SAM: Cache Space Allocation in Collaborative Edge-caching Network

Qinghua Ding*, Haitian Pang*, Lifeng Sun*

* National Laboratory for Information Science and Technology
Tsinghua University, Beijing, China
Email: {dqh14,pht14,sunlf}@mails.tsinghua.edu.cn

***Abstract*—Edge caching technique is becoming more and more important in 5G network architecture, which shows great performance in terms of mitigating the backbone network congestion and reducing the delivery latency. The key idea of this technique is to cache popular contents in the small base stations(SBSs) which locate at the edge of the wireless network, and deliver these contents to nearby mobile users(MUs) from the cache directly. The collaborative in-cluster prefetching is improving the network performance even more. In this paper, we study the case where one monopolistic state-owned mobile network operator (MNO) offers edge caching service to multiple Content Providers(CPs), and focus on the mechanism design for the cache space allocation in this market. As the caching policy on each edge node is coupled with each other due to in-cluster collaboration, one time auction or flat pricing will both harm the social welfare. We propose a 2-optimal sequential auction mechanism (SAM) to solve this problem, which improves the social welfare by 20%∼70% compared to the traditional Vickery-Clarke-Groves(VCG) mechanism.**

***Index Terms*—Edge Caching, Collaborative Caching, Auction Mechanism.**

I. Introduction

Mobile data offloading has increased 16-fold since 2011, according to the technical report of Cisco [1], and 4G data accounts for 69% of all mobile traffic in 2016. However, the traditional wireless network architecture is inappropriate in dealing with the massive content requests, and many new techniques are developed to solve this problem. One promising technique adopted by the 5G is edge caching – caching popular contents at the edge of the wireless network to improve the data delivery performance [2]. In practice, the small base stations (SBSs) are usually enabled to cache contents, and once the associated MUs' requested content is stored in the cache, the content will be delivered from the cache directly to the user. When the user request does not hit the cache, the content will still be delivered to the user from the core network, which may incur long transmission latency.

Another important technique for improving the performance of this caching system is the in-cluster collaboration [3]. The SBSs within the same cluster will collaborate with each other through wired connection, thus the missed requested of some SBS can also be served via other SBSs through well-designed routing. This technique then further promotes the efficiency of content caching.

However, along with the collaborative caching technique come the challenges in the field of economics, engineering and politics. In our work, we focus on the mechanism design in a 5G -based edge caching market, where one monopolistic state-owned MNO sells cache spaces to multiple myopic CPs. The state-owned MNO is supposed to focus on the maximization of the social welfare rather than private profits. But neither flat pricing scheme nor one-time auction is beneficial to the social welfare due to the in-cluster collaboration. Thus in our study, we focus on the mechanism design in such edge caching market and study the following questions systematically.

- How should the MNO allocate the cache spaces and what pricing strategy should the MNO take? How efficienct could this mechanism be?
- How should MNO design the mechanism fitting the collaborative caching system and thus improve the overall performance?

In the perspective of economics, resource allocation in the edge caching market has some intrinsic difficulties. Firstly, the CP's evaluation of cache spaces on different nodes is coupled with each other due to the collaborative fetching scheme, thus the values of caching the same content on different nodes is non-additive. Specifically, the mechanisms that determine the allocation at one time could cause reduction to the social welfare because of disregarding the collaboration between nodes.

Secondly, the economic robustness of this mechanism should be well guaranteed. Thus the individual rationality and incentive compatibility is satisfied that everyone is willing to participate in this market and truthfully reports the demand. Moreover, the mechanism should be relatively high in efficiency, thus the social welfare needs to be guaranteed.

In this paper, we design a sequential auction mechanism(SAM) for this collaborative caching allocation problem. The auction mechanism processes in a sequential way, thus we decouple the correlation between the values of the allocated items and the unallocated items in this way. We further prove that a weakly dominant strategy exists for the myopic bidders so that they'll report their bid truthfully. Finally, we prove that our mechanism is 2-optimal thus provides good guarantee for the social welfare.

Our main contributions are as follows.

978-1-4673-8999-0/17/$31.00

- *A sequential auction mechanism(SAM) for the edge caching market* We developed an economic model for the single-MNO multi-CP edge caching market, and we propose a sequential auction mechanism(SAM) for cache space allocation of the MNO. The proposed mechanism takes collaborative caching into consideration and the decouples the correlation of values for different cache spaces in a progressive way.
- *A 2-optimal guarantee on the social welfare* We prove the economic robustness and effiiency of our proposed mechanism. We demonstrate that the SAM satisfies IR, IC and provide a 50% bound on the social welfare. Through simulation, we show that the SAM improves the social welfare by 20%∼70% than the traditional VCG mechanism on average.

The remainder of this paper is organized as follows. In Section II, we refer to related works in the fields of edge caching and mechanism design. In Section III, we present the economic model of the edge caching network. In Sec. IV, we show the sequential auction mechanism for the cache space allocation. In Sec. V, we study the performance of our mechanism via simulation. Finally, we conclude our work in Sec. VI.

