A QoE-Driven Personalized Incentive Mechanism Design for AIGC Services in Resource-Constrained Edge Networks

Hongjia Wu, Minrui Xu, Zehui Xiong, Lin Gao, *Senior Member, IEEE*, Haoyuan Pan, *Member, IEEE*, Dusit Niyato, *Fellow, IEEE*, and Tse-Tin Chan, *Member, IEEE*

***Abstract*—With rapid advancements in large language models (LLMs), AI-generated content (AIGC) has emerged as a key driver of technological innovation and economic transformation. Personalizing AIGC services to meet individual user demands is essential but challenging for AIGC service providers (ASPs) due to the subjective and complex demands of mobile users (MUs), as well as the computational and communication resource constraints faced by ASPs. To tackle these challenges, we first develop a novel multi-dimensional quality-of-experience (QoE) metric. This metric comprehensively evaluates AIGC services by integrating accuracy, token count, and timeliness. We focus on a mobile edge computing (MEC)-enabled AIGC network, consisting of multiple ASPs deploying differentiated AIGC models on edge servers and multiple MUs with heterogeneous QoE requirements requesting AIGC services from ASPs. To incentivize ASPs to provide personalized AIGC services under MEC resource constraints, we propose a QoE-driven incentive mechanism. We formulate the problem as an equilibrium problem with equilibrium constraints (EPEC), where MUs as leaders determine rewards, while ASPs as followers optimize resource allocation. To solve this, we develop a dual-perturbation reward optimization algorithm, reducing the implementation complexity of adaptive pricing. Experimental results demonstrate that our proposed mechanism achieves a reduction of approximately 64.9% in average computational and communication overhead, while the average service cost for MUs and the resource consumption of ASPs decrease by 66.5% and 76.8%, respectively, compared to state-of-the-art benchmarks.**

***Index Terms*—Generative AI (GAI), incentive mechanism, mobile edge computing (MEC), multi-dimensional quality of experience (QoE), resource allocation.**

This work was supported in part by the National Natural Science Foundation of China under Grant 62371302; in part by the National Key Research and Development Program of China under Grant 2021YFB2900300; in part by the Natural Science Foundation of Guangdong Province under Grant 2024A1515010178; in part by the Shenzhen Science and Technology Program under Grants KQTD20190929172545139, GXWD20231129103946001, KJZD20240903095402004, ZDSYS20210623091808025, and JCYJ20250604181549064; and in part by the Research Matching Grant Scheme from the Research Grants Council of Hong Kong. *(Corresponding authors: Tse-Tin Chan and Haoyuan Pan.)*

H. Wu and T.-T. Chan are with the Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong, China (e-mail: whongjia@eduhk.hk; tsetinchan@eduhk.hk).

Z. Xiong is with the School of Electronics, Electrical Engineering and Computer Science (EEECS), Queen's University Belfast, Belfast, BT7 1NN, U.K. (e-mail: z.xiong@qub.ac.uk).

L. Gao is with the School of Electronics and Information Engineering and the Guangdong Provincial Key Laboratory of Aerospace Communication and Networking Technology, Harbin Institute of Technology, Shenzhen, China (e-mail: gaol@hit.edu.cn).

H. Pan is with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China (e-mail: hypan@szu.edu.cn).

M. Xu and D. Niyato are with the College of Computing and Data Science, Nanyang Technological University, Singapore (e-mail: minrui001@e.ntu.edu.sg; dniyato@ntu.edu.sg).

I. Introduction

A. Background and Motivation

Advancements in machine learning algorithms and large language models (LLMs) have significantly enhanced the quality and diversity of AI-generated content (AIGC), driving innovation across various domains and transforming the economic landscape [1]. AIGC services, such as OpenAI's ChatGPT, have demonstrated remarkable capabilities in tasks like language understanding, text generation, summarization, and conversational AI. With the widespread adoption of AIGC services, users across diverse applications have increasingly varied expectations for service performance, making personalized service a critical requirement. However, providing personalized AIGC services is hindered by the subjective and complex demands of mobile users (MUs), as well as the limited computational and communication resources of AIGC service providers (ASPs). Addressing these barriers in mobile AIGC networks requires overcoming both technical and economic challenges.

Challenge 1: Resource Constraints. Requesting real-time AIGC services directly from the cloud often incurs significant latency and network congestion. Furthermore, the inference of LLMs necessary for AIGC services is resource-intensive and time-consuming, making it impractical for deployment on mobile devices due to high costs. To this end, current schemes [2], [3] deploy pre-trained AIGC models on edge servers, leveraging the infrastructure of wireless edge networks. When MUs request AIGC services, they can send the prompts to edge servers. The prompts are then executed using generative AI models, and the results are sent back to MUs. This approach alleviates the computational burden on mobile devices and enables flexible and scalable AIGC service provisioning in mobile edge networks. However, when numerous MUs simultaneously request services, edge servers with limited computational and communication resources can become overwhelmed, resulting in increased latency and degraded performance.

Challenge 2: Personalized Demands. MUs exhibit diverse service demands depending on their applications. For example, MUs in vehicular networks prioritize concise and rapid decision-making responses [4], whereas those involved in academic essay writing require detailed and profound content. To meet MUs' expectations, aligning AIGC services with their personalized demands is essential, considering the unique characteristics of different application areas. However, achieving this goal faces two main problems. First, the stochastic nature of LLMs, which relies on next-token

prediction [5], may result in incoherent outputs, factual inaccuracies, or contradictions. This uncertainty makes it difficult for LLMs to consistently generate highly personalized content that accurately matches MUs' expectations. Second, existing service assessment metrics, such as designer-perceivable quality of experience (QoE) [6], joint perpetual similarity and quality (JPSQ) [7], and user-perceived quality [8], fail to adequately consider the complex trade-offs between accuracy, token count, and timeliness in mobile AIGC networks.

Challenge 3: Self-Interested Behaviors of ASPs and MUs. ASPs are driven by profit and tend to reduce resource investment without sufficient incentives. Meanwhile, MUs prefer receiving high-quality services at minimal cost. This misalignment of goals can lead to a structural imbalance between service provision and economic incentives. In response to this contradiction, existing research has proposed incentive strategies using mechanisms such as multi-agent deep reinforcement learning [9], game-theoretic [10], and contract theory [11], [12]. However, these studies fail to explicitly model the multi-dimensional demands of individual users, making it challenging for ASPs to optimize resource allocation for diverse user needs. Furthermore, they often focus on a single type of resource and overlook the complexity involved in designing incentives for multi-ASP scenarios and the inherent competition among MUs.

To address these challenges, we propose a dual analytical framework from the perspectives of both MUs and ASPs, leading to the following two key research questions:

i) *How can MUs optimize service costs while obtaining personalized AIGC services that meet their heterogeneous requirements in a resource-competitive network?*

ii) *How can ASPs optimally allocate computational and communication resources to minimize resource consumption while delivering personalized AIGC services?*

B. Solution and Contributions

As illustrated in Fig. 1, we consider a mobile edge computing (MEC)-enabled AIGC network, consisting of multiple ASPs with differentiated AIGC models deployed on edge servers and multiple MUs with heterogeneous QoE requirements requesting AIGC services from ASPs. To address the above problems, we first design a multi-dimensional quality of experience (QoE) metric to quantify MUs' personalized demands. Building on this metric, we propose a QoE-driven incentive mechanism that aligns the interests of both MUs and ASPs. In our framework, MUs act as leaders by driving the trading process through differentiated reward strategies, while ASPs function as followers by dynamically adjusting resource allocation in response. This mechanism achieves equilibrium through bidirectional optimization: Competing MUs strategically determine rewards in a non-cooperative game, each aiming to obtain better personalized AIGC services at lower costs (problem i)), while ASPs jointly optimize the allocation of computational and communication resources to balance QoE fulfillment and resource consumption based on these rewards (problem ii)). In summary, the key contributions of this work are summarized as follows.

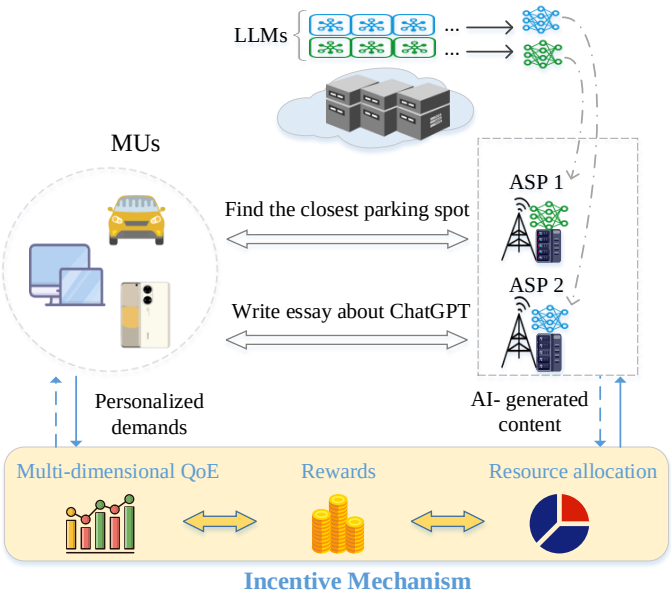


Fig. 1. Illustration of the proposed QoE-driven incentive mechanism in an MEC-enabled AIGC network. The network comprises multiple ASPs deploying differentiated AIGC models on edge servers, while multiple MUs with heterogeneous QoE requirements request services from these ASPs.

- *Incentive mechanism design for on-demand service provisioning in resource-constrained AIGC edge networks.* This mechanism integrates the perspectives of MUs and ASPs, optimizing service costs for MUs to access personalized AIGC services while ensuring that ASPs can efficiently deliver on-demand services without over-provisioning resources. To our knowledge, this is the first analytical study on incentive mechanism design explicitly tailored to the unique characteristics of LLM-based text services in multi-MU and multi-ASP scenarios.
- *Novel design of multi-dimensional QoE metric to quantify service quality.* To quantify MUs' personalized demands, we design a metric integrating three critical dimensions: accuracy, token count, and timeliness. The accuracy is evaluated through CoT reasoning performance gaps, capturing discrepancies between the expected and actual model outputs. Token count is determined by the total number of input and output tokens, while timeliness is constrained by the maximum tolerable latency. Through comprehensive analysis and validation of CoT reasoning steps and examples, the proposed QoE metric effectively models diverse service requirements while accounting for the limited resources of ASPs, supporting on-demand service provision.
- *Theoretical analysis and algorithm designs.* We propose a two-level game-theoretic model to analyze the MU's optimal reward and the ASP's optimal resource allocation. Given multiple ASPs with differentiated AIGC models deployed on edge servers and multiple MUs with heterogeneous QoE requirements, the optimization problems of the ASP and the MU are of high complexity (e.g., inter-user competition and resource constraints). We solve these problems by reformulating them as an equilibrium

problem with equilibrium constraints (EPEC) framework, theoretically proving the existence and uniqueness of the equilibrium, and designing a dual-perturbation reward optimization algorithm to reduce the implementation complexity of adaptive pricing.
- *Performance evaluation*. Numerical results based on real-world scenarios show that our proposed mechanism achieves approximately 64.9% reductions in average computational and communication overhead, 66.5% decreases in average MU service cost, and 76.8% savings in resource consumption of ASPs on average, compared to state-of-the-art benchmarks.

The rest of this paper is organized as follows. Section II discusses related work. In Section III, we present the system overview and design a multi-dimensional QoE metric. Section IV gives the game formulation, and Section V analyzes the existence of the equilibria at two levels. Section VI shows numerical experiments to demonstrate the advantages of the proposed incentive mechanism, and finally, Section VII concludes the paper.

II. Related Work

In this section, we discuss fundamental research in three areas: AIGC services in networks, multimodal prompt engineering, and incentive mechanisms for AIGC services. First, we explore issues and methods for implementing AIGC in resource-constrained edge environments, laying the foundation for our study. Building on this, we investigate multi-modal prompt engineering techniques essential for quantitatively assessing content accuracy in the QoE design. Finally, we highlight our unique contributions to incentive mechanism designs, addressing critical gaps in existing studies.

A. AIGC Services in Networks

To alleviate the computational burden on the cloud and reduce high latency caused by long transmission distances, Du *et al.* [2], [8] proposed an AIGC-as-a-Service (AaaS) architecture and introduced a dynamic ASP selection scheme for optimal user-provider connections.

To ensure secure and trustworthy AIGC services, Lin *et al.* [13] designed a decentralized trust mechanism using smart contract-based verification to prevent unreliable outcomes in the Metaverse. Similarly, Liu *et al.* [14] proposed a systematic approach incorporating MASP selection, payment schemes, and fee-ownership transfer, presenting a blockchain framework to protect mobile AIGC services. For collaborative inference, several studies [15]–[17] explored distributed task allocation frameworks. These works achieved efficient AIGC services by optimizing workload distribution across multiple mobile devices, edge servers, and distant data centers. Moreover, the authors in [18], [7], and [19] explored the integration of semantic communication and AIGC services, focusing on optimizing bandwidth and resource utilization while enhancing content quality in wireless networks. In addition to the above works, several studies focus on performance improvement in edge networks. For instance, Wang *et al.* [20] proposed the WP-AIGC framework, integrating wireless perception and AIGC to optimize computing resources at the edge server. Wang *et al.* [21] proposed a joint optimization scheme for offloading, computation time, and diffusion steps in the reverse diffusion stage, using average error as a metric to evaluate the quality of generated results.

These works, from diverse perspectives, have significantly advanced the deployment and optimization of AIGC services in networks, providing valuable insights for future research. Notably, distinct from existing studies, our work focuses on designing an incentive mechanism that motivates ASPs to provide services on-demand, meeting MUs' personalized demands in multi-user and multi-ASP scenarios.

B. Multimodal Prompt Engineering

Prompt engineering in large language models (LLMs) involves designing and optimizing prompts or instructions to guide the model's output generation [22]. This technique enhances LLM performance across diverse tasks, including language modeling, question answering, and image captioning. Wei *et al.* [23] first demonstrated how generating a Chain-of-Thought (CoT) improves LLMs' ability to perform complex reasoning tasks. Their approach, which uses a few CoT demonstrations as exemplars in prompts, significantly improved performance on arithmetic, commonsense, and symbolic reasoning tasks. Building upon this, Wang *et al.* [24] proposed the Plan-and-Solve (PS) prompting strategy for multi-step reasoning tasks. By dividing a task into smaller subtasks and executing them sequentially, PS prompting addresses missing-step errors and enhances the quality of generated reasoning steps. Incorporating multimodal reasoning further enriches LLM capabilities, enabling the synthesis of consistent and coherent CoTs across different modalities. Lu *et al.* [25] introduced ScienceQA, a multimodal reasoning benchmark comprising approximately 21k multiple-choice questions on diverse science topics. Similarly, Zhang *et al.* in [26] proposed a new approach called Multimodal-CoT that incorporates both language (text) and vision (images) modalities into a two-stage framework for CoT prompting.

While these studies have advanced CoT prompting and multimodal reasoning, they have not studied how to improve the quality of LLM-based services by capturing the user's demands from the theoretical properties of CoT and reducing resource consumption from the service and optimization perspectives.

C. Incentive Mechanisms for AIGC Services

Mechanisms incentivizing MU participation are critical in scenarios like federated learning (FL) and data trading [32]. For example, Chen *et al.* [27] proposed a data quality assessment method for AIGC-generated data samples and designed a reward mechanism to incentivize MUs to participate in FL. Du *et al.* [11] developed an incentive mechanism based on generative AI and contract theory to encourage MUs to share semantic information. Zhan *et al.* [28] designed an AIGC task assignment framework for automated difficulty assessment using a visual language model (VLM) with the aid of contract theory. Fan *et al.* [9] developed a decentralized incentive

TABLE I
INCENTIVE MECHANISMS FOR AIGC SERVICES.

Ref.	Multi-MU Multi-ASP	Optimization objective	New metric	On demand
[9]	✓	Service allocation	×	×
[10]	×	Service provision	Blind image spatial quality	×
[11]	✓	Information sharing	Bit error probability	×
[12]	✓	Service cost	×	×
[27]	×	Service performance	Data quality	×
[28]	✓	Task allocation	Difficulty assessment	×
[29]	×	Service cost	×	×
[30]	×	Service performance	Quality and latency of image generation	×
[31]	✓	Service performance	Age of thought	×
Our work	✓	Service performance	QoE	✓

TABLE II
SUMMARY OF MAIN NOTATIONS.

