This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2024.3515002

Distributed Optimal Contract for Data Rewarding with Network Effects

Alireza Baneshi, Mina Montazeri, Hamed Kebriaei*, *Senior Member, IEEE*

***Abstract*—Data rewarding is a novel business model that leads to an economic trend in mobile networks. In this scheme, the advertiser incentivizes mobile users (MUs) to watch advertisement (ads) and, in return, receive a reward in the form of mobile data. In this work, we model the interaction between an advertiser who has asymmetric information about MUs and MUs who are connected to each other under a network, using the contract theory approach. We obtain the necessary and sufficient conditions for an optimal and practical contract to motivate MUs to participate in the data rewarding scheme and encourage them to declare their private information truthfully. The formulation of this contract is a non-convex-constrained optimization problem. Using lemmas and propositions, we reformulate the initial optimization problem that is challenging to solve as an optimization problem with convex constraints and prove that these two problems are equivalent. Then, with the help of a distributed and non-convex algorithm, we obtain the amount of ads demand and incentive reward.**

***Index Terms*—Data rewarding, contract theory, asymmetric information, incentive compatibility, distributed algorithm.**

I. INTRODUCTION

Owing to the remarkable expansion of mobile traffic volume, the global revenue generated by mobile phone services is approaching a saturation point [1]. Consequently, network operators are seeking innovative data pricing strategies in their marketing efforts to establish new revenue streams and augment their earnings [2], [3]. One method of achieving this objective is by providing greater flexibility to MUs and deriving additional revenues from third-party sources such as advertising companies. To this end, network operators have proposed two schemes for data markets: data rewarding and sponsored content [4], [5]. These schemes offer a promising business approach that ensures the profitability of network operators [4]. Extensive research has been conducted on the design of sponsored content, as evidenced by various scholarly papers [6], [7], [8]. However, our particular focus in this paper lies in exploring the data rewarding scheme.

The data rewarding scheme operates by enabling MUs to view advertisements offered by network operators, in exchange, MUs receive rewards from the advertisers, typically in the form of mobile data [9]. MUs have the freedom to utilize the rewarded data for any mobile content of their choice, while the advertiser's profit solely stems from delivering advertisements to MUs, rather than how MUs utilize the rewarded data. This scheme guarantees the profit of the network operator.

Incentives play a crucial role in encouraging MUs to participate and increasing advertisement viewership within the data rewarding scheme. The advertiser designs an incentive mechanism that offers rewards to MUs, with the aim of maximizing her profit. Research studies have consistently demonstrated the impact of incentives on attracting viewers. For example, a survey conducted by Bangera et al. showed that 76% of respondents are interested in watching ads in exchange for mobile data [10]. In a separate experiment, Sen et al. investigated the effect of financial incentives on viewership and found that users positively respond to such incentives, leading to increased ad views [11].

In [12], interactions between the advertiser and MUs in the data rewarding scheme are modeled with the Stackelberg game. However, this paper relies on the assumption of complete information, which is not applicable in real-world scenarios. To address this limitation, contract theory provides a framework for designing mechanisms under asymmetric information, where the designer lacks knowledge of the agent's private information and aims to encourage truthful disclosure [13], [14]. In the context of data rewarding, [9] utilized contract theory to design incentive rewards. The authors modeled the interaction between the advertiser and MUs, who have multi-dimensional private information, using a multi-dimensional contract. By using the structural features of the problem, they converted the multi-dimensional contract into a one-dimensional contract and calculated incentive rewards.

However, the aforementioned research assumes that MUs are not influenced by each other, which may not accurately reflect real-world dynamics. In reality, individuals' decisions are often influenced by their social connections and interactions. When people see their friends or peers engaging with ads, it can create a social influence that motivates them to watch ads [15]. Research by [16] exemplifies how friends' preferences within a social network influence purchasing decisions and engagement with ads. Word-of-mouth recommendations can amplify the network effect and encourage more people to engage with ads [17]. Also, social media influencers can reinforce the network effect by endorsing or promoting certain ads, which motivates their followers to engage with those ads [15]. Social networks like Facebook exemplify platforms where the behaviors of MUs mutually influence one another [18]. In such systems, the interactions among MUs can be represented as a graph, where nodes correspond to MUs and edges signify

Corresponding Author: Hamed Kebriaei.
This work was supported in part by the Institute for Research in Fundamental Sciences (IPM) under Grant CS 1403-4-340.
Alireza Baneshi, Mina Montazeri and Hamed Kebriaei are with the School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran. Emails: {alirezabaneshi@ut.ac.ir, montazeri70@gmail.com, kebriaei@ut.ac.ir}.
H. Kebriaei is also with the School of Computer Science, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran. Mina Montazeri is also with the Urban Energy Systems Laboratory, Swiss Federal Laboratories for Materials Science and Technology, Dübendorf, Switzerland. Email: {mina.montazeri@empa.ch}.

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2024.3515002

the influence of their connections. These interdependencies are commonly referred to as "network effects" [19].

In this research, we propose an optimal incentive mechanism in the framework of contract theory for the data rewarding scheme when MUs interact with each other and have private information. Contract items are designed to achieve the following three aims: 1) maximizing the advertiser's profit, 2) ensuring MU participation (Individual Rationality (IR)), and 3) encouraging MUs to disclose their private information truthfully (Incentive Compatibility (IC)). To accurately reflect the real setups in the data rewarding scheme, we assume that MUs can influence each other's behavior through social interactions represented by a graph. Furthermore, we assume that MUs' private information, also known as their "type", can be modeled by a discrete distribution function. This private information indicates the strength of social network relationships between MUs [20], [6].

We model the data rewarding scheme as a non-convex constrained optimization problem. The advertiser seeks to obtain incentive rewards that maximize her profit while satisfying the incentive compatibility and individual rationality constraints. The network effects term in this optimization problem causes double complexity, making it non-convex and having coupling variables. As a result, designing and solving the contract for this problem differs from contracts of other applications such as [21], [22], [23]. While past research has tried to simplify the issue by making assumptions and turning it into a convex problem, these assumptions can be restrictive. As a result, we tackle the issue in its true form as a non-convex problem. In addition to the complexity caused by the coupling variables, traditional centralized methods are not suitable for solving these problems due to the distributed nature of the investigated system. The centralized framework has some performance limitations such as high communication requirements, significant computational load, limited flexibility and scalability, and most importantly, lack of user privacy [24]. Therefore, we derive the ads demand and incentive reward using a non-convex distributed algorithm. We analytically reformulate the contract optimization problem with complex non-convex objective and non-convex constraints into an equivalent simpler optimization problem with non-convex objective but convex constraints. Proofs demonstrate that the proposed contract guarantees MUs' participation and encourages them to truthfully disclose their private information (IC and IR constraints). Through rigorous evaluation, we validate the data rewarding scheme's performance efficiency compared to benchmark schemes. In summary, the main contributions of this paper are as follows:

- We propose an incentive mechanism for the data rewarding problem under information asymmetry, where the social interaction of MUs is represented by a graph.
- We reformulate the "ads demand" and "incentive reward" allocation problem from an optimization with numerous non-convex constraints into a tractable optimization.
- Due to the network nature of the system and the advantages of the distributed framework and the non-convexity of the problem, we use a non-convex distributed algorithm to obtain the ads demand and incentive reward.
- We derive a reformulation for the contract problem.This allows us to simultaneously maximize the advertiser's profit while ensuring MUs' participation and truthful disclosure of private information.

II. System Model and Problem Formulation

In this section, we first describe the data rewarding ecosystem. Subsequently, in order to formulate the optimal contract within the data rewarding scheme between the advertiser and MUs, we delve into the utility functions of each party.

The data rewarding ecosystem comprises three entities: advertiser, network operator, and a set of MUs $\mathcal{N} = \{1, 2, \ldots, N\}$ that interact with the advertiser. Each MU has private information, known as the type of MU_j. This type, denoted as γ_j, represents the "true type" of MU_j and its value belongs to the set $\mathbf{B} = \left\{\beta^1, \beta^2, \ldots, \beta^M\right\}$.

In this scheme, the network operator provides mobile data to MUs. This is typically done through a flat-rate data plan (C, D) that is set for a specific period (e.g., monthly). Here, the parameter $C > 0$ represents the subscription fee, while $D > 0$ signifies the amount of data included in the subscription. Then, each MU watches the ads provided by the advertiser and in turn, the advertiser provides an incentive reward to MU so that MU is encouraged to watch more ads. In this way, the advertiser earns revenue from providing advertisements. Also, the advertiser pays the ad display fee to the network operator. The schematic of this market is demonstrated in Figure 1.

