

Selfish Caching Games on Directed Graphs

Qian Ma, *Member, IEEE*, Edmund Yeh, *Senior Member, IEEE*, and Jianwei Huang, *Fellow, IEEE*

Abstract—Caching networks can reduce the routing costs of accessing contents by caching contents closer to users. However, cache nodes may belong to different entities and behave selfishly to maximize their own benefits, which often lead to performance degradation for the overall network. While there has been extensive literature on allocating contents to caches to maximize the social welfare, the analysis of selfish caching behaviors remains largely unexplored. In this paper, we model the selfish behaviors of cache nodes as selfish caching games on arbitrary directed graphs with heterogeneous content popularity. We study the existence of a pure strategy Nash equilibrium (PSNE) in selfish caching games, and analyze its efficiency in terms of social welfare. We show that a PSNE does not always exist in arbitrary-topology caching networks. However, if the network does not have a mixed request loop, i.e., a directed loop in which each edge is traversed by at least one content request, we show that a PSNE always exists and can be found in polynomial time. Furthermore, we can avoid mixed request loops by properly choosing request forwarding paths. We then show that the efficiency of Nash equilibria, captured by the price of anarchy (PoA), can be arbitrarily poor if we allow arbitrary content request patterns, and adding extra cache nodes can make the PoA worse, i.e., cache paradox happens. However, when cache nodes have homogeneous request patterns, we show that the PoA is bounded even allowing arbitrary topologies. We further analyze the selfish caching games for cache nodes with limited computational capabilities, and show that an approximate PSNE exists with bounded PoA in certain cases of interest. Simulation results show that increasing the cache capacity in the network improves the efficiency of Nash equilibria, while adding extra cache nodes can degrade the efficiency of Nash equilibria.

Index Terms—Caching networks, selfish caching games, Nash equilibrium, price of anarchy.

I. INTRODUCTION

CACHING networks can reduce the routing costs for accessing contents by caching the requested contents as close to the requesting users as possible. Prevailing caching networks include content delivery networks (CDN) [2, 3], information-centric networks (ICN) [4], femtocell networks [5], web caching networks [6], and peer-to-peer networks [7]. There has been extensive previous work (e.g., [8–10]) on how to optimally allocate contents to available caches.

Qian Ma is with the School of Intelligent Systems Engineering, Sun Yat-sen University. E-mail: maqian25@mail.sysu.edu.cn.

Edmund Yeh is with the Department of Electrical and Computer Engineering, Northeastern University, Boston, USA 02215. E-mail: eyeh@ece.neu.edu.

Jianwei Huang is with the Shenzhen Institute of Artificial Intelligence and Robotics for Society, the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, Shenzhen 518172, China. E-mail: jianweihuang@cuhk.edu.cn. (*Corresponding author: Jianwei Huang.*)

This work is supported by the National Natural Science Foundation of China under Grant 62002399, the Shenzhen Institute of Artificial Intelligence and Robotics for Society, the Presidential Fund from the Chinese University of Hong Kong, Shenzhen, the National Science Foundation grant CNS-1718355 and a grant from Intel Corp.

Part of this work has been published at the Twentieth ACM International Symposium on Mobile Ad Hoc Networking and Computing (Mobihoc '19), July 2–5, 2019, Catania, Italy [1].

However, most existing work assumes that cache nodes are altruistic and cooperate with each other to optimize an overall network performance objective.

In practice, cache nodes may belong to different entities [11]. For example, in wireless community mesh networks such as Google WiFi [12] and Guifi [13], individual users contribute their wireless routers (as caches) to the community. On the Internet, different operators and providers deploy their own caching infrastructures and services. Examples include AT&T Content Delivery Network Service, Google Global Cache, Netflix Open Connect, and Akamai.¹

In caching networks where different entities operate their own caches, cache nodes may behave selfishly to maximize their own benefits. For example, in multi-hop wireless community mesh networks [14], a cache node has an incentive to cache the content items to minimize its own routing cost, which may not always maximize the social welfare. This motivates us to study the selfish caching behaviors through a game-theoretic approach.

To our best knowledge, this is the first paper to examine *selfish caching games* on arbitrary directed graphs with heterogeneous content popularity. We focus on the pure strategy Nash equilibrium (PSNE),² and address two fundamental questions. *First, is a PSNE guaranteed to exist in any selfish caching game? Second, if a PSNE exists, does it have a guaranteed efficiency in terms of social welfare?* The short answers to the above two questions are “No” and “No”. In other words, the selfish caching game does not always admit a PSNE. Even if a PSNE exist, its efficiency in terms of social welfare can be very poor.

In this paper, we characterize the conditions under which (i) a PSNE exists, and (ii) a PSNE has a guaranteed efficiency. We characterize the efficiency of PSNE by the *price of anarchy* (PoA), which is the ratio of the social welfare achieved by the worst PSNE to that achieved by a socially optimal strategy [16, 17]. The analysis of PSNE and PoA takes into account the asymmetric and node-specific interdependencies among cache nodes, which reflect the network topology and content request patterns. Our analysis will help the network designer understand when the network behaves with certain performance guarantees, and how to create these conditions in the network.

We analyze the selfish caching game in two scenarios. We first consider a scenario where all contents have equal sizes, which corresponds to practical applications such as video-on-demand services using harmonic broadcasting that

¹In this paper, we treat one provider as one cache node. Another example is a caching network where each provider owns multiple cache nodes. In such a case, we can model the interactions among multiple cache nodes belonging to the same provider as a coalitional game.

²The main reason for implementing PSNE in practice is simplicity [15].

divide each video into segments of equal size [18]. We then consider a scenario where contents have unequal sizes, which corresponds to practical applications such as video streaming services over HTTP (e.g., Netflix and Hulu) that split each video into segments of lengths from 2 to 10 seconds [19].

Our primary contributions are:

- *Selfish Caching Game*: To the best of our knowledge, this is the *first* work that studies the selfish caching game on directed graphs with arbitrary topologies and heterogeneous content popularity.
- *Pure Strategy Nash Equilibrium (PSNE)*: For selfish caching games with equal-sized content items, we first show that a PSNE does not always exist. We then show that a PSNE exists if the network does not have a mixed request loop, i.e., a directed loop in which each edge is traversed by at least one content request. Furthermore, we propose a polynomial-time algorithm to find a PSNE for the selfish caching game with no mixed request loop.
- *Price of Anarchy*: We show that the PoA in general can be arbitrarily poor if we allow arbitrary content request patterns. Furthermore, adding extra cache nodes can make the PoA worse, a phenomenon which we call the *cache paradox*. However, when cache nodes have homogeneous request patterns, we show that the selfish caching game is an α -scalable valid utility game and the PoA is bounded in arbitrary-topology caching networks.
- *Approximate PSNE*: For selfish caching games with unequal-sized content items, each node's payoff maximization problem is NP-hard. When cache nodes have limited computational capability, we show that their selfish caching behaviors lead to an approximate PSNE with bounded PoA in certain cases of interest.

The rest of the paper is organized as follows. In Section II, we review related literature. In Section III, we introduce our system model. In Section IV, we model the selfish caching game and analyze the PSNE. In Section V, we study the PoA. In Section VI, we analyze selfish caching games with unequal-sized content items. In Section VII, we provide simulation results. We conclude in Section VIII.

II. RELATED WORK

There has been a rich body of previous work on caching, many of which are summarized in an excellent recent survey [20]. In the following, we introduce related work regarding caching optimization and selfish caching game, respectively.

Caching Optimization. There is considerable recent literature on a variety of caching optimization problems, including proactive caching [21, 22], optimal caching under queuing models [23, 24], optimal caching under unknown content popularities [25, 26], distributed adaptive algorithms for optimal caching [8–10, 27], caching at the edges [28–32], TTL (time-to-live) caches [33, 34], optimal caching in evolving networks [35], joint caching and routing optimization [36–38], optimal cache partitioning [39], and collaborative caching [40–45]. All the above work assumes that all cache nodes aim to maximize the social welfare.

Selfish Caching Game. There are several papers which study selfish caching behaviors in simple settings. In [46],

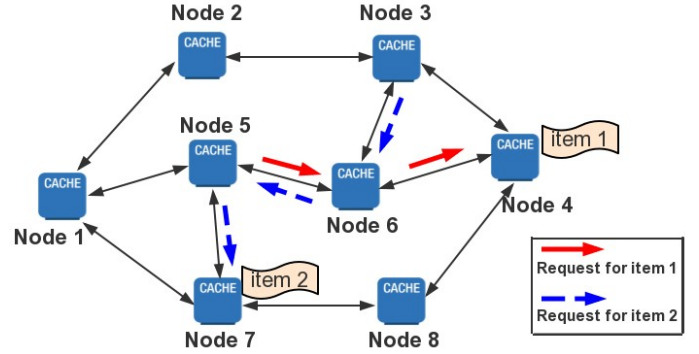


Figure 1. A caching network with $|V| = 8$ nodes and $|\mathcal{I}| = 2$ content items, where node 4 (node 7, respectively) is the designated server of item 1 (item 2, respectively). The request forwarding paths are fixed in our model. For example, the path of node 5 requesting item 1 is $p^{(5,1)} = (5, 6, 4)$, and the path of node 3 requesting item 2 is $p^{(3,2)} = (3, 6, 5, 7)$.

Chun *et al.* study the selfish caching game on undirected graphs with a single content item, assuming homogeneous content popularity across users. In [47], Goemans *et al.* study the content market sharing game, where users get rewards for caching content items. The paper assumes that any node which caches a requested item can serve the request with same cost, without considering network topology. The authors in [6] and [48] study a distributed selfish replication game in an undirected complete graph, where the distance between any two nodes is the same. In [49], Gopalakrishnan *et al.* study the capacitated selfish replication game in an undirected network, where users are equally interested in a set of content items.

The analysis in the above literature is applicable to undirected graphs, and some are restricted to homogeneous content popularity. In this work, we study the selfish caching game on directed graphs with arbitrary topologies and heterogeneous content popularity.

III. SYSTEM MODEL

We consider a network of selfish caches, represented by a directed caching graph $G(V, E)$ with an *arbitrary topology*, where V is the set of cache nodes and E is the set of bidirectional edges which enable ARQ with asymmetric edge costs (see an example in Figure 1). Each cache node requests one or more content items (e.g., movies) from the set $\mathcal{I} = \{1, \dots, |\mathcal{I}|\}$. For each content item $i \in \mathcal{I}$, there is a fixed set of *designated server nodes* $\mathcal{D}^i \subseteq V$, $|\mathcal{D}^i| > 0$, that store i in their permanent storage (outside of their caches).³ We consider equal-sized content items in Sections III–V, which correspond to applications such as video-on-demand services using harmonic broadcasting that divide each video into segments of equal size [18].⁴ We will consider the case of unequal-sized items in Section VI.

³For example, designated server nodes can be content providers' caches.

⁴Without loss of generality, we normalize the size of each item to be one.

