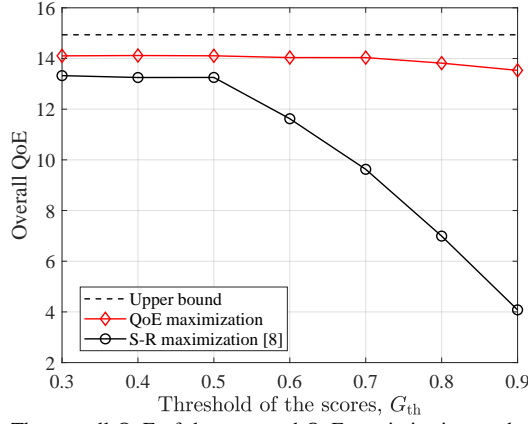


Fig. 1. The overall QoE of the proposed QoE maximization method and the S-R maximization method [8] with $(N_{Si}, N_{Bi}) = (6, 6)$ and $|\mathcal{M}| = 6$.

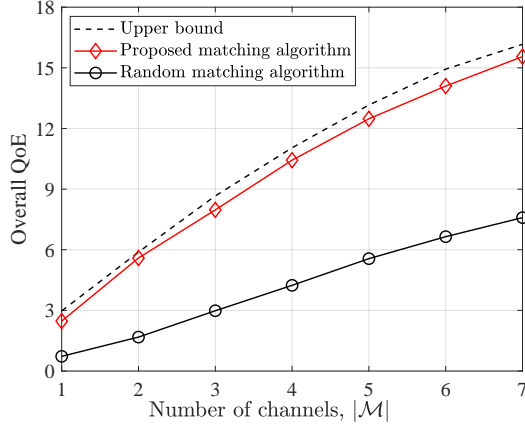


Fig. 2. The overall QoE of different algorithms with $(N_{Si}, N_{Bi}) = (6, 6)$.

G_{th} increases, the overall QoE of both methods decreases due to the tighter restriction. However, the overall QoE of the proposed method keep above 13 while the S-R maximization method shows worse performance, especially when $G_{th} > 0.5$. This implies that the proposed method can better adapt to the changes in the user requirements.

Fig. 2 illustrates the overall QoE versus the number of channels for different algorithms. Due to the unacceptable computation complexity of the exhaustive searching method as analysed in Section IV-B, we do not compare the proposed matching algorithm with that but with the upper bound of overall QoE, along with the random matching method. From this figure, the proposed matching algorithm outperforms the random method significantly and is very close to the upper bound.

The proposed method with and without multi-cell cooperation are compared in Fig. 3. The method with multi-cell cooperation yields better performance than the one without multi-cell cooperation, which verifies the effectiveness of the proposed method in coping with the inter-cell interference.

VI. CONCLUSION

In this paper, we studied the semantic-aware resource allocation in multi-task networks with multi-cell cooperation. Specifically, a novel quality of experience (QoE) model was developed for semantic communications in terms of semantic

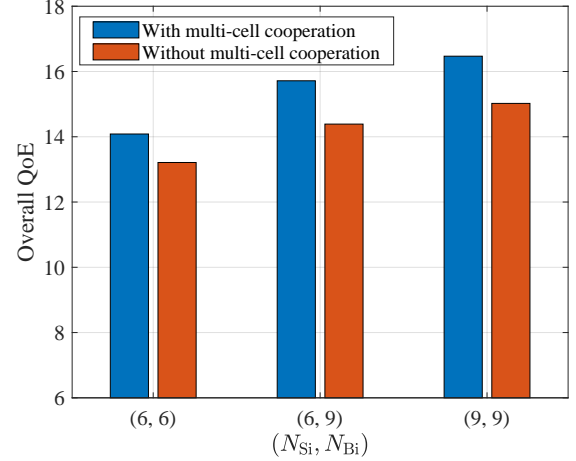


Fig. 3. The overall QoE of the proposed method with and without multi-cell cooperation with $|\mathcal{M}| = 6$.

rate and semantic accuracy. Then we formulated a QoE maximization problem and solved it by a matching theory based solution. Simulation results showed that the developed formulation can characterize the user requirements for semantic communications more properly and provide higher user satisfaction than the semantic transmission rate (S-R) based formulation. Moreover, the proposed algorithm outperforms all baselines significantly.

REFERENCES

- [1] Z. Qin, X. Tao, J. Lu, and G. Y. Li, "Semantic communications: Principles and challenges," Dec. 2021, *arXiv: 2201.01389*.
- [2] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep learning enabled semantic communication systems," *IEEE Trans. Signal Process.*, vol. 69, no. 1, pp. 2663–2675, Apr. 2021.
- [3] M. Jankowski, D. Gündüz, and K. Mikołajczyk, "Wireless image retrieval at the edge," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 89–100, 2021.
- [4] C. Liu, C. Guo, Y. Yang, and J. Chen, "Bandwidth and power allocation for task-oriented semantic communication," Jan. 2022, *arXiv: 2201.10795*.
- [5] Y. Wang, M. Chen, W. Saad, T. Luo, S. Cui, and H. V. Poor, "Performance optimization for semantic communications: An attention-based learning approach," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Madrid, Spain, Dec. 2021, pp. 1–6.
- [6] C. Liu, C. Guo, Y. Yang, and N. Jiang, "Adaptable semantic compression and resource allocation for task-oriented communications," Apr. 2022, *arXiv:2204.08910*.
- [7] L. Xia, Y. Sun, X. Li, G. Feng, and M. A. Imran, "Wireless resource management in intelligent semantic communication networks," Feb. 2022, *arXiv:2202.07632*.
- [8] L. Yan, Z. Qin, R. Zhang, Y. Li, and G. Y. Li, "Resource allocation for text semantic communications," *IEEE Wireless Commun. Lett.*, pp. 1–1, Apr. 2022.
- [9] G. Shi, Y. Xiao, Y. Li, and X. Xie, "From semantic communication to semantic-aware networking: Model, architecture, and open problems," *IEEE Commun. Mag.*, vol. 59, no. 8, pp. 44–50, Sep. 2021.
- [10] J. Cui, Y. Liu, Z. Ding, P. Fan, and A. Nallanathan, "QoE-based resource allocation for multi-cell NOMA networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 9, pp. 6160–6176, July 2018.
- [11] X. Lu, Q. Ni, W. Li, and H. Zhang, "Dynamic user grouping and joint resource allocation with multi-cell cooperation for uplink virtual MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3854–3869, Mar. 2017.
- [12] A. Chattopadhyay, B. D. Haeffele, D. Geman, and R. Vidal, "Quantifying task complexity through generalized information measures," <https://openreview.net/pdf?id=vcKVhY7AZqK>, 2020 (accessed 24th Dec. 2021).

- [13] E. Bodine-Baron, C. Lee, A. Chong, B. Hassibi, and A. Wierman, "Peer effects and stability in matching markets," in *Proc. Int. Symp. Algorithmic Game Theory*. Berlin, Germany: Springer, 2011, pp. 117–129.