The advertiser designs a set of contracts for different types of MUs. This contract includes a set of pairs of "Ads Demand Volume (x)" and "Reward Amount (R)". The advertiser announces the contracts to the MUs, and then the MUs announce their type to the advertiser and receive the corresponding contract. MUs are free to announce their type to the advertiser.

Let $\hat{\gamma}_j$ be the "announced type" by MU_j. Notably, the announced type may differ from the true type of MU, as MUs may strategically misrepresent their type to maximize utility. Then, we define $x_j\left(\hat{\gamma}_j = \beta^q\right)$, and $R_j\left(\hat{\gamma}_j = \beta^q\right)$ as the duration of ads (ads demand) and the incentive reward (reward amount) received by MU_j with announced type β^q, respectively. For simplicity, we use $x_j^q = x_j\left(\hat{\gamma}_j = \beta^q\right)$ and $R_j^q = R_j\left(\hat{\gamma}_j = \beta^q\right)$. In other words, in x_j^q and R_j^q, the superscript index q signifies that the value of the announced type of MU_j is the q-th element of the set $\mathbf{B}$.

Therefore, we can express the utility of MU_j with $\gamma_j = \beta^i$ and $\hat{\gamma}_j = \beta^q$, $\forall i, q \in \mathcal{M} = \{1, 2, \ldots, M\}$ as follows [9], [20].

$$
\begin{aligned}
U_j(x_j^q, R_j^q, \mathbf{X}_{-j}, \beta^i) = S(x_j^q) - p\phi\left(x_j^q\right) + \theta R_j^q - C \\
+\beta^i x_j^q \sum_{k=1}^{N} \sum_{l=1}^{M} g_{jk} \lambda_k^l x_k^l. \quad (1)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{X}_{-j} \triangleq \big[x_1^1, \ldots, x_1^M, \ldots, x_{j-1}^1, \ldots, x_{j-1}^M, \\
x_{j+1}^1, \ldots, x_{j+1}^M, \ldots, x_N^1, \ldots, x_N^M\big]^\top
\end{aligned}
$$

is the ads demand of other MUs excluding MU_j. $S(x_j^q)$ models the total benefit of MU_j with $\hat{\gamma}_j = \beta^q$ from utilizing the mobile services. Inspired by [25], [9], we model this function

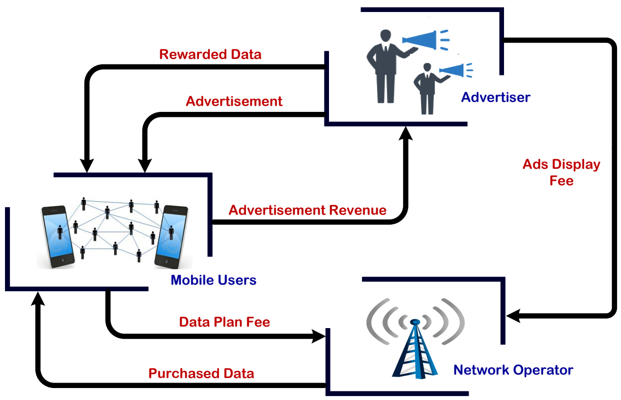


Fig. 1: An interactive framework for Data Rewarding scheme.

as $S(x_j^q) = \theta[D + \tau\phi(x_j^q)]$, where the parameter $\theta > 0$ represents valuation from one unit of content consumption. The increasing function $\phi(x_j^q)$ indicates the amount of data used to watch ads by MU_j with $\hat{\gamma}_j = \beta^q$, and τ is a scaling factor. The second term $p\phi(x_j^q)$ indicates the disutility resulting from watching ads, and the constant p is the MU's average disutility from watching ads [25].

Inspired by [6], [26], the term $\beta^i x_j^q \sum_{k=1}^{N} \sum_{l=1}^{M} g_{jk}\lambda_k^l x_k^l$ represents the external benefits due to the network effects exhibited by the rewarding platform. In this formulation, g_{jk} indicates the strength of MU_k to influence MU_j. For example, the larger the g_{jk}, the more participation of MU_k, the more profit MU_j makes. We consider that for all j and k, $g_{jk} \in [0, 1]$ and for all j, $g_{jj} = 0$. All g_{jk} are commonly known [20], [27].

Earlier, we mentioned that the parameter γ_j, which represents the MU's type, controls the strength of the network effect for each MU [20]. The type of each MU is private information, meaning that only the MU itself knows it, and neither the advertiser nor other MUs know about it. However, it is reasonable to assume that the probability distribution over the types of all MUs is commonly known [28], [27], [29]. This is because the advertiser can estimate statistical information about the distribution of MU private information by learning from MU's historical behavior or by conducting a MU survey. The probability that MU_j belonging to β^i is denoted as λ_j^i with $\sum_{i=1}^{M} \lambda_j^i = 1$, $\forall j \in \mathcal{N}$. Note that since each MU has uncertainty about the ads demand of its neighbors, the last term in (1) has appeared as the expected value of $\mathbf{X}_{-j}$. Also without loss of generality, we consider $\beta^1 > \beta^2 > \cdots > \beta^M > 0$.

The profit of the advertiser is advertising revenue minus the cost to be paid to the network operator for buying ads slot and the incentive reward cost offered to MUs [25]. The expected profit of advertiser from all types is given as follows,

$$\Pi(\mathbf{X}, \mathbf{R}) = \sum_{j=1}^{N} \sum_{i=1}^{M} \lambda_j^i \left[\sigma h\left(x_j^i\right) - q x_j^i - \theta R_j^i\right], \qquad (2)$$

where $\sigma h(x_j^i)$ represents the ads revenue from the volume of ads seen by MU_j with $\hat{\gamma}_j = \beta^i$. σ is the advertisement revenue coefficient. $q x_j^i$ represents the cost to be paid to the network operator for delivering the ads to MU_j with $\hat{\gamma}_j = \beta^i$. Also, we define, $\mathbf{X} = \left[x_1^1, \dots, x_1^M, \dots, x_N^1, \dots, x_N^M\right]^\top$ and $\mathbf{R} = \left[R_1^1, \dots, R_1^M, \dots, R_N^1, \dots, R_N^M\right]^\top$.

III. Contract Between Advertiser and MUs

The advertiser aims to design a data rewarding strategy without knowing MUs' private information, to maximize its expected profit. The contract designed by the advertiser is intended to fulfill three objectives: maximize the advertiser's profit, motivate MUs to watch ads and participate in this scheme, and ensure that MUs truthfully announce their type to the advertiser. Such a contract is deemed feasible and optimal. The second and third objectives are met through the implementation of IR and IC, respectively. Thus, a contract is guaranteed to be feasible if it fulfills both IR and IC. If the advertiser's profit is also met along with these conditions, the contract is optimal. In the following, we define the IR and IC.

Definition 1. *A mechanism is individually rational if each MU can achieve a non-negative utility by announcing her type truthfully, i.e.,*

$$U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq 0 \qquad \forall j \in \mathcal{N}, \ \forall i \in \mathcal{M}. \qquad (3)$$

Definition 2. *A mechanism is incentive compatibility if each MU can achieve equal or higher utility by announcing her type truthfully, i.e.,*

$$U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq U_j(x_j^q, R_j^q, \mathbf{X}_{-j}, \beta^i) \qquad (4)$$
$$\forall j \in \mathcal{N}, \ \forall i, q \in \mathcal{M}, \ i \neq q.$$

Thus, the feasible set of all contracts is defined as $\mathcal{S} = \left\{\left(x_j^i, R_j^i\right) \mid (3)\,,\ (4)\right\}$ .From the contract theory point of view, the advertiser needs to maximize its profit while the contract is feasible (i.e., satisfying both IR and IC constraints). Hence, the optimal contract is the solution to the following optimization

$$\max_{x_j^i, R_j^i} \ \Pi = \sum_{j=1}^{N} \sum_{i=1}^{M} \lambda_j^i \left[\sigma h\left(x_j^i\right) - q x_j^i - \theta R_j^i\right] \qquad (5)$$
$$\begin{aligned} s.t. \quad & U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq 0 \qquad \forall j \in \mathcal{N}, \ \forall i \in \mathcal{M}, \\ & U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq U_j(x_j^q, R_j^q, \mathbf{X}_{-j}, \beta^i) \\ & \forall j \in \mathcal{N}, \quad \forall i, q \in \mathcal{M}, \quad i \neq q, \\ & \bar{x} \geq x_j^i \geq 0 \qquad \forall j \in \mathcal{N}, \forall i \in \mathcal{M}. \end{aligned}$$

A. Reformulation as Equivalent Optimization Problem

Based on (5), we face a non-convex optimization with NM IR constraints and $NM(M-1)$ IC constraints, all of which are non-convex. In what follows, we aim to derive a tractable reformulation of the problem with convex set of constraints which is equivalent to optimization problem (5).