A. Caching Strategies

Each node $s \in V$ has a cache of capacity $c_s \in \mathbb{N}$, i.e., node s can store exactly c_s equal-sized content items. We denote the caching strategy of node $s \in V$ by $\mathbf{x}_s = \{x_{si} : \forall i \in \mathcal{I}\} \in \{0, 1\}^{|\mathcal{I}|}$, where

$$x_{si} \in \{0, 1\}, \text{ for all } i \in \mathcal{I},$$

indicates whether node s stores content item i , and satisfies

$$\sum_{i \in \mathcal{I}} x_{si} \leq c_s, \text{ for all } s \in V.$$

We let $\mathbf{x}_{-s} = \{\mathbf{x}_1, \dots, \mathbf{x}_{s-1}, \mathbf{x}_{s+1}, \dots, \mathbf{x}_{|V|}\}$ denote the caching strategy of nodes other than node s , and let $\mathbf{x} = \{\mathbf{x}_s, \mathbf{x}_{-s}\}$ denote the global caching strategy. Given \mathbf{x}_s , we let $Z_s = \{i : x_{si} = 1, i \in \mathcal{I}\}$ denote the set of items cached by node $s \in V$.

B. Content Requests

We describe each content request by a pair (s, i) , where the request source⁵ $s \in V$ requests content item $i \in \mathcal{I}$. We assume that each request (s, i) arrives according to a stationary ergodic process [50, 51] with arrival rate $\lambda_{(s,i)} \geq 0$ for all $s \in V$ and $i \in \mathcal{I}$, which reflects *heterogeneous content popularity* across items and request nodes.⁶

Request (s, i) is forwarded over a pre-determined fixed *request forwarding path*⁷ $p^{(s,i)}$, from request source s to one of content item i 's designated server nodes in \mathcal{D}^i . Specifically, the path $p^{(s,i)}$ of length $K \leq |V|$ is a sequence (p_1, \dots, p_K) of nodes $p_k \in V$ such that $p_1 = s$, $p_K \in \mathcal{D}^i$, and $(p_k, p_{k+1}) \in E$ for all $k \in \{1, \dots, K-1\}$. We require that $p^{(s,i)}$ contains no loops ($p_k \neq p_l$ for all $1 \leq k < l \leq K$) and no node other than the terminal node on $p^{(s,i)}$ is a designated server for content item i ($p_k \notin \mathcal{D}^i$ for all $1 \leq k < K$). For request (s, i) , we let $V_{(s,i)} = \{v : v \in p^{(s,i)}, v \neq s, v \notin \mathcal{D}^i\}$ denote the set of intermediate nodes on path $p^{(s,i)}$. We denote $V_s = \cup_{i \in \mathcal{I}} V_{(s,i)}$ as the set of intermediate nodes on all the request forwarding paths of node s .⁸

Request (s, i) travels along path $p^{(s,i)}$ until either (i) the request reaches a node $v \in p^{(s,i)}$ such that node v caches content item i , i.e., $x_{vi} = 1$ or, (ii) if $x_{vi} = 0$ for all $v \in p^{(s,i)} \setminus \{p_K\}$, the request reaches $p_K \in \mathcal{D}^i$. Having found the closest copy of content item i , the network generates a *response message* carrying the requested content item i . The response message is propagated in the reverse direction along

⁵We consider a request source to be a point of aggregation which combines many network users. While a single user may request a given content item only once over a time period, an aggregation point is likely to submit many requests for a given content item over a time period.

⁶We consider selfish caching behaviors under complete information, where cache nodes know all other nodes' content request patterns [11, 46, 47]. Specifically, cache nodes can estimate content request patterns through historical information or long-term learning [21].

⁷Similar as in the named data networks, we assume that the request forwarding path is determined in a longer timescale compared with caching. And we consider selfish caching behaviors under complete information where cache nodes know the request forwarding paths [11, 46, 47].

⁸Note that each cache node can play some or all of the following roles: a designated server of content items, a source of requests, and an intermediate node on request forwarding paths.

Table I
KEY NOTATION

$G(V, E)$	Caching graph, with nodes in V and edges in E
c_s	Cache capacity of node $s \in V$
w_{uv}	Cost on edge $(u, v) \in E$
\mathcal{I}	Set of content items
\mathcal{D}^i	Set of designated servers for content item $i \in \mathcal{I}$
(s, i)	Request for item i from node s
$\lambda_{(s,i)}$	Arrival rate of request (s, i)
$p^{(s,i)}$	Request forwarding path of request (s, i)
$V_{(s,i)}$	The set of intermediate nodes on path $p^{(s,i)}$
V_s	The set of intermediate nodes on node s ' paths
x_{si}	Caching strategy of node $s \in V$ for item $i \in \mathcal{I}$
\mathbf{x}_s	Caching strategy of node $s \in V$
Z_s	The set of content items cached by node $s \in V$
\mathbf{x}	Global caching strategy of all nodes
$h_{(s,i)}$	The routing cost to serve request (s, i)
h_s	The routing cost of node s
g_s	The caching gain of node s
G	The aggregate caching gain in the network

the request forwarding path, i.e., from the closest node with content item i back to the request source node s .⁹

C. Routing Costs

Transferring a content item across edge $e = (u, v) \in E$ incurs a cost (e.g., delay or financial expense) denoted by $w_{uv} \geq 0$.¹⁰ Since the size of each request message is relatively small compared with the content item, we assume that costs are only due to content item transfers, and the costs of forwarding requests are negligible [9]. To serve the request (s, i) , the routing cost depends on the caching decision x_{si} of the request source node s , as well as the caching decisions $x_{vi}, \forall v \in V_{(s,i)}$, of all the intermediate nodes on the request forwarding path $p^{(s,i)}$. Specifically, the routing cost of transferring item i over the reverse direction of $p^{(s,i)}$ is

$$\begin{aligned} & h_{(s,i)}(x_{si}, \{x_{vi} : v \in V_{(s,i)}\}) \\ &= \sum_{k=1}^{|p^{(s,i)}|-1} w_{p_{k+1}p_k} \prod_{k'=1}^k (1 - x_{p_{k'}i}) \\ &= \sum_{k=1}^{|p^{(s,i)}|-1} w_{p_{k+1}p_k} (1 - x_{si}) \prod_{k'=2}^k (1 - x_{p_{k'}i}). \end{aligned}$$

Note that $h_{(s,i)}(\cdot)$ includes the cost on edge (p_{k+1}, p_k) , i.e., $w_{p_{k+1}p_k}$, if and only if none of the nodes from p_1 to p_k on path $p^{(s,i)}$ has cached content item i . For example, in Figure 1, $p^{(3,2)} = (3, 6, 5, 7)$ and the routing cost of request $(3, 2)$ depends on x_{32} , x_{62} and x_{52} . If $(x_{32}, x_{62}, x_{52}) = (0, 0, 1)$, then $h_{(3,2)}(x_{32}, x_{62}, x_{52}) = w_{63} + w_{56}$.

D. Selfish Caching Behavior

Each selfish cache node $s \in V$ seeks a caching strategy to optimize its own benefit, i.e., minimizing the aggregate

⁹In this paper, we assume that forwarding and transmission follow standard network protocols. In some settings, forwarding and transmission incur a service cost to the cache node due to the consumption of the transmit power and communication resource. We will consider such costs in the future work.

¹⁰We do not model the congestion effect on each edge. How to jointly consider cost and throughput issues is an interesting open problem.

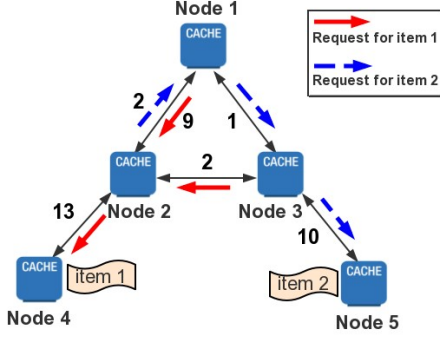


Figure 2. An example where PSNE does not exist. The caching network has $|V| = 5$ nodes and $|\mathcal{I}| = 2$ content items, where node 4 (node 5, respectively) is the designated server of item 1 (item 2, respectively). The cache capacity is 1 at each node. The request arrival rates satisfy $\lambda_{(v,i)} = \lambda_i, \forall v \in V, i \in \mathcal{I}$, where $\lambda_1 = 10$ and $\lambda_2 = 14$. The request forwarding paths are fixed, for example, $p^{(3,1)} = (3, 2, 4)$.

expected cost for serving all its own requests, calculated as follows:

$$h_s(\mathbf{x}_s, \{\mathbf{x}_v : v \in V_s\}) = \sum_{i \in \mathcal{I}} \lambda_{(s,i)} \cdot h_{(s,i)}(\mathbf{x}_{si}, \{\mathbf{x}_{vi} : v \in V_{(s,i)}\}). \quad (1)$$

For notation simplicity, we write $h_s(\cdot)$ as $h_s(\mathbf{x}_s, \mathbf{x}_{-s})$. In the absence of caching, i.e., $\mathbf{x} = \mathbf{0}$, the aggregate expected cost of node s is:

$$h_s(\mathbf{0}) = \sum_{i \in \mathcal{I}} \lambda_{(s,i)} \sum_{k=1}^{|p^{(s,i)}|-1} w_{p_{k+1}p_k}.$$

We define the *caching gain* of node s as

$$g_s(\mathbf{x}_s, \mathbf{x}_{-s}) = h_s(\mathbf{0}) - h_s(\mathbf{x}_s, \mathbf{x}_{-s}). \quad (2)$$

Intuitively, the caching gain is the cost reduction enabled by caching. Since $h_s(\mathbf{0})$ is a constant, minimizing the aggregate expected cost in (1) is equivalent to maximizing the caching gain in (2). Hence, the caching gain in (2) serves as node s 's payoff function.

IV. SELFISH CACHING GAME

In this section, we model the interactions among selfish cache nodes by a selfish caching game on directed graphs. We construct an example where the pure strategy Nash equilibrium (PSNE) does not exist for such a game. We then identify the condition under which a PSNE exists, and propose a polynomial-time algorithm to find a PSNE under the condition.

A. Game Modeling

We define the selfish caching game as follows:

Game 1 (Selfish Caching Game on Directed Graphs).

- *Players*: the set V of cache nodes on the caching graph;
- *Strategies*: the caching strategy $\mathbf{x}_s = \{x_{si} : \forall i \in \mathcal{I}\}$ for each cache node $s \in V$, where $x_{si} \in \{0, 1\}$ and $\sum_{i \in \mathcal{I}} x_{si} \leq c_s$;
- *Payoffs*: the caching gain $g_s(\mathbf{x}_s, \mathbf{x}_{-s})$ for each $s \in V$.

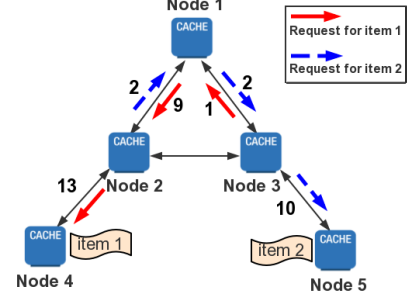


Figure 3. An example where there is no mixed request loop. The request forwarding path is $p^{(3,1)} = (3, 1, 2, 4)$.