Notation	Definition
n, N	Index of an ASP, number of ASPs
m, M	Index of an MU, number of MUs
a_0^{nm}	Input message or task provided by MU m to ASP n
a_i^{nm}	Message chain provided by ASP n to MU m
B_{nm}	ASP n's communication resource allocated for MU m
c_n^f	Cost factor per unit of computational resources
c_n^B	Cost factor per unit of communication resources
f_{nm}	ASP n's computational resource allocated for MU m
K	Number of CoT examples
O_k^{nm}	CoT examples utilized by ASP n in providing the service to MU m
$\mathcal{Q}_{nm}$	Quantified QoE value for MU m provided by ASP n
R_{nm}	Reward for unit QoE value offered by MU m to ASP n
x_{nm}^{in}	Input token from MU m to ASP n
x_{nm}^{out}	Output token requested from ASP n by MU m
γ_{nm}	SNR for the communication between MU m and ASP n
$\hat{\theta}_{nm}$	Upper bound on the performance gap of the service provided by ASP n to MU m
$\Upsilon(c_n^*)$	Maximum deviation between the context c_n owned by ASP n and the distribution of true contexts
$\zeta(a_i^{nm})$	Ambiguity of chain $(a_i^{nm})_{0\leq i\leq l}$
μ_m	Profit conversion factor for MU m's AIGC services
κ_n	Maximum tolerable latency for ASP n's service.
ξ_n	Computational resource required per token of ASP n

mechanism for mobile AIGC service allocation that balances service provision with MU demand. These studies predominantly adopt server-driven pricing models to incentivize MU participation in service provision.

In contrast to mechanisms focused on MU participation, another research direction, similar to our work, focuses on incentivizing ASPs to allocate resources and provide AIGC services. Wen *et al.* [29] proposed a user-centric incentive mechanism to motivate ASPs to provide AIGC services while incorporating prospect theory to model the subjective utility of clients. Du *et al.* [12] extended the scope to multi-MU and multi-ASP scenarios employing contract theory to reward ASPs based on their resource contributions. Although these mechanisms improve flexibility and applicability, they focus primarily on computational resources and overlook joint resource optimization, as well as the unique requirements of AIGC services. To address the need for AIGC-specific metrics, researchers have proposed mechanisms targeting particular applications. For example, Wang *et al.* [10] introduced metrics to evaluate the precision of wireless perception and the quality of AIGC, linking resource allocation to service quality through a diffusion-based pricing strategy. Ye *et al.* [30] modeled the relationship between prompt optimization, diffusion denoising steps, and AIGC quality, designing a quality-based contract to improve AIGC generation while reducing latency. While these works improve the quality of service in AIGC, they are limited to single-MU scenarios and primarily focus on image generation. In addition, the authors in [31] proposed the age-of-thought (AoT) metric and a deep Q-network-based auction to incentivize the provisioning of LLM agents, but focused on cached resources without addressing on-demand services for personalized demands.

Unlike the above works, our study focuses on AIGC services based on LLMs in multi-MU and multi-ASP scenarios. We design a comprehensive QoE metric that captures MUs' personalized demands across multiple dimensions, tailored to the unique characteristics of text-based content generation. By jointly optimizing computational and communication resources, our mechanism incentivizes ASPs to allocate resources on demand, enabling scalable, efficient, and personalized service provision. A detailed comparison of these approaches is provided in Table I, highlighting the key contributions of this study.

III. SYSTEM OVERVIEW

We first present the system model of the incentive mechanism for personalized AIGC services in Section III-A. We then discuss the ambiguity of language employed in LLMs in Section III-B. Following this, we design a novel QoE metric to model MU demands across three dimensions in Section III-C. To simplify the presentation, we summarize the essential notations in Table II.

A. System Model

We consider a scenario involving a set $\mathcal{M}=\{1,\ldots,M\}$ of MUs and a set $\mathcal{N}=\{1,\ldots,N\}$ of ASPs. As depicted in Fig. 2, each MU has unique service requirements that reflect individual preferences. The requirements for AIGC services vary according to specific applications. For example, in vehicular environments, MUs aim for fast, accurate, and concise responses from AIGC services to enhance safety and convenience. Conversely, when engaged in content creation tasks, MUs prioritize content quality, depth, and relevance over response time. ASPs are resource-limited edge servers equipped with different types of LLMs. For example, ASP 1 is dedicated to vehicular networking, while ASP 2 specializes in academic writing. The objective is to maximize the respective benefits of MUs and ASPs in personalized AIGC trading networks through a flexible user-driven rewarding mechanism

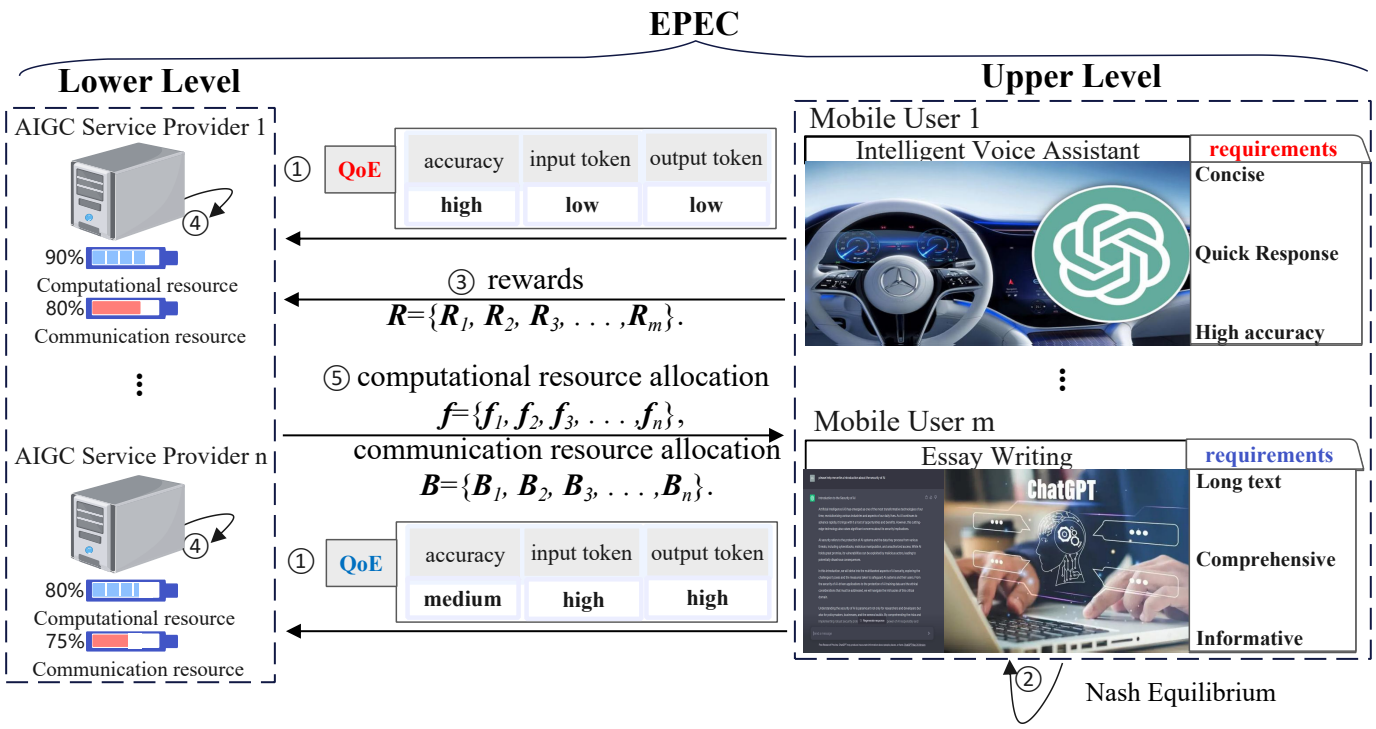


Fig. 2. The workflow of the QoE-driven incentive mechanism. Each MU has distinct service requirements reflecting individual preferences. For example, vehicles demand concise and accurate responses to ensure safety, whereas content creators require richer inputs and outputs to support content generation.

and ASPs' on-demand resource allocation. In this mechanism, MUs act as upper-level leaders determining rewards, while ASPs serve as lower-level followers making resource allocation decisions. This structure ensures the provision of personalized AIGC services that effectively meet the diverse demands of MUs. First, MUs broadcast their specific demands (accuracy, input and output tokens) defined in QoE to ASPs (①). Subsequently, the MUs determine the reward per QoE given to ASPs based on a Nash equilibrium (②), with the initial rewards set as random values, and transmit them to ASPs (③). The ASPs then make resource allocation decisions based on the QoE, cost, and rewards given by the MUs (④). Based on the feedback from ASPs (⑤), MUs adjust the rewards to maximize their benefits (②). Steps ② and ⑤ are repeated until an agreement is reached between MUs and ASPs. The iterative process ensures mutual benefits and ultimately enhances the AIGC service experience for all MUs.

B. Ambiguity of Language

Chain-of-Thought (CoT) prompting [33] is a technique designed to enhance the reasoning capabilities of LLMs. By guiding the model to break down complex problems into logical steps, CoT prompts enable incremental solution derivation that improves contextual understanding and logical consistency. CoT examples play a crucial role in this process. They are a set of exemplary reasoning chains that demonstrate how the model can complete a specific task through step-by-step reasoning. These examples not only help the model learn how to structure its reasoning process but also guide it in generating more accurate and consistent outputs. However, the inherent ambiguity in natural language poses a significant challenge to LLMs' reasoning abilities. This ambiguity affects their ability to determine precise intentions or contexts behind ambiguous messages or sequences.

To investigate the impact of this ambiguity on inference performance, we explore and analyze the theoretical foundations underlying CoT prompting in the inference process. When serving MU m, ASP n is provided with K CoT examples of varying lengths, denoted as $O_k^{nm} = (o_{k,r}^{nm})_{0\le r\le l_k}$, where l_k represents the length of the chain O_k^{nm}, and each $o_{k,r}^{nm}$ is a sequence of tokens representing a thought or reasoning step. These examples are designed to help ASP n produce correct answers via CoT generation for MU m. For all k, O_k^{nm} are generated with true intentions $(I^{nm})^*$ and context c_n^*. Following this, ASP n generates a series of messages denoted as $(a_i^{nm})_{0\le i\le l}$ based on the provided CoT examples O_k^{nm} and the initial task a_0^{nm}. Based on the CoT prompting process, we can adapt the following definition of ambiguity from [34].

Definition 1. ***(The ambiguity of the chain)*** *When ASP n performs the reasoning task for MU m, the ambiguity of a message chain $(a_i^{nm})_{0\le i\le l}$, derived from true intentions $(I^{nm})^*$ and true context c_n^*, is defined as the complement of the likelihood $\hat{q}$ of the context c_n^* and intentions $(I^{nm})^*$ conditioned on $(a_i^{nm})_{0\le i\le l}$, i.e.,*

$$\zeta((a_i^{nm})_{0\le i\le l}) = 1 - \hat{q}(c_n^*, (I^{nm})^* | (a_i^{nm})_{0\le i\le l}). \quad (1)$$

Similarly, the ambiguity of the message chain $\zeta((O_k^{nm})_{1\le k\le K})$ can be defined as the complement of the likelihood of the context c_n^* and intentions $(I^{nm})^*$ conditioned on $(O_k^{nm})_{1\le k\le K}$, i.e.,

$$\zeta((O_k^{nm})_{1\le k\le K}) = 1 - \hat{q}(c_n^*, (I^{nm})^* | (O_k^{nm})_{1\le k\le K}). \quad (2)$$

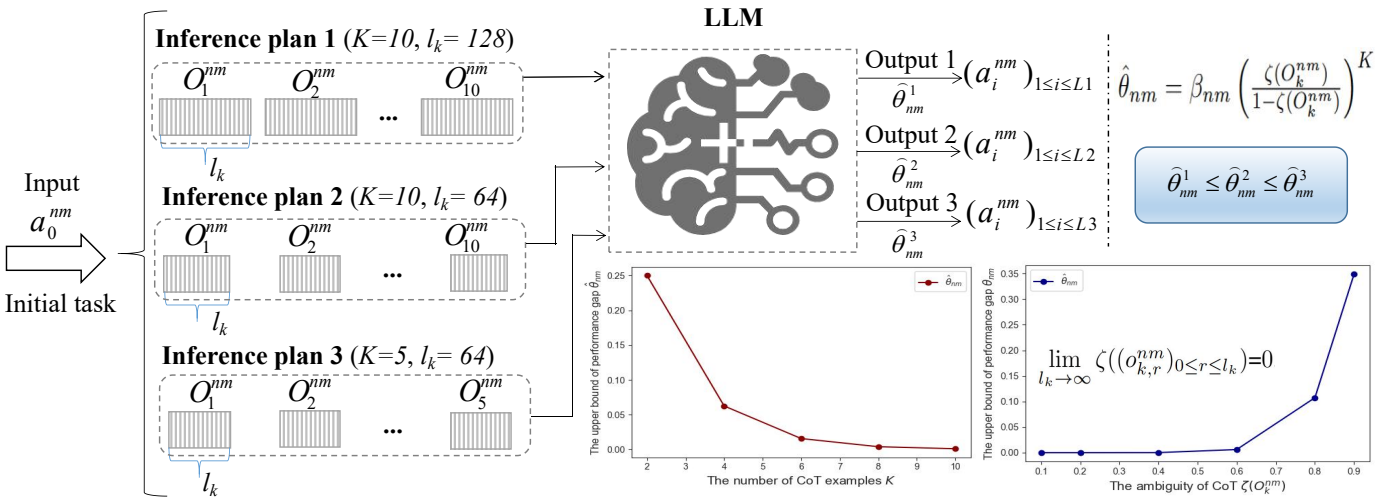


Fig. 3. Impact of different CoT examples on performance gap $\hat{\theta}_{nm}$.

After defining the ambiguity of the chain, we focus on evaluating the overall performance of the model using CoT prompting technology. To achieve this, we use the concept of ambiguity to assess the model's performance by comparing the likelihood of the results obtained through CoT examples with those derived from the true natural language distribution. This comparison quantifies the performance gap between the service provided by ASP n for MU m and the ideal result for MU m, denoted as

$$\begin{aligned}\theta_{nm} = |&p_n\left((a_i^{nm})_{1\le i\le l}|a_0^{nm}, O_k^{nm}\right) \\ &- \hat{q}\left((a_i^{nm})_{1\le i\le l}|a_0^{nm}, c_n^*\right)|.\end{aligned} \tag{3}$$

To further understand this performance gap, we explore a theoretical framework that considers the distribution of contexts in the training datasets, aiming to gain insights into the upper bound of the gap.

Theorem 1. ***(The upper bound of the performance gap)*** *When ASP n infers the task of MU m based on CoT prompting technology, consider a set of K CoT examples $O_k^{nm} = (o_{k,r}^{nm})_{0\le r\le l_k}$, generated from the intention $(I^{nm})^*$ with the optimal context $c_n^* \sim q(c_n)$. Then, for any message sequence $(a_i^{nm})_{1\le i\le l}$, we have [33]:*

$$\theta_{nm} \le 2\frac{(\Upsilon(c_n^*))^K \zeta(a_0^{nm})}{1-\zeta(a_0^{nm})}\prod_{k=1}^{K}\frac{\zeta(O_k^{nm})}{1-\zeta(O_k^{nm})}, \tag{4}$$

where a_0^{nm}, sampled from $q(\cdot|(I_0^{nm})^*)$, is the input message or task generated from $(I_0^{nm})^*$, which is sampled from $q(\cdot|c_n^*)$. $\Upsilon(c_n^*) = \sup_{c\in C}\frac{\hat{q}(c_n^*)}{\hat{q}(c_n)}$ is a skewness parameter, indicating the maximum deviation between the context c_n owned by ASP n and the distribution of true contexts.

If $\Upsilon(c_n^*) = 1$, it implies that natural language does not discriminate among specific contexts. This condition holds when dealing with sufficiently large and well-balanced datasets [35], [36]. Moreover, under certain conditions, a geometrical convergence rate can be obtained as follows.

Lemma 1. *Consider a set of CoT examples $O_k^{nm} = (o_{k,r}^{nm})_{1\le r\le l_k}$ satisfying the following conditions [31]:*

1) The CoT examples O_k^{nm} are carefully selected such that the true c_n^* can be recovered from O_k^{nm} with relatively high certainty.

2) As the length of the sequence grows, the associated ambiguity measure $\zeta(O_k^{nm})$ diminishes.