Lemma 1. *Under the feasible contract, the IR constraint is satisfied for all β^i if and only if the IR constraint is binding (or active) on the lowest type, which means,*

$$U_j(x_j^M, R_j^M, \mathbf{X}_{-j}, \beta^M) = 0 \quad \forall j \in \mathcal{N}. \qquad (6)$$

Proof. See Appendix A. □

Lemma 1 allows us to infer that the IR constraint in (5) can be replaced with its equivalent constraint, namely, (6). Furthermore, the equation (6) serves as a necessary and sufficient condition for the IR constraint.

Definition 3. *Local Upward Incentive Constraint (LUIC):* $LUIC\left(\hat{\gamma}_j=\beta^i, \hat{\gamma}_j=\beta^{i+1}\right) \; : \; U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq U_j(x_j^{i+1}, R_j^{i+1}, \mathbf{X}_{-j}, \beta^i)$, $\forall j \in \mathcal{N}$, $\forall i \in \{1,2,\ldots,M-1\}$

Definition 4. *Local Downward Incentive Constraint (LDIC):* $LDIC\left(\hat{\gamma}_j=\beta^i, \hat{\gamma}_j=\beta^{i-1}\right) \; : \; U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq U_j(x_j^{i-1}, R_j^{i-1}, \mathbf{X}_{-j}, \beta^i)$, $\forall j \in \mathcal{N}$, $\forall i \in \{2,3,\ldots,M\}$

Lemma 2. *The solution of (x_j^i, R_j^i) in (5) is IC if and only if all of the following conditions hold:*

$$U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq U_j(x_j^{i+1}, R_j^{i+1}, \mathbf{X}_{-j}, \beta^i), \quad \forall j \in \mathcal{N},\ \forall i \in \{1,2,\ldots,M-1\} \tag{7}$$

$$U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq U_j(x_j^{i-1}, R_j^{i-1}, \mathbf{X}_{-j}, \beta^i), \quad \forall j \in \mathcal{N},\ \forall i \in \{2,3,\ldots,M\} \tag{8}$$

$$x_j^1 \geq x_j^2 \geq \cdots \geq x_j^M, \qquad \forall j \in \mathcal{N} \tag{9}$$

Proof. See Appendix B. □

Lemma 2 indicates necessary and sufficient conditions for the IC constraint, meaning that in (5), the IC constraints can be replaced by the LUIC, LDIC, and monotonicity constraints. Next, we demonstrate that the optimal contract occurs in set $\mathcal{S}' \subseteq \mathcal{S}$, which is defined as follows:

$$\begin{aligned}\mathcal{S}' = \big\{ (x_j^i, R_j^i) \,\big|\, & x_j^1 \geq x_j^2 \geq \cdots \geq x_j^M, \\ & U_j(x_j^M, R_j^M, \mathbf{X}_{-j}, \beta^M) = 0, \\ & U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) = U_j(x_j^{i+1}, R_j^{i+1}, \mathbf{X}_{-j}, \beta^i) \big\}\end{aligned}$$

To demonstrate this, it is enough to prove that in the optimal contract, we can replace the IC constraint with $x_j^1 \geq x_j^2 \geq \cdots \geq x_j^M$ and $U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) = U_j(x_j^{i+1}, R_j^{i+1}, \mathbf{X}_{-j}, \beta^i)$. Because we have proved in Lemma 1 that the IR constraint can be replaced by $U_j(x_j^M, R_j^M, \mathbf{X}_{-j}, \beta^M) = 0$.

Lemma 3. *In the optimal contract, the IC constraint can be replaced by*

$$U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) = U_j(x_j^{i+1}, R_j^{i+1}, \mathbf{X}_{-j}, \beta^i), \quad \forall j \in \mathcal{N}, \quad \forall i \in \{1,2,\ldots,M-1\}, \tag{10}$$

and

$$x_j^1 \geq x_j^2 \geq \cdots \geq x_j^M, \qquad \forall j \in \mathcal{N}. \tag{11}$$

Proof. See Appendix C. □

From Lemma 3, we conclude that in the optimal contract, IC constraint is guaranteed if the conditions $U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) = U_j(x_j^{i+1}, R_j^{i+1}, x_j^l, \beta^i)$, $\forall j \in \mathcal{N}$, $\forall i \in \{1,2,\ldots,M-1\}$ are met. This means that if MU$_j$'s announced type is β^{i+1} and its true type is β^i, it has the same utility as when its announced type is β^i. Thus, we can use Lemma 3 to replace $NM(M-1)$ IC constraints with only $N(M-1)$ constraints. With the help of Lemmas 1 and 3, we reformulate the optimization (5) as the following optimization.

$$\max_{x_j^i, R_j^i} \Pi = \sum_{j=1}^{N}\sum_{i=1}^{M} \lambda_j^i \left[\sigma h\left(x_j^i\right) - q x_j^i - \theta R_j^i\right] \tag{12}$$
$$\begin{aligned} s.t. \quad & U_j(x_j^M, R_j^M, \mathbf{X}_{-j}, \beta^M) = 0 \qquad \forall j \in \mathcal{N}, \\ & U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) = U_j(x_j^{i+1}, R_j^{i+1}, \mathbf{X}_{-j}, \beta^i) \\ & \forall j \in \mathcal{N}, \quad \forall i \in \{1,2,\ldots,M-1\}, \\ & x_j^1 \geq x_j^2 \geq \cdots \geq x_j^M \qquad \forall j \in \mathcal{N}, \\ & \bar{x} \geq x_j^i \geq 0 \qquad \forall j \in \mathcal{N},\ \forall i \in \mathcal{M}. \end{aligned}$$

The optimization problem (12) is equivalent to the optimization problem (5). In (12) compared to (5), the constraints have become simpler and their number has decreased significantly. However, the problem's constraints are still non-convex, and we are faced with two optimization variables. Hence, we use the following proposition to deal with these challenges and to investigate the solution to the optimization problem (12).

Proposition 1. *The optimal solution of the optimization (12) is the same as the solution to the following optimization.*

$$\max_{x_j^i} \quad \sum_{j=1}^{N}\sum_{i=1}^{M} \Bigg\{ \lambda_j^i \left[\sigma h\left(x_j^i\right) - q x_j^i - p\phi\left(x_j^i\right) - C + \theta\left[D + \tau\phi\left(x_j^i\right)\right]\right] + Z_j^i\, x_j^i \sum_{k=1}^{N}\sum_{l=1}^{M} g_{jk}\lambda_k^l x_k^l \Bigg\} \tag{13a}$$
$$s.t. \quad x_j^1 \geq x_j^2 \geq \cdots \geq x_j^M \qquad \forall j \in \mathcal{N}, \tag{13b}$$
$$\bar{x} \geq x_j^i \geq 0 \qquad \forall j \in \mathcal{N},\ \forall i \in \mathcal{M}. \tag{13c}$$

where, for all, $j \in \mathcal{N}$

$$Z_j^i = \begin{cases} \beta^i \sum_{k=1}^{i} \lambda_j^k - \beta^{i-1} \sum_{k=1}^{i-1} \lambda_j^k & \forall i \in \{2,3,\ldots,M\}, \\ \beta^1 \lambda_j^1 & for \ i = 1. \end{cases}$$

Proof. See Appendix 1. □

In this way, we arrive at the optimization (13). This optimization is precisely equivalent to optimization (5), with two key distinctions: the constraints in this formulation are convex, and the number of constraints has been reduced. These constraints are linear and easier to deal with. Also, the number of optimization variables in the problem has been reduced, and it has only one optimization variable.