Since the selfish caching game is a finite game, there exists at least one mixed strategy Nash equilibrium (including pure strategy Nash equilibrium as a special case). However, since it is difficult to implement random caching strategies in practical caching networks, we focus on analyzing pure strategy Nash equilibria in this paper, as defined below.

Definition 1 (Pure Strategy Nash Equilibrium). A *pure strategy Nash equilibrium* of the selfish caching game is a caching strategy profile \mathbf{x}^{NE} such that for every cache node $s \in V$,

$$g_s(\mathbf{x}_s^{\text{NE}}, \mathbf{x}_{-s}^{\text{NE}}) \geq g_s(\mathbf{x}_s, \mathbf{x}_{-s}^{\text{NE}}), \text{ for all feasible } \mathbf{x}_s. \quad (3)$$

B. An Example with No PSNE

In the following, we first show that the PSNE does not always exist.

Theorem 1. *There exists a selfish caching game for which the pure strategy Nash equilibrium does not exist.*

Proof. Figure 2 is an example with no PSNE. For node 4 (node 5, respectively), caching item 2 (item 1, respectively) is its dominant strategy. Now we analyze the selfish behaviors of nodes 1, 2, and 3. It is easy to verify that for all 8 feasible caching strategy profiles, there always exists one cache node that can improve its caching gain by changing its caching strategy unilaterally. For example, if all the three nodes cache item 1, then node 3 has the incentive to cache item 2 to improve its caching gain assuming that the other two nodes do no change their caching strategies. Hence there is no strategy profile where everyone is achieving its maximum payoff assuming other nodes do not change their strategies. Hence, the PSNE does not exist. \square

C. Existence of a PSNE

Deciding the existence of a PSNE for games on graphs is NP-hard in general [52]. However, we identify the condition under which a PSNE of the selfish caching game exists and can be found in polynomial time. To proceed, we first introduce the definition below.

Definition 2 (Mixed Request Loop). A *mixed request loop* on a directed graph is a directed loop $(p_1, p_2, \dots, p_K, p_{K+1} =$

p_1) involving $3 \leq K \leq |V|$ nodes, where $p_k \in V$ for $1 \leq k \leq K$, $p_k \neq p_l$ for all $1 \leq k < l \leq K$, and at least one content request traverses edge $(p_k, p_{k+1}) \in E$ for all $1 \leq k \leq K$.

In Figure 2, $(1, 3, 2, 1)$ forms a mixed request loop, where requests for item 1 traverse edge $(3, 2)$, and requests for item 2 traverse edges $(2, 1)$ and $(1, 3)$.

Note that a loop on graph is not always a mixed request loop. For example, $(1, 3, 2, 1)$ in Figure 3 is a loop. However, the request forwarding path is $p^{(3,1)} = (3, 1, 2, 4)$ rather than $(3, 2, 4)$, meaning no request traverses edge $(3, 2)$. Hence, loop $(1, 3, 2, 1)$ is not a mixed request loop. In other words, we can avoid mixed request loops by properly choosing the request forwarding paths.

Next, we will show that a PSNE exists in the selfish caching game on caching graphs with no mixed request loop.¹¹

Theorem 2. *A PSNE always exists in the selfish caching game on caching graphs with no mixed request loop.*

Proof. See Appendix A. \square

Theorem 2 holds in caching networks with arbitrary topologies and heterogeneous content popularity. We prove the existence¹² of the PSNE by finding a PSNE in polynomial time.

D. Polynomial-Time Algorithm to Find a PSNE

In this section, we present a polynomial-time algorithm to find a PSNE for the selfish caching game. Specifically, for each selfish caching game, we can define a *state graph* [47] as follows.

Recall that given node s ' caching strategy x_s , the set $Z_s = \{i : x_{si} = 1, i \in \mathcal{I}\}$ is the set of content items cached by node $s \in V$. Hence we can use $x = \{x_s : \forall s \in V\}$ and $Z = \{Z_s : \forall s \in V\}$ interchangeably to represent the caching strategy profile.

Definition 3 (State Graph [47]). *A state graph is a directed graph where each vertex corresponds to a strategy profile Z . There is a directed arc from vertex Z to vertex Z' with label v if the only difference between Z and Z' is the strategy of player v and the payoff of player v in Z is strictly less than its payoff in Z' .*

A PSNE corresponds to a vertex on the state graph without any outgoing arc, i.e., a sink. Hence identifying a PSNE of the selfish caching game is equivalent to identifying a sink on the corresponding state graph.

We propose a polynomial-time algorithm (Algorithm 1 [47]) to find a sink on the state graph. The algorithm proceeds in rounds. The first round starts at the vertex $Z = \emptyset$, corresponding to the strategy profile where none of the cache nodes cache any content item (Line 1 of Algorithm 1). In each round, the first arc traversed on the state graph corresponds to an *add arc* where a player, say s , changes from Z_s

Algorithm 1: Find PSNE on State Graph [47]

Input: $G(V, E), \mathcal{I}, w_{uv}, \forall (u, v) \in E, \lambda_{(s,i)}$ and $p^{(s,i)}$, for all $s \in V, i \in \mathcal{I}$
Output: Z^{NE}

```

1 Set  $Z = \emptyset$ ;
2 repeat
3   Randomly pick a node  $s \in V$  where  $|Z_s| < c_s$ ;
4   Add item  $i^*$  where
      $i^* \in \arg \max_{i \in \mathcal{I} \setminus Z_s} g_s(Z_s \cup \{i\}, Z_{-s})$  to node  $s$ , i.e.,
      $Z_s \leftarrow Z_s \cup \{i^*\}$ ;
5   while  $\exists v \in V, j \notin Z_v, t \in Z_v$ , such that
      $g_v(Z_v \cup \{j\} \setminus \{t\}, Z_{-v}) > g_v(Z_v, Z_{-v})$  do
6     Set  $Z_v \leftarrow Z_v \cup \{j\} \setminus \{t\}$ ;
7   end
8 until  $\forall s \in V$  satisfies  $|Z_s| = c_s$ ;
9 Set  $Z^{\text{NE}} = Z$ ;
```

to $Z_s \cup \{i^*\}$. Intuitively, player s adds only one content item i^* to its cache, where we select i^* among all possible content items not currently in Z_s to maximize player s ' caching gain (Lines 3-4 of Algorithm 1). After the first arc, subsequent arcs in the same round correspond to *change arcs*. Specifically, a change arc corresponds to a player, say v , replacing Z_v by $Z_v \cup \{j\} \setminus \{t\}$, where $j \notin Z_v$ and $t \in Z_v$. Intuitively, player v replaces content item t for content item j if $g_v(Z_v \cup \{j\} \setminus \{t\}, Z_{-v}) > g_v(Z_v, Z_{-v})$ (Lines 5-7 of Algorithm 1). When the current vertex on the state graph has no change arcs, one round ends. For the vertex where a round ends, if there is an add arc outgoing from it, a new round starts; otherwise, it is a sink and the algorithm terminates. Such a sink corresponds to the PSNE.

In the following theorem, we show that Algorithm 1 can find a sink on the state graph in polynomial time.

Theorem 3. *For the selfish caching game on arbitrary-topology caching graphs with no mixed request loop, Algorithm 1 computes a PSNE in polynomial time by traversing a path of length at most $|V||\mathcal{I}|^2(|V| - 2)^2$ on the corresponding state graph.*

Proof. See Appendix B. \square

Note that for any given selfish caching game, Algorithm 1 does not require the construction of the whole state graph. At any given vertex of the state graph, Algorithm 1 only requires one to find the next arc to traverse, which takes $\mathcal{O}(|V|)$ time. Hence, the total maximum running time of Algorithm 1 is $\mathcal{O}(|V|^2|\mathcal{I}|^2(|V| - 2)^2)$. Furthermore, different random choices of the next arc to traverse in Algorithm 1 correspond to different outcomes if there is more than one PSNE in the selfish caching game.

Since each cache node maximizes its own benefit, a PSNE of the selfish caching game does not in general optimize the social welfare. We will quantify the efficiency of the Nash equilibria in terms of social welfare next.

V. PRICE OF ANARCHY

To evaluate the efficiency of Nash equilibria, we analyze the price of anarchy (PoA) [16], i.e., the ratio of the social welfare

¹¹Note that no mixed request loop is a sufficient (but not necessary) condition for a PSNE to exist.

¹²The selfish caching game generally admits multiple PSNEs, depending on system parameters such as edge weights and request arrival rates.

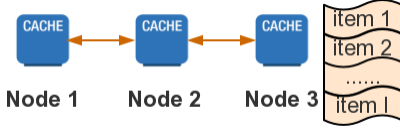


Figure 4. An example where PoA approaches 0. Consider a caching network with $|V| = 3$ nodes where node 3 is the designated server of content items in set $\mathcal{I} = \{1, 2, \dots, I\}$. The request arrival rates at node 1 satisfy $\lambda_{(1,1)} > 0$, $\lambda_{(1,i)} = \lambda_{(1,1)} - \epsilon > 0$, where $\epsilon > 0$, for $i \in I \setminus \{1\}$. Node 2 does not generate request, i.e., $\lambda_{(2,i)} = 0, \forall i \in I$. The cache capacities are $c_1 = 1$ and $c_2 = I - 1$.

achieved by the worst Nash equilibrium to that achieved by a socially optimal strategy.

In this paper, we define the social welfare as the aggregate caching gain in the network. Specifically, the social welfare maximization problem is

$$\begin{aligned} \max \quad & G(\mathbf{x}) \triangleq \sum_{s \in V} g_s(\mathbf{x}_s, \mathbf{x}_{-s}) \\ \text{s.t.} \quad & \sum_{i \in \mathcal{I}} x_{si} \leq c_s, \quad x_{si} \in \{0, 1\}, \quad \forall s \in V, i \in \mathcal{I}. \end{aligned} \quad (4)$$

Problem (4) is NP-hard [9]. It is challenging to calculate the socially optimal solution and analyze the PoA in general.

In the following, we first show that the PoA can be arbitrarily poor if we allow any content request patterns. We then identify the cache paradox where adding extra cache nodes can make the PoA worse. Under reasonable constraints of request patterns and paths, however, we can show that the PoA is bounded in general caching networks. Furthermore, for given caching networks with known network topology and parameters, we can derive a better bound for PoA.

A. An Example with an Arbitrarily Poor PoA

Next, we show that the PoA can be arbitrarily close to 0, indicating that the selfish caching behaviors can lead to unboundedly poor performance in terms of social welfare.