Then, for any fixed $\eta \in [0, \frac{1}{2})$, there exists a length threshold $l_{k,\eta}^* \in \mathbb{N}$, such that for any $l_k \ge l_{k,\eta}^*$, we can have:

$$\zeta(O_k^{nm}) \le \eta. \tag{5}$$

Proof. In practice, satisfying condition 1) is challenging as there is no rigorous procedure to quantify ambiguity for a given sequence of thoughts. However, under condition 2), we can obtain $\lim_{l_k\to\infty}\zeta((o_{k,r}^{nm})_{0\le r\le l_k})=0$. Hence, there exists $l_{k,\eta} \in \mathbb{N}$ such that for any $l_k \ge l_{k,\eta}^*$, we have $\zeta(O_k^{nm}) \le \eta$. □

As the number of CoT examples satisfying the conditions increases, a geometrical convergence rate for the gap θ_{nm} can be established as

$$\theta_{nm} \le \beta_{nm}\left(\frac{\eta}{1-\eta}\right)^K, \tag{6}$$

where $\beta_{nm} = 2\frac{(\Upsilon(c_n^*))^K\zeta(a_0^{nm})}{1-\zeta(a_0^{nm})}$ and $\frac{\eta}{1-\eta} \in [0, 1)$.

According to the analysis of **Theorem 1** and **Lemma 1**, the performance gap is primarily influenced by the number of CoT examples K and the reasoning step length l_k. As illustrated in Fig. 3, there is a clear functional relationship between the number of CoT examples K and the performance gap θ, demonstrating that an increased number of CoT examples significantly reduces the performance gap, leading to higher accuracy in AIGC. In particular, $\theta \to 0$ indicates high-quality AIGC, while $\theta \to \infty$ implies a significant performance gap and low accuracy. However, the impact of the reasoning step length l_k on the performance gap does not exhibit a clear functional relationship. To further validate the theoretical analysis, we conducted experiments on an open platform to empirically examine how reasoning step length affects the accuracy of inference results in real-world scenarios. As illustrated in Fig. 4, the ASP can choose CoT examples with different

reasoning step lengths for approaching the final answer[1]. The results indicate that as the reasoning step length increases, the associated ambiguity measure $\zeta(O_k^{nm})$ decreases, leading to a smaller $\hat{\theta}_{nm}$ and thus reducing the performance gap. In conclusion, when using CoT with a larger number of steps, LLMs often produce more accurate results, as revealed in **Lemma 1** and literature [37].

C. QoE of AIGC Services

AIGC services exhibit a complex interdependence among accuracy, token count, and service latency. Due to the autoregressive nature of LLMs, computational demand scales quadratically with the number of tokens, as self-attention requires each token to compute relationships with all preceding tokens. Furthermore, improving accuracy incurs growing resource costs that further amplify latency. Compared to traditional services, AIGC faces more intricate multi-dimensional trade-offs, requiring an effective balance between cost and performance.

To this end, we propose a novel QoE metric specifically designed for AIGC services, integrating three key dimensions: accuracy, token count, and response timeliness. By explicitly modeling the trade-offs among these factors, this metric provides a unified framework for evaluating both technical performance and user experience. We utilize the derived upper bound $\hat{\theta}_{nm} = \beta_{nm}(\frac{\eta}{1-\eta})^K,\ n \in \mathcal{N},\ m \in \mathcal{M}$ to evaluate the output accuracy. As demonstrated in Section III-B, a smaller $\hat{\theta}_{nm}$ indicates that the model's output is closer to the ideal result, implying higher accuracy. Achieving higher accuracy typically requires more computational effort, and the computational demand for inference is influenced by the number of tokens processed.

Moreover, ASP n needs to allocate computational resource $\boldsymbol{f}_n \triangleq (f_{n1}, f_{n2}, \ldots, f_{nm})$ and bandwidth $\boldsymbol{B}_n \triangleq (B_{n1}, B_{n2}, \ldots, B_{nm})$ for content generation and transmission. Considering these factors jointly, the QoE of the AIGC service provided by ASP n to MU m is defined as

$$\begin{aligned}\mathcal{Q}_{nm} = \kappa_n &- \frac{\xi_n \ln(1/\hat{\theta}_{nm}) \sum_{i=0}^{x_{nm}^{\text{out}}-1} \left(x_{nm}^{\text{in}} + i\right)}{f_{nm}} \\ &- \frac{32(x_{nm}^{\text{in}} + x_{nm}^{\text{out}})}{B_{nm}\log_2(1+\gamma_{nm})},\end{aligned} \tag{7}$$

where κ_n denotes the maximum latency threshold for the service. This threshold is crucial for on-demand service provision, allowing the model to flexibly adapt to real-world application constraints. The computational cost of processing x_{nm}^{in} input tokens (1 token $\approx$ 4 bytes = 32 bits) and generating x_{nm}^{out} output tokens is $\xi_n \sum_{i=0}^{x_{nm}^{\text{out}}-1} \left(x_{nm}^{\text{in}} + i\right)$, where ξ_n denotes the computational resource required per token. This quadratic growth underscores the necessity of efficient resource management. The $\ln(1/\hat{\theta}_{nm})$ is defined as the accuracy cost. As the performance gap $\hat{\theta}_{nm}$ decreases, indicating a higher accuracy demand for AIGC, the corresponding accuracy cost increases logarithmically [31]. We adopt orthogonal frequency division multiple access (OFDMA) for data transmission to avoid interference between communication links [38]. $\gamma_{nm} = \frac{g_{nm}p_n^t}{\sigma_{nm}^2}$ denotes the signal-to-noise ratio (SNR) for the communication between MU m and ASP n, where g_{nm} is the channel gain, p_n^t is the transmission power of ASP n, and σ is the additive white Gaussian noise (AWGN).

The parameters $\hat{\theta}_{nm}$, x_{nm}^{in}, and x_{nm}^{out} in the QoE correspond to "accuracy", "input token", and "output token" in Fig. 2, respectively. MUs can adjust accuracy and input/output token count to achieve personalized AIGC services within the maximum tolerable latency set by the ASP. This flexibility enables the QoE metric to adapt to diverse application scenarios, ensuring optimized service quality tailored to both system constraints and user preferences.

IV. Problem Formulation

We first define the utility functions of ASPs and MUs in Section IV-A and IV-B. Then, we model the interaction between MUs and ASPs as an EPEC in Section IV-C.

A. Utility of ASPs

ASPs aim to provide AIGC services that satisfy MUs' demands while maximizing their rewards. The utility function of ASP n comprises two key components: the rewards received from MUs and the costs associated with generating and transmitting AIGC. The optimization problem for ASP n can be formulated as

$$\begin{aligned}&\max\ U_n^{asp}(\boldsymbol{f}_n, \boldsymbol{B}_n) = \sum_{m=1}^{M}(R_{nm}\mathcal{Q}_{nm} - c_n^f f_{nm} - c_n^B B_{nm}), \\ &s.t.\ C1: \mathcal{Q}_{nm} \geq 0, \\ &\qquad C2: \sum_{m=1}^{M} f_{nm} \leq f_n^{\max},\ \sum_{m=1}^{M} B_{nm} \leq B_n^{\max},\end{aligned} \tag{8}$$

where R_{nm} is the reward for unit QoE value, and $\mathcal{Q}_{nm}$ denotes the quantified QoE provided by ASP n to MU m. The decision variables f_{nm} and B_{nm} are the computational and communication resources allocated by ASP n for MU m, respectively. The cost parameters c_n^f and c_n^B represent the unit costs of these resources. Constraint C1 ensures that the services provided by ASP n are meaningful and viable, which is achieved by controlling the allocation of computational and communication resources[2]. The constraint C2 ensures that the total resources of ASP n remain within their respective limits, where $f_n^{\max}$ and $B_n^{\max}$ represent the maximum available computational and communication resources. The utility function is designed to capture the trade-off between the rewards received from MUs and the costs incurred in providing AIGC services.

[1]The experiments are conducted on the ScienceQA benchmark [26]. This large-scale multimodal science question dataset contains 21k multiple-choice questions across 3 subjects, 26 topics, 127 categories, and 379 skills, divided into training, validation, and test splits with 12726, 4241, and 4241 examples, respectively. This setting aimed to demonstrate the efficacy of multimodal-CoT by leveraging both language and vision modalities in a two-stage reasoning framework.

[2]While negative utility may arise from service constraints $\mathcal{Q}_{nm} \geq 0$, it reflects rational short-term tradeoffs (e.g., maintaining market presence). In practice, non-negativity can be guaranteed by adding a baseline cost compensation $c_n^f f_n^{\max} + c_n^B B_n^{\max}$ to U_n^{asp}. This compensation reflects the maximum costs that an ASP may incur, ensuring that its utility function remains non-negative even in the worst-case scenario with minimal rewards. Since this term is a constant offset, it does not affect the optimal strategies or equilibrium analysis [39] and is therefore omitted for simplicity.

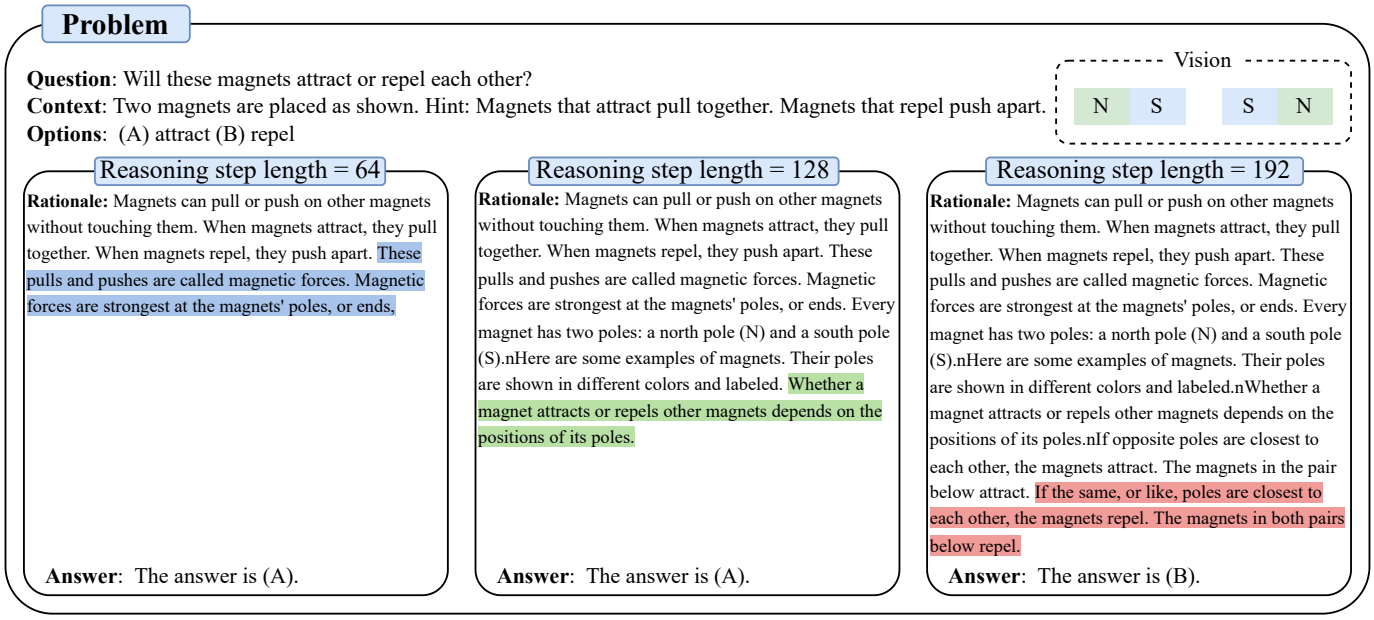


Fig. 4. Example of different reasoning step lengths in CoT. When the number of CoT examples is fixed, using longer reasoning steps within each example generally improves accuracy. ASPs can therefore select CoT examples with varying step lengths to better approach the final answer.

By maximizing this utility, the ASP determines the optimal joint resource allocation, i.e., $\boldsymbol{f}_n$ and $\boldsymbol{B}_n$, which increases profitability and reduces resource consumption.

B. Utility of MUs

Higher QoE improves user satisfaction, but offering higher rewards to ASPs increases costs. The utility function of MU m captures this trade-off between service satisfaction and cost. To quantify service satisfaction, we use a gain function that models the benefits derived from AIGC services. Specifically, we adopt a logarithmic gain function [40], defined as $\mathcal{G}_{nm} = \mu_m ln(1+\sum_{n=1}^{N} \mathcal{Q}_{nm})$, where μ_m is the profit conversion coefficient. The logarithmic form captures diminishing returns, meaning that as QoE increases, user satisfaction improves but at a decreasing rate. This reflects the realistic observation that users experience higher marginal gains at lower QoE levels, while improvements beyond a certain threshold yield smaller incremental benefits. Building on this gain function, the optimization problem for MU m is formulated as

$$\begin{aligned} &\max\ U_m^{mu}(\boldsymbol{R}_m) = \mu_m ln(1+\sum_{n=1}^{N} \mathcal{Q}_{nm}) - \sum_{n=1}^{N} R_{nm}\mathcal{Q}_{nm}, \\ &s.t.\ R_m^{\min} \le R_{nm} \le R_m^{\max},\ n \in \mathcal{N}, \end{aligned} \tag{9}$$

where $\boldsymbol{R}_m \triangleq (R_{1m}, \ldots, R_{Nm})$ is the reward profile of MU m, with $R_m^{\min}$ and $R_m^{\max}$ denoting its minimum and maximum payable rewards, respectively. By optimizing R_{nm}, the MU can balance the service cost with the satisfaction derived, ensuring cost-effective access to personalized AIGC services. Given the competition for limited ASP resources, a higher reward from one MU can lead to better service, potentially reducing the QoE for others. Since MUs act selfishly and independently, their interactions naturally form a non-cooperative game, termed the multi-MU reward game ψ, where MUs adjust their rewards to maximize satisfaction while minimizing costs.

Definition 2. *A multi-MU reward game is a tuple* $\psi = \{\mathcal{M}, \boldsymbol{R}, \boldsymbol{U^{mu}}\}$ *defined by*

- *Players: The set of MUs.*
- *Strategies: The reward decisions* $\boldsymbol{R}_m$ *of any MU* m*.*
- *Utilities: The vector* $\boldsymbol{U^{mu}} = \{U_1^{mu}, \ldots, U_M^{mu}\}$ *contains the utility functions of all the MUs defined in (9).*

ASPs optimize resource allocation based on the rewards provided by MUs to maximize their benefits. Meanwhile, MUs strategically adjust these rewards to obtain AIGC services that meet their requirements. This interdependent decision-making process creates a hierarchical game structure, where each entity optimizes its utility, leading to a complex equilibrium dynamic. Such interactions naturally follow a multi-leader, multi-follower framework, with MUs as the leaders and ASPs as the followers. In the following section, we formalize this relationship as an EPEC problem, establishing the foundation for equilibrium analysis.

C. Equilibrium Problem with Equilibrium Constraints

We model the incentive mechanism between ASPs and MUs as an equilibrium problem with equilibrium constraints (EPEC) [41], where MUs are leaders and ASPs are followers. At the upper level, MUs determine the rewards based on ASPs' responses and decisions of other MUs. At the lower level, each ASP optimizes the allocation of computational and communication resources by considering its constraints, costs, and rewards from MUs. This setup results in a Stackelberg equilibrium, where ASPs (followers) choose their best responses, and MUs (leaders) maximize their utilities. Our objective is to achieve a hierarchical equilibrium, characterized by a Nash equilibrium among MUs and a Stackelberg

equilibrium between MUs and ASPs. The equilibria for the two levels are defined as follows.

Definition 3. *Let $(\boldsymbol{f}_n^*, \boldsymbol{B}_n^*)$ and $\boldsymbol{R}_m^*$ denote the optimal resource allocation of ASP n and the optimal reward decision of MU m, respectively. Then, the points $(\boldsymbol{f}_n^*, \boldsymbol{B}_n^*)$ and $\boldsymbol{R}_m^*$ are the equilibria at two levels if the following conditions hold:*

$$\begin{aligned} &U_n^{asp}\left((\boldsymbol{f}_n^*, \boldsymbol{B}_n^*), \boldsymbol{R}_m^*\right) \geq U_n^{asp}((\boldsymbol{f}_n, \boldsymbol{B}_n), \boldsymbol{R}_m^*), \\ &U_m^{mu}\left(\boldsymbol{f}_n^*(\boldsymbol{R}_m^*, \boldsymbol{R}_{-m}^*), \boldsymbol{B}_n^*(\boldsymbol{R}_m^*, \boldsymbol{R}_{-m}^*), \boldsymbol{R}_m^*, \boldsymbol{R}_{-m}^*\right) \geq \\ &U_m^{mu}\left(\boldsymbol{f}_n^*(\boldsymbol{R}_m^*, \boldsymbol{R}_{-m}^*), \boldsymbol{B}_n^*(\boldsymbol{R}_m^*, \boldsymbol{R}_{-m}^*), \boldsymbol{R}_m, \boldsymbol{R}_{-m}^*\right), \\ &\forall n \in \mathcal{N},\ m \in \mathcal{M}, \end{aligned} \tag{10}$$

where $\boldsymbol{R}_{-m}$ denotes the reward profile of all MUs except MU m.