II. Related Works

Edge caching is an original idea in 5G wireless network. A number of previous works have considered the feasibility of wireless caching [4]–[7]. Under the wireless caching network architecture, many content caching and network routing methods are proposed [8]. And researchers in [9] provided a comprehensive study of the wireless caching network performance.

Mechanism design in edge caching network and other similar networks architectures has been a hot topic these years. A double auction mechanism is put forward in [10] for the cache space allocation between the MNO and AP owners. Michele Mangilli et al, developed a greedy auction mechanism in [11] which attains relatively high social welfare. In [12], the author uses the trade reduction approach to design an auction mechanism for the allocation problem. A CP differentiation mechanism is proposed in [13] using Stackelberg game for economics modeling and stochastic geometry for network modeling. While in [14], a contract-based mechanism is proposed for the caching trade market. Pang et al. [15] proposed a framework to jointly allocate the resource in edge caching and data sponsor.

As we consider resource allocation in a collaborative edge caching context, the previously designed flat pricing mechanisms or one-time auction mechanisms do not capture the features of this network architecture, and will cause reduction in the social welfare. [16] investigated the D2D network via auction theory. Inspired by the combinatorial auction method put forward by Benny Lehmann in 2002 [17], we developed a sequential auction mechanism to allocate the cache spaces on different nodes, which is more adaptive to this network architecture.

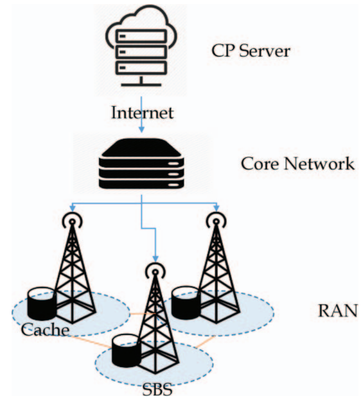

Fig. 1: Architecture of Collaborative Edge Caching Network.

III. System Model

A. Network Model

In our model, a monopolistic MNO owns a set of SBSs, $\mathcal{S} = \{1, 2, ..., S\}$. As widely used in caching techniques, we suppose that the contents are chunked to some unit size c_0, and we further suppose the capacity of jth SBS to be c_j. We suppose that each MU is associated to some SBS, and the MU's request will be responded directly if the content is cached in that SBS, causing relatively short delay in delivery. Otherwise, the content can be collaboratively fetched from peer nodes if the content is cached in other SBSs, but with a relatively longer delay. If the content is not cached in the SBSs, a backhaul process will be called, which also means an even longer delay.

However, one of the main purposes of the edge caching network is to shorten the content delivery latency. Other benefits of edge caching include mitigated bandwidth and improved service quality. But in fact, the most important benefit that the CPs are concerned with is the shortened delivery latency. We formulate the delay for the request of the k^{th} content of the i^{th} CP received by the j^{th} node as below.

$$d_{ij}^{(k)} = \begin{cases} t_0 & \text{If } x_{ij}^{(k)} = 1, \\ t_1 & \text{If } x_{ij}^{(k)} = 0, \text{ and } S_i^{(k)} = \Phi \\ t_{jj'} & \text{Else, here } j^{'} = \arg \min\limits_{s \in \mathcal{S}_i^{(k)}} t_{js} \end{cases}$$

Here $x_{ij}^{(k)}$ indicates whether the kth content of CP i is cached in node j or note. In the formulation above, the latency for the respond is discussed in three cases. If $x_{ij}^{(k)} = 1$, that is, the content is directly stored in the SBS cache that received the request, then it is delivered from the cache with a latency t_0. If the content is not stored in node j or any peer node, then it must be retrieved through the backhaul with a latency t_1. When the content is not stored in node i and $S_i^{(k)} \neq \Phi$, the content can be collaboratively retrieved from the peer node $j^{'}$ with a latency $t_{jj'}$, and usually it is retrieved from the node that minimises the retrieval time. Here $S_i^{(k)}$ represents the set of nodes that caches the k^{th} content of CP i. Usually, we have $t_0 \leq t_{jj'} \leq t_1$ since direct response is usually faster than collaborative fetching and even faster than the backhaul.

B. Economic Languages

Suppose that in our network, there are N CPs, $\mathcal{N} = \{1, 2, ..., N\}$. The CP i has F_i contents for delivery, $\mathcal{F}_i = \{1, 2, ..., F_i\}$. And these CPs compete with each other to lease the cache space of the MNO maximising their own utilities. Here we introduce the caching profile based on the network model constructed above.

***Definition** 1 (caching profile):* The caching profile $\boldsymbol{x}$ is a three-dimensional array, with each dimension indicating the number of CP, the number of node and the number of content. Thus the elements $x_{ij}^{(k)} \in \{0, 1\}$ of $\boldsymbol{x}$ indicates whether to cache the k^{th} content of CP i on node j or not.