Lemma 1. *There exists a selfish caching game for which the PoA is arbitrarily close to 0.*

Proof. We construct an example where PoA approaches 0, as shown in Figure 4. In this example, the socially optimal caching strategy is for node 1 to cache content item 1 and for node 2 to cache content items 2 to I . The optimal social welfare, i.e., aggregate caching gain, is¹³

$$G^{\text{SO}} = \lambda_{(1,1)}(w_{21} + w_{32}) + \sum_{i=2}^I \lambda_{(1,i)} w_{32}.$$

There may exist more than one PSNE. One is that node 1 caches item 1 and node 2 caches none of the content items (since node 2 has no request of its own). The social welfare achieved by this PSNE is $G^{\text{NE}} = \lambda_{(1,1)}(w_{21} + w_{32})$. We have

$$\frac{G^{\text{NE}}}{G^{\text{SO}}} = \frac{1}{1 + \sum_{i=2}^I \frac{\lambda_{(1,i)} w_{32}}{\lambda_{(1,1)}(w_{21} + w_{32})}}.$$

When $w_{32} \gg w_{21}$ and $\epsilon \rightarrow 0$, we have $\frac{\lambda_{(1,i)} w_{32}}{\lambda_{(1,1)}(w_{21} + w_{32})} \rightarrow 1$ and $\frac{G^{\text{NE}}}{G^{\text{SO}}} \rightarrow \frac{1}{I}$, which goes to 0 as I becomes very large. Since PoA measures the worst case ratio between any PSNE

¹³The superscript “SO” represents socially optimal.

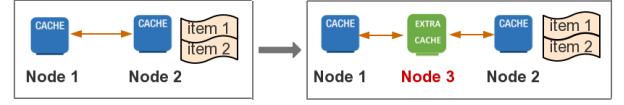


Figure 5. An example where adding an extra cache node makes the PoA worse.

and the social optimal solution, the PoA will be no larger than $G^{\text{NE}}/G^{\text{SO}}$ and hence can be arbitrarily close to 0. \square

B. Cache Paradox

In practice, one way to improve the aggregate caching gain in the network is to add extra cache nodes. However, we identify the following cache paradox.

Lemma 2. *In the selfish caching game, adding extra cache nodes can make the PoA worse.*

Proof. Consider a caching network with two nodes in Figure 5 (left subfigure), where node 2 is the designated server for two content items. Assume $c_1 = 1$ and $\lambda_{(1,1)} > \lambda_{(1,2)} > 0$. At the equilibrium, node 1 caches item 1, which is also socially optimal. Hence, $\text{PoA} = 1$.

Now we add an extra cache node, i.e., node 3 (see the right subfigure in Figure 5), where $c_3 = 1$ and $\lambda_{(3,1)} = \lambda_{(3,2)} = 0$. Assume $w_{21} = w_{31} + w_{23}$, $w_{31} > 0$, and $w_{23} > 0$. Then one equilibrium is that node 1 caches item 1 and node 3 caches nothing. However, the socially optimal strategy is that node 1 caches item 1 and node 3 caches item 2. Hence, the PoA with node 3 satisfies

$$\text{PoA}' = \frac{\lambda_{(1,1)}(w_{31} + w_{23})}{\lambda_{(1,1)}(w_{31} + w_{23}) + \lambda_{(1,2)} w_{23}} < 1.$$

Intuitively, adding node 3 does not change the social welfare achieved at the equilibrium, but increases the optimal social welfare, and hence makes the PoA worse. \square

C. Bound on PoA

In this section, we show that under reasonable constraints of request patterns and paths, the selfish caching game belongs to a class of games that we call α -scalable valid utility games, and the PoA is bounded by the length of the longest request forwarding path in the network.

Recall that $Z_s = \{i : x_{si} = 1, i \in \mathcal{I}\}$ represents the set of content items cached by node $s \in V$. For convenience, we express the caching gain of node s as $g_s(Z_s, Z_{-s})$, and the aggregate caching gain of the network as $G(Z) = \sum_{s \in V} g_s(Z_s, Z_{-s})$.

We first define the *valid utility games* (for general games not restricted to selfish caching games), introduced by Vetta in [15].

Definition 4 (Valid Utility Game [15]). *A game (with social function $\gamma(\cdot)$ and individual payoff functions $f_s(\cdot), \forall s \in V$)¹⁴*

¹⁴Note that the social function can be any objective that the network aims to optimize, and may not be the summation of individual players' payoff functions.

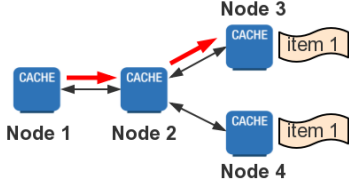


Figure 6. An example that satisfies the path overlap property. Here, $p^{(2,1)} = (2, 3)$ and $p^{(1,1)} = (1, 2, 3)$.

is a valid utility game if the following three properties are satisfied:

- 1) The social function $\gamma(\cdot)$ is non-decreasing and submodular. Mathematically, for every content item $i \in \mathcal{I}$ and for any subsets Z, Z' such that $Z \subseteq Z'$,

$$\gamma(Z) \leq \gamma(Z'), \quad (5)$$

$$\gamma(Z \cup \{i\}) - \gamma(Z) \geq \gamma(Z' \cup \{i\}) - \gamma(Z'). \quad (6)$$

- 2) The sum of players' payoff functions $f_s(\cdot)$ for any strategy profile \mathbf{x} should be no larger than the social function $\gamma(\cdot)$:

$$\sum_{s \in V} f_s(\mathbf{x}_s, \mathbf{x}_{-s}) \leq \gamma(\mathbf{x}). \quad (7)$$

- 3) The payoff of a player is no less than the difference between the social function when the player participates and that when it does not participate

$$f_s(\mathbf{x}_s, \mathbf{x}_{-s}) \geq \gamma(\mathbf{x}_s, \mathbf{x}_{-s}) - \gamma(\mathbf{0}, \mathbf{x}_{-s}). \quad (8)$$

Vetta in [15] proved that the PoA of a valid utility game is bounded by 2. In the following, we define a new class of games called α -scalable valid utility games, which generalizes the notation of valid utility games.

Definition 5 (α -Scalable Valid Utility Game). A game is an α -scalable valid utility game if it satisfies the two properties in (5), (6), and (7), and the payoff of a player is no less than the product of a positive constant α and the difference between the social function when the player participates and that when it does not participate:

$$f_s(\mathbf{x}_s, \mathbf{x}_{-s}) \geq \frac{1}{\alpha} \cdot [\gamma(\mathbf{x}_s, \mathbf{x}_{-s}) - \gamma(\mathbf{0}, \mathbf{x}_{-s})]. \quad (9)$$

Note that the valid utility game is a special case of the α -scalable valid utility games with $\alpha = 1$. We show that the selfish caching game is an α -scalable valid utility game with $\alpha = \max_{v \in V, i \in \mathcal{I}} |p^{(v,i)}| - 1$ when the following two properties are satisfied.

Definition 6 (Homogeneous Request Pattern Property). The request arrival processes for content item $i \in \mathcal{I}$ at different nodes are the same, i.e.,

$$\lambda_{(s,i)} = \lambda_i, \forall s \in V, i \in \mathcal{I}. \quad (10)$$

The homogeneous request pattern property implies that each content item has a global popularity. Note that even under the homogeneous request pattern property, the popularity of different content items can be different, i.e., $\lambda_i \neq \lambda_j, i \neq j, i, j \in \mathcal{I}$.

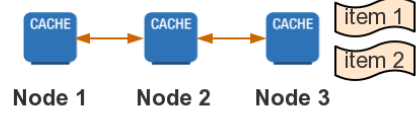


Figure 7. An example to calculate the value of $\delta(G)$. We assume $\lambda_{(v,i)} = \lambda_i, \forall v \in V, i \in \mathcal{I}$. According to (14), we have $\delta(G) = \frac{1}{1+w_{21}/w_{32}} \in [0, 1]$. When $w_{21}/w_{32} \rightarrow 0$, we have $\delta(G) \rightarrow 1$; when $w_{21}/w_{32} \rightarrow \infty$, we have $\delta(G) \rightarrow 0$.

Definition 7 (Path Overlap Property). If node s is on path $p^{(v,i)}$, then starting from node s , path $p^{(v,i)}$ overlaps with path $p^{(s,i)}$, i.e.,

$$s \in p^{(v,i)} \Rightarrow p^{(s,i)} \subseteq p^{(v,i)}. \quad (11)$$

Figure 6 shows an example that satisfies the path overlap property. Note that the path overlap property is naturally satisfied when each node chooses a unique shortest path to fetch content items.

Theorem 4. The selfish caching game with the homogeneous request pattern and path overlap properties on caching graphs with no mixed request loop is an α -scalable valid utility game where

$$\alpha = \max_{v \in V, i \in \mathcal{I}} |p^{(v,i)}| - 1. \quad (12)$$

Proof. See Appendix C. \square

In the following theorem, we show that when the selfish caching game is an α -scalable valid utility game, the PoA is bounded by the length of the longest request forwarding path in the network.

Theorem 5. When the selfish caching game is an α -scalable valid utility game, the PoA satisfies

$$PoA \geq \frac{1}{1 + \alpha} = \frac{1}{\max_{v \in V, i \in \mathcal{I}} |p^{(v,i)}|}. \quad (13)$$

Proof. See Appendix D. \square

The PoA bound decreases with α . The intuition is that as the length of the request forwarding path increases, the selfish behaviors of the intermediate nodes on a request forwarding path affect more succeeding nodes. The above performance guarantee is true for general caching networks with an arbitrary topology. However, given a caching network with a known topology and network parameters, we can further explore the network structure and derive a better bound for PoA. This is achieved by characterizing the discrete curvature of the social function, as discussed next.

D. PoA and the Discrete Curvature of the Social Function

To understand how the discrete curvature [15] of the social function will affect our PoA analysis, we first introduce the discrete derivative. For a set function $G(\cdot)$, we define the discrete derivative at Y in the direction Z as

$$G'_Z(Y) = G(Y \cup Z) - G(Y).$$

We define the *discrete curvature* of a non-decreasing, submodular social function $G(\cdot)$ to be

$$\delta(G) = \max_{\forall s \in V: G'_{Z_s}(\emptyset) > 0} \frac{G'_{Z_s}(\emptyset) - G'_{Z_s}(\mathcal{I}^{|V|} - Z_s)}{G'_{Z_s}(\emptyset)} \in [0, 1], \quad (14)$$

where $\mathcal{I}^{|V|} - Z_s$ represents the caching strategy profile (which can be infeasible) under which node s caches content items in set $\mathcal{I} \setminus Z_s$ while all other nodes cache all content items in set \mathcal{I} . Figure 7 shows an example to calculate the value of $\delta(G)$.

Given the discrete curvature of the social function, we can obtain a better bound for the PoA of the selfish caching game.

Theorem 6. *Given the discrete curvature $\delta(G)$ of the social function, for any selfish caching game with the homogeneous request pattern and path overlap properties on caching graphs with no mixed request loop, the PoA satisfies*

$$\text{PoA} \geq \frac{1}{\alpha + \delta(G)}. \quad (15)$$

Proof. See Appendix E. \square

The PoA bound decreases with $\delta(G)$. The intuition is that under a larger $\delta(G)$, the selfish behavior of a cache node has a greater impact on the achieved social welfare. Since $\delta(G) \in [0, 1]$ exploits the curvature property of the given network structure, the performance guarantee in (15) is better than the one in (13).