IV. Distributed Algorithm for Contract Problem

Due to the coupling variables (x_j^i) in the optimization problem (13) and the distributed nature of the investigated networked system, traditional centralized strategies are not suitable solution methods. In addition, the centralized framework has some performance limitations, such as high communication requirements, substantial computation burden, limited flexibility and scalability, and most importantly, lack of privacy [24]. In what follows, we reformulate the problem in distributed form and propose an algorithm to solve it. We rewrite the optimization problem (13) as below,

$$\max_{\mathbf{X}} \qquad F = \sum_{j=1}^{N} F_j\left(\mathbf{X}\right) \tag{14}$$
$$s.t. \qquad Q\mathbf{X} \leq \mathbf{0}, \quad \mathbf{0} \leq \mathbf{X} \leq \bar{\mathbf{x}},$$

where $\mathbf{X}$ is the coupling and global variable of the optimization problem. $\mathbf{0}$, $\bar{\mathbf{x}}$ denotes the vector of all zeros and all $\bar{x}$, respectively. $Q\mathbf{X} \leq 0$ is the compact form of $N(M-1)$ constraints (13b), with $Q \in \mathbb{R}^{N(M-1)\times NM}$. Also, without loss of generality, we denote the set containing all these constraints by $\mathbb{Q}$. The objective function of MU_j is as follows

$$F_j\left(\mathbf{X}\right) = \sum_{i=1}^{M}\Bigg\{\lambda_j^i\left[\sigma h\left(x_j^i\right) - q x_j^i - p\phi\left(x_j^i\right) - C + \theta\left[D + \tau\phi\left(x_j^i\right)\right]\right] + Z_j^i\, x_j^i \sum_{k=1}^{N}\sum_{l=1}^{M} g_{jk}\lambda_k^l x_k^l\Bigg\}. \tag{15}$$

The objective function of (14) is not concave, but the constraints are convex. Therefore, we are facing a non-convex problem. In the following, we adopt a distributed algorithm to solve the non-convex optimization (14).

Let $\mathcal{G} = (\mathcal{N}, \mathcal{E}, \mathsf{G})$ denotes the directed communication graph of MUs, where $\mathcal{N} = \{1, 2, \ldots, N\}$ corresponds to a set of MUs and the set of edges $\mathcal{E} \subseteq \mathcal{N}\times\mathcal{N}$ corresponds to social relations. G is a weighted adjacency matrix for graph. We have denoted the jk-th element of the weighted adjacency matrix by g_{jk}. The weighted adjacency matrix is common knowledge, meaning that all g_{jk} are commonly known [27].

In the proposed algorithm, MU_j starts with an initial estimation of the solution to the problem, i.e., the global variable $\mathbf{X}$, where we denote it by $\mathbf{X}^{(j)}(0) \in \mathbb{Q}$. Then, in the update and projection step, MU_j generates a temporary estimate $\tilde{\mathbf{X}}^{(j)}$. Next, MU_j observes the values of $\tilde{\mathbf{X}}^{(r)}$ by communicating with its neighbors and computes the weighted average.

The weight matrix A is defined as $\mathsf{A} = [a_r^j]_{j,r\in\mathcal{N}} \in \mathbb{R}^{N\times N}$, where the scalar a_r^j is the weight that MU_j assigns to the information $\tilde{\mathbf{X}}^{(r)}$ obtained from a neighboring MU_r and denotes the jr-th element of the weight matrix A.

Algorithm 1 summarizes the procedure for calculating $\mathbf{X}^{(j)}$. In Algorithm 1, α is the gradient step size and $d_j\left(k\right)$ is the gradient of F_j at point $\mathbf{X}^{(j)}\left(k-1\right)$. Also, $\mathbf{Pr}_{\mathbb{Q}}\left(.\right)$ is the projection operator onto the set $\mathbb{Q}$.

The sequence $\mathbf{X}^{(j)}$ converges to the set of Karush-Kuhn-Tucker (KKT) points of the objective function F on the set of constraint $\mathbb{Q}$, for all $j \in \mathcal{N}$. These KKT points are not necessarily the optimal solution, nevertheless, Theorem 1 in [30] ensures that the constraints of the problem, which are the IR and IC constraints in our case, are satisfied.

Assumption 1. *The gradient step size is a sequence such that $\alpha_k = \alpha_0/k^{\zeta}$, where $\alpha_0 > 0$, $\frac{1}{2} < \zeta \leq 1$. Also, the spectral radius ρ of matrix $\mathsf{A}^\top\left(I_N - \mathbf{1}\mathbf{1}^\top/N\right)\mathsf{A}$ satisfies $\rho < 1$, where $\mathbf{1}$ denotes to the $N\times 1$ vector whose components are all equal to one, and $^\top$ denotes transposition.*

Algorithm 1: Non-Convex Distributed Projected Gradient Algorithm

```
Initialize: k ← 1 , set a gradient step size α_k
        X^(j)(0) ∈ ℚ    ∀j ∈ 𝒩
do
    for j = 1 : N do
        Pick the ascent direction as: d_j(k)
        Update and Projection:
            X̃^(j)(k) ← Pr_ℚ( X^(j)(k−1) + α_k d_j(k) )
        Consensus:
            X^(j)(k) ← Σ_{r=1}^{N} a_r^j X̃^(r)(k)
    end
    δ ← max_{j∈𝒩} ‖X^(j)(k) − X^(j)(k−1)‖
    k ← k + 1
while δ ≥ ε_stop;
```

Proposition 2. *Under Assumptions 1, weights rule, doubly stochastic of the matrix $\mathbf{A}$, and Periodically Strongly Connected graph $\mathcal{G}$, the estimation sequence $\left(\mathbf{X}^{(j)}\left(k\right)\right)_{k\geq 0}$, for $j \in \mathcal{N}$, converges to the set of KKT points of the objective function F on the set of constraint $\mathbb{Q}$ [30].*

Proof. See Appendix E. □

Therefore, using Algorithm 1, the ads demand is obtained by solving the optimization (13). Then, using (33), the incentive rewards corresponding to each ads demand are calculated. Hence, the contract items are obtained for each MU.

Remark 1. *It's worth mentioning that from a technical point of view, the heterogeneity in MUs' parameters (i.e., the uniqueness of the parameter value for each MU) doesn't pose a problem. This is because each Lemma and its corresponding proof hold true for MU_j, irrespective of the other MUs.*

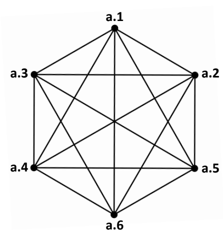
Remark 2. *The distributed approach lets MUs handle computations on their smartphones, cutting costs and complexity. Advertisers save on expenses, gain scalability, and improve fault tolerance, which can lead to lower service fees, better data plans, and more rewards for MUs.*

V. SIMULATION

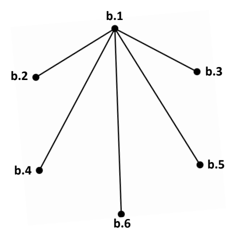
In this section, we investigate the performance of our proposed contract in the data rewarding scheme through extensive numerical simulations. For the simulation setup, we assume $g_{jr} = 0.1$ when agents j and r are connected, otherwise $g_{jr} = 0$. We consider 6 different types in the system, i.e., $M = 6$. The true and announced types are chosen from the set $\mathbf{B} = \{100, 80, 60, 40, 20, 1\}$ ¢$/(\mathbf{Min})^2$. As is customary in the literature [7], [9], the distribution of type for MUs is assumed to be a uniform distribution, i.e. $\lambda_j^i = 1/M$. Other parameters of the system are set according to Table I unless otherwise stated. We consider the linear formulation of $\phi(x)$

TABLE I: Simulated data rewarding platform parameters.

Parameter	Value
τ	35
C	1250 ¢
D	1000 $\mathbf{MB}$
θ	1.1 ¢/$\mathbf{MB}$
σ	400 ¢/$\mathbf{MB}$
p	120 ¢/$\mathbf{MB}$
q	50 ¢/$\mathbf{Min}$
$\bar{x}$	5 $\mathbf{Min}$



(a) Fully connected



(b) Star

Fig. 2: Bidirectional Network Structure Between MUs.

in the simulation, i.e., $\phi(x) = x$ [9]. We also consider $h(x)$ in the form $bx - \frac{a}{2}x^2$, where $a = 0.3$ and $b = 1$ [20], [27].