In summary, the MUs' optimization problems can be formulated as the following EPEC problems:

$$\begin{aligned} \max_{\boldsymbol{R}_m}\ & U_m^{mu} = \mu_m ln(1+\sum_{n=1}^{N} \mathcal{Q}_{nm}^*) - \sum_{n=1}^{N} R_{nm}\mathcal{Q}_{nm}^*, \\ \text{s.t.}\ & \begin{cases} R_m^{\min} \leq R_{nm} \leq R_m^{\max},\ n \in \mathcal{N}, \\ \mathcal{Q}_{nm}^* = \kappa_n - \dfrac{\xi_n \ln(1/\hat{\theta}_{nm}) \sum_{i=0}^{x_{nm}^{\text{out}}-1}\left(x_{nm}^{\text{in}} + i\right)}{f_{nm}^*} \\ \qquad\qquad - \dfrac{32(x_{nm}^{\text{in}} + x_{nm}^{\text{out}})}{B_{nm}^* \log_2(1+\gamma_{nm})}, \\ (\boldsymbol{f}_n^*, \boldsymbol{B}_n^*) = \arg\max U_n^{asp}(\boldsymbol{f}_n, \boldsymbol{B}_n), \\ \qquad\qquad \text{subject to}\, C1, C2. \end{cases} \end{aligned} \tag{11}$$

To solve the above EPEC, we use backward induction to address the lower-level (utility maximization for ASPs) and upper-level (non-cooperative game among MUs) problems.

V. Equilibrium Analysis And Solutions

In this section, we use backward induction to investigate the existence and uniqueness of equilibria for the EPEC, i.e., the optimal resource allocation of ASPs and reward strategies of MUs.

A. Lower Level: Optimal Resource Allocation for ASPs

In the lower level of EPEC, for any reward decisions $\boldsymbol{R}$ given by MUs, ASP n aims to determine its optimal computational and communication resource allocations, i.e., $\boldsymbol{f}_n^*$ and $\boldsymbol{B}_n^*$, to maximize its utility $U_n^{asp}(\boldsymbol{f}_n, \boldsymbol{B}_n)$. In the following, we analyze and derive the unique optimal allocation.

Theorem 2. *The utility maximization problem for ASP n has a unique optimal solution $(\boldsymbol{f}_n^*, \boldsymbol{B}_n^*)$.*

Proof. We first examine the Hessian matrix of ASP n's utility $U_n^{asp}(\boldsymbol{f}_n, \boldsymbol{B}_n)$ with respect to f_{nm} and B_{nm}. Let this Hessian matrix be denoted by $\boldsymbol{H}_n$, which can be block-diagonalized as

$$\boldsymbol{H}_n = \begin{bmatrix} \boldsymbol{H}_n^f & 0 \\ 0 & \boldsymbol{H}_n^B \end{bmatrix}. \tag{12}$$

The block matrix $\boldsymbol{H}_n^f$ can be computed as the second-order partial derivative of $U_n^{asp}(\boldsymbol{f}_n, \boldsymbol{B}_n)$ with respect to f_{nm}, i.e.,

$$\begin{aligned} \boldsymbol{H}_n^f = -\text{diag}\Bigg[& \frac{2R_{n1}\sum_{i=0}^{x_{n1}^{\text{out}}-1}\left(x_{n1}^{\text{in}} + i\right)\xi_n \ln(1/\hat{\theta}_{n1})}{f_{n1}^3}, \\ & \frac{2R_{n2}\sum_{i=0}^{x_{n2}^{\text{out}}-1}\left(x_{n2}^{\text{in}} + i\right)\xi_n \ln(1/\hat{\theta}_{n2})}{f_{n2}^3}, \\ & \cdots, \\ & \frac{2R_{nM}\sum_{i=0}^{x_{nM}^{\text{out}}-1}\left(x_{nM}^{\text{in}} + i\right)\xi_n \ln(1/\hat{\theta}_{nM})}{f_{nM}^3}\Bigg] < \mathbf{0}. \end{aligned} \tag{13}$$

Similarly, the block matrix $\boldsymbol{H}_n^B$ can be calculated by

$$\begin{aligned} \boldsymbol{H}_n^B = -\text{diag}\Bigg[& \frac{64(x_{n1}^{\text{in}} + x_{n1}^{\text{out}})R_{n1}}{B_{n1}^3 \log_2(1+\gamma_{n1})}, \\ & \frac{64(x_{n2}^{\text{in}} + x_{n2}^{\text{out}})R_{n2}}{B_{n2}^3 \log_2(1+\gamma_{n2})}, \\ & \cdots, \\ & \frac{64(x_{nM}^{\text{in}} + x_{nM}^{\text{out}})R_{nM}}{B_{nM}^3 \log_2(1+\gamma_{nM})}\Bigg] < \mathbf{0}. \end{aligned} \tag{14}$$

Since $R_{nm} \geq R_{nm}^{\min} > 0$, it can be shown that both $\boldsymbol{H}_n^f$ and $\boldsymbol{H}_n^B$ are diagonal matrices with strictly negative diagonal elements. Thus, the block-diagonal Hessian matrix $\boldsymbol{H}_n$ is negative definite, which implies that the utility function $U_n^{\text{asp}}(\boldsymbol{f}_n, \boldsymbol{B}_n)$ is strictly concave and continuous. Furthermore, based on the uniqueness condition established in [42], it follows that ASP n has a unique optimal solution $(\boldsymbol{f}_n^*, \boldsymbol{B}_n^*)$. □

B. Upper Level: Optimal Reward Equilibrium among MUs

In the upper level of the EPEC, each MU m competes with other MUs and determines its reward vector $\boldsymbol{R}_m$. Given the ASPs' responses $(\boldsymbol{f}, \boldsymbol{B})$ and other MUs' decisions $\boldsymbol{R}_{-m}$, MU m selects its optimal reward $\boldsymbol{R}_m$ by maximizing its utility $U_m^{\text{mu}}(\boldsymbol{R}_m)$.

Theorem 3. *A unique Nash equilibrium (NE) exists in the multi-MU reward game ψ.*

Proof. We define the Hessian matrix of U_m^{mu} with respect to $\boldsymbol{R}_m$ as $(\boldsymbol{\Lambda}_m + \boldsymbol{H}_m)$. The matrix $\boldsymbol{\Lambda}_m = diag\left(\frac{\partial^2 U_m^{mu}}{\partial R_{1m}^2}, \ldots, \frac{\partial^2 U_m^{mu}}{\partial R_{Nm}^2}\right)$ and the second-order partial derivative matrix $\boldsymbol{H}_m$ is expressed by

$$\boldsymbol{H}_m = \begin{bmatrix} 0 & \frac{\partial^2 U_m^{mu}}{\partial R_{1m}\partial R_{2m}} & \cdots & \frac{\partial^2 U_m^{mu}}{\partial R_{1m}\partial R_{Nm}} \\ \frac{\partial^2 U_m^{mu}}{\partial R_{2m}\partial R_{1m}} & 0 & \cdots & \frac{\partial^2 U_m^{mu}}{\partial R_{2m}\partial R_{Nm}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 U_m^{mu}}{\partial R_{Nm}\partial R_{1m}} & \frac{\partial^2 U_m^{mu}}{\partial R_{Nm}\partial R_{2m}} & \cdots & 0 \end{bmatrix}, \tag{15}$$

where

$$\begin{aligned} \frac{\partial^2 U_m^{mu}}{\partial R_{nm}^2} = \mu_m \frac{\mathcal{Q}_{nm}''(1+\sum_{n=1}^{N}\mathcal{Q}_{nm}) - (\mathcal{Q}_{nm}')^2}{(1+\sum_{n=1}^{N}\mathcal{Q}_{nm})^2} - 2\mathcal{Q}_{nm}' \\ - R_{nm}\mathcal{Q}_{nm}'' < 0,\ n \in \mathcal{N}, \end{aligned} \tag{16}$$

$$\frac{\partial^2 U_m^{mu}}{\partial R_{nm}\partial R_{n'm}} = -\mu_m \frac{\mathcal{Q}'_{nm}\mathcal{Q}'_{n'm}}{(1+\sum_{n=1}^{N}\mathcal{Q}_{nm})^2} < 0, \quad \forall n \in \mathcal{N},\ n \neq n'. \tag{17}$$

The proof of $\frac{\partial^2 U_m^{mu}}{\partial R_{nm}^2} < 0$ and $\frac{\partial^2 U_m^{mu}}{\partial R_{nm}\partial R_{n'm}} < 0$ are omitted due to space limits. We randomly choose a vector $\boldsymbol{h} \in \mathbb{R}^{N\times 1}$ with elements not all 0. Then, we have $\boldsymbol{h}^T(\boldsymbol{\Lambda}_m + \boldsymbol{H_m})\boldsymbol{h} = \sum_{n=1}^{N}(h^n)^2(\frac{\mu_m \mathcal{Q}''_{nm}}{1+\sum_{n=1}^{N}\mathcal{Q}_{nm}} - 2\mathcal{Q}'_{nm} - R_{nm}\mathcal{Q}''_{nm}) - \frac{\mu_m}{(1+\sum_{n=1}^{N}\mathcal{Q}_{nm})^2}(\sum_{n=1}^{N} h^n \mathcal{Q}'_{nm})^2 < 0$, indicating that the utility function U_m^{mu} is strictly concave. According to [43], there exists a unique Nash equilibrium in the multi-MU reward game ψ. □

We verify the effectiveness of the incentive mechanism by proving the optimality and uniqueness of MUs' reward decisions and ASPs' resource allocation through **Theorem 2** and **Theorem 3**.

C. The Process of the QoE-driven Incentive Mechanism

The QoE-driven incentive process dynamically adjusts the reward strategies for MUs based on the ASPs' responses. This approach effectively aligns the interests of both ASPs and MUs, facilitating the delivery of personalized AIGC services. The detailed process of this incentive mechanism is outlined in Algorithm 1.

1) The Solution Process: The algorithm initializes each MU m with a reward vector $\boldsymbol{R}_m^{(0)}$ and a step size Δ. At each iteration t, given the current rewards, each ASP $n \in \mathcal{N}$ determines its optimal computational and communication resource allocation $(\boldsymbol{f}_n^*, \boldsymbol{B}_n^*)$. Each MU $m \in \mathcal{M}$ updates its reward $R_{nm}^{(t)}$ offered to ASP n by adjusting it by Δ_t, while keeping the rewards $\boldsymbol{R}_{-n,m}^{(t)}$ fixed for all other ASPs. Specifically, if increasing the reward by Δ_t maximizes the utility $U_m^{mu}(\boldsymbol{R}_m)$, the reward is updated to $R_{nm}^{(t)}+\Delta_t$; if decreasing by Δ_t yields the highest utility, it is set to $R_{nm}^{(t)}-\Delta_t$; otherwise, the reward remains unchanged. The process continues until no MU can improve its utility by unilaterally adjusting its reward, which is verified when the total change in utilities between consecutive iterations falls below a predefined threshold ϵ. At this point, the algorithm converges to obtain the final reward strategy $\boldsymbol{R}^*$ and the corresponding resource allocation $(\boldsymbol{f}^*, \boldsymbol{B}^*)$.

2) Execution Flow: At the beginning of each update cycle, MUs report their personalized demands, and ASPs provide current resource and network conditions. Based on this information, the central decision agent runs Algorithm 1 to determine optimal rewards and resource allocations. This framework integrates the real-time sensing capability of distributed nodes with centralized control, ensuring both efficiency and consistency in decision-making. In contrast, fully distributed methods require frequent strategy exchanges among nodes during each iteration, resulting in significant communication overhead. Similarly, gradient-based distributed optimization techniques rely on each MU to accurately estimate or jointly compute the partial derivatives of its utility function, which introduces additional computational and coordination complexity. Algorithm 1 leverages a zeroth-order perturbation mechanism [44] to streamline this process. In each iteration, the central decision agent perturbs each MU's reward slightly in both directions and observes the resulting utility changes. Based on this directional feedback, it simulates how each MU would react and updates the rewards accordingly. This approach eliminates the need for gradient computation and extensive message passing, thereby reducing overall system overhead.

Algorithm 1 Dual-Perturbation Reward Optimization

Require: Personalized demands of MUs and system parameters of ASPs
Ensure: Optimal reward $\boldsymbol{R}^*$ and resource allocation $(\boldsymbol{f}^*, \boldsymbol{B}^*)$

Initialize: Rewards $\boldsymbol{R}_m^{(0)}$ for MU m, step size Δ, and convergence threshold ϵ.
for $t = 1, 2, \ldots, T$ **do**
  for each ASP $n \in \mathcal{N}$ **do**
    Based on the rewards provided by all MUs, each ASP determines its optimal computational and communication resource allocation $(\boldsymbol{f}_n^*, \boldsymbol{B}_n^*)$;
  end for
  for each MU $m \in \mathcal{M}$ **do**
    Store current rewards: $\boldsymbol{R}_m^{(t)} = \boldsymbol{R}_m$;
    Each MU tries to increase and decrease its rewards with the step size Δ_t, and calculates its own utility based on the ASPs' optimal strategies;
    for each ASP n **do**
      if $U_m^{mu}(R_{nm}^{(t)}, \boldsymbol{R}_{-n,m}^{(t)}) \leq U_m^{mu}(R_{nm}^{(t)} + \Delta_t, \boldsymbol{R}_{-n,m}^{(t)})$ and $U_m^{mu}(R_{nm}^{(t)} - \Delta_t, \boldsymbol{R}_{-n,m}^{(t)}) \leq U_m^{mu}(R_{nm}^{(t)} + \Delta_t, \boldsymbol{R}_{-n,m}^{(t)})$ **then**
        $R_{nm} = \min\{R_{nm}^{(t)} + \Delta_t, R_m^{\max}\}$;
      else if $U_m^{mu}(R_{nm}^{(t)}, \boldsymbol{R}_{-n,m}^{(t)}) \leq U_m^{mu}(R_{nm}^{(t)} - \Delta_t, \boldsymbol{R}_{-n,m}^{(t)})$ and $U_m^{mu}(R_{nm}^{(t)} + \Delta_t, \boldsymbol{R}_{-n,m}^{(t)}) \leq U_m^{mu}(R_{nm}^{(t)} - \Delta_t, \boldsymbol{R}_{-n,m}^{(t)})$ **then**
        $R_{nm} = \max\{R_m^{\min}, R_{nm}^{(t)} - \Delta_t\}$;
      else
        $R_{nm} = R_{nm}^{(t)}$
      end if
    end for
  end for
  if $\sum_{m=1}^{M} |U_m^{mu}(\boldsymbol{R}_m^{(t)}) - U_m^{mu}(\boldsymbol{R}_m^{(t-1)})| \leq \epsilon$ **then**
    $\mathbf{R}^* = \mathbf{R}$
    $(\mathbf{f}^*, \mathbf{B}^*) = \texttt{ASP\_response}(\mathbf{R}^*)$
    break
  end if
end for

3) Convergence Analysis: The convergence complexity of Algorithm 1 depends on the step-size strategy [45], as detailed in Appendix A. With a diminishing step size $\Delta_t = \frac{u}{t}$, Algorithm 1 converges to an ϵ-Nash equilibrium at a rate of $T_\epsilon = \mathcal{O}(\exp(\frac{\ddot{\kappa}}{\epsilon}))$, where $\ddot{\kappa} = (\frac{N}{u}\sum_{m=1}^{M}(R_m^{\max})^2 + MN\frac{u\pi^2}{6})$. The analysis indicates that finding the NE in multi-MU games is typically challenging, even with a small number of MUs [46]. For a fixed number of MUs and ASPs, the convergence speed is primarily influenced by the initial reward vector $\mathbf{R}_m^{(0)}$ for each MU $m \in \mathcal{M}$ and the step size parameter

u. If $\|\mathbf{R}_m^{(0)} - \mathbf{R}_m^*\|_2 \approx u$, rapid convergence to the equilibrium is achieved; however, suboptimal choices of $\mathbf{R}_m^{(0)}$ or u may significantly slow convergence. Due to the determinism of the update rule, given an initialization and fixed parameters, each iteration t produces a unique reward vector $\mathbf{R}_m^{(t)}$, and thus the (initialization-dependent) iterative trajectory is unique and reproducible. In contrast, using a constant step size $\Delta_t = \hat{\Delta}$, the algorithm approaches a neighborhood of the NE at a rate of $\mathcal{O}(\frac{\kappa'}{\epsilon'})$, where $\epsilon' = \epsilon - \frac{\hat{\Delta}MN}{2}$ and $\kappa' = \frac{N\sum_{m=1}^{M}(R_m^{\max})^2}{\hat{\Delta}}$. This setting achieves faster, polynomial convergence in $1/\epsilon'$, but only ensures proximity to the equilibrium within a fixed neighborhood. In practice, diminishing step sizes are preferred when high solution accuracy is critical and longer convergence times are acceptable, whereas constant step sizes are more suitable when a slight accuracy loss is tolerable (e.g., due to quantized decision variables) and rapid convergence is prioritized.