VI. SELFISH CACHING GAMES WITH UNEQUAL-SIZED ITEMS

In this section, we analyze more general selfish caching games with unequal-sized items, which correspond to the practical applications such as video streaming services over HTTP (e.g., Netflix and Hulu) that split each video into segments of lengths from 2 to 10 seconds [19]. We show that the caching gain maximization problem for each cache node is NP-hard. We further generalize the model by considering that each cache node has limited computational capability to solve its caching gain maximization problem. This may lead to an approximate PSNE. We analyze the existence and efficiency of an approximate PSNE under these two generalizations.

A. Game Modeling

Let L_i denote the size of content item i , for all $i \in \mathcal{I}$. We define the selfish caching game with unequal-sized items as follows:

Game 2 (Selfish Caching Game with Unequal-Sized Items).

- *Players:* the set V of cache nodes on the caching graph $G(V, E)$;
- *Strategies:* the caching strategy $\mathbf{x}_s = \{x_{si} : \forall i \in \mathcal{I}\}$ for each cache node $s \in V$, where $x_{si} \in \{0, 1\}$ and $\sum_{i \in \mathcal{I}} L_i x_{si} \leq c_s$;
- *Payoffs:* the caching gain $g_s(\mathbf{x}_s, \mathbf{x}_{-s})$ for each $s \in V$.

In the following, we will show that for each cache node $s \in V$, given fixed \mathbf{x}_{-s} , its caching gain maximization problem is

equivalent to a knapsack problem. For node $s \in V$, the caching gain in (2) can be equivalently written as

$$g_s(\mathbf{x}_s, \mathbf{x}_{-s}) = \sum_{i \in \mathcal{I}} \lambda_{(s,i)} \sum_{k=1}^{|p^{(s,i)}|-1} w_{p_{k+1}p_k} (1 - \prod_{k'=2}^k (1 - x_{p_{k'}i})) \\ + \sum_{i \in \mathcal{I}} x_{si} \cdot \lambda_{(s,i)} \sum_{k=1}^{|p^{(s,i)}|-1} w_{p_{k+1}p_k} \prod_{k'=2}^k (1 - x_{p_{k'}i}). \quad (16)$$

Given fixed \mathbf{x}_{-s} , the first term in (16) is a constant, while the second term in (16) depends on \mathbf{x}_s . Define weight

$$q_{si}(\mathbf{x}_{-s}) = \lambda_{(s,i)} \sum_{k=1}^{|p^{(s,i)}|-1} w_{p_{k+1}p_k} \prod_{k'=2}^k (1 - x_{p_{k'}i}). \quad (17)$$

Intuitively, $q_{si}(\mathbf{x}_{-s})$ represents the routing cost for request (s, i) under \mathbf{x}_{-s} if $x_{si} = 0$. Given fixed \mathbf{x}_{-s} , the caching gain maximization problem of node s is equivalent to the following knapsack problem:

$$\max_{\mathbf{x}_s} \sum_{i \in \mathcal{I}} x_{si} \cdot q_{si}(\mathbf{x}_{-s}) \quad (18) \\ \text{s.t.} \quad \sum_{i \in \mathcal{I}} L_i x_{si} \leq c_s, x_{si} \in \{0, 1\}, \forall i \in \mathcal{I}.$$

Solving the knapsack problem (18) is NP-hard.¹⁵ In practice, each cache node has limited computational capability in a short time period (e.g., minutes or hours for which the request patterns remain unchanged [25]), and can only solve the knapsack problem (18) to an approximate solution $\hat{\mathbf{x}}_s = \{\hat{x}_{si} : \forall i \in \mathcal{I}\}$. There is extensive literature on the polynomial-time approximation algorithms for the knapsack problem [53]. We present one such algorithm in Lines 4-12 of Algorithm 2, which achieves a $1/2$ approximation ratio (see Section 9.4.2 of [53]).

Now we consider the general case where cache nodes obtain only a $1/\beta$ approximate solution with $\beta > 1$ for problem (18). This leads to the β -approximate PSNE of Game 2.

B. Existence of an Approximate PSNE

A β -approximate PSNE is a strategy profile for which no player can improve its caching gain by a factor more than β of its current caching gain by unilaterally changing its strategy.¹⁶

Definition 8 (β -Approximate PSNE [55]). *A pure strategy profile $\mathbf{x}^{\beta-NE}$ is a β -approximate PSNE if no player can find an alternative pure strategy with a payoff which is more than β times its current payoff. That is for any player $s \in V$,*

$$g_s(\mathbf{x}'_s, \mathbf{x}_{-s}^{\beta-NE}) \leq \beta \cdot g_s(\mathbf{x}_s^{\beta-NE}, \mathbf{x}_{-s}^{\beta-NE}), \text{ for all feasible } \mathbf{x}'_s. \quad (19)$$

Next, we show that a β -approximate PSNE exists when the following property is satisfied:

¹⁵When $L_i = L_j, \forall i, j \in \mathcal{I}$, problem (18) is a max-weight knapsack problem, which is easy to solve and corresponds to the scenario with equal-sized items in Sections III–V.

¹⁶An alternative notion of approximate PSNE (see, e.g., [54]) is based on an additive error, rather than the multiplicative error. Our definition is equally natural, and indeed more in line with the notion of price of anarchy in game theory [15, 16, 55].

Definition 9 (Cloud Property). *All content items are stored in the same designated server node, i.e.,*

$$|\mathcal{D}^i| = 1 \text{ and } \mathcal{D}^i = \mathcal{D}^j, \forall i \neq j, i, j \in \mathcal{I}. \quad (20)$$

Furthermore, for each cache node $s \in V$, its request forwarding path for different content items is the same, i.e.,

$$p^{(s,i)} = p^s, \forall i \in \mathcal{I}. \quad (21)$$

In practice, a network in which all content items are stored in the cloud server satisfies (20). Note that (21) is naturally satisfied when each node chooses the unique shortest path to fetch content items. Furthermore, (21) is naturally satisfied in a tree topology.

In the following, we show that a β -approximate PSNE exists in Game 2 with the cloud property and the path overlap property in (11).¹⁷ Note that in the caching graph satisfying the cloud property and the path overlap property, there is no mixed request loop.

Theorem 7. *A β -approximate PSNE always exists in Game 2 with the cloud property and the path overlap property.*

Proof. See Appendix F. \square

The result holds in arbitrary-topology networks with heterogeneous content popularity. However, the complexity for finding an approximate PSNE may grow exponentially with the number of nodes and their strategies in general.

C. Polynomial-Time Algorithm to Find an Approximate PSNE

In this section, we propose a polynomial-time algorithm to find an approximate PSNE of Game 2.

With the cloud property in (20) and (21), and given designated server node u , the caching gain of each cache node $s \in V$ depends on not only \mathbf{x}_s but also $\{\mathbf{x}_v : v \in V_s\}$, where $V_s = \{v : v \in p^s, v \neq s, v \neq u\}$ is the set of intermediate nodes on node s ' request forwarding path p^s . We group nodes with the same number of intermediate nodes into one set, i.e., denote the set of nodes with $|V_s| = m$ by $\mathcal{V}^m = \{s \in V : |V_s| = m\}$ where $0 \leq m \leq |V| - 2$. Note that since node u stores all content items, its caching strategy does not affect other nodes.

If the path overlap property in (11) is satisfied, we know that if node $v \in V_s$, then $s \notin V_v$. That is, if node v is on path p^s , then node s is not on path p^v . Hence, for each node $s \in \mathcal{V}^m$ with $|V_s| = m$ intermediate nodes in set $V_s = \{v : v \in p^s, v \neq s, v \neq u\}$, every intermediate node $v \in V_s$ has a smaller number of intermediate nodes, i.e., $|V_v| < m$. This motivates us to find the equilibrium strategies for nodes in sets $\mathcal{V}^m, 0 \leq m \leq |V| - 2$, according to the increasing order of m .

We propose a polynomial-time algorithm (Algorithm 2) to find a β -approximate PSNE. Specifically, We find the equilibrium strategies of nodes in sets \mathcal{V}^m for $0 \leq m \leq |V| - 2$ sequentially (Lines 1-2 of Algorithm 2). For example, for

¹⁷The cloud property and the path overlap property are sufficient conditions for existence. Analyzing the sufficient and necessary conditions for existence of (approximate) PSNE on graphs is an open problem, and we will consider it in the future work.

Algorithm 2: Find β -Approximate PSNE

Input: $G(V, E), \mathcal{I}, w_{uv}, \forall (u, v) \in E, \lambda_{(s,i)}$ and $p^{(s,i)}$, for all $s \in V, i \in \mathcal{I}$
Output: $\mathbf{x}^{\beta-NE}$

```

1 Classify nodes into sets  $\mathcal{V}^m$  for  $0 \leq m \leq |V| - 2$ ;
2 for  $m = 0 : |V| - 2$  do
3   for  $s \in \mathcal{V}^m$  do
4     Set  $\hat{\mathbf{x}}_s = \mathbf{0}$ , i.e.,  $\hat{x}_{si} = 0, \forall i \in \mathcal{I}$ ;
5     Relax problem (18) to a linear programming problem
       by relaxing  $\mathbf{x}_s \in \{0, 1\}^{|\mathcal{I}|}$  to  $\tilde{\mathbf{x}}_s \in [0, 1]^{|\mathcal{I}|}$ ;
6     Compute an optimal solution  $\tilde{\mathbf{x}}_s^*$  of the
       LP-relaxation;
7     Set  $I_s = \{i : \tilde{x}_{si}^* = 1\}$  and  $F_s = \{i : 0 < \tilde{x}_{si}^* < 1\}$ ;
8     if  $\sum_{i \in I_s} q_{(s,i)}(\mathbf{x}_{-s}) > \max_{i \in F_s} q_{(s,i)}(\mathbf{x}_{-s})$  then
9       Set  $\hat{x}_{si} = 1, \forall i \in I_s$ ;
10    else
11      Set  $\hat{x}_{sj} = 1$  for  $j = \arg \max_{i \in F_s} q_{(s,i)}(\mathbf{x}_{-s})$ ;
12    end
13  end
14 end
15 Set  $\mathbf{x}^{\beta-NE} = \hat{\mathbf{x}}$ ;
```

node $s \in \mathcal{V}^0$ such that $V_s = \emptyset$ (Line 3 of Algorithm 2), its β -approximate equilibrium strategy $\mathbf{x}_s^{\beta-NE}$ is the β -approximate solution to problem (18), calculated by Lines 4-12 of Algorithm 2. Note that for a node $s \in \mathcal{V}^0$, $q_{si}(\mathbf{x}_{-s}) = q_{si}$ is a constant independent of other nodes' strategies. For node $s \in \mathcal{V}^m$ with $1 \leq m \leq |V| - 2$ (Line 3 of Algorithm 2), its equilibrium strategy is the β -approximate solution to problem (18), calculated by Lines 4-12 of Algorithm 2.¹⁸ Note that its equilibrium strategy depends only on the caching strategies of nodes $v \in V_s$ in its intermediate node set with $|V_v| < m$, and hence $q_{si}(\mathbf{x}_{-s}) = q_{si}(\{\mathbf{x}_v^{\beta-NE} : v \in V_s\})$. We continue this sequential process until all nodes decide their β -approximate equilibrium caching strategies. The resulting caching strategy profile is a β -approximate PSNE of Game 2.