Here we denote $\boldsymbol{y_i}$ as the set of caches allocated to CP i.

$$\boldsymbol{y_i} = \{x_{ij}^{(k)} | x_{ij}^{(k)} = 1\}$$

Thus the caches in $\boldsymbol{y_i}$ will shorten the content delivery latency of CP i. And we suppose the CP's value of set $\boldsymbol{y_i}$ to be the reduction of retrieval time due to these caches. Since we have the total latency for CP i given certain cache profile $\boldsymbol{y_i}$ as

$$D_i(\boldsymbol{y_i}) = \sum_{j=1}^{S} \sum_{k=1}^{F_i} n_{ij}^{(k)} d_{ij}^{(k)}$$

Here $n_{ij}^{(k)}$ is the expected hit rate at node j for the k^{th} content. Then we can express the *value function* of CP i as

$$v_i(\boldsymbol{y_i}) = D_i(\Phi) - D_i(\boldsymbol{y_i})$$

In fact, this is a function over the set of all feasible cache decisions. And the value of different cache decisions is non-additive. The definition of *additivity of value function* is shown below.

***Definition** 2 (Additivity of value function):* The value function over a set of items $\boldsymbol{T}$ is additive if for any subset $\boldsymbol{U} \in \boldsymbol{T}$, the following proposition is satisfied

$$v(\boldsymbol{U}) = \sum_{u \in \boldsymbol{U}} v(u)$$

.

Here is a simple example of the non-additive value function of the CPs. Suppose there are two nodes for caching $S = 1, 2$, each with a capacity of an unit cache. Then for some CP who has one content to cache which has unit visit on each node, the total delay without caching is $2t_1$. Then caching one content on either node will reduce the overall delivery time to $t_0 + t_{12}$. Thus the value of either cache decision is $v(x_{11}^{(1)}) = v(x_{12}^{(1)}) = 2t_1 - (t_0 + t_{12})$. However, the value of caching on both nodes is $v(x_{11}^{(1)}, x_{12}^{(1)}) = 2t_1 - 2t_0$. It is clear that $v(x_{11}^{(1)}, x_{12}^{(1)}) - (v(x_{11}^{(1)}) + v(x_{12}^{(1)})) = 2t_{12} - 2t_1 \leq 0$. Thus $v(x_{11}^{(1)}, x_{12}^{(1)}) \leq v(x_{11}^{(1)}) + v(x_{12}^{(1)})$ and the additivity does not hold. Through this example, we can also find out that the value of caching decision on different nodes is coupled.

Denote CP i's payment as p_i, we can express its *utility* as

$$U_{CP_i} = \gamma \cdot v_i(\boldsymbol{y_i}) - p_i$$

Here γ is a scaler indicating the CP's evaluation of the delivery delay. Without loss of generality, we suppose $\gamma = 1$ in the following discussion. Suppose that the cost of crowdsourcing APs is C^{MNO}, which is a constant value in our problem. Thus the utility of the MNO is

$$U_{MNO} = \sum_{i=1}^{N} p_i - C^{MNO}$$

Based on the discussion above, we have the *social welfare* as the sum of the CPs' and the MNO's utilities.

$$U_{SW} = \sum_{i=1}^{N} v_i(\boldsymbol{y_i}) - C^{MNO}$$

C. Problem Formulation

In the social-caring MNO's perspective, the auction mechanism needs to attain high social welfare. We can formalise the social welfare maximisation(SWM) problem[1] as follows.

$$
\begin{aligned}
\max \quad & \sum_{i=1}^{N} v_i(\boldsymbol{y_i}) && \text{(SWM)} \\
\text{s.t.} \quad & \sum_{i=1}^{N} \sum_{k=1}^{F_i} x_{ij}^{(k)} \leq c_j, \quad \forall j \in \mathcal{S} && (1) \\
& \boldsymbol{y_i} = \{x_{ij}^{(k)} | x_{ij}^{(k)} = 1\}, \quad \forall i \in \mathcal{N} && (2) \\
& x_{ij}^{(k)} \in \{0, 1\}, \quad \forall i, j, k \in \mathcal{N}, \mathcal{S}, \mathcal{F}_i \\
& IR, IC
\end{aligned}
$$

This is an optimisation problem which is NP-hard according to Benny Lehmann's work [17]. Directly solving this problem could lead to exponential time of calculation. And in reality, the MNO and CPs are selfish utility maximisers having different optimisation goals. Thus we need to design a mechanism that will drive the system to approximate the maximum social welfare point. To solve this problem, we devised a Sequential Auction Mechanism(SAM), which solves this problem in polynomial time and still preserve the economic property of individual rational(IR), incentive compatible(IC) and 2-optimal efficiency.

IV. Sequential Auction Mechanism

A. Sequential Auction Mechanism

In this section, we propose a sequential auction mechanism(SAM) to achieve high efficiency by decoupling the correlated values of caches due to collaboration. By allocating the cache spaces in a sequential way, the CPs can determine the value of unallocated caches based on the already allocated ones. Thus the CP's evaluation of the caches is more accurate due to the extra information derived from previously revealed allocations.