VI. Simulation Results

In this section, we first show the impact of different demands on incentives through QoE quantitative analysis and a case study. Then, we demonstrate the effectiveness and advantages of the proposed incentive through comparative experiments with existing schemes.

A. Analysis of Critical Factors

To investigate the incentive process, we adopt a quantitative approach to examine how varying the personalized demands of MU 1, while keeping those of other MUs fixed, influences the trading outcomes. We precisely manipulate three critical parameters associated with $\mathcal{Q}_{11}$: the accuracy $\hat{\theta}_{11}$ of AIGC[3], with values $[10^{-11}, 10^{-9}, 10^{-7}, 10^{-5}, 10^{-3}]$ depending on $K = [10, 8, 6, 4, 2]$; the number of output tokens x_{11}, varying across $[300, 600, 900, 1200, 1500]$; and the maximum tolerable latency κ_{11}, ranging from $[300, 600, 900, 1200, 1500]$ms. Furthermore, we analyze the average utility of MUs and ASPs by adjusting the number of MUs and ASPs. By systematically varying these parameters, we aim to investigate their effects on the decision-making process comprehensively. The main experimental parameters and values are summarized in Table III.

1) Accuracy of AIGC $\hat{\theta}$: As the value of $\hat{\theta}_{11}$ increases (indicating lower accuracy requirements), the reward R_{11} per unit of $\mathcal{Q}_{11}$ offered by MU 1 to ASP 1 decreases (Fig. 5(a)). This reduction in R_{11} stems from MU 1's rational response to relaxed accuracy constraints, which lowers the marginal value of additional computational efforts (Fig. 5(b)), thus enabling MU 1 to enjoy the service at a lower cost. However, the QoE shows an upward trend as $\hat{\theta}_{11}$ increases. This upward trend in QoE can be attributed to its definition: a higher QoE value indicates the ability to meet the requirements of MUs more quickly. Notably, as accuracy requirements become easier to satisfy, the QoE value rises accordingly. Furthermore, MU 1 experiences an increase in utility (Fig. 5(c)). This is because the relaxed accuracy requirements decrease the total incentive rewards $R_{11}\mathcal{Q}_{11}$ for ASP 1, allowing the service to be obtained at a lower cost. Meanwhile, ASP 1 adapts its resource allocation to optimize its utility, ensuring that reduced total incentive rewards do not result in a loss of utility. These findings show that relaxing AIGC accuracy requirements reduces computational demand and incentive costs while improving QoE, thereby validating the dynamic trade-off between quality and cost.

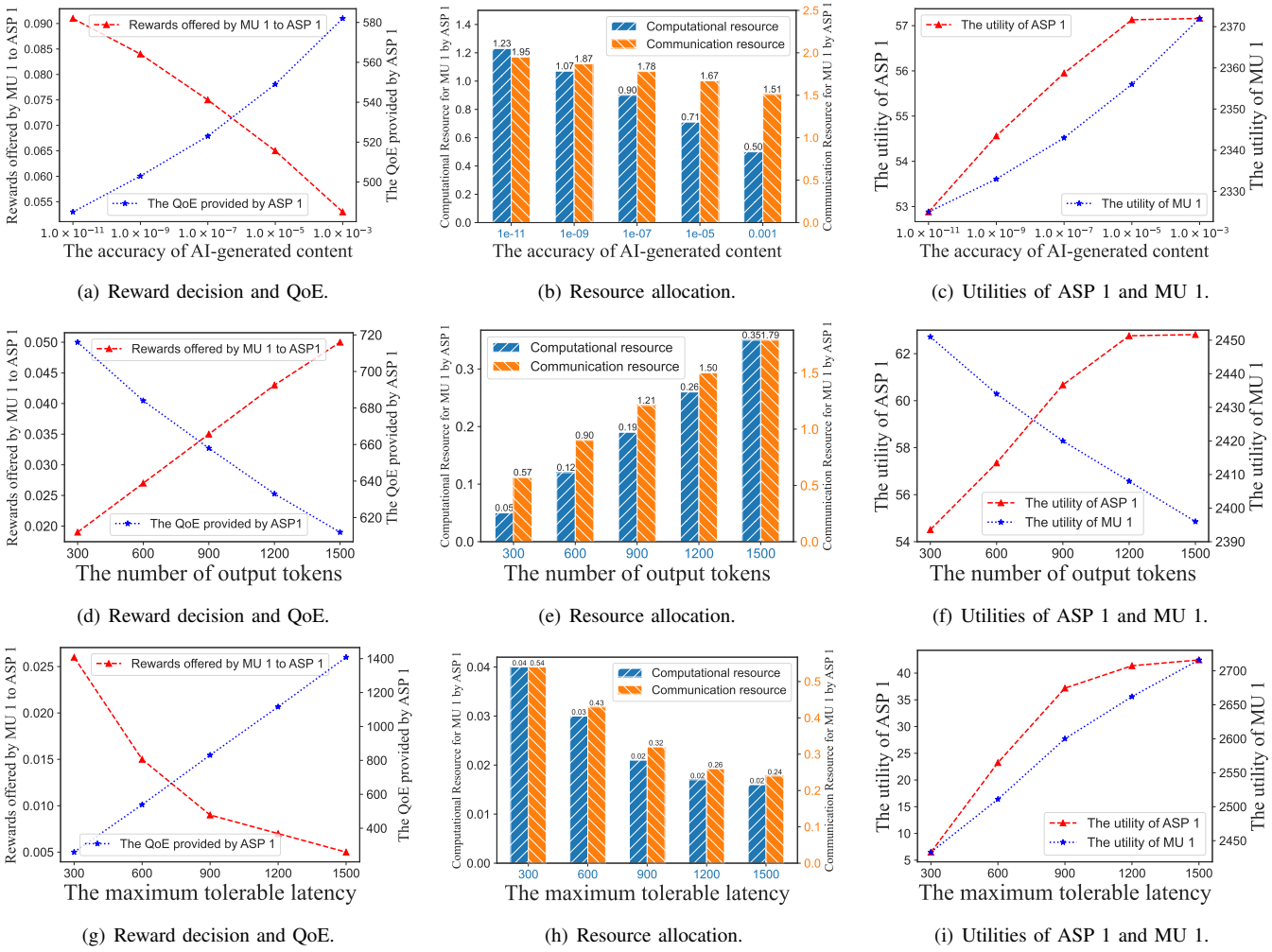
[3]The ASP derives the value of K based on the MUs' accuracy requirements through $\hat{\theta}_{nm} = \beta_{nm}(\frac{\eta}{1-\eta})^K$, thereby guiding the inference service to meet the demands of MUs.

TABLE III
Experimental Parameter Settings.

Parameters	Values
Number of ASPs	[1,2,3,4,5]
Number of MUs	[5,10,15,20,25]
K (CoT Examples)	[2, 4, 6, 8, 10]
Input/Output Token (x^{in}/x^{out})	100–2000 tokens
Maximum Computational Resource ($f^{\max}$)	5-30 TFLOPS
Maximum Bandwidth ($B^{\max}$)	100-500 MHz
Maximum Tolerable Latency (κ)	300-1500 ms
SNR	10-30 dB

2) Number of Output Tokens x^{out}: As the required number of output tokens increases, the reward R_{11} per unit of $\mathcal{Q}_{11}$ provided by MU 1 to ASP 1 tends to rise (Fig. 5(d)). This is because each additional output token requires processing more tokens due to the autoregressive nature of the model, where each output depends on all previous inputs. To meet MU 1's increasing demands, more computational and communication resources are allocated to MU 1 (Fig. 5(e)). Furthermore, MU 1 experiences a decrease in utility, whereas ASP 1's utility increases with the rise in output tokens (Fig. 5(f)). This trend indicates that as the demand for x^{out} grows, MU 1 compensates ASP 1 with higher total incentive rewards, leading to increased costs and reduced utility. However, ASP 1 strategically optimizes its resource allocation to efficiently meet the personalized demands of MU 1, maximizing its utility within the constraints of limited resources and latency thresholds. Ultimately, these findings underscore that higher token demand increases resource consumption, making it the primary driver of cost growth while enhancing the profitability of ASPs.

3) Maximum Tolerable Latency κ: As the value of κ increases, the reward R_{11} per unit of $\mathcal{Q}_{11}$ provided by MU 1 to ASP 1 gradually decreases (Fig. 5(g)). This occurs because the increase in κ makes MU 1's demand less urgent, allowing ASP 1 to meet MU 1's requirements with fewer resources (Fig. 5(h)), thereby lowering the cost per unit of QoE. Moreover, both the utilities of MU 1 and ASP 1 show an upward trend (Fig. 5(i)). For MU 1, lower service requirements naturally result in cost savings and higher utility. For ASP 1, a larger κ facilitates more cost-efficient service provision, thereby enhancing its utility. Overall, the increase in κ allows ASPs to reduce resource consumption and increase benefits. However, to ensure fairness and consistency across tasks of the same type, the κ value should be uniformly set by ASPs. A reasonable relaxation of latency constraints can lower costs while ensuring service quality, ultimately yielding mutual benefits for both MUs and ASPs.



(a) Reward decision and QoE. (b) Resource allocation. (c) Utilities of ASP 1 and MU 1.

(d) Reward decision and QoE. (e) Resource allocation. (f) Utilities of ASP 1 and MU 1.

(g) Reward decision and QoE. (h) Resource allocation. (i) Utilities of ASP 1 and MU 1.

Fig. 5. The impact of three parameters on the decisions of MU 1 and ASP 1.

4) Number of MUs and ASPs: The increasing trends observed in Fig. 6(a) for ASPs are attributed to continued market expansion. With the increase in MUs, the demand for their services expands, thereby driving the profitability of ASPs. This growth reflects the positive correlation between the number of MUs and the average utility of ASPs. In contrast, the stability of the average MU utility (fluctuation range $< 5\%$) stems from the emergence of a Nash equilibrium in the non-cooperative game. When new MUs enter the system, existing MUs strategically adjust their reward strategies to respond to competition and maximize their utilities. As the number of ASPs increases, the average utility of MUs gradually rises, while the average utility of ASPs decreases (Fig. 6(b)). This is because when new ASPs are added, MUs reduce the total incentive rewards for each ASP to lower costs, thereby increasing their utilities. However, this reward reduction directly affects the ASP's revenue, leading to a decline in its average utility. Furthermore, MUs' benefits are based on the QoE evaluation of all tasks, making MUs more concerned with the overall benefits of all tasks rather than an individual task. The above findings indicate that, as the market size expands, the average utility of MUs remains relatively stable, while the benefits for ASPs are influenced by the number of MUs and ASPs.

TABLE IV
PERSONALIZED DEMANDS OF MUS.

	AIGC Service Provider 1	AIGC Service Provider 2
MU 1	$(\theta = 1e^{-7}, x^{\text{out}} = 200)$	$(\theta = 1e^{-5}, x^{\text{out}} = 1400)$
MU 2	$(\theta = 1e^{-9}, x^{\text{out}} = 500)$	$(\theta = 1e^{-7}, x^{\text{out}} = 1200)$
MU 3	$(\theta = 1e^{-8}, x^{\text{out}} = 800)$	$(\theta = 1e^{-8}, x^{\text{out}} = 1000)$

Note: The maximum tolerable latency of ASP 1 and ASP 2 are 500ms and 1000ms, respectively.

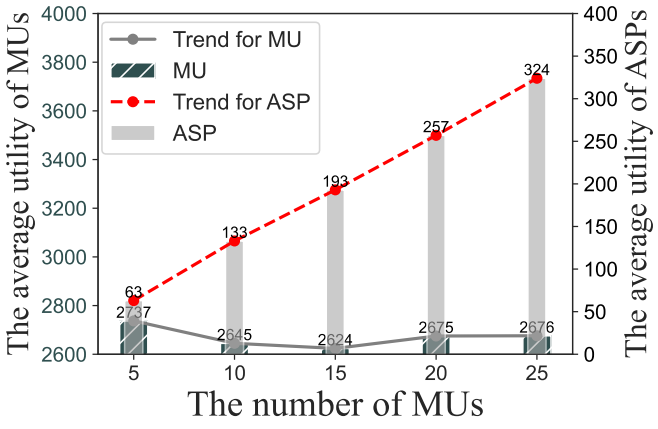
B. Case of Two ASPs and Three MUs

Based on a set of predefined personalized demands, the QoE-driven incentive mechanism optimally determines rewards for each MU and allocates resources for each ASP. This study highlights the varied AIGC service demands of three distinct MUs, including voice-guided navigation for autonomous driving and academic writing services. As shown in Table IV, MUs prioritize accuracy and concise responses to enhance safety and overall experience in autonomous driving.

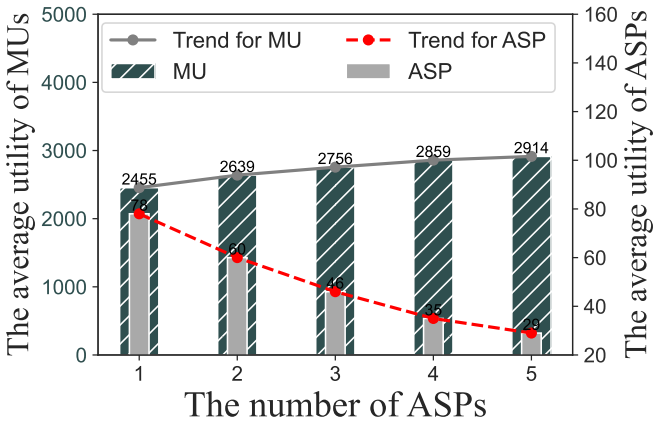
TABLE V
RESULTS OF THE QOE-DRIVEN INCENTIVE MECHANISM:
REWARD DECISIONS FOR MUS AND RESOURCE ALLOCATION DECISIONS FOR ASPS.

	AIGC Service Provider 1				AIGC Service Provider 2			
	computational resource	communication resource	rewards	QoE	computational resource	communication resource	rewards	QoE
MU 1	0.4	1.58	0.06	353	0.62	1.73	0.056	730
MU 2	0.93	2.13	0.1	298	1.23	2.71	0.086	643
MU 3	1.44	2.82	0.126	255	0.99	2.26	0.076	674

Note: The units of computational and communication resources are TFLOPs and MHz, respectively.



(a) The average utility versus number of MUs.



(b) The average utility versus number of ASPs.

Fig. 6. Impacts of the number of MUs/ASPs on utilities.

Conversely, academic writing services focus on the richness and granularity of AIGC.

Table V presents the results obtained through the QoE-driven incentive mechanism. The two ASPs adopt different resource allocation strategies based on the specific service demands of the MUs. For example, MU 3 has high demands for both accuracy and the number of output tokens from ASP 1, leading to the highest rewards allocated to ASP 1. In response, ASP 1 allocates more computational and communication resources to meet MU 3's service requests. However, despite the higher reward, the QoE for MU 3 is the lowest. This is because meeting MU 3's more stringent requirements requires more resources, increasing the total cost and pushing the completion time closer to the maximum tolerable limit. Under our QoE formulation, a lower value directly corresponds to a reduced latency margin, signifying a poorer quality of experience in terms of service timeliness.

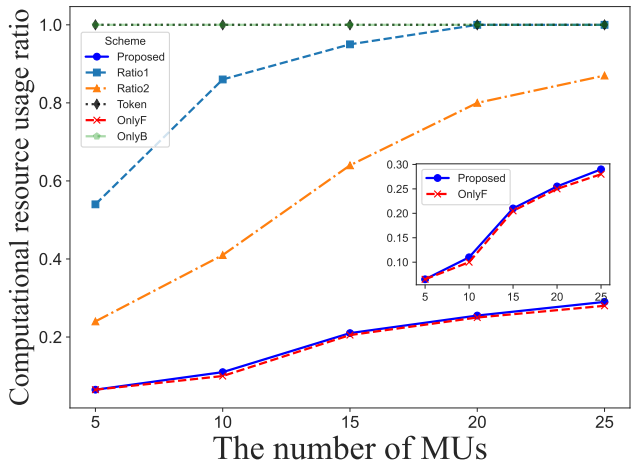
The case study illustrates the practical value of the incentive mechanism by effectively addressing the personalized demands of the MUs. This not only alleviates the resource allocation challenges faced by ASPs but also enhances MUs' satisfaction, thereby improving the overall operational efficiency and competitiveness of the service ecosystem.