In the following theorem, we show that Algorithm 2 can find a β -approximate PSNE of Game 2 in polynomial time.

Theorem 8. *For Game 2 with the cloud property and the path overlap property, Algorithm 2 computes a β -approximate PSNE in $\mathcal{O}(|V||\mathcal{I}|)$ time.*

Proof. See Appendix G. \square

We next analyze the PoA of Game 2.

D. Price of Anarchy

we show that the PoA for the β -approximate PSNE is bounded under the homogeneous request pattern property in (10).

Theorem 9. *For Game 2 with the cloud property, the path overlap property, and the homogeneous request pattern property, the PoA for the β -approximate PSNE satisfies*

$$PoA^\beta \geq \frac{1}{1 + \alpha \cdot \beta} = \frac{1}{1 + \beta \cdot (\max_{v \in V, i \in \mathcal{I}} |p^{(v,i)}| - 1)}. \quad (22)$$

¹⁸Lines 4-12 of Algorithm 2 achieve a $1/2$ approximation ratio of problem (18), and hence $\beta = 2$. Note that β is identical across all nodes.

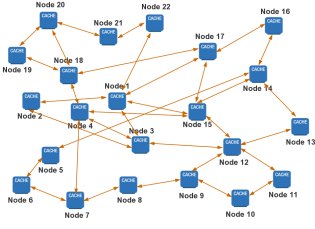


Figure 14. GEANT network.

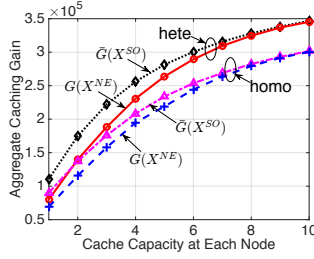
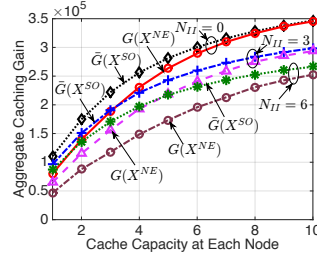
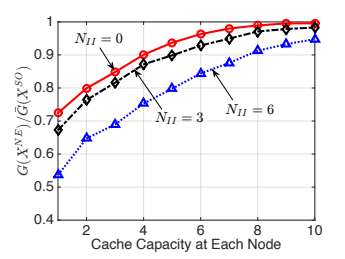
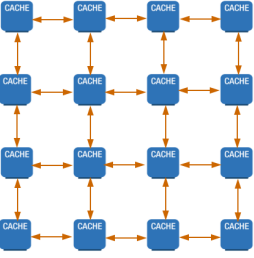
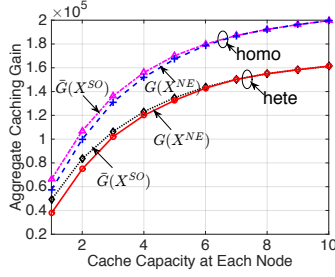
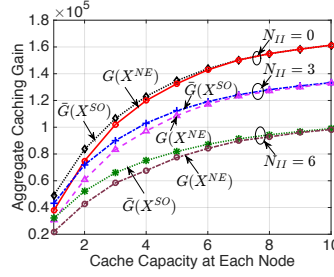
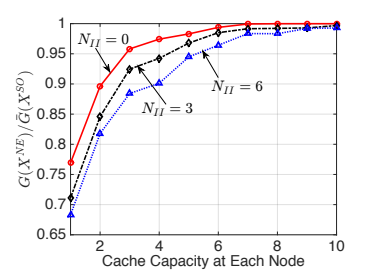
Figure 15. $G(\cdot)$ vs. c_v , under both heterogeneous and homogeneous request patterns.Figure 16. $G(\cdot)$ vs. c_v , under different no. of Type-II nodes N_{II} .Figure 17. $G(x^{NE})/\bar{G}(x^{SO})$ vs. c_v , under different N_{II} .

Figure 18. Grid topology.

Figure 19. $G(\cdot)$ vs. c_v , under both heterogeneous and homogeneous request patterns.Figure 20. $G(\cdot)$ vs. c_v , under different no. of Type-II nodes N_{II} .Figure 21. $G(x^{NE})/\bar{G}(x^{SO})$ vs. c_v , under different N_{II} .

We generate the arrival rates $\lambda_{(s,i)}, \forall s \in V, i \in \mathcal{I}$ uniformly at random in the interval $[0, 10]$.

Results in the Abilene Network. Figure 10 shows the aggregate caching gain $G(x^{NE})$ and $\bar{G}(x^{SO})$ under different cache capacities at each node,²¹ for the case with heterogeneous request patterns $\lambda_{(s,i)}$ (the upper two curves) and for the case with homogeneous request patterns $\lambda_{(s,i)} = \lambda_i, \forall s \in V, i \in \mathcal{I}$ (the lower two curves)²², respectively. We can see that the gap between $G(x^{NE})$ and $\bar{G}(x^{SO})$ under homogeneous λ_i is smaller than the gap under heterogeneous $\lambda_{(s,i)}$. Thus, the homogeneous request pattern leads to better performance achieved by selfish caching behaviors in the Abilene network.

In practice, some cache nodes are intermediate routers which do not request for any content items. We define nodes with positive request rates as Type-I nodes (with a total number N_I), and nodes with no request as Type-II nodes (with a total number N_{II}). We show the impact of N_{II} in Figure 11. We can see that the gap between $G(x^{NE})$ and $\bar{G}(x^{SO})$ decreases with the cache capacity at each node, while the gap increases with N_{II} . This implies that the impact of the selfish behaviors is mitigated when the cache resource increases, and the selfish behaviors of Type-II nodes degrade the (relative) performance of Nash equilibria (since the selfish Type-II nodes will not cache content items at equilibrium).

To understand the impact of the randomness of request arrival rates $\lambda_{(s,i)}$, we perform simulations on 100 sets of randomly generated $\{\lambda_{(s,i)} : \forall s \in V, i \in \mathcal{I}\}$, and show the average ratios $G(x^{NE})/\bar{G}(x^{SO})$ of the 100 trials in Figure

²¹We show the results for the case where cache nodes may have different cache capacities in Appendix J.

²²We take $\lambda_i = \sum_{s \in V} \lambda_{(s,i)} / V$ given the heterogeneous $\lambda_{(s,i)}, \forall s \in V, i \in \mathcal{I}$.

12, where the error bars represent the standard deviations. As is consistent with our observation from Figure 11, the performance of the Nash equilibria increases with the cache capacity, while decreases with N_{II} .

In practice, one direct way to improve the aggregate caching gain in the network is to add extra cache nodes. To check the impact of extra caches on the performance of Nash equilibria, we sequentially add node 12, node 13, until node 21, shown in Figure 9. We show the ratio $G(x^{NE})/\bar{G}(x^{SO})$ with different number of extra nodes in Figure 13. We can see that adding more extra caches makes PoA worse. The reason is that adding extra cache nodes can improve the optimal social welfare, while it cannot improve the social welfare achieved by Nash equilibria due to the selfish nature of cache nodes. Hence the “relative” performance of the Nash equilibria (measured in terms of PoA) reduces.

Results in the GEANT Network. We perform simulations on the GEANT network shown in Figure 14. We consider a set $\mathcal{I} = \{1, \dots, 20\}$ of content items. We generate the cost on each edge uniformly at random from the interval $[1, 100]$. We show the performances corresponding to selfish behaviors in Figures 15–17. As in the Abilene network, the homogeneous request pattern leads to better (relative) performance achieved by selfish caching behaviors, and the ratio $G(x^{NE})/\bar{G}(x^{SO})$ increases with c_v and decreases with N_{II} .

Results in the Grid Topology. We perform simulations on the Grid topology shown in Figure 18. We consider a set $\mathcal{I} = \{1, \dots, 16\}$ of content items, and generate the cost on each edge uniformly at random from the interval $[1, 100]$. We show the performance corresponding to selfish behaviors in Figures 19–21. Different from the Abilene and GEANT networks, we observe in Figure 19 that the homogeneous

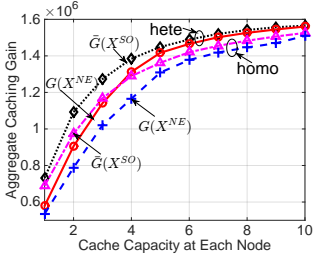


Figure 22. $G(\cdot)$ vs. c_v , under both heterogeneous and homogeneous request patterns.

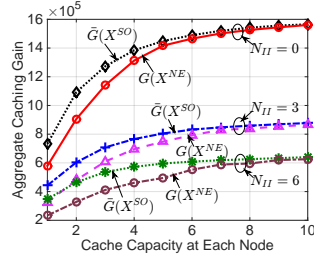


Figure 23. $G(\cdot)$ vs. c_v , under different no. of Type-II nodes N_{II} .

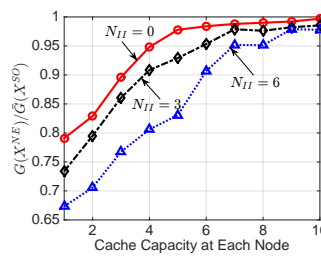


Figure 24. $G(x^{NE})/\bar{G}(x^{SO})$ vs. c_v , under different N_{II} .

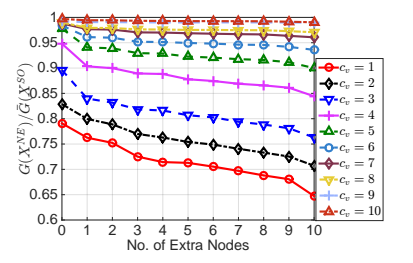


Figure 25. $G(x^{NE})/\bar{G}(x^{SO})$ vs. no. of extra cache nodes, under different c_v .

request pattern leads to a larger aggregate caching gain but a smaller ratio $G(x^{NE})/\bar{G}(x^{SO})$ than that under the heterogeneous request pattern. As in the Abilene and GEANT networks, $G(x^{NE})/\bar{G}(x^{SO})$ increases with c_v , and decreases with N_{II} .