In the sequential auction mechanism(SAM), the auctioneer will run an sub-auction for each cache space sequentially. When allocating some cache space, the auctioneer will collect

[1]The constant C^{MNO} is omitted in the optimisation problem.

Algorithm 1 Sequential Auction Mechanism

procedure SAM
    (i) Initiation
    $x_{ij}^{(k)} \leftarrow \Phi, \forall i, j, k \in \mathcal{N}, \mathcal{S}, \mathcal{F}_i$
    $\boldsymbol{y_i} \leftarrow \Phi$
    $p_i \leftarrow 0, \forall i \in \mathcal{N}$
    (ii) Resource Allocation
    for all $j \in \mathcal{S}$ **do**
        for all $s = 1$ to $\frac{c_j}{c_0}$ **do**
            $\boldsymbol{b_i} \leftarrow$ CP i's bid for a unit space on node j
            $i \leftarrow \arg\max_{s \in \mathcal{N}} b_s$
            $i^{'} \leftarrow \arg\max_{s \in \mathcal{N} | \{i\}} b_s$
            $x_{ij}^{(k_i)} \leftarrow 1$
            $\boldsymbol{y_i} \leftarrow \boldsymbol{y_i} \cup \{x_{ij}^{(k)}\}$
            $p_i \leftarrow p_i + r_{i'}$
        end for
    end for
end procedure

each CP's bid for this cache space. Here the bid of CP i $\boldsymbol{b_i}$ consists of two components, $\boldsymbol{b_i} = (r_i, k_i)$. r_i is the price he is willing to pay for the cache space, and k_i is the rank of the content he is willing to cache for this space. After receiving the bids, the auctioneer will allocate this cache space to the bidder i who offers the highest price and will add the second highest price to his total payment p_i. The VCG-based pricing scheme is used for keeping the truthfulness of the bidders.

Now we consider the optimal bid for CP i when bidding for some cache space on node j. In reality, the already revealed allocation will help determine the content to cache this turn. In order to maximise his profit, CP i will cache the most valuable content that maximises his utility given $\boldsymbol{y_i}$, which is the caches already allocated for CP i. Thus we have the determination of k_i as

$$k_i = \arg \max_{k \in \mathcal{F}_i, x_{ij}^{(k)} = 0} v_i(\boldsymbol{y_i} \cup \{x_{ij}^{(k)}\})$$

Hence we can determine the CP i's value of this single cache space as

$$\Delta v_i = v_i(\boldsymbol{y_i} \cup \{x_{ij}^{(k_i)}\}) - v_i(\boldsymbol{y_i})$$

Δv_i can be regarded as the marginal value of storing content k_i on this node for given allocated caches, $\boldsymbol{y_i}$. We will prove that the CP will remain truthful and bid as their value in the next section, which means $r_i = \Delta v_i$.

***Theorem** 1:* The complexity of SAM is $\mathcal{O}(|\mathcal{S}| \cdot n \log(|\mathcal{N}|))$.

Proof: First above, the double loops in SAM contributes a factor of $\mathcal{O}(|\mathcal{S}| \cdot n)$ to the overall complexity. Then inside the loop, finding the maximum requires complexity of $\mathcal{O}(\log(|\mathcal{N}|))$. Thus the overall complexity is $\mathcal{O}(|\mathcal{S}| \cdot n \cdot \log(|\mathcal{N}|))$ ∎

B. Proofs of Properties of SAM

Here we prove the economic properties of the SAM. We'll demonstrate that the SAM is individual rational, incentive compatible and 2-optimal. Before demonstration, let us consider the bidders' bidding strategies first. When determining their bids, it's sure that they will make use of the information of the already allocated items. And here we further suppose that the bidders are myopic, thus won't take future profits into consideration.

***Assumption** 1:* The bidders are myopic, and will use the information of revealed allocation rather than taking the unallocated cache spaces into consideration.

In fact, this assumption of myopic bidders is relatively reasonable. For one thing, in this incomplete information game, each bidder lacks the information of other bidders' values. Moreover, take unallocated items into consideration when determining one's strategy will definitely incur more uncertainty and risks. Due to the lack of information and uncertainty, it can be hard for CPs to manipulate the game considering future profits. Thus this assumption is basically held during the actual auction process.

***Theorem** 2:* The SAM is individual rational(IR).

Proof: At the beginning of the auction, every participants' utility is zero, $u_i = 0, \forall i \in \mathcal{N}$. Then the auction of caches is promoted in a sequential way. Suppose at the turn of allocating some cache on node j, the already allocated cache space for certain bidder i is $\boldsymbol{y_i^{'}}$, and the already determined payment of this bidder is $p_i^{'}$. Thus the total payoff of this bidder is $u_i^{'} = v_i(\boldsymbol{y_i^{'}}) - p_i^{'}$ up to now.