C. Comparative Analysis

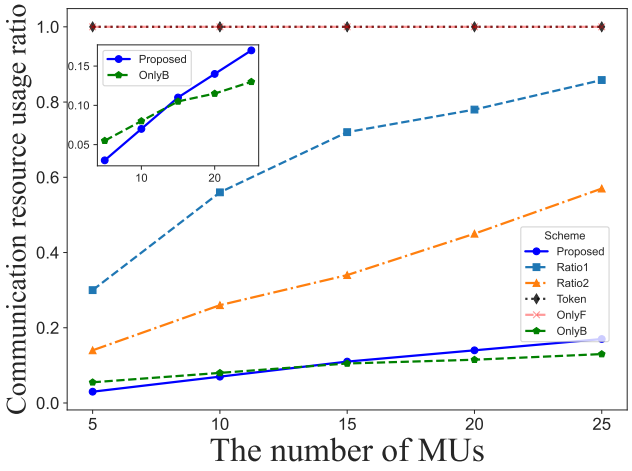
We conduct a comparative analysis with four incentive design schemes, which can be divided into two categories: (1) joint resource optimization schemes under different reward mechanisms, and (2) different resource allocation schemes under the same reward mechanism. This comparison demonstrates that our proposed scheme effectively reduces service access costs for MUs while minimizing resource consumption of ASPs. The four schemes are summarized as follows:

- ***Ratio-based (fixed total reward)***: Applying the ratio idea from [47], MUs decide the reward of unit QoE proportionally based on a fixed total price R_{total}. The total price can be flexibly decided, e.g., $R_{total} = 5$ or $R_{total} = 1$. ASPs jointly optimize the computational and communication resource allocation based on the rewards and resource constraints to satisfy the personalized demands of MUs.
- ***Token-based***: MUs dynamically adjust rewards based on the number of input and output tokens, similar to the pricing model used by ChatGPT[4]. ASPs jointly optimize resource allocation based on the rewards and token consumption by MUs.
- ***OnlyF***: Inspired by the optimization ideas of [10], [21], each MU determines its rewards based on the expected QoE. ASPs optimize computational resource allocation, while communication resources are equally distributed among MUs.
- ***OnlyB***: Similar to *OnlyF*, the reward in *OnlyB* is based on QoE. However, ASPs focus on optimizing communication resource allocation [7], [9], while computational resources are equally divided among MUs.

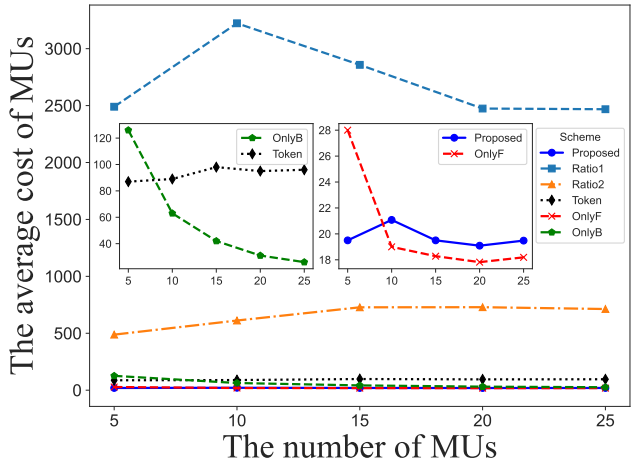
[4]https://openai.com/api/pricing/



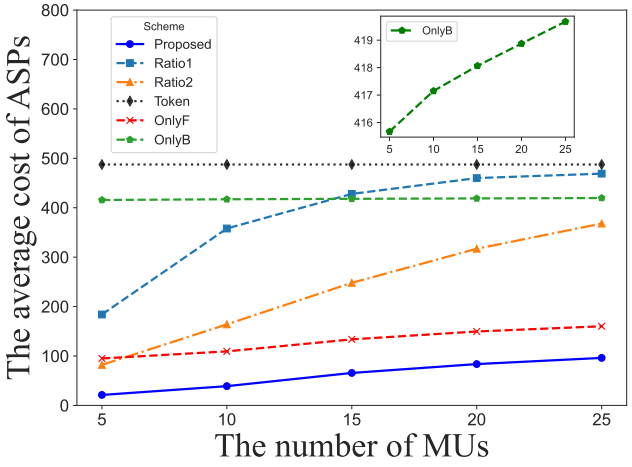
(a) Computational overhead.



(b) Communication overhead.



(c) The average cost of MUs.



(d) The average cost of ASPs.

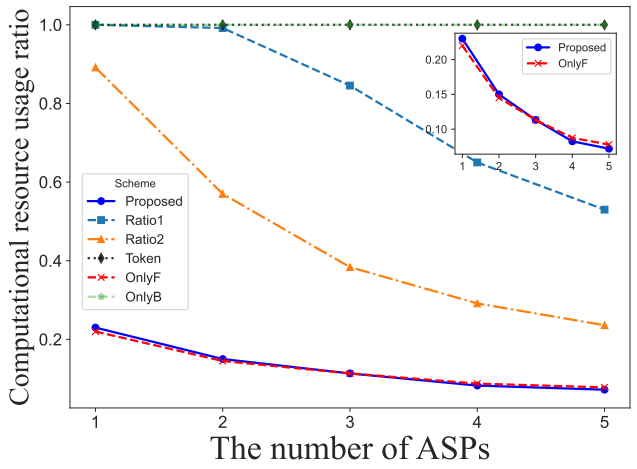
Fig. 7. The impact of the number of MUs on different schemes.

We evaluate the *proposed* incentive mechanism against the above schemes across resource usage ratio and average cost of MUs (ASPs). The results demonstrate that the *proposed* scheme achieves superior efficiency, scalability, and cost-effectiveness through a QoE-driven incentive mechanism and joint resource optimization.

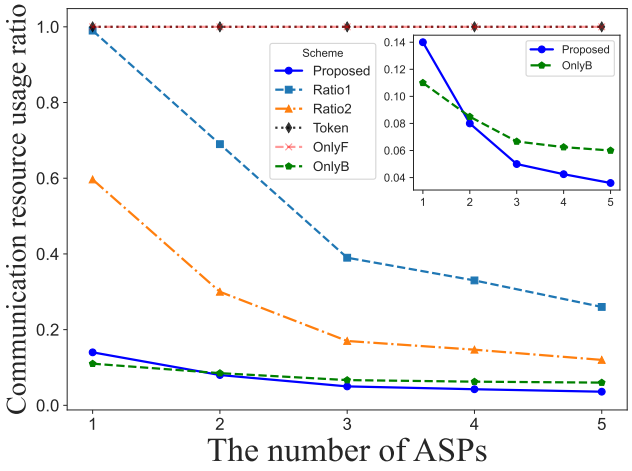
As shown in Figs. 7(a) and 7(b), the *proposed* scheme reduces average resource usage ratio by approximately 75%, 68% and 77% compared to *Ratio1* ($R_{total} = 5$), *Ratio2* ($R_{total} = 1$) and *Token* schemes. This improvement is reflected in the average cost of the ASPs, as illustrated in Fig. 7(d), where the *proposed* consistently maintains the lowest cost. Furthermore, the non-cooperative game between MUs leads to a stable utility, with the *proposed*'s average MU cost fluctuating narrowly between 19.09 and 21.08 (Fig. 7(c)), avoiding the extreme volatility seen in *Ratio1* ($2,468-3,223$) and *Ratio2* ($488-729$). These results stem from the ability of the *proposed* mechanism to dynamically adjust rewards based on QoE, enabling cost control in the competition. In contrast, Ratio-based schemes have high unit reward due to the high R_{total} setting, forcing ASPs to over-provision resources (e.g., *Ratio1*'s computational resource usage ratio is 1.0 for 20+ MUs). The *Token* scheme determines the unit reward based solely on the total number of input and output tokens without considering the personalized demands of MUs. Such a model leads to resource wastage, especially when demand for certain resources is lower than expected, and it fails to adapt to dynamic load changes.

For both *OnlyF* and *OnlyB*, the average MU cost and utilization of the optimized resource decrease. *OnlyF* fully utilizes communication resources (1.0) while incurring a lower average MU cost and consuming less computational resources than *proposed*. However, this increases the average ASP cost due to over-provisioning communication resources. Similarly, in *OnlyB*, the average MU cost decreases from 126.52 to 26.03, but the ASP cost surges to 419.68. This is because the computational resources are allocated equally without considering actual needs, leading to resource wastage and increased costs. By jointly optimizing both computational and communication resources, the *proposed* avoids these pitfalls and achieves resource allocation on demand.

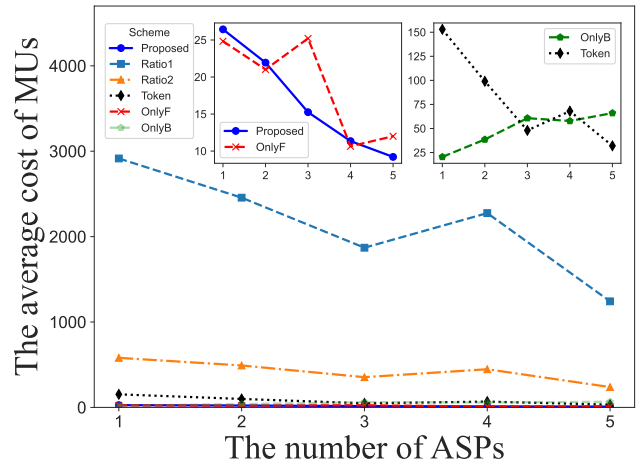
As shown in Fig. 8, as the number of ASPs (i.e., parallel tasks) increases, the *proposed* saves approximately 73% in



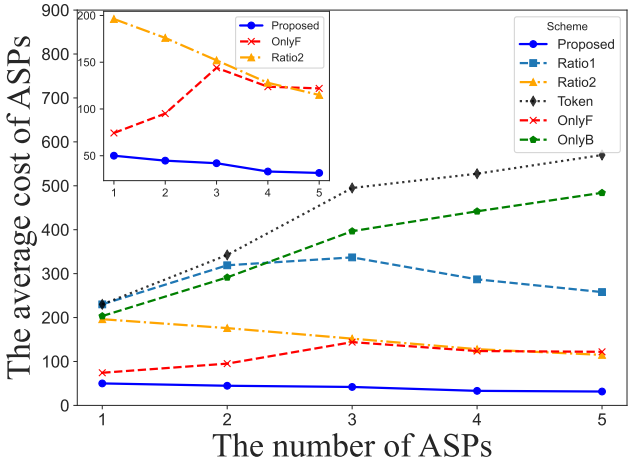
(a) Computational overhead.



(b) Communication overhead.



(c) The average cost of MUs.



(d) The average cost of ASPs.

Fig. 8. The impact of the number of ASPs on different schemes.

computational and communication overhead, 75% in average cost of MUs, and 84% in average cost of ASPs compared to the other schemes. This is because, when new ASPs are added, MUs, aiming to maximize their benefits, reduce the rewards given to each ASP, resulting in a decrease in both the average cost of MUs and ASPs. This phenomenon reflects the *proposed* scheme's adaptability and on-demand resource allocation capabilities. *Ratio2* clearly outperforms *Ratio1*, primarily due to the impact of the fixed total price R_{total}. If R_{total} is set too high, it can lead to overuse of resources; conversely, if set too low, it may fail to meet the service requirements of MUs. This method limits flexibility and adaptability, causing unnecessary resource waste or ineffective service. In addition, the average cost of MUs and ASPs for *OnlyF* and *OnlyB* exhibit variations influenced by the total resources and cost factors, indicating that single resource optimization is insufficient to maximize benefits effectively. For *Token*, the average cost of ASPs continues to increase because the resource usage ratio is always kept at 1.0. It cannot be dynamically adjusted according to the actual load, resulting in inefficient use of resources.

In summary, existing schemes, including fixed total reward-based approaches (*Ratio1*/*Ratio2*), single-dimensional optimization strategies (*OnlyF*/*OnlyB*), and token-based rewards (*Token*), exhibit limitations in balancing the interests of both ASPs and MUs. In contrast, the *proposed*, leveraging QoE-driven dynamic rewards and joint resource optimization, enables efficient on-demand resource allocation. As a result, it achieves average reductions of approximately 64.9% in computational and communication overhead, 66.5% in the service cost for MUs, and 76.8% in the resource consumption of ASPs.

VII. Conclusion

In this paper, we addressed the challenge of delivering personalized AIGC services in resource-constrained edge networks. Our key contribution was the design of a novel multi-dimensional QoE metric that effectively captured the personalized demands of MUs by incorporating AIGC accuracy, token count, and service latency. We quantified model accuracy by measuring the performance gap in chain-of-thought (CoT) reasoning and validated its effectiveness using the ScienceQA benchmark. To cope with limited ASP resources and the competition among MUs, we developed a QoE-driven incentive

mechanism for personalized service provisioning. We reformulated the incentive process as an EPEC, proved the existence and uniqueness of the equilibrium, and proposed a dual-perturbation reward optimization algorithm to compute the optimal solution. Extensive experiments demonstrated that our mechanism achieves efficient, on-demand resource allocation while significantly reducing the overall cost for both ASPs and MUs.

APPENDIX A
PROOF OF CONVERGENCE

Theorem. *Suppose the following assumptions [45] hold:*

- **(A1)** *Each utility $U_m^{mu}(\boldsymbol{R}_m)$ is strictly concave and continuously differentiable in $\boldsymbol{R}_m$,*
- **(A2)** *The perturbation directions are bounded, i.e., $\|\mathbf{d}_m^{(t)}\|_2^2 \leq d_m^+$, for all $m \in \mathcal{M}$,*
- **(A3)** *The difference of rewards between two iterations is bounded, i.e., $\|\mathbf{R}_m^{(t)} - \mathbf{R}_m^{(s)}\|_2^2 \leq R_m^+,\ \forall t \neq s,\ t, s \in \{1, \ldots, T\}$, where R_m^+ is a constant.*

Then, under different step size settings, the convergence properties of Algorithm 1 are as follows:

- **(i)** *If the step sizes Δ_t satisfy $\sum_{t=1}^{\infty} \Delta_t = \infty, \sum_{t=1}^{\infty} \Delta_t^2 < \infty$, then Algorithm 1 converges to an ϵ-Nash equilibrium (NE), i.e.,*

$$\frac{1}{T_\epsilon} \sum_{t=1}^{T_\epsilon} \left[\sum_{m=1}^{M} \left(U_m^{mu}(\mathbf{R}_m^*) - U_m^{mu}(\mathbf{R}_m^{(t)}) \right) \right] \leq \epsilon,$$

within $T_\epsilon = \mathcal{O}\left(\exp\left(\frac{\ddot{\kappa}}{\epsilon}\right)\right)$ iterations, where $\ddot{\kappa} = (\frac{N}{u} \sum_{m=1}^{M} (R_m^{max})^2 + MN\frac{u\pi^2}{6})$.

- **(ii)** *If the step size is fixed as $\Delta_t = \hat{\Delta}$, then Algorithm 1 approaches a $\left[\frac{\hat{\Delta} d_1^+}{2}, \frac{\hat{\Delta} d_2^+}{2}, \ldots, \frac{\hat{\Delta} d_M^+}{2}\right]$ neighborhood of a NE at a rate of $\mathcal{O}\left(\frac{\kappa'}{\epsilon'}\right)$, where $\epsilon' = \epsilon - \frac{\hat{\Delta} MN}{2}, \quad \kappa' = \frac{N \sum_{m=1}^{M} (R_m^{\max})^2}{\hat{\Delta}}$, valid for all $\epsilon > \frac{\hat{\Delta} MN}{2}$.*

Proof. We prove convergence by analyzing the total utility gap and the distance to the NE point $\mathbf{R}^* = (\mathbf{R}_1^*, \ldots, \mathbf{R}_M^*)$, where $\mathbf{R}_m^*$ uniquely maximizes $U_m^{mu}(\mathbf{R}_m, \mathbf{R}_{-m}^*)$ for each MU $m \in \mathcal{M}$. The proof proceeds in four steps.