Results for the Scenario with Unequal-Sized Items. We perform simulations in the Abilene network for the case where different content items have different sizes in Figures 22 – 25. We assume that the sizes of the $|I| = 10$ content items are $L = \{0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0\}$, and node 1 is the designated server of the 10 content items. We compare the performance achieved by the approximate Nash equilibria of Game 2 and that by the socially optimal solution. Figure 22 shows that the homogeneous request pattern leads to larger gaps between $G(x^{NE})$ and $\bar{G}(x^{SO})$, and hence a worse performance achieved by selfish caching behaviors at the approximate Nash equilibrium. Figures 23 – 25 show that the gap between $G(x^{NE})$ and $\bar{G}(x^{SO})$ decreases with the cache capacity at each node, while the gap increases with N_{II} . Furthermore, adding more extra caches makes the PoA worse.

VIII. CONCLUSION

In this paper, we analyze selfish caching games on directed graphs, which can yield arbitrary bad performance. We show that a PSNE exists and can be found in polynomial time if there is no mixed request loop, and we can avoid mixed request loops by properly choosing request forwarding paths. We then show that although cache paradox happens, i.e., adding extra cache nodes does not improve the performance of PSNE, with the homogeneous request pattern property and the path overlap property, the PoA is bounded in arbitrary-topology networks. We further show that the selfish caching game with unequal-sized items admits an approximate PSNE with bounded PoA in special cases.

There are several interesting directions to explore in the future, such as analyzing the impact of the congestion effect on each edge, analyzing the joint caching and routing decisions of selfish nodes, analyzing the privacy issue, analyzing the dynamic selfish caching game under incomplete information, and analyzing the coalitional game for the caching network with multiple cache providers where each provider owns several cache nodes.

REFERENCES

- [1] Q. Ma, E. Yeh, and J. Huang, “How bad is selfish caching?” in *The Twentieth ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '19)*. ACM, 2019.
- [2] S. Borst, V. Gupta, and A. Walid, “Distributed caching algorithms for content distribution networks,” in *IEEE INFOCOM*, 2010, pp. 1–9.
- [3] M. Dehghan, A. Seetharam, B. Jiang, T. He, T. Salonidis, J. Kurose, D. Towsley, and R. Sitaraman, “On the complexity of optimal routing and content caching in heterogeneous networks,” in *IEEE INFOCOM*, 2015, pp. 936–944.
- [4] E. Yeh, T. Ho, Y. Cui, M. Burd, R. Liu, and D. Leong, “Vip: A framework for joint dynamic forwarding and caching in named data networks,” in *ACM ICN*, 2014, pp. 117–126.
- [5] K. Shanmugam, N. Golrezaei, A. G. Dimakis, A. F. Molisch, and G. Caire, “Femtocaching: Wireless content delivery through distributed caching helpers,” *IEEE TIT*, vol. 59, no. 12, pp. 8402–8413, 2013.
- [6] N. Laoutaris, O. Telelis, V. Zissimopoulos, and I. Stavrakakis, “Distributed selfish replication,” *IEEE TPDS*, vol. 17, no. 12, pp. 1401–1413, 2006.
- [7] E. Cohen and S. Shenker, “Replication strategies in unstructured peer-to-peer networks,” in *ACM SIGCOMM*, vol. 32, no. 4, 2002, pp. 177–190.
- [8] K. Poularakis, G. Iosifidis, A. Argyriou, I. Koutsopoulos, and L. Tassiulas, “Distributed caching algorithms in the realm of layered video streaming,” *IEEE TMC*, 2018.
- [9] S. Ioannidis and E. Yeh, “Adaptive caching networks with optimality guarantees,” in *ACM SIGMETRICS*, vol. 44, no. 1, 2016, pp. 113–124.
- [10] W. C. Ao and K. Psounis, “Distributed caching and small cell cooperation for fast content delivery,” in *ACM MobiHoc*, 2015, pp. 127–136.
- [11] L. Wang, G. Tyson, J. Kangasharju, and J. Crowcroft, “Milking the cache cow with fairness in mind,” *IEEE/ACM TON*, vol. 25, no. 5, pp. 2686–2700, 2017.
- [12] M. Afanasyev, T. Chen, G. M. Voelker, and A. C. Snoeren, “Usage patterns in an urban wifi network,” *IEEE/ACM TON*, vol. 18, no. 5, pp. 1359–1372, 2010.
- [13] D. Vega, L. Cerda-Alabern, L. Navarro, and R. Meseguer, “Topology patterns of a community network: Guifi. net,”

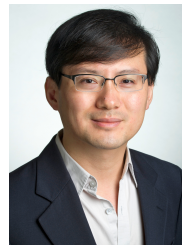
- in *IEEE WiMob*, 2012, pp. 612–619.
- [14] R. Draves, J. Padhye, and B. Zill, “Routing in multi-radio, multi-hop wireless mesh networks,” in *ACM MobiCom*, 2004, pp. 114–128.
 - [15] A. Vetta, “Nash equilibria in competitive societies, with applications to facility location, traffic routing and auctions,” in *IEEE FOCS*, 2002, pp. 416–425.
 - [16] C. Papadimitriou, “Algorithms, games, and the internet,” in *ACM STOC*, 2001, pp. 749–753.
 - [17] T. Roughgarden and É. Tardos, “How bad is selfish routing?” *Journal of the ACM*, vol. 49, no. 2, pp. 236–259, 2002.
 - [18] L.-S. Juhn and L.-M. Tseng, “Harmonic broadcasting for video-on-demand service,” *IEEE transactions on broadcasting*, vol. 43, no. 3, pp. 268–271, 1997.
 - [19] T.-Y. Huang, N. Handigol, B. Heller, N. McKeown, and R. Johari, “Confused, timid, and unstable: picking a video streaming rate is hard,” in *Proceedings of the 2012 Internet Measurement Conference*. ACM, 2012, pp. 225–238.
 - [20] G. S. Paschos, G. Iosifidis, M. Tao, D. Towsley, and G. Caire, “The role of caching in future communication systems and networks,” *arXiv preprint arXiv:1805.11721*, 2018.
 - [21] S. Shukla, O. Bhardwaj, A. A. Abouzeid, T. Salonidis, and T. He, “Hold’em caching: Proactive retention-aware caching with multi-path routing for wireless edge networks,” in *ACM MobiHoc*, 2017, p. 24.
 - [22] J. Tadrous, A. Eryilmaz, and H. El Gamal, “Joint smart pricing and proactive content caching for mobile services,” *IEEE/ACM TON*, vol. 24, no. 4, pp. 2357–2371, 2016.
 - [23] Y. Li and S. Ioannidis, “Universally stable cache networks,” in *IEEE INFOCOM*, 2020.
 - [24] M. Mahdian, A. Moharrer, S. Ioannidis, and E. Yeh, “Kelly cache networks,” in *IEEE INFOCOM*, 2019, pp. 217–225.
 - [25] M. Garetto, E. Leonardi, and S. Traverso, “Efficient analysis of caching strategies under dynamic content popularity,” in *IEEE INFOCOM*, 2015, pp. 2263–2271.
 - [26] J. Zhang, X. Lin, and X. Wang, “Coded caching under arbitrary popularity distributions,” *IEEE TIT*, vol. 64, no. 1, pp. 349–366, 2018.
 - [27] J. Li, T. Khoa Phan, W. Koong Chai, D. Tuncer, G. Pavlou, D. Griffin, and M. Rio, “Dr-cache: Distributed resilient caching with latency guarantees,” in *IEEE INFOCOM*, 2018, pp. 441–449.
 - [28] T. Zhao, I.-H. Hou, S. Wang, and K. Chan, “Red/led: An asymptotically optimal and scalable online algorithm for service caching at the edge,” *IEEE JSAC*, vol. 36, no. 8, pp. 1857–1870, 2018.
 - [29] X. Li, X. Wang, P.-J. Wan, Z. Han, and V. C. Leung, “Hierarchical edge caching in device-to-device aided mobile networks: Modeling, optimization, and design,” *IEEE JSAC*, 2018.
 - [30] X. Zhao, P. Yuan, and S. Tang, “Collaborative edge caching in context-aware device-to-device networks,” *IEEE TVT*, vol. 67, no. 10, pp. 9583–9596, 2018.
 - [31] J. Kwak, Y. Kim, L. B. Le, and S. Chong, “Hybrid content caching in 5g wireless networks: Cloud versus edge caching,” *IEEE TWC*, vol. 17, no. 5, pp. 3030–3045, 2018.
 - [32] X. Cao, J. Zhang, and H. V. Poor, “An optimal auction mechanism for mobile edge caching,” in *IEEE ICDCS*, 2018, pp. 388–399.
 - [33] A. Ferragut, I. Rodriguez, and F. Paganini, “Optimizing ttl caches under heavy-tailed demands,” in *ACM SIGMETRICS*, 2016, p. 101–112.
 - [34] M. Dehghan, L. Massoulié, D. Towsley, D. S. Menasché, and Y. C. Tay, “A utility optimization approach to network cache design,” *IEEE/ACM TON*, vol. 27, no. 3, pp. 1013–1027, 2019.
 - [35] Z. Qin, X. Gan, L. Fu, X. Di, J. Tian, and X. Wang, “Content delivery in cache-enabled wireless evolving social networks,” *IEEE TWC*, 2018.
 - [36] M. Dehghan, A. Seetharam, B. Jiang, T. He, T. Salonidis, J. Kurose, D. Towsley, and R. Sitaraman, “On the complexity of optimal routing and content caching in heterogeneous networks,” in *IEEE INFOCOM*, 2015, pp. 936–944.
 - [37] M. M. Amble, P. Parag, S. Shakkottai, and L. Ying, “Content-aware caching and traffic management in content distribution networks,” in *IEEE INFOCOM*, 2011.
 - [38] S. Ioannidis and E. Yeh, “Jointly optimal routing and caching for arbitrary network topologies,” *IEEE JSAC*, vol. 36, no. 6, pp. 1258–1275, 2018.
 - [39] W. Chu, M. Dehghan, D. Towsley, and Z.-L. Zhang, “On allocating cache resources to content providers,” in *ACM ICN*, 2016, pp. 154–159.
 - [40] A. Gharaibeh, A. Khreishah, B. Ji, and M. Ayyash, “A provably efficient online collaborative caching algorithm for multicell-coordinated systems,” *IEEE TMC*, vol. 15, no. 8, pp. 1863–1876, 2016.
 - [41] K. Shin, C. Joe-Wong, S. Ha, Y. Yi, I. Rhee, and D. S. Reeves, “T-chain: A general incentive scheme for cooperative computing,” *IEEE/ACM TON*, vol. 25, no. 4, pp. 2122–2137, 2017.
 - [42] P. Rahimzadeh, C. Joe-Wong, K. Shin, Y. Im, J. Lee, and S. Ha, “Svc-tchain: Incentivizing good behavior in layered p2p video streaming,” in *IEEE INFOCOM*, 2017, pp. 1–9.
 - [43] R. Yu, S. Qin, M. Bennis, X. Chen, G. Feng, Z. Han, and G. Xue, “Enhancing software-defined ran with collaborative caching and scalable video coding,” in *IEEE ICC*, 2016, pp. 1–6.
 - [44] A. T. Ip, J. Lui, and J. Liu, “A revenue-rewarding scheme of providing incentive for cooperative proxy caching for media streaming systems,” *ACM TOMM*, vol. 4, no. 1, p. 5, 2008.
 - [45] P. Maillé, G. Simon, and B. Tuffin, “Impact of revenue-driven cdn on the competition among network operators,” in *IEEE CNSM*, 2015, pp. 163–167.
 - [46] B. G. Chun, K. Chaudhuri, H. Wee, M. Barreno, C. H. Papadimitriou, and J. Kubiawicz, “Selfish caching in distributed systems: a game-theoretic analysis,” in *ACM PODC*, 2004, pp. 21–30.