After this cache space is allocated, the payoff of the winner bidder should be $u_i = v_i(\boldsymbol{y_i^{'}} \cup \{x_{ij}^{k_i}\}) - (p_i^{'} + r_j)$, here we suppose that bidder j is the second highest bidder. Then the increment of bidder i's payoff this turn is $Deltau_i = u_i - u_i^{'} = v_i(\boldsymbol{y_i^{'}} \cup \{x_{ij}^{k_i}\}) - v_i(\boldsymbol{y_i^{'}}) - r_j = r_i - r_j \geq 0$. Here the second equality is ensured by the truthfulness of bidders. Thus the increment of the winner bidder's payoff is non-negative. And the increment of the loser bidders' payoff is zero. Thus the sub-auction this turn won't introduce negative increment to any CPs' utility.

Considering zero initial payoff and non-negative increment each turn, the payoff of each bidder in the end should also be non-negative. Then the IR property holds. ∎

***Theorem** 3:* The SAM is incentive compatible(IC).

Proof: Given the myopic bidder assumption, the bidders won't take future steps into consideration in the allocation turn of some unit space on node j. Thus for CP i, he will guarantee the maximisation of his payoff in this single turn based on the already allocated items $\boldsymbol{y_i^{'}}$. If all bidders bid truthfully, then the increment of the winner bidder i's payoff is $\Delta u_i = r_i - r_j$ after this turn. Here bidder j is the second highest bidder of all.

However, if winner bidder i bid untruthfully, we consider the following two cases. If he bid less than v_j, he will lose the sub-auction and get no profit increment. If he bid an untruthful bid $b_i^{'} > r_j$, then the increment of his payoff of bidding untruthfully is $\Delta u_i^{'} = r_i - r_j = \Delta u_i$, which stays unchanged. Thus he won't gain more if he bid untruthfully.

Then for a loser bidder k, if he bid less than r_i, he still won't change the outcome. If he bid more than r_i, then his increment of payoff is $\Delta u_k^{'} = r_k - r_i \leq 0$, thus the utility increment will be negative. On the whole, the bidders cannot get more utility increment if he bids untruthfully in some turn, they will report their true marginal value. And the IC property holds. ∎

To prove the efficiency lower bound of 50%, we first introduce the notion of submodular value function as below.

***Definition 3** (submodular value function):* A value function $v(\boldsymbol{U})$ over a set of items $\boldsymbol{T}$ is submodular if and only if the following proposition holds.

$$\forall \boldsymbol{U} \subset \boldsymbol{T}, x, y \in \boldsymbol{T} \text{ and } x, y \notin \boldsymbol{U}, v(x|\boldsymbol{U}) \geq v(x|\boldsymbol{U} \cup \{y\})$$

Here $v(x|\boldsymbol{U}) = v(\boldsymbol{U} \cup \{x\}) - v(\boldsymbol{U})$ is the marginal value of caching content x given the already allocated set $\boldsymbol{U}$.

We can prove that the CP's value function is submodular. An extension of the illustration in Section III will give some insight into this property. The value of the fist node given set $\boldsymbol{U} = \{2\}$ is $v(x_{11}^{(1)}|\{x_{12}^{(1)}\}) = v(\{x_{11}^{(1)}, x_{12}^{(1)}\}) - v(\{2\}) = t_{12} - t_0$. However, the value of the first node given $\boldsymbol{U} = \Phi$ is $v(x_{11}^{(1)}|\Phi) = v(\{x_{11}^{(1)}\}) - v(\Phi) = 2t_1 - (t_0 + t_{12})$. Then $v(x_{11}^{(1)}|\Phi) \geq v(x_{11}^{(1)}|\{x_{12}^{(1)}\})$ because $\Delta v = 2t_1 - 2t_{12} \geq 0$. Thus the submodularity holds in this case. Here we demonstrate the submodularity of the CP's value function as below.

Theorem 4: The CP's value function is submodular.

Proof: The users' value function is determined previously as $v_i(\boldsymbol{y_i}) = D_i(\Phi) - D_i(\boldsymbol{y_i})$, thus we have that for any single cache item $x_{im}^{(p)}$ and $x_{in}^{(q)}$, the marginal value of $\boldsymbol{x_{im}^{(p)}}$ given $\boldsymbol{y_i}$ can be represented as $v_i(x_{im}^{(p)}|\boldsymbol{y_i}) = v_i(\boldsymbol{y_i} \cup \{x_{im}^{(p)}\}) - v_i(\boldsymbol{y_i}) = D_i(\boldsymbol{y_i}) - D_i(\boldsymbol{y_i} \cup \{x_{im}^{(p)}\})$. Similarly, $v_i(x_{im}^{(p)}|\boldsymbol{y_i} \cup x_{in}^{(q)}) = D_i(\boldsymbol{y_i} \cup \{x_{in}^{(q)}\}) - D_i(\boldsymbol{y_i} \cup \{x_{in}^{q}, x_{im}^{(p)}\})$. Then we have $\Delta v_i = v_i(x_{im}^{(p)}|\boldsymbol{y_i}) - v_i(x_{im}^{(p)}|\boldsymbol{y_i} \cup x_{in}^{(q)}) = D_i(\boldsymbol{y_i}) + D_i(\boldsymbol{y_i} \cup \{x_{im}^{(p)}, x_{in}^{(q)}\}) - D_i(\boldsymbol{y_i} \cup \{x_{im}^{(p)}\}) - D_i(\boldsymbol{y_i} \cup \{x_{in}^{(q)}\})$. Through some transformation, we get $\Delta v_i = v_i(\{x_{im}^{(p)}, x_{in}^{(q)}\}|\boldsymbol{y_i}) - v_i(x_{im}^{(p)}|\boldsymbol{y_i}) - v_i(x_{in}^{(q)}|\boldsymbol{y_i})$. In order to determine the sign of Δv_i, we discuss the following two cases.