We first evaluate the feasibility of the proposed contract by examining the validity of IR and IC. We illustrate this with a basic example involving six MUs who can interact via a fully connected graph (as shown in Figure 2a). In this network, five MUs truthfully announce their type, while the sixth MU has the option to announce her type untruthfully. Figure 3 shows the utility of the sixth MU when she announces values of a type different from those of her true type to the advertiser. Each curve in Figure 3 corresponds to different values of the true type of MU a.6, while the announced type of the other five MUs is fixed in all curves. The figure reveals that MU a.6 achieves maximum utility when she announces her type, or her private information, truthfully to the advertiser, as marked by the black stars on the curves. Consequently, the mechanism satisfies the IC constraint. Additionally, the utility of MU a.6 is positive when she announces her information truthfully, indicating that the IR constraint is also satisfied. This makes MU a.6 willing to participate in the scheme. In particular, the utility of MU a.6 with $\gamma_{a.6} = 1$ and $\hat{\gamma}_{a.6} = 1$ is zero, as confirmed in Lemma 1.

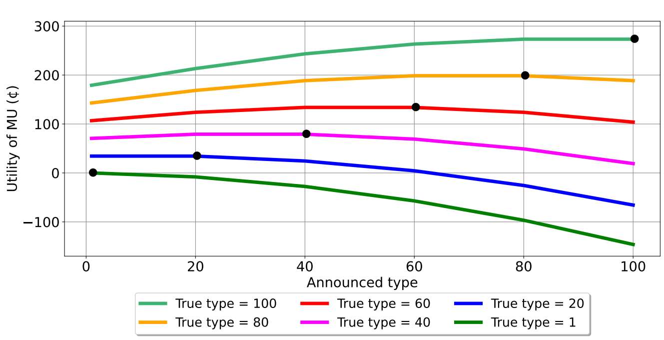


Fig. 3: MU_6 Utility when announcing different values of type.

In the subsequent step, we assess how the network structure affects MU's utility. We examine three MU types: fully connected, branch, and central MUs. Our evaluation involves two network models: a fully connected network (as shown in Figure 2a) and a star network (depicted in Figure 2b). We presume that MUs in both these networks have the same parameters and private information, with their connectivity being the only difference. Consequently, all MUs in Figure 2a have the same utilities. To study the behavior of fully connected MUs, we use MU a.1's utility as an example, which mirrors the utilities of all other MUs in Figure 2a. Furthermore, since MUs b.2, b.3, b.4, b.5, and b.6 have the same parameters and connections, we focus on MU b.2's utility to explore the characteristics of branch MUs. We also designate MU b.1 as the central MU. As demonstrated in Figure 4, fully connected MUs have the highest utility among the three MU types due to the maximum mutual influence in such a network. Additionally, in the star network, the central MU's utility surpasses that of the branch MU because the central MU exerts a stronger mutual influence on the other MUs compared to the branch MU. Therefore, we can infer that the more influence a MU has on others, the more advertisement is allocated to that MU, resulting in more utility. It's important to note that we also assessed the connection strength among MUs. All MUs are assumed to have the same strength of connectivity.

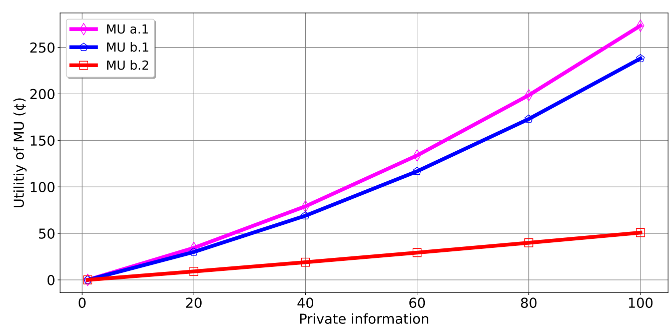


Fig. 4: Utility of MUs a.1 (fully connected MU), b.1 (central MU), and b.2 (branch MU).

In the third step, we explore how the advertiser's profit is affected when a MU untruthfully announces her type. The bidirectional network structure for this scenario is illustrated in Figure 5. The profit of the advertiser when one MU announces

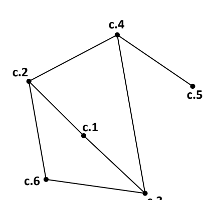


Fig. 5: The bidirectional network structure between MUs.

her type untruthfully is displayed in Table II. The first row displays the outcome when all MUs truthfully announce their type. In the subsequent rows, it's assumed that only one MU announces the information untruthfully. As we can see, the untruthful announced type of MUs c.2, c.3, and c.4, who are each connected to three other MUs, have the most significant impact on the advertiser's profit. Conversely, MU c.5, who is connected to just one MU, has the least impact on the advertiser's profit. Therefore, we deduce that the advertiser should prioritize designing a mechanism that ensures incentive compatibility for the MU with the most influence on others.

TABLE II: Impact of an untruthful MU on advertiser's profit.

Untruthful MU	**Advertiser's profit (¢)**
No of the MU	962.68
MU c.1	807.71
MU c.2	794.16
MU c.3	794.16
MU c.4	794.16
MU c.5	815.51
MU c.6	807.71

Next, Figure 6 presents a comparison of MU's utility in networks with varying numbers of MUs. As depicted, a MU in a network with a greater number of nodes enjoys higher utility compared to MUs in networks with fewer nodes. This is due to the reason that the influence of MUs on each other increases as the number of MUs in the network grows.

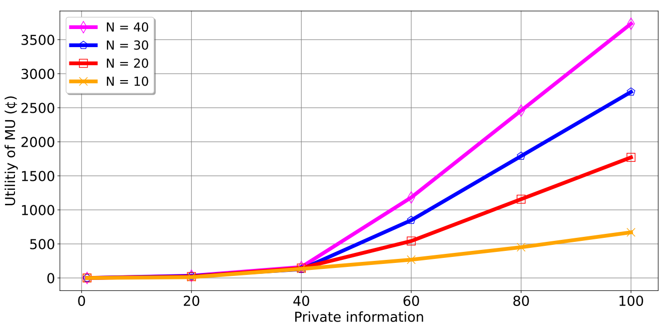


Fig. 6: MU Utility with different numbers of MUs.

Moreover, we compare the performance of the proposed data rewarding scheme with other benchmark schemes. The benchmark schemes that are examined in this work are the discriminatory data rewarding scheme [7] and the data rewarding scheme without considering IC constraints [31]. The discriminatory data rewarding scheme is obtained from the optimal contract under information symmetry, i.e. the advertiser is aware of the types of MUs. Theoretically, this leads to the desired result in terms of advertiser profit and acts as an upper bound. Thus, in the optimization problem, there is no IC constraint and only IR constraint. In the data rewarding scheme without considering the IC constraints, the advertiser does not know the type of each MU and wrongly assumes that MU announces her true type. However, MU as a selfish and rational agent may announce another type to gain more profit. The bidirectional network structure for this case study is depicted in Figure 5. As shown in Figure 7, under the discriminatory data rewarding scheme where the contract design is done with complete information, the advertiser maximizes his utility by minimizing the utility of MU, thus assigning zero utility to MU. In the data rewarding scheme without considering IC constraints, MU announces her type in such a way as to maximize her utility by deceiving the advertiser. For this reason, as shown in Figure 7, MU achieves higher utility than other schemes. According to Table III, if MUs deceive the advertiser and untruthfully announce their type, the total utility of MUs reaches its maximum value and the advertiser earns less profit than other schemes. In contrast, in the discriminatory data rewarding scheme, the total utility of MUs remains as low as possible, i.e. zero, and the advertiser gets the most profit compared to other schemes.

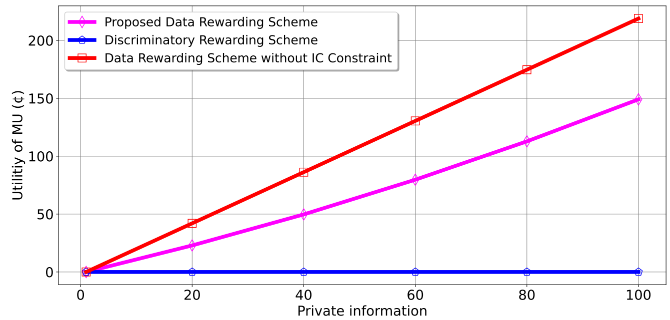


Fig. 7: Utility of MU under three different schemes.