Step 1: Direction Selection via Two-Sided Finite Difference

For each $R_{nm}^{(t)}$, the direction $\mathbf{d}_{nm}^{(t)}$ is designed to approximate the gradient sign using a two-sided finite difference. Given the utility function $U_m^{mu}(\mathbf{R}_m)$ is continuously differentiable (A1), we evaluate the utility at perturbed rewards $R_{nm}^{(t)} \pm \Delta_t$:

$$U_m^{mu}(R_{nm}^{(t)} + \Delta_t, \mathbf{R}_{-n,m}^{(t)}) \approx U_m^{mu}(R_{nm}^{(t)}, \mathbf{R}_{-n,m}^{(t)}) + \Delta_t \frac{\partial U_m^{mu}}{\partial R_{nm}^{(t)}} + \frac{\Delta_t^2}{2} \frac{\partial^2 U_m^{mu}}{\partial (R_{nm}^{(t)})^2} + O(\Delta_t^3), \tag{18}$$

$$U_m^{mu}(R_{nm}^{(t)} - \Delta_t, \mathbf{R}_{-n,m}^{(t)}) \approx U_m^{mu}(R_{nm}^{(t)}, \mathbf{R}_{-n,m}^{(t)}) - \Delta_t \frac{\partial U_m^{mu}}{\partial R_{nm}^{(t)}} + \frac{\Delta_t^2}{2} \frac{\partial^2 U_m^{mu}}{\partial (R_{nm}^{(t)})^2} + O(\Delta_t^3). \tag{19}$$

Subtracting the above equations gives the utility difference:

$$U_m^{mu}(R_{nm}^{(t)} + \Delta_t, \mathbf{R}_{-n,m}^{(t)}) - U_m^{mu}(R_{nm}^{(t)} - \Delta_t, \mathbf{R}_{-n,m}^{(t)}) \approx 2\Delta_t \frac{\partial U_m^{mu}}{\partial R_{nm}^{(t)}} + O(\Delta_t^3). \tag{20}$$

Thus, the gradient estimate becomes:

$$\hat{g}_{nm}^{(t)} = \frac{U_m^{mu}(R_{nm}^{(t)} + \Delta_t, \mathbf{R}_{-n,m}^{(t)}) - U_m^{mu}(R_{nm}^{(t)} - \Delta_t, \mathbf{R}_{-n,m}^{(t)})}{2\Delta_t} \approx \frac{\partial U_m^{mu}}{\partial R_{nm}^{(t)}} + O(\Delta_t^2). \tag{21}$$

The corresponding direction is then selected as:

$$\mathbf{d}_{nm}^{(t)} = \text{sign}(\hat{g}_{nm}^{(t)}) = \begin{cases} 1 & \text{if } \hat{g}_{nm}^{(t)} > 0, \\ -1 & \text{if } \hat{g}_{nm}^{(t)} < 0, \\ 0 & \text{if } \hat{g}_{nm}^{(t)} \approx 0. \end{cases} \tag{22}$$

Since $\hat{g}_{nm}^{(t)} \approx \frac{\partial U_m^{mu}}{\partial R_{nm}^{(t)}}$, this yields:

$$d_{nm}^{(t)} \approx \text{sign}\left(\frac{\partial U_m^{mu}}{\partial R_{nm}^{(t)}}\right), \mathbf{d}_m^{(t)} \approx \text{sign}(\nabla U_m^{mu}(\mathbf{R}_m^{(t)})), \tag{23}$$

with an approximation error of $O(\Delta_t^2)$. This two-sided finite difference approach provides a more accurate gradient sign estimate than single-sided comparisons, enhancing robustness and aligning with zeroth-order optimization techniques.

Step 2: Monotonic Utility Improvement and Utility Gap

Since $\mathbf{d}_m^{(t)} \approx \text{sign}(\nabla U_m^{mu}(\mathbf{R}_m^{(t)}))$, we use $\mathbf{d}_m^{(t)}$ as the ascent direction for U_m^{mu} in the following analysis. The update:

$$\mathbf{R}_m^{(t+1)} = \mathbf{R}_m^{(t)} + \Delta_t \mathbf{d}_m^{(t)}, \tag{24}$$

implements a sign-based gradient ascent on U_m^{mu}. Since U_m^{mu} is strictly concave (A1), if $\mathbf{d}_{nm}^{(t)} \neq 0$, then $U_m^{mu}(\mathbf{R}_m^{(t+1)}) > U_m^{mu}(\mathbf{R}_m^{(t)})$. If $\mathbf{d}_{nm}^{(t)} = 0$, the utility remains unchanged. Thus, the sequence $\{U_m^{mu}(\mathbf{R}_m^{(t)})\}_{t=1}^{\infty}$ is non-decreasing for each m. The algorithm aims to minimize the utility gap:

$$D_t = \sum_{m=1}^{M} \left(U_m^{mu}(\mathbf{R}_m^*) - U_m^{mu}(\mathbf{R}_m^{(t)}) \right) \geq 0, \tag{25}$$

where $\mathbf{R}_m^*$ is the reward vector maximizing $U_m^{mu}(\mathbf{R}_m)$. For concave functions (A1), the subgradient inequality holds [48]:

$$U_m^{mu}(\mathbf{R}_m^*) \leq U_m^{mu}(\mathbf{R}_m^{(t)}) + \nabla U_m^{mu}(\mathbf{R}_m^{(t)})^\top (\mathbf{R}_m^* - \mathbf{R}_m^{(t)}). \tag{26}$$

Summing over all MUs:

$$D_t \leq \sum_{m=1}^{M} \nabla U_m^{mu}(\mathbf{R}_m^{(t)})^\top (\mathbf{R}_m^* - \mathbf{R}_m^{(t)}). \tag{27}$$

Using $\mathbf{d}_m^{(t)} \approx \text{sign}(\nabla U_m^{mu}(\mathbf{R}_m^{(t)}))$, we approximate:

$$\sum_{m=1}^{M} \nabla U_m^{mu}(\mathbf{R}_m^{(t)})^\top (\mathbf{R}_m^* - \mathbf{R}_m^{(t)}) \approx (\mathbf{d}^{(t)})^\top (\mathbf{R}^* - \mathbf{R}^{(t)}), \tag{28}$$

where $\mathbf{d}^{(t)} = (\mathbf{d}_1^{(t)}, \ldots, \mathbf{d}_M^{(t)})$ and $\mathbf{R}^{(t)} = (\mathbf{R}_1^{(t)}, \ldots, \mathbf{R}_M^{(t)})$. Therefore, the utility gap is upper bounded by:

$$D_t \leq (\mathbf{d}^{(t)})^\top (\mathbf{R}^* - \mathbf{R}^{(t)}) \implies (\mathbf{d}^{(t)})^\top (\mathbf{R}^{(t)} - \mathbf{R}^*) \leq -D_t. \tag{29}$$

Step 3: Distance to Optimal Point

The feasible set is compact, with $\|\mathbf{R}_m^{(t)} - \mathbf{R}_m^{(s)}\|_2^2 \leq R_m^+$, where $R_m^+ = N(R_m^{\max})^2$(A3). Thus, the total squared distance across all MUs is upper bounded:

$$\|\mathbf{R}^{(t)} - \mathbf{R}^{(s)}\|_2^2 \leq R^+,\ R^+ = \sum_{m=1}^{M} R_m^+ = N \sum_{m=1}^{M} (R_m^{\max})^2. \tag{30}$$

Next, consider the squared Euclidean distance between the current reward profile and the optimal $\mathbf{R}^*$:

$$\begin{aligned}\|\mathbf{R}^{(t+1)} - \mathbf{R}^*\|_2^2 &= \|\mathbf{R}^{(t)} + \Delta_t \mathbf{d}^{(t)} - \mathbf{R}^*\|_2^2 \\ &= \|\mathbf{R}^{(t)} - \mathbf{R}^*\|_2^2 + 2\Delta_t (\mathbf{d}^{(t)})^\top (\mathbf{R}^{(t)} - \mathbf{R}^*) + \Delta_t^2 \|\mathbf{d}^{(t)}\|_2^2.\end{aligned} \tag{31}$$

Using the inequality $(\mathbf{d}^{(t)})^\top (\mathbf{R}^{(t)} - \mathbf{R}^*) \leq -D_t$ from **Step 2**, we obtain:

$$\|\mathbf{R}^{(t+1)} - \mathbf{R}^*\|_2^2 \leq \|\mathbf{R}^{(t)} - \mathbf{R}^*\|_2^2 - 2\Delta_t D_t + \Delta_t^2 \|\mathbf{d}^{(t)}\|_2^2. \tag{32}$$

Step 4: Convergence Rate

By averaging the inequality in **Step 3** over iterations $t = 1, \ldots, T$:

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} \|\mathbf{R}^{(t+1)} - \mathbf{R}^*\|_2^2 \leq \frac{1}{T}\sum_{t=1}^{T} \|\mathbf{R}^{(t)} - \mathbf{R}^*\|_2^2 \\ - \frac{2}{T}\sum_{t=1}^{T} \Delta_t D_t + \frac{1}{T}\sum_{t=1}^{T} \Delta_t^2 \|\mathbf{d}^{(t)}\|_2^2.\end{aligned} \tag{33}$$

This can be rewritten recursively as:

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} \|\mathbf{R}^{(t+1)} - \mathbf{R}^*\|_2^2 \leq \|\mathbf{R}^{(1)} - \mathbf{R}^*\|_2^2 - \frac{2}{T}\sum_{t=1}^{T}\sum_{l=1}^{t} \Delta_l D_l \\ + \frac{1}{T}\sum_{t=1}^{T}\sum_{l=1}^{t} \Delta_l^2 \|\mathbf{d}^{(l)}\|_2^2.\end{aligned} \tag{34}$$

Since $\sum_{l=1}^{t} \frac{1}{l} \geq \frac{1}{T}\sum_{t=1}^{T}\sum_{l=1}^{t} \frac{1}{l}$, the second term in the right-hand-side of (34) can be expressed as:

$$\begin{aligned}\frac{2}{T}\sum_{t=1}^{T}\sum_{l=1}^{t} \Delta_l D_l &\geq \frac{2}{T}\sum_{t=1}^{T}\left(\sum_{l=1}^{t} \Delta_l\right) \min_{l\in[1,t]} D_l \\ &\geq \left(\frac{2}{T}\sum_{t=1}^{T}\sum_{l=1}^{t} \Delta_l\right) \frac{1}{T}\sum_{t=1}^{T} \min_{l\in[1,t]} D_l,\end{aligned} \tag{35}$$

Substituting (35) into (34), we derive the following bound:

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} \min_{l\in[1,t]} D_l \leq \frac{\|\mathbf{R}^{(1)} - \mathbf{R}^*\|_2^2 + \frac{1}{T}\sum_{t=1}^{T}\sum_{l=1}^{t} \Delta_l^2 \|\mathbf{d}^{(l)}\|_2^2}{\frac{2}{T}\sum_{t=1}^{T}\sum_{l=1}^{t} \Delta_l} \\ - \frac{\frac{1}{T}\sum_{t=1}^{T} \|\mathbf{R}^{(t+1)} - \mathbf{R}^*\|_2^2}{\frac{2}{T}\sum_{t=1}^{T}\sum_{l=1}^{t} \Delta_l}.\end{aligned} \tag{36}$$

Since $\frac{1}{T}\sum_{t=1}^{T} \|\mathbf{R}^{(t+1)} - \mathbf{R}^*\|_2^2 \geq 0$ and $\|\mathbf{R}^{(1)} - \mathbf{R}^*\|_2^2 \leq R^+$, it follows that:

$$\frac{1}{T}\sum_{t=1}^{T} \min_{l\in[1,t]} D_l \leq \frac{R^+ + \frac{1}{T}\sum_{t=1}^{T}\sum_{l=1}^{t} \Delta_l^2 \|\mathbf{d}^{(l)}\|_2^2}{\frac{2}{T}\sum_{t=1}^{T}\sum_{l=1}^{t} \Delta_l}. \tag{37}$$

Substituting the step size $\Delta_l = \frac{u}{l}$ into (37) , and using the facts that $\|\mathbf{d}^{(t)}\|_2^2 \leq MN$ (A2), $\Delta_l^2 = \frac{u^2}{l^2}$, and the inequality $\sum_{l=1}^{t} \frac{1}{l^2} < \sum_{l=1}^{\infty} \frac{1}{l^2} = \frac{\pi^2}{6}$, we obtain:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{l=1}^{t} \Delta_l^2 \|\mathbf{d}^{(l)}\|_2^2 \leq \frac{u^2 MN}{T}\sum_{t=1}^{T}\sum_{l=1}^{t} \frac{1}{l^2} \leq MNu^2 \frac{\pi^2}{6}. \tag{38}$$

Moreover, leveraging the properties of the harmonic number [48]: $\sum_{l=1}^{t} \frac{1}{l} = \log t + \frac{1}{2t} - \frac{1}{12t^2} + O(t^{-4}) \approx \log t + \frac{1}{2t} + \dot{\kappa}$, where $\dot{\kappa}$ denotes a small residual constant. Based on this approximation, the following result can be derived:

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} \min_{l\in[1,t]} D_l &\leq \frac{\ddot{\kappa}}{\frac{2}{T}\sum_{t=1}^{T}(\log t + \frac{1}{2t} + \dot{\kappa})} \\ &\leq \frac{\ddot{\kappa}}{2\log T + 2\dot{\kappa}} \approx \frac{\ddot{\kappa}}{2\log T},\end{aligned} \tag{39}$$

where $\ddot{\kappa} = (\frac{N}{u}\sum_{m=1}^{M}(R_m^{max})^2 + MN\frac{u\pi^2}{6})$.

Since $\sum_{t=1}^{T} \log t = \Theta(T \log T)$ and hence $\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \log t \to \infty$, which implies that Algorithm 1 converges to a stationary point. Assume that for $T_\epsilon \geq T$ and T_ϵ is a very large number, we try to achieve the accuracy of $\frac{1}{T_\epsilon}\sum_{t=1}^{T_\epsilon} \min_{l\in[1,t]} D_l = \epsilon$. From (39) and concavity of logarithm function, we can obtain:

$$\epsilon \leq \frac{\ddot{\kappa}}{2\log T_\epsilon} \implies \log T_\epsilon \leq \frac{\ddot{\kappa}}{\epsilon} \implies T_\epsilon = \mathcal{O}\left(\exp\left(\frac{\ddot{\kappa}}{\epsilon}\right)\right). \tag{40}$$

Therefore, Algorithm 1 guarantees convergence to an ϵ-NE at a speed of $\mathcal{O}\left(\exp\left(\frac{\ddot{\kappa}}{\epsilon}\right)\right)$.

To reduce the convergence complexity of Algorithm 1, one practical strategy is to relax the required solution accuracy. Specifically, the variable R_{nm} can only be selected from a finite discrete set, i.e., $R_{nm} \in \left\{R_m^{min}, \frac{R_m^{max} - R_m^{min}}{L} + R_m^{min}, \frac{2(R_m^{max} - R_m^{min})}{L} + R_m^{min}, \ldots, R_m^{\max}\right\}$, where L is a fixed quantization level. This discretization can be effectively modeled by using a constant step size $\hat{\Delta} = \frac{R_m^{max} - R_m^{min}}{L}$ in the update rule, which transforms the action space into a finite set.

Substituting $\Delta_l = \hat{\Delta}$ into (37), and using the bounded direction $\|\mathbf{d}^{(l)}\|_2^2 \leq MN$ (A2), the numerator of (37) is bounded as:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{l=1}^{t} \Delta_l^2 \|\mathbf{d}^{(l)}\|_2^2 \leq \hat{\Delta}^2 MN \frac{1}{T}\sum_{t=1}^{T} t = \frac{\hat{\Delta}^2 MN(T+1)}{2}. \tag{41}$$

Similarly, the denominator becomes:

$$\frac{2}{T}\sum_{t=1}^{T}\sum_{l=1}^{t} \Delta_l = 2\hat{\Delta}\frac{T(T+1)}{2T} = \hat{\Delta}(T+1). \tag{42}$$

Therefore, (37) yields:

$$\frac{1}{T}\sum_{t=1}^{T} \min_{l\in[1,t]} D_l \leq \frac{R^+}{\hat{\Delta}(T+1)} + \frac{\hat{\Delta} MN}{2}. \tag{43}$$

Assuming that $T_\epsilon \geq T$ and T_ϵ is a very large number, we try to achieve the accuracy of $\frac{1}{T_\epsilon}\sum_{t=1}^{T_\epsilon} \min_{l\in[1,t]} D_l = \epsilon$. Then, we can derive the following inequality:

$$\epsilon - \frac{\hat{\Delta}MN}{2} \leq \frac{R^+}{\hat{\Delta}(T_\epsilon+1)}. \tag{44}$$

Let $\epsilon' = \epsilon - \frac{\hat{\Delta}MN}{2}$, valid when $\epsilon > \frac{\hat{\Delta}MN}{2}$, we obtain:

$$\epsilon' \leq \frac{R^+}{\hat{\Delta}(T_\epsilon+1)} \Rightarrow T_\epsilon \leq \frac{R^+}{\hat{\Delta}\epsilon'} - 1 \implies T_\epsilon = \mathcal{O}\left(\frac{\kappa'}{\epsilon'}\right), \tag{45}$$

where $\kappa' = \frac{N\sum_{m=1}^{M}(R_m^{\max})^2}{\hat{\Delta}}$, for all $\epsilon > \frac{\hat{\Delta}MN}{2}$.