Case I. $p = q$.

For $v_i(x_{im}^{(p)}|\boldsymbol{y_i}) = v_i(\boldsymbol{y_i} \cup \{x_{im}^{(p)}\}) - v_i(\boldsymbol{y_i}) = D_i(\boldsymbol{y_i}) - D_i(\boldsymbol{y_i} \cup \{x_{im}^{(p)}\})$, we have $D_i(\boldsymbol{y_i})$ as $D_i(\boldsymbol{y_i}) = \sum_{j=1}^{S}\sum_{k=1}^{F_i} n_{ij}^{(k)} d_{ij}^{(k)}$. The change of total delay after adding $x_{im}^{(p)}$ into set $\boldsymbol{y_i}$ is due to the shortened delivery time of content p on some nodes $\boldsymbol{S_p}$. Thus we have $v_i(x_{im}^{(p)}|\boldsymbol{y_i}) = \sum_{j \in \boldsymbol{S_p}} n_{ij}^{(p)}(d_{ij}^{(p)} - t_{jm})$, here $\boldsymbol{S_p} = \{j \in \mathcal{S}|t_{jm} < d_{ij}^{(p)}\}$. Similarly, $v_i(x_{in}^{(q)}|\boldsymbol{y_i}) = \sum_{j \in \boldsymbol{S_q}} n_{ij}^{(q)}(d_{ij}^{(q)} - t_{jn})$, with $\boldsymbol{S_q} = \{j \in \mathcal{S}|t_{jm} < d_{ij}^{(q)}\}$. Further we have $v_i(\{x_{im}^{(p)}, x_{in}^{(q)}\}|\boldsymbol{y_i}) = \sum_{j \in \boldsymbol{S_p^{'}}} n_{ij}^{(p)}(d_{ij}^{(p)} - t_{jm}) + \sum_{j \in \boldsymbol{S_q^{'}}} n_{ij}^{(q)}(d_{ij}^{(q)} - t_{jn})$, and $\boldsymbol{S_p^{'}} = \{j \in \mathcal{S}|t_{jm} < d_{ij}^{(p)}, t_{jm} < t_{jn}\}$, $\boldsymbol{S_q^{'}} = \{j \in \mathcal{S}|t_{jn} < d_{ij}^{(q)}, t_{jn} < t_{jm}\}$. Then $\Delta v_i = \sum_{j \in \boldsymbol{S_p^*}} n_{ij}^{(p)}(d_{ij}^{(p)} - t_{jm}) + \sum_{j \in \boldsymbol{S_q^*}} n_{ij}^{(q)}(d_{ij}^{(q)} - t_{jn}) \geq 0$, here $\boldsymbol{S_p^*} = \{j \in \mathcal{S}|t_{jm} < d_{ij}^{(p)}, t_{jm} > t_{jn}\}$, $\boldsymbol{S_q^*} = \{j \in \mathcal{S}|t_{jn} < d_{ij}^{(q)}, t_{jn} > t_{jm}\}$. So $\Delta v_i \geq 0$ in this case.

Case II. $p \neq q$.

In this case, the expression of $v_i(x_{im}^{(p)}|\boldsymbol{y_i})$ and $v_i(x_{in}^{(q)}|\boldsymbol{y_i})$ stay unchanged while $v_i(\{x_{im}^{(p)}, x_{in}^{(q)}\}|\boldsymbol{y_i}) = \sum_{j \in \boldsymbol{S_p}} n_{ij}^{(p)}(d_{ij}^{(p)} - t_{jm}) + \sum_{j \in \boldsymbol{S_q}} n_{ij}^{(q)}(d_{ij}^{(q)} - t_{jn})$. Then we have $\Delta v_i = 0$ in this case.

With consideration over the above two cases, $\Delta v_i \geq 0$ always exists. Thus the value function is submodular. ∎

The submodularity assembles the convexity property in continuous functions and it will lead to some good economic properties. Here we prove that the submodular value function of CPs will ensure the 50% lower bound on efficiency of this mechanism as below.

Theorem 5: The mechanism will achieve at least 50% of the social efficiency.