VI. Conclusion

In conclusion, this paper introduces a data rewarding scheme that economically incentivizes MUs while considering network effects. We employ contract theory to model the interaction between advertiser and MUs under information asymmetry. The contract design is formulated as a non-convex constrained optimization problem, and we derive the necessary and sufficient conditions for the problem's constraints. To overcome the challenges posed by non-convex constraints, we simplify the initial optimization problem into a tractable equivalent problem with convex constraints, which we then solve using a distributed algorithm. Our simulations demonstrate that our proposed scheme outperforms the benchmark.

Appendix A
Proof of Lemma 1

We divide the proof of this lemma into two parts, the forward direction "**If**", and the backward direction, "**Only If**".

To show the "**If**" part, we have to prove: $U_j(x_j^M, R_j^M, \mathbf{X}_{-j}, \beta^M) = 0\,,\ \forall j \in \mathcal{N} \implies$ IR constraint
To prove this part, from IC constraint, we have $U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq U_j(x_j^M, R_j^M, \mathbf{X}_{-j}, \beta^i)$, and also using $\beta^i \geq \beta^M$, for all $j \in \mathcal{N}$, we obtain

$$U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq \theta\left[D + \tau\phi\left(x_j^M\right)\right] - p\phi\left(x_j^M\right)$$
$$+\,\theta R_j^M - C + \beta^M x_j^M \sum_{k=1}^{N}\sum_{l=1}^{M} g_{jk}\lambda_k^l x_k^l.$$

Thus, we can conclude that $U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq U_j(x_j^M, R_j^M, \mathbf{X}_{-j}, \beta^M)$. Furthermore, it proves that IR constraint for β^M is binding, i.e., $U_j(x_j^M, R_j^M, \mathbf{X}_{-j}, \beta^M) = 0$.

TABLE III: Profit of advertiser and total utilities of MUs under different schemes.

Scheme	**Advertiser's profit (¢)**	**MUs' total utilities (¢)**
Proposed data rewarding scheme	962.68	376.63
Discriminatory data rewarding scheme	1437.20	0
Data rewarding scheme without considering IC constraints	877.32	580.23

Otherwise, we could decrease θR_j^i for $\beta^i \in \mathbf{B}$ by $\epsilon > 0$, which would preserve all the constraints of the optimization (5), and at the same time increase the profit of the advertiser. In this way, the proof of "**If**" part is then completed.
To show the "**Only If**" part, we have to prove:
IR constraint $\Longrightarrow U_j(x_j^M, R_j^M, \mathbf{X}_{-j}, \beta^M) = 0\ ,\ \forall j \in \mathcal{N}$.

The proof of this part is self-evident since $U_j(x_j^M, R_j^M, \mathbf{X}_{-j}, \beta^M) = 0$ is a specific case of the IR constraint. If the IR constraint is met, it's certain that this constraint is also valid. Consequently, the proof of "**Only If**" part is established, and the proof of this lemma is completed.

Appendix B
Proof of Lemma 2

We divide the proof of this lemma into two parts, the forward direction "**If**", and the backward direction, "**Only If**".

To show the "**If** " part, we have to prove:
LUIC($\hat{\gamma}_j = \beta^i, \hat{\gamma}_j = \beta^{i+1}$), LDIC($\hat{\gamma}_j = \beta^i, \hat{\gamma}_j = \beta^{i-1}$), and $x_j^1 \geq x_j^2 \geq \cdots \geq x_j^M \Longrightarrow$ IC constraint. To prove this part, it suffices to prove the following statements.

$$\begin{aligned} &1)\ \ If \quad LUIC\left(\hat{\gamma}_j = \beta^i, \hat{\gamma}_j = \beta^{i+1}\right),\ x_j^1 \geq x_j^2 \geq \cdots \geq x_j^M \\ &\quad then \quad U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq U_j(x_j^q, R_j^q, \mathbf{X}_{-j}, \beta^i), \\ &\qquad\qquad \forall j \in \mathcal{N},\ \forall i, q \in \{1, 2, \ldots, M-1\}\,, \quad q > i, \end{aligned} \tag{16}$$

$$\begin{aligned} &2)\ \ If \quad LDIC\left(\hat{\gamma}_j = \beta^i, \hat{\gamma}_j = \beta^{i-1}\right),\ x_j^1 \geq x_j^2 \geq \cdots \geq x_j^M \\ &\quad then \quad U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq U_j(x_j^q, R_j^q, \mathbf{X}_{-j}, \beta^i), \\ &\qquad\qquad \forall j \in \mathcal{N},\ \forall i, q \in \{2, 3, \ldots, M\}\,, \quad q < i. \end{aligned} \tag{17}$$

First, we prove (16). By replacing the utility of MU$_j$ with $\gamma_j = \beta^{i+1}$ and $\hat{\gamma}_j = \beta^{i+1}$, $\hat{\gamma}_j = \beta^{i+2}$ from equation (1) in $LUIC\left(\hat{\gamma}_j = \beta^{i+1}, \hat{\gamma}_j = \beta^{i+2}\right)$, we have

$$\begin{aligned} &\theta\left[D + \tau\phi\left(x_j^{i+1}\right)\right] - p\phi\left(x_j^{i+1}\right) + \theta R_j^{i+1} - C + \\ &\beta^{i+1} x_j^{i+1} \sum_{k=1}^{N}\sum_{l=1}^{M} g_{jk}\lambda_k^l x_k^l \geq \theta\left[D + \tau\phi\left(x_j^{i+2}\right)\right] - \\ &p\phi\left(x_j^{i+2}\right) + \theta R_j^{i+2} - C + \beta^{i+1} x_j^{i+2} \sum_{k=1}^{N}\sum_{l=1}^{M} g_{jk}\lambda_k^l x_k^l. \end{aligned} \tag{18}$$

Given $\beta^i > \beta^{i+1} > \beta^{i+2}$ and $x_j^i \geq x_j^{i+1} \geq x_j^{i+2}$ for all $j \in \mathcal{N}$, we have,

$$\left(\sum_{k=1}^{N}\sum_{l=1}^{M} g_{jk}\lambda_k^l x_k^l\right)\left(\beta^i - \beta^{i+1}\right)\left(x_j^{i+1} - x_j^{i+2}\right) \geq 0. \tag{19}$$

By adding (18) and (19) together, we have

$$U_j(x_j^{i+1}, R_j^{i+1}, \mathbf{X}_{-j}, \beta^i) \geq U_j(x_j^{i+2}, R_j^{i+2}, \mathbf{X}_{-j}, \beta^i). \tag{20}$$

Combining equation (20) and $U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq U_j(x_j^{i+1}, R_j^{i+1}, \mathbf{X}_{-j}, \beta^i)$, we obtain $U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq U_j(x_j^{i+2}, R_j^{i+2}, \mathbf{X}_{-j}, \beta^i)$. In the same way, it is proven for higher types as well. Thus, we can further prove (16). Also, we can prove (17) in a way similar to that of (16). The proof of the forward direction "**If**" is then completed.

To show the "**Only If**", we prove the following statements:

1) IC constraint $\Longrightarrow LUIC\left(\hat{\gamma}_j = \beta^i, \hat{\gamma}_j = \beta^{i+1}\right)$, (21)
2) IC constraint $\Longrightarrow LDIC\left(\hat{\gamma}_j = \beta^i, \hat{\gamma}_j = \beta^{i-1}\right)$, (22)
3) IC constraint $\Longrightarrow x_j^1 \geq x_j^2 \geq \cdots \geq x_j^M$. (23)

Statements (21) and (22) are clearly established and proven. The LUIC and LDIC constraints are specific cases of the IC constraint. Therefore, if the IC constraint is established, it's certain that these constraints are also established. Thus, we just need to prove (23) which is as follows:

From the IC constraint for MU$_j$ with $\gamma_j = \beta^i$ and MU$_j$ with $\gamma_j = \beta^q$, we have respectively

$$U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) \geq U_j(x_j^q, R_j^q, \mathbf{X}_{-j}, \beta^i), \tag{24}$$
$$U_j(x_j^q, R_j^q, \mathbf{X}_{-j}, \beta^q) \geq U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^q), \tag{25}$$

By expanding (24) and (25) using equation (1) and then adding their sides together, we obtain $\left(\beta^i - \beta^q\right)\left(x_j^i - x_j^q\right) \geq 0$. Hence, if $\beta^i > \beta^q$, it follows that $x_j^i \geq x_j^q$. This implies that if $\beta^1 > \beta^2 > \cdots > \beta^M$, then $x_j^1 \geq x_j^2 \geq \cdots \geq x_j^M$, $\forall j \in \mathcal{N}$. This condition, also known as the monotonicity condition, implies that if MU$_j$ announces a higher type, she has a stronger inclination towards viewing ads, which makes sense. With this, the proof is successfully concluded.