In this case, if the initial point is not appropriately chosen, Algorithm 1 will be unstable for the first few iterations, i.e., fluctuating between different neighborhood of NEs. However, when the number of iterations is large, the Algorithm 1 approaches a $[\frac{\hat{\Delta}d_1^+}{2}, \frac{\hat{\Delta}d_2^+}{2}, \ldots, \frac{\hat{\Delta}d_M^+}{2}]$ neighborhood of a NE at a rate of $\mathcal{O}(\frac{\kappa'}{\epsilon'})$.

□

REFERENCES

[1] M. Xu *et al.*, "Unleashing the power of edge-cloud generative ai in mobile networks: A survey of aigc services," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 2, pp. 1127–1170, Second quarter, 2024.

[2] H. Du, Z. Li, D. Niyato, J. Kang, Z. Xiong, H. Huang, and S. Mao, "Diffusion-based reinforcement learning for edge-enabled ai-generated content services," *IEEE Trans. Mobile Comput.*, vol. 23, no. 9, pp. 8902–8918, Sep. 2024.

[3] M. Xu, D. Niyato, H. Zhang, J. Kang, Z. Xiong, S. Mao, and Z. Han, "Sparks of generative pretrained transformers in edge intelligence for the metaverse: Caching and inference for mobile artificial intelligence-generated content services," *IEEE Veh. Technol. Mag.*, vol. 18, no. 4, pp. 35–44, Dec. 2023.

[4] R. Zhang, K. Xiong, H. Du, D. Niyato, J. Kang, X. Shen, and H. V. Poor, "Generative AI-enabled vehicular networks: Fundamentals, framework, and case study," *IEEE Network*, vol. 38, no. 4, pp. 259–267, Jul. 2024.

[5] M. Xu *et al.*, "When large language model agents meet 6g networks: Perception, grounding, and alignment," *IEEE Wirel. Commun.*, vol. 31, no. 6, pp. 63–71, Dec. 2024.

[6] Y. Liu, H. Du, D. Niyato, J. Kang, S. Cui, X. Shen, and P. Zhang, "Optimizing mobile-edge AI-generated everything (AIGX) services by prompt engineering: Fundamental, framework, and case study," *IEEE Network*, vol. 38, no. 5, pp. 220–228, Sep. 2024.

[7] Y. Liu *et al.*, "Cross-modal generative semantic communications for mobile AIGC: Joint semantic encoding and prompt engineering," *IEEE Trans. Mobile Comput.*, vol. 23, no. 12, pp. 14 871–14 888, Dec. 2024.

[8] H. Du, Z. Li, D. Niyato, J. Kang, Z. Xiong, X. Shen, and D. I. Kim, "Enabling AI-generated content services in wireless edge networks," *IEEE Wirel. Commun.*, vol. 31, no. 3, pp. 226–234, Jun. 2024.

[9] J. Fan, M. Xu, Z. Liu, H. Ye, C. Gu, D. Niyato, and K.-Y. Lam, "A learning-based incentive mechanism for mobile AIGC service in decentralized internet of vehicles," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Hong kong, Oct. 2023.

[10] J. Wang, H. Du, D. Niyato, J. Kang, Z. Xiong, D. Rajan, S. Mao, and X. Shen, "A unified framework for guiding generative AI with wireless perception in resource constrained mobile edge networks," *IEEE Trans. Mobile Comput.*, vol. 23, no. 11, pp. 10 344–10 360, Nov. 2024.

[11] H. Du, J. Wang, D. Niyato, J. Kang, Z. Xiong, and D. I. Kim, "Ai-generated incentive mechanism and full-duplex semantic communications for information sharing," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 9, pp. 2981–2997, Sep. 2023.

[12] H. Du *et al.*, "Enhancing deep reinforcement learning: A tutorial on generative diffusion models in network optimization," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 4, pp. 2611–2646, Fourth quarter, 2024.

[13] Y. Lin, Z. Gao, H. Du, D. Niyato, J. Kang, Z. Xiong, and Z. Zheng, "Blockchain-based efficient and trustworthy AIGC services in metaverse," *IEEE Trans. Serv. Comput.*, vol. 17, no. 5, pp. 2067–2079, Mar. 2024.

[14] Y. Liu, H. Du, D. Niyato, J. Kang, Z. Xiong, A. Jamalipour, and X. Shen, "Prosecutor: Protecting mobile AIGC services on two-layer blockchain via reputation and contract theoretic approaches," *IEEE Trans. Mobile Comput.*, vol. 23, no. 12, pp. 10 966–10 983, Apr. 2024.

[15] H. Du *et al.*, "Exploring collaborative distributed diffusion-based ai-generated content (AIGC) in wireless networks," *IEEE Network*, vol. 38, no. 3, pp. 178–186, May 2024.

[16] C. Xu, J. Guo, J. Zeng, S. Meng, X. Chu, J. Cao, and T. Wang, "Enhancing AI-generated content efficiency through adaptive multi-edge collaboration," in *Proc. IEEE Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jersey City, USA, Jul. 2024, pp. 960–970.

[17] S. Zhang, M. Xu, W. Y. Bryan Lim, and D. Niyato, "Sustainable aigc workload scheduling of geo-distributed data centers: A multi-agent reinforcement learning approach," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Kuala Lumpur, Malaysia, Dec. 2023, pp. 3500–3505.

[18] R. Cheng, Y. Sun, D. Niyato, L. Zhang, L. Zhang, and M. A. Imran, "A wireless ai-generated content (AIGC) provisioning framework empowered by semantic communication," *IEEE Trans. Mobile Comput.*, vol. 24, no. 3, pp. 2137–2150, Mar. 2025.

[19] Y. Lin, Z. Gao, H. Du, D. Niyato, J. Kang, A. Jamalipour, and X. S. Shen, "A unified framework for integrating semantic communication and AI-generated content in metaverse," *IEEE Network*, vol. 38, no. 4, pp. 174–181, Jul. 2024.

[20] J. Wang, H. Du, D. Niyato, Z. Xiong, J. Kang, S. Mao, and X. Shen, "Guiding AI-generated digital content with wireless perception," *IEEE Wirel. Commun.*, vol. 31, no. 4, pp. 147–154, Apr. 2024.

[21] Y. Wang, C. Liu, and J. Zhao, "Offloading and quality control for AI generated content services in 6G mobile edge computing networks," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Singapore, Jun. 2024.

[22] M. Xu, D. Niyato, J. Kang, Z. Xiong, S. Guo, Y. Fang, and D. I. Kim, "Generative AI-enabled mobile tactical multimedia networks: Distribution, generation, and perception," *IEEE Commun. Mag.*, vol. 62, no. 10, pp. 96–102, Oct. 2024.

[23] J. Wei *et al.*, "Chain-of-thought prompting elicits reasoning in large language models," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 35, Louisiana, US, Nov. 2022, pp. 24 824–24 837.

[24] L. Wang *et al.*, "Plan-and-solve prompting: Improving zero-shot chain-of-thought reasoning by large language models," in *Proc. Annu. Meet. Assoc. Comput. Linguist. (ACL)*, Toronto, Canada, May 2023, pp. 2609–2634.

[25] P. Lu *et al.*, "Learn to explain: Multimodal reasoning via thought chains for science question answering," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 35, Louisiana, US, Nov. 2022, pp. 2507–2521.

[26] Z. Zhang, A. Zhang, M. Li, hai zhao, G. Karypis, and A. Smola, "Multimodal chain-of-thought reasoning in language models," *Trans. Mach. Learn. Res.*, pp. 2835–8856, Jun. 2024.

[27] G. Huang, Q. Wu, J. Li, and X. Chen, "IMFL-AIGC: incentive mechanism design for federated learning empowered by artificial intelligence generated content," *IEEE Trans. Mobile Comput.*, vol. 23, no. 12, pp. 12 603–12 620, Dec. 2024.

[28] Z. Zhan, Y. Dong, D. M. Doe, Y. Hu, S. Li, S. Cao, and Z. Han, "Vision language model-empowered contract theory for AIGC task allocation in teleoperation," *IEEE Trans. Mobile Comput.*, vol. 24, no. 8, pp. 7742–7756, Aug. 2025.

[29] J. Wen, J. Nie, Y. Zhong, C. Yi, X. Li, J. Jin, Y. Zhang, and D. Niyato, "Diffusion-model-based incentive mechanism with prospect theory for edge aigc services in 6G IoT," *IEEE Internet Things J.*, vol. 11, no. 21, pp. 34 187–34 201, Nov. 2024.

[30] D. Ye, S. Cai, H. Du, J. Kang, Y. Liu, R. Yu, and D. Niyato, "Optimizing AIGC services by prompt engineering and edge computing: A generative diffusion model-based contract theory approach," *IEEE Trans. Veh. Technol.*, vol. 74, no. 1, pp. 571–586, Jan. 2025.

[31] M. Xu, D. Niyato, H. Zhang, J. Kang, Z. Xiong, S. Mao, and Z. Han, "Cached model-as-a-resource: Provisioning large language model agents for edge intelligence in space-air-ground integrated networks," May 2024, *arXiv: 2403.05826*. [Online]. Available: https://arxiv.org/abs/2403.05826

[32] Q. Zhang *et al.*, "Exploring edge-driven collaborative fine-tuning toward customized AIGC services," *IEEE Network*, vol. 39, no. 3, pp. 293–301, May 2025.

[33] R. Tutunov, A. Grosnit, J. Ziomek, J. Wang, and H. Bou-Ammar, "Why can large language models generate correct chain-of-thoughts?" Jun. 2024, *arXiv: 2310.13571*. [Online]. Available: https://arxiv.org/abs/2310.13571

[34] H. Jiang, "A latent space theory for emergent abilities in large language models," Sep. 2023, *arXiv: 2304.09960*. [Online]. Available: https://arxiv.org/abs/2304.09960

[35] H. Liu, J. Z. HaoChen, A. Gaidon, and T. Ma, "Self-supervised learning is more robust to dataset imbalance," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, Virtual, Apr. 2022.

[36] T. Yao *et al.*, "Self-supervised learning for large-scale item recommendations," in *Proc. Int. Conf. Inf. Knowl. Manag. (CIKM)*, Queensland, Australia, Oct. 2021, p. 4321–4330.

[37] M. Jin *et al.*, "The impact of reasoning step length on large language models," in *Proc. Annu. Meet. Assoc. Comput. Linguist. (ACL)*, Bangkok, Thailand, Aug. 2024.

[38] C. Singhal and S. De, *Resource Allocation in Next-Generation Broadband Wireless Access Networks*, ser. Advances in Wireless Technologies and Telecommunication (2327-3305). IGI Global, Feb. 2017. [Online]. Available: https://books.google.com.sg/books?id=8BwhDgAAQBAJ

[39] E. Tewolde and V. Conitzer, "Game transformations that preserve nash equilibria or best-response sets," in *Proc. Int. Conf. Auton. Agents Multiagent Syst. (AAMAS)*, Auckland, New Zealand, May 2024, p. 2513–2515.

[40] Y. Zhan, C. H. Liu, Y. Zhao, J. Zhang, and J. Tang, "Free market of multi-leader multi-follower mobile crowdsensing: An incentive mechanism design by deep reinforcement learning," *IEEE Trans. Mobile Comput.*, vol. 19, no. 10, pp. 2316–2329, Oct. 2020.

[41] B. S. Mordukhovich, "Equilibrium problems with equilibrium constraints via multiobjective optimization," *Optimization Methods and Software*, vol. 19, no. 5, pp. 479–492, Feb. 2004.

[42] Y. Zhang and M. Guizani, *Game theory for wireless communications and networking*. CRC press, Jun. 2011.

[43] Z. Chen, Q. Ma, L. Gao, and X. Chen, "Price competition in multi-server edge computing networks under saa and siq models," *IEEE Trans. Mobile Comput.*, vol. 23, no. 1, pp. 754–768, Jan. 2024.

[44] H. Zhang, Y. Xiao, L. X. Cai, D. Niyato, L. Song, and Z. Han, "A multi-leader multi-follower stackelberg game for resource management in lte unlicensed," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 348–361, Jan. 2017.

[45] Y. Xiao, G. Bi, and D. Niyato, "A simple distributed power control algorithm for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3594–3600, Nov. 2011.

[46] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, Eds., *Algorithmic Game Theory*. Cambridge University Press, 2007. [Online]. Available: https://EconPapers.repec.org/RePEc:cup:cbooks:9780521872829

[47] B. Liao *et al.*, "Reward-guided speculative decoding for efficient LLM reasoning," in *Proc. Int. Conf. Mach. Learn. (ICML)*, Vancouver, Canada, Jul. 2025.

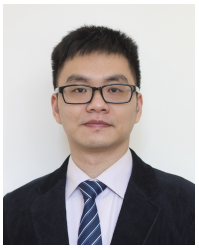
[48] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

Hongjia Wu received the Ph.D. degree in computer science from the National University of Defense Technology, China, in 2023. She was a visiting student at the Singapore University of Technology and Design. She is currently a Post-doctoral Fellow at the Education University of Hong Kong in China. Her research interests mainly focus on computational offloading, game theory, dispersed computing and internet of things.

Minrui Xu (Graduated Student Member, IEEE) received the B.S. degree from Sun Yat-Sen University, Guangzhou, China, in 2021. He is currently working toward the Ph.D. degree in the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His research interests mainly focus on Metaverse, deep reinforcement learning, and mechanism design.

Zehui Xiong is currently a Full Professor with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, United Kingdom. Prior to that, he was with Singapore University of Technology and Design, and Alibaba-NTU Singapore Joint Research Institute. He received his Ph.D. degree from Nanyang Technological University and was a visiting scholar with Princeton University and University of Waterloo. Recognized as a Clarivate Highly Cited Researcher, he has published over 250 peer-reviewed research papers in leading journals, with numerous Best Paper Awards from international flagship conferences. His research interests include 5G/B5G/6G and cellular systems, autonomous vehicles and intelligent transportation, cooperative communication and green communication, land transportation, vehicular communications and connected vehicles, wireless networks.

Lin Gao (Senior Member, IEEE) is a Professor at the School of Electronics and Information Engineering, Harbin Institute of Technology, Shenzhen, China. He received the Ph.D. degree in Electronic Engineering from Shanghai Jiao Tong University, Shanghai, China, in 2010. His main research interests are in the interdisciplinary area between game theory, optimization, and machine learning, with particular focuses on reinforcement learning, federated learning, crowd/edge intelligence, mobile edge computing, cognitive communication and networking. He is the co-recipient of 5 Best Paper Awards from leading conference proceedings on wireless communications and networking. He received the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2016.



Haoyuan Pan (Member, IEEE) received the B.E. and Ph.D. degrees in Information Engineering from The Chinese University of Hong Kong (CUHK), Hong Kong, in 2014 and 2018, respectively.

He was a Post-Doctoral Fellow with the Department of Information Engineering, CUHK, from 2018 to 2020. He is currently an assistant professor with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China. His research interests include wireless communications and networks, Internet of Things (IoT), semantic communications, and age of information (AoI).

Dusit Niyato (Fellow, IEEE) is a professor in the School of Computer Science and Engineering, at Nanyang Technological University, Singapore. He received B.Eng. from King Mongkuts Institute of Technology Ladkrabang (KMITL), Thailand in 1999 and Ph.D. in Electrical and Computer Engineering from the University of Manitoba, Canada in 2008. His research interests are in the areas of sustainability, edge intelligence, decentralized machine learning, and incentive mechanism design.

Tse-Tin Chan (Member, IEEE) received the B.Eng. and Ph.D. degrees in Information Engineering from The Chinese University of Hong Kong (CUHK), Hong Kong, China, in 2014 and 2020, respectively. From 2020 to 2022, he was an Assistant Professor with the Department of Computer Science, The Hang Seng University of Hong Kong (HSUHK), Hong Kong. He is currently an Assistant Professor with the Department of Mathematics and Information Technology, The Education University of Hong Kong (EdUHK), Hong Kong. His research interests include wireless communications and networking, the Internet of Things (IoT), age of information (AoI), and artificial intelligence (AI)-native wireless communications.