Proof: Let Q be the original problem and $Q^{'}$ be the subproblem of allocating the caches on the nodes excluding one cache on the first node. Suppose the first cache space is assigned to bidder m for storing content p, then his value function in the subproblem is refreshed as $v_m^{'}(\boldsymbol{y_m}) = v_i(\boldsymbol{y_m} \cup \{x_{m1}^{(p)}\}) - v_i(\{x_{m1}^{(p)}\})$, with $\boldsymbol{y_m} \subset \boldsymbol{A_m}|\{x_{m1}^{(p)}\}$, here $\boldsymbol{A_m} = \{x_{mj}^{(k)}|\forall j \in \mathcal{S}, \forall k \in \mathcal{F}_i\}$. And others' value functions won't change in the subproblem. Let us denote by ALG the total value achieved by this algorithm, and $p = v_i(\{x_{m1}^{(p)}\})$. Then we have $ALG(Q) = ALG(Q^{'}) + p$. Moreover, we denote OPT the optimal value of social efficiency, and $\boldsymbol{y}^* = \{\boldsymbol{y_1^*}, \boldsymbol{y_2^*}, ..., \boldsymbol{y_N^*}\}$ the optimal allocation profile. Suppose in the optimal allocation, there exists one cache on the first node that is allocated to n to store content q, Then $\boldsymbol{y}^{'} = \{\boldsymbol{y_1^*}, ..., \boldsymbol{y_n}|\{x_{n1}^{(q)}\}, ..., \boldsymbol{y_N^*}\}$ can be an allocation for the subproblem. which achieves lower social welfare than the optimal social welfare of the subproblem. Then the value achieved by $S^{'}$ is only different from S^{opt} in that bidder m has different value function and bidder n has different allocation. Thus the value loss due to bidder n's different allocation is at most $v_j(\{x_{n1}^{(q)}\})$ because the value function is submodular. What's more, the value loss due to bidder n's change of value function is $v_n(\boldsymbol{y_n^*}) - v_n^{'}(\boldsymbol{y_n^*}) = v_i(\boldsymbol{y_i^*}) - v_i(\boldsymbol{y_i^*} \cup \{x_{n1}^{(p)}\}) + v_i(\{x_{n1}^{(p)}\}) \leq v_i(\{x_{n1}^{(p)}\}) = p$. Thus the total gap between the optimal value of subproblem and the whole problem is at most $2p$. That is, $OPT(Q) - OPT(Q^{'}) \leq 2p$. Thus we have $OPT(Q) \leq OPT(Q^{'}) + 2p \leq 2ALG(Q^{'}) + 2p \leq 2ALG(Q)$. Thus the efficiency of this mechanism is at least 50%. ∎

V. Simulation

A. Experiment Setup

We study the performance of this mechanism in the basic settings of $N = 10, S = 10, C = 5, F = 10$, here $C = \frac{c_j}{c_0}, \forall j \in \mathcal{S}$. And we use the traditional VCG mechanism as our baseline. Since the caches on certain node are homogeneous, the allocation of the caches on one node can be solved by a multi-unit VCG auction. And by running a multi-unit VCG auction for each node simultaneously, the VCG mechanism

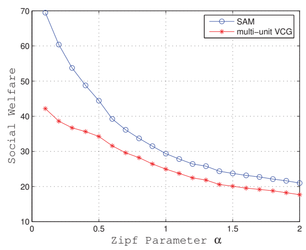

Fig. 2: Social welfare with different α

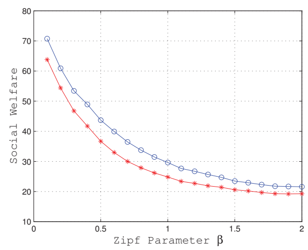

Fig. 3: Social welfare with different β

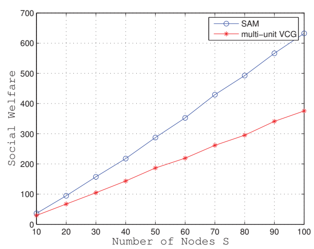

Fig. 4: Social welfare under different S

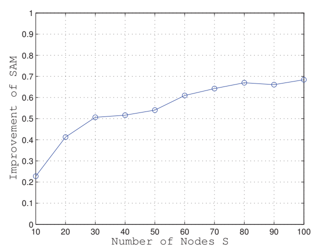

Fig. 5: Improvement under different S

can solve this allocation problem. The improvement of social welfare over the VCG mechanism is an important index for evaluation of the performance of SAM. We further suppose the popularity distribution over CPs and contents follows the Zipf's law with different Zipf parameter α and β.

B. Discussion over N and S

In order to determine the performance of our mechanism, we value the improvement of our mechanism with different numbers of CPs and cache nodes. In this experiment, we set the Zipf parameters as $\alpha = \beta = 0.7$. We observed that the performance of SAM is relatively stable with respect to the change of numbers of CPs N. This could be illustrated as the market is mainly determined by several most important buyers, and the increasing the number of less powerful CPs won't influence the market that much.

However, we observe that with the increment of the number of nodes S will enlarge the performance gap. When the node number is relatively large, the improvement attains nearly 70%. And within a wide range of node numbers, our mechanism improves the social welfare by more than 50% compared with VCG. In fact, the increasing node number will add to the weight of inter-node connections, thus the VCG which concentrates on utility maximization of single node will lead to worse overall performance, while the SAM focusing on collaborative caching performs much better. The experiments show that our mechanism improves the VCG mechanism by 20%-70% in a wide range of parameter setting over α,N and S.