Appendix C
Proof of Lemma 3

Based on Lemma 2, we have to prove the following

$$\begin{aligned} &1)\ U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) = U_j(x_j^{i+1}, R_j^{i+1}, \mathbf{X}_{-j}, \beta^i),\ \forall j \in \mathcal{N}, \\ &\qquad \forall i \in \{1, 2, \ldots, M-1\} \text{ and } x_j^1 \geq x_j^2 \geq \cdots \geq x_j^M, \\ &\quad \Longrightarrow LDIC\left(\hat{\gamma}_j = \beta^i, \hat{\gamma}_j = \beta^{i-1}\right) : \forall i \in \{2, 3, \ldots, M\} \end{aligned} \tag{26}$$

2) The $LUIC\left(\hat{\gamma}_j = \beta^i, \hat{\gamma}_j = \beta^{i+1}\right)$ constraint is binding on the optimal contract, i.e.,

$$\begin{aligned} &U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) = U_j(x_j^{i+1}, R_j^{i+1}, \mathbf{X}_{-j}, \beta^i), \\ &\qquad \forall j \in \mathcal{N},\ \forall i \in \{1, 2, \ldots, M-1\} \end{aligned} \tag{27}$$

We begin with proving (26). First, we write the $LUIC\left(\hat{\gamma}_j = \beta^i, \hat{\gamma}_j = \beta^{i+1}\right)$ constraint in its equality state (binding), which becomes $U_j(x_j^i, R_j^i, \mathbf{X}_{-j}, \beta^i) =$

$U_j(x_j^{i+1}, R_j^{i+1}, \mathbf{X}_{-j}, \beta^i) : \forall j \in \mathcal{N}, \forall i \in \{1, 2, \ldots, M-1\}$. Using (1) we have

$$\theta\left[D+\tau\phi\left(x_j^i\right)\right]-p\phi\left(x_j^i\right)+\theta R_j^i-C+ \beta^i x_j^i \sum_{k=1}^N \sum_{l=1}^M g_{jk}\lambda_k^l x_k^l = \theta\left[D+\tau\phi\left(x_j^{i+1}\right)\right]-p\phi\left(x_j^{i+1}\right) + \theta R_j^{i+1} - C + \beta^i x_j^{i+1} \sum_{k=1}^N \sum_{l=1}^M g_{jk}\lambda_k^l x_k^l. \quad (28)$$

Second, by replacing the utility of MU$_j$ with $\gamma_j = \beta^{i+1}$ and $\hat{\gamma}_j = \beta^{i+1}$, $\hat{\gamma}_j = \beta^i$ from equation (1) in $LDIC\left(\hat{\gamma}_j = \beta^{i+1}, \hat{\gamma}_j = \beta^i\right)$, we have

$$\theta\left[D+\tau\phi\left(x_j^{i+1}\right)\right]-p\phi\left(x_j^{i+1}\right)+\theta R_j^{i+1}-C+ \beta^{i+1} x_j^{i+1} \sum_{k=1}^N \sum_{l=1}^M g_{jk}\lambda_k^l x_k^l \geq \theta\left[D+\tau\phi\left(x_j^i\right)\right] - p\phi\left(x_j^i\right)+\theta R_j^i - C + \beta^{i+1} x_j^i \sum_{k=1}^N \sum_{l=1}^M g_{jk}\lambda_k^l x_k^l. \quad (29)$$

From (28), $\beta^i > \beta^{i+1}$, and $x_j^i \geq x_j^{i+1}$ we obtain:

$$\begin{aligned} &\theta\left[D+\tau\phi\left(x_j^{i+1}\right)\right]-p\phi\left(x_j^{i+1}\right)+\theta R_j^{i+1}-C \\ &-\theta\left[D+\tau\phi\left(x_j^i\right)\right]+p\phi\left(x_j^i\right)-\theta R_j^i+C \\ &= \beta^i\left(x_j^i - x_j^{i+1}\right)\sum_{k=1}^N \sum_{l=1}^M g_{jk}\lambda_k^l x_k^l \\ &\geq \beta^{i+1}\left(x_j^i - x_j^{i+1}\right)\sum_{k=1}^N \sum_{l=1}^M g_{jk}\lambda_k^l x_k^l. \end{aligned} \quad (30)$$

Rearranging the terms in (30), we obtain the desired inequality in (29). Hence, (26) is proved.

We now prove (27). By writing the $LUIC\left(\hat{\gamma}_j = \beta^i, \hat{\gamma}_j = \beta^{i+1}\right)$ constraint and substituting the MU$_j$'s utility function from equation (1) in it, we have:

$$\theta\left[D+\tau\phi\left(x_j^i\right)\right]-p\phi\left(x_j^i\right)+\theta R_j^i-C+ \beta^i x_j^i \sum_{k=1}^N \sum_{l=1}^M g_{jk}\lambda_k^l x_k^l \geq \theta\left[D+\tau\phi\left(x_j^{i+1}\right)\right] - p\phi\left(x_j^{i+1}\right)+\theta R_j^{i+1} - C + \beta^i x_j^{i+1} \sum_{k=1}^N \sum_{l=1}^M g_{jk}\lambda_k^l x_k^l, \quad (31)$$

which can be written as

$$\theta R_j^i \geq (\tau\theta - p)\left(\phi\left(x_j^{i+1}\right)-\phi\left(x_j^i\right)\right)+\theta R_j^{i+1} \quad \beta^i\left(x_j^{i+1}-x_j^i\right)\sum_{k=1}^N \sum_{l=1}^M g_{jk}\lambda_k^l x_k^l. \quad (32)$$

According to Lemma 1, it is evident that in the optimal contract, each MU's lowest type has a utility of zero. Thus, in the optimal contract, the advertiser sets $\theta R_j^M = -\theta\left[D+\tau\phi\left(x_j^M\right)\right] + p\phi\left(x_j^M\right) + C - \beta^M x_j^M \sum_{k=1}^N \sum_{l=1}^M g_{jk}\lambda_k^l x_k^l$. Given R_j^M, for $i = M-1$ in (32), the advertiser reduces the value of the data reward to some extent to preserve the equal sign in (32). Given R_j^{M-1}, for $i = M-2$ in (32), the advertiser reduces the value of the data reward to some extent to preserve the equal sign in (32). Hence, in the optimal contract, the advertiser reduces the value of data reward R_j^i to take the equal sign of (32) for $i \in \{1, 2, \ldots, M-1\}$. Therefore, $LUIC\left(\hat{\gamma}_j = \beta^i, \hat{\gamma}_j = \beta^{i+1}\right)$ for $i \in \{1, 2, \ldots, M-1\}$ are binding and (27) is proved. In this way, the proof is completed.

Appendix D
Proof of Proposition 1

First, by expanding the first and second constraints of (12), we obtain a closed form for data reward R_j^i as follows

$$R_j^i = \frac{1}{\theta}\left(f_j(x_j^M, \mathbf{X}_{-j}, \beta^M) + \sum_i^M w_j^i\right) \quad \forall j \in \mathcal{N}, \quad \forall i \in \{1, 2, \ldots, M-1\}, \quad (33)$$

so that in equation (33), for all $j \in \mathcal{N}$: $w_j^i = f_j(x_j^i, \mathbf{X}_{-j}, \beta^i) - f_j(x_j^{i+1}, \mathbf{X}_{-j}, \beta^i)$ $\forall i \in \{1, 2, \ldots, M-1\}$, and $w_j^M = 0$, and $f_j(x_j^i, \mathbf{X}_{-j}, \beta^i) = -\theta\left[D+\tau\phi\left(x_j^i\right)\right]+p\phi\left(x_j^i\right)+C-\beta^i x_j^i \sum_{k=1}^N \sum_{l=1}^M g_{jk}\lambda_k^l x_k^l$. Next, we substitute (33) in the objective function of the optimization (12) and remove its dependence on $\mathbf{R}$, specifically

$$\sum_{j=1}^N \sum_{i=1}^M \left\{ \lambda_j^i\left[\sigma h\left(x_j^i\right) - q x_j^i + \theta D - C + (\theta\tau - p)\,\phi\left(x_j^i\right)\right] + Z_j^i\, x_j^i \sum_{k=1}^N \sum_{l=1}^M g_{jk}\lambda_k^l x_k^l \right\}, \quad (34)$$

where, for all $j \in \mathcal{N}$: $Z_j^i = \beta^i \sum_{k=1}^i \lambda_j^k - \beta^{i-1}\sum_{k=1}^{i-1}\lambda_j^k$ $\forall i \in \{2, 3, \ldots, M\}$ and $Z_j^1 = \beta^1\lambda_j^1$. Thus, we reach the optimization (13), and the proof is completed.