VI. Conclusion

In this paper, we developed a framework for studying the one-ISP multi-CP cache space allocation problem in the edge-caching network. We propose a sequential auction mechanism which takes the collaboration between nodes into consideration, and our mechanism achieves relatively high social efficiency. The mechanism has good economic robustness and guarantees a 50% social welfare lower bound. Through simulation, our mechanism improves 20-70% of social welfare than the VCG mechanism. Future works are needed to study the CPs, ISPs and APs as an ecosystem, studying their behaviors in a larger background with more complex interactions among these three parts.

References

[1] Cisco Visual Networking Index. Cisco visual networking index: global mobile data traffic forecast update, 2014–2019. *Tech. Rep*, 2015.

[2] Xiaofei Wang, Min Chen, Tarik Taleb, Adlen Ksentini, and Victor Leung. Cache in the air: exploiting content caching and delivery techniques for 5g systems. *IEEE Communications Magazine*, 52(2):131–139, 2014.

[3] Amitabha Ghosh, Nitin Mangalvedhe, Rapeepat Ratasuk, Bishwarup Mondal, Mark Cudak, Eugene Visotsky, Timothy A Thomas, Jeffrey G Andrews, Ping Xia, Han Shin Jo, et al. Heterogeneous cellular networks: From theory to practice. *IEEE Communications Magazine*, 50(6), 2012.

[4] Franck Le, Mudhakar Srivatsa, and Arun K Iyengar. Byte caching in wireless networks. In *Distributed Computing Systems (ICDCS), 2012 IEEE 32nd International Conference on*, pages 265–274. IEEE, 2012.

[5] Karthikeyan Shanmugam, Negin Golrezaei, Alexandros G Dimakis, Andreas F Molisch, and Giuseppe Caire. Femtocaching: Wireless content delivery through distributed caching helpers. *IEEE Transactions on Information Theory*, 59(12):8402–8413, 2013.

[6] Hasti Ahlehagh and Sujit Dey. Video-aware scheduling and caching in the radio access network. *IEEE/ACM Transactions on Networking (TON)*, 22(5):1444–1462, 2014.

[7] Jingjing Wang, Chunxiao Jiang, Zhu Han, Yong Ren, and Lajos Hanzo. Network association strategies for an energy harvesting aided super-wifi network relying on measured solar activity. *IEEE Journal on Selected Areas in Communications*, 34(12):3785–3797, 2016.

[8] Ejder Bastug, Jean-Louis Guénégo, and Mérouane Debbah. Proactive small cell networks. In *Telecommunications (ICT), 2013 20th International Conference on*, pages 1–5. IEEE, 2013.

[9] Ejder BaştuÇğ, Mehdi Bennis, Marios Kountouris, and Mérouane Debbah. Cache-enabled small cell networks: Modeling and tradeoffs. *EURASIP Journal on Wireless Communications and Networking*, 2015(1):1–11, 2015.

[10] George Iosifidis, Lin Gao, Jianwei Huang, and Leandros Tassiulas. A double-auction mechanism for mobile data-offloading markets. *IEEE/ACM Transactions on Networking (TON)*, 23(5):1634–1647, 2015.

[11] Michele Mangili, Fabio Martignon, Stefano Paris, and Antonio Capone. Bandwidth and cache leasing in wireless information centric networks: a game theoretic study. *IEEE Transactions on Vehicular Technology*, 2016.

[12] Yifeng Zhong, Ke Xu, Xiang-Yang Li, Hui Su, and Qingyang Xiao. Estra: Incentivizing storage trading for edge caching in mobile content delivery. In *Global Communications Conference (GLOBECOM), 2015 IEEE*, pages 1–6. IEEE, 2015.

[13] Jun Li, He Chen, Youjia Chen, Zihuai Lin, Branka Vucetic, and Lajos Hanzo. Pricing and resource allocation via game theory for a small-cell video caching system. *IEEE Journal on Selected Areas in Communications*, 34(8):2115–2129, 2016.

[14] Kenza Hamidouche, Walid Saad, and Mérouane Debbah. Breaking the economic barrier of caching in cellular networks: Incentives and contracts. *arXiv preprint arXiv:1607.03683*, 2016.

[15] Haitian Pang, Lin Gao, and Lifeng Sun. Joint optimization of data sponsoring and edge caching for mobile video delivery. *arXiv preprint arXiv:1611.00211*, 2016.

[16] Jingjing Wang, Chunxiao Jiang, Zhi Bie, Tony QS Quek, and Yong Ren. Mobile data transactions in device-to-device communication networks: Pricing and auction. *IEEE Wireless Communications Letters*, 5(3):300–303, 2016.

[17] Benny Lehmann, Daniel Lehmann, and Noam Nisan. Combinatorial auctions with decreasing marginal utilities. In *Proceedings of the 3rd ACM conference on Electronic Commerce*, pages 18–28. ACM, 2001.