Appendix E
Proof of Proposition 2

Considering that the set of constraints $\mathbb{Q}$ is nonempty, compact, and convex, it is proved in Theorem 1 of [30] that the sequence of each MU's estimation of the problem's solution $\left(\mathbf{X}^{(j)}\right)$ converges to KKT points of the objective function F on the set of constraint $\mathbb{Q}$.

References

[1] S. Analytics, "Worldwide cellular user forecasts 2018-2023," Tech. Rep, Tech. Rep., 2018.

[2] Z. Wang, L. Gao, and J. Huang, "Multi-cap optimization for wireless data plans with time flexibility," *IEEE Transactions on Mobile Computing*, vol. 19, no. 9, pp. 2145–2159, 2019.

[3] ——, "Exploring time flexibility in wireless data plans," *IEEE Transactions on Mobile Computing*, vol. 18, no. 9, pp. 2048–2061, 2018.

[4] "SOCIFI: Sponsored data & data rewards," https://www.socifi.com/sponsored-data-and-data-rewards.

[5] "Datami: Sponsored data, data rewards, mobile engagement," https://www.datami.com/.

[6] M. Montazeri, P. Rokhforoz, H. Kebriaei, and O. Fink, "Incentive mechanism in the sponsored content market with network effects," *IEEE Transactions on Computational Social Systems*, 2023.

[7] Z. Xiong, J. Zhao, Z. Yang, D. Niyato, and J. Zhang, "Contract design in hierarchical game for sponsored content service market," *IEEE Transactions on Mobile Computing*, 2020.

[8] Z. Xiong, S. Feng, D. Niyato, P. Wang, A. Leshem, and Z. Han, "Joint sponsored and edge caching content service market: A game-theoretic approach," *IEEE Transactions on Wireless Communications*, vol. 18, no. 2, pp. 1166–1181, 2019.

[9] Z. Xiong, J. Kang, D. Niyato, P. Wang, H. V. Poor, and S. Xie, "A multi-dimensional contract approach for data rewarding in mobile networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 9, pp. 5779–5793, 2020.

[10] P. Bangera, S. Hasan, and S. Gorinsky, "An advertising revenue model for access isps," in *2017 IEEE Symposium on Computers and Communications (ISCC)*. IEEE, 2017, pp. 582–589.

[11] S. Sen, G. Burtch, A. Gupta, and R. Rill, "Incentive design for ad-sponsored content: Results from a randomized trial," in *2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2017, pp. 826–831.

[12] H. Yu, E. Wei, and R. A. Berry, "Monetizing mobile data via data rewards," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 4, pp. 782–792, 2020.

[13] T. Roughgarden, "Algorithmic game theory," *Communications of the ACM*, vol. 53, no. 7, pp. 78–86, 2010.

[14] A. Sinha and A. Anastasopoulos, "Mechanism design for resource allocation in networks with intergroup competition and intragroup sharing," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1098–1109, 2017.

[15] K. Lajnef, "The effect of social media influencers' on teenagers behavior: an empirical study using cognitive map technique," *Current Psychology*, pp. 1–14, 2023.

[16] R. Iyengar, S. Han, and S. Gupta, "Do friends influence purchases in a social network?" *Harvard Business School Marketing Unit Working Paper*, no. 09-123, 2009.

[17] " How Social Media Impacts Consumer Buying - Forbes," https://www.forbes.com/sites/forbesagencycouncil/2022/04/28/how-social-media-impacts-consumer-buying/.

[18] Y. Zhuang and O. Yağan, "Multistage complex contagions in random multiplex networks," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 1, pp. 410–421, 2019.

[19] A. Fazeli, A. Ajorlou, and A. Jadbabaie, "Competitive diffusion in social networks: Quality or seeding?" *IEEE Transactions on Control of Network Systems*, vol. 4, no. 3, pp. 665–675, 2016.

[20] A. Jadbabaie and A. Kakhbod, "Optimal contracting in networks," *Journal of Economic Theory*, vol. 183, pp. 1094–1153, 2019.

[21] M. Diamanti, P. Charatsaris, E. E. Tsiropoulou, and S. Papavassiliou, "Incentive mechanism and resource allocation for edge-fog networks driven by multi-dimensional contract and game theories," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 435–452, 2022.

[22] Y. Liu, M. Tian, Y. Chen, Z. Xiong, C. Leung, and C. Miao, "A contract theory based incentive mechanism for federated learning," in *Federated and Transfer Learning*. Springer, 2022, pp. 117–137.

[23] M. Wu, D. Ye, J. Ding, Y. Guo, R. Yu, and M. Pan, "Incentivizing differentially private federated learning: A multidimensional contract approach," *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10 639–10 651, 2021.

[24] T. Yang, X. Yi, J. Wu, Y. Yuan, D. Wu, Z. Meng, Y. Hong, H. Wang, Z. Lin, and K. H. Johansson, "A survey of distributed optimization," *Annual Reviews in Control*, vol. 47, pp. 278–305, 2019.

[25] H. Yu, E. Wei, and R. A. Berry, "A business model analysis of mobile data rewards," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019, pp. 2098–2106.

[26] J. Nie, J. Luo, Z. Xiong, D. Niyato, and P. Wang, "A stackelberg game approach toward socially-aware incentive mechanisms for mobile crowdsensing," *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 724–738, 2018.

[27] O. Candogan, K. Bimpikis, and A. Ozdaglar, "Optimal pricing in networks with externalities," *Operations Research*, vol. 60, no. 4, pp. 883–905, 2012.

[28] Y. Zhang, L. Song, W. Saad, Z. Dawy, and Z. Han, "Contract-based incentive mechanisms for device-to-device communications in cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2144–2155, 2015.

[29] F. Bloch and N. Quérou, "Pricing in social networks," *Games and economic behavior*, vol. 80, pp. 243–261, 2013.

[30] P. Bianchi and J. Jakubowicz, "Convergence of a multi-agent projected stochastic gradient algorithm for non-convex optimization," *IEEE transactions on automatic control*, vol. 58, no. 2, pp. 391–405, 2012.

[31] M. Montazeri, H. Kebriaei, B. N. Araabi, and D. Niyato, "Optimal mechanism design in the sponsored content service market," *IEEE Communications Letters*, vol. 25, no. 9, pp. 3051–3054, 2021.

Alireza Baneshi received his B.S. degree in Electrical Engineering at Shiraz University of Technology, Shiraz, Iran, 2019, and his M.S. degree in Electrical Engineering from the University of Tehran, Tehran, Iran, in 2023. His research interests lie in game theory, mechanism design, contract theory, and distributed optimization.

Mina Montazeri received her Ph.D. degree in control systems from the University of Tehran, Iran, in 2023. Currently, she holds the position of Post-Doctoral Researcher at the Urban Energy Systems Laboratory, Empa, Dübendorf, Switzerland. Her research interests include reinforcement learning, mechanism design, data-driven control, optimization and smart grid.

Hamed Kebriaei (Senior Member, IEEE) is an Associate Professor of Control Systems at the School of Electrical and Computer Engineering, University of Tehran. His research interests include optimization and learning in control systems, game theory, distributed optimization, and multi-agent reinforcement learning. He was honored with the Outstanding Reviewer Award from IEEE Transactions on Cybernetics in 2022 and the Outstanding Young Researcher Award from the University of Tehran in 2023. He served as a Guest Editor for IEEE Control Systems Letters in 2023. He is also an Associate Editor for EUCA Conference Editorial Board (EUCA-CEB) 2025.