- [47] M. Goemans, L. E. Li, V. S. Mirrokni, and M. Thottan, "Market sharing games applied to content distribution in ad-hoc networks," in *ACM MobiHoc*, 2004, pp. 55–66.
- [48] G. G. Pollatos, O. A. Telelis, and V. Zissimopoulos, "On the social cost of distributed selfish content replication," in *International Conference on Research in Networking*. Springer, 2008, pp. 195–206.
- [49] R. Gopalakrishnan, D. Kanoulas, N. N. Karuturi, C. P. Rangan, R. Rajaraman, and R. Sundaram, "Cache me if you can: capacitated selfish replication games," in *Latin American Symposium on Theoretical Informatics*. Springer, 2012, pp. 420–432.
- [50] B. Jiang, P. Nain, and D. Towsley, "On the convergence of the ttl approximation for an lru cache under independent stationary request processes," *ACM TOMPECS*, vol. 3, no. 4, p. 20, 2018.
- [51] N. K. Panigrahy, J. Li, and D. Towsley, "Network cache design under stationary requests: Exact analysis and poisson approximation," in *IEEE MASCOTS*, 2018, pp. 251–263.
- [52] Y.-K. Wang, Y. Yin, and S. Zhong, "Belief propagation for spatial spectrum access games," in *Proceedings of the 15th ACM international symposium on Mobile ad hoc networking and computing*. ACM, 2014, pp. 225–234.
- [53] H. Kellerer, U. Pferschy, and D. Pisinger, "Introduction to np-completeness of knapsack problems," in *Knapsack problems*. Springer, 2004, pp. 483–493.
- [54] C. Daskalakis, A. Mehta, and C. Papadimitriou, "Progress in approximate nash equilibria," in *ACM EC*, 2007, pp. 355–358.
- [55] S. Chien and A. Sinclair, "Convergence to approximate nash equilibria in congestion games," *Games and Economic Behavior*, vol. 71, no. 2, pp. 315–327, 2011.
- [56] A. Li, X. Yang, and D. Wetherall, "Safeguard: safe forwarding during route changes," in *PACM CoNEXT*, 2009, pp. 301–312.



Qian Ma is an Associate Professor of School of Intelligent Systems Engineering, Sun Yat-sen University. She worked as a Postdoc Research Associate at Northeastern University during 2018-2019. She received the Ph.D. degree in the Department of Information Engineering from the Chinese University of Hong Kong in 2017, and the B.S. degree from Beijing University of Posts and Telecommunications (China) in 2012. Her research interests lie in the field of network optimization and economics. She is the recipient of the Best Student Paper Award from the

IEEE International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt) in 2015.



Edmund Yeh received the B.S. degree in Electrical Engineering with Distinction and Phi Beta Kappa from Stanford University in 1994, the M. Phil degree in Engineering from Cambridge University on the Winston Churchill Scholarship in 1995, and the Ph.D. in Electrical Engineering and Computer Science from MIT under Professor Robert Gallager in 2001. He is a Professor of Electrical and Computer Engineering at Northeastern University with a courtesy appointment in Khoury School of Computer Sciences. He was previously Assistant and Associate

Professor of Electrical Engineering, Computer Science, and Statistics at Yale University.

Professor Yeh is an IEEE Communications Society Distinguished Lecturer. He serves as TPC Co-Chair for ACM MobiHoc 2021 and served as General Chair for ACM SIGMETRICS 2020. He is the recipient of the Alexander von Humboldt Research Fellowship, the Army Research Office Young Investigator Award, the Winston Churchill Scholarship, the National Science Foundation and Office of Naval Research Graduate Fellowships, the Barry M. Goldwater Scholarship, the Frederick Emmons Terman Engineering Scholastic Award, and the President's Award for Academic Excellence (Stanford University). Professor Yeh serves as Treasurer of the Board of Governors of the IEEE Information Theory Society. He has served as an Associate Editor for IEEE Transactions on Networking, IEEE Transactions on Mobile Computing, and IEEE Transactions on Network Science and Engineering, as Guest Editor-in-Chief of the Special Issue on Wireless Networks for Internet Mathematics, and Guest Editor for IEEE Journal on Selected Areas in Communications - Special Series on Smart Grid Communications. He has received three Best Paper Awards, including awards at the 2017 ACM Conference on Information-Centric Networking (ICN), and at the 2015 IEEE International Conference on Communications (ICC) Communication Theory Symposium.



Jianwei Huang is a Presidential Chair Professor and the Associate Dean of the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen. He is also the Vice President of Shenzhen Institute of Artificial Intelligence and Robotics for Society. He received the Ph.D. degree from Northwestern University in 2005, and worked as a Postdoc Research Associate at Princeton University during 2005-2007. He has been an IEEE Fellow, a Distinguished Lecturer of IEEE Communications Society, a Clarivate Analytics Highly Cited

Researcher in Computer Science, and Associate Editor-in-Chief of IEEE Open Journal of the Communications Society. He is the incoming Editor-in-Chief of IEEE Transactions on Network Science and Engineering, effective 01/2021. Dr. Huang has published 290 papers in leading international journals and conferences in the area of network optimization and economics, with a total Google Scholar citations of 12,700+ and an H-index of 58. He is the co-author of 9 Best Paper Awards, including IEEE Marconi Prize Paper Award in Wireless Communications in 2011. He has co-authored seven books, including the textbook on "Wireless Network Pricing." He received the CUHK Young Researcher Award in 2014 and IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2009. He has served as an Associate Editor of IEEE Transactions on Mobile Computing, IEEE/ACM Transactions on Networking, IEEE Transactions on Network Science and Engineering, IEEE Transactions on Wireless Communications, IEEE Journal on Selected Areas in Communications - Cognitive Radio Series, and IEEE Transactions on Cognitive Communications and Networking. He has served as the Chair of IEEE ComSoc Cognitive Network Technical Committee and Multimedia Communications Technical Committee. He is the recipient of IEEE ComSoc Multimedia Communications Technical Committee Distinguished Service Award in 2015 and IEEE GLOBECOM Outstanding Service Award in 2010. More detailed information can be found at <https://sse.cuhk.edu.cn/en/faculty/huangjianwei>.

APPENDIX A PROOF OF THEOREM 2

A PSNE of the selfish caching game corresponds to a vertex on the state graph without any outgoing arc, i.e., a sink. By showing the existence of a sink on the state graph, we can show the existence of the PSNE of the selfish caching game. Since Algorithm 1 can find a sink on the corresponding state graph, the PSNE of the selfish caching game exists.

APPENDIX B PROOF OF THEOREM 3

As one player adds a content item to its cache at the first arc of each round, the number of rounds is no greater than $|V||\mathcal{I}|$.

We next show that each round ends after traversing at most $|\mathcal{I}|(|V| - 2)^2$ arcs. We focus on one round where player s adds content item i . The add step will only affect nodes $v \in V$ where $s \in p^{(v,i)}$. If $x_{vi} = 0$, i.e., $i \notin Z_v$, the add step will not change node v 's behavior, as the add step will only decrease the value of content item i to node v . If $x_{vi} = 1$, i.e., $i \in Z_v$, since the add step will decrease the value of content item i to node v , node v may replace item i with item j . Subsequently, the change step where node v replaces item i with item j will lead to the change step where node u (such that $v \in p^{(u,j)}$ and $x_{uj} = 1$) replaces item j with item n . The sequence of change steps caused by node s adding item i ends in at most $|V| - 2$ steps, and node s will keep item i in its cache during the sequence of change steps, due to the assumption on no mixed request loop. Note that the number of such sequences of change steps is at most $|\mathcal{I}|(|V| - 2)$. Hence, each round ends in at most $|\mathcal{I}|(|V| - 2)^2$ arcs.

This proves our bound of $|V||\mathcal{I}|^2(|V| - 2)^2$ on the length of the traversal to reach a PSNE.

APPENDIX C PROOF OF THEOREM 4

The first property (5) and (6), i.e., the social function $G(\cdot)$ is nondecreasing and submodular, is proved in [9].

By the definition of our social function, we have

$$G(\mathbf{x}) = \sum_{s \in V} g_s(\mathbf{x}_s, \mathbf{x}_{-s}),$$

and therefore the second property (7) is satisfied.

In the following, we prove that the third property (9) is satisfied. For the social function $G(\cdot)$, we have

$$\begin{aligned} & G(\mathbf{x}_s, \mathbf{x}_{-s}) - G(\mathbf{0}, \mathbf{x}_{-s}) \\ \stackrel{(a)}{=} & \sum_{i \in Z_s} \lambda_{(s,i)} \sum_{k=1}^{|p^{(s,i)}|-1} w_{p_{k+1}p_k} \prod_{k'=2}^k (1 - x_{p_{k'}i}) \\ & + \sum_{i \in Z_s} \sum_{v \in V \setminus \{\mathcal{D}^i, s\}} \lambda_{(v,i)} \sum_{k=k_p^{(v,i)}(s)}^{|p^{(v,i)}|-1} w_{p_{k+1}p_k} \prod_{k'=1}^k (1 - x_{p_{k'}i}) \end{aligned}$$

$$\begin{aligned} \stackrel{(b)}{=} & \sum_{i \in Z_s} \lambda_i \sum_{k=1}^{|p^{(s,i)}|-1} w_{p_{k+1}p_k} \prod_{k'=2}^k (1 - x_{p_{k'}i}) \cdot \sum_{v \in V \setminus \mathcal{D}^i} \mathbb{1}_{\left\{v=s \text{ or } \left[s \in p^{(v,i)} \text{ and } \prod_{k'=1}^k p_{p_{k'}i}^{(v,i)(s)-1} (1 - x_{p_{k'}i}) = 1\right]\right\}} \\ \stackrel{(c)}{\leq} & \sum_{i \in Z_s} \lambda_i \sum_{k=1}^{|p^{(s,i)}|-1} w_{p_{k+1}p_k} \prod_{k'=2}^k (1 - x_{p_{k'}i}) \cdot \left(\max_{v \in V, i \in \mathcal{I}} |p^{(v,i)}| - 1\right) \end{aligned}$$

Step (a) is from the definition of $G(\cdot)$. Step (b) holds under the homogeneous request pattern and path overlap properties.

Step (c) is due to $\prod_{k'=1}^k p_{p_{k'}i}^{(v,i)(s)-1} (1 - x_{p_{k'}i}) \leq 1$ and $|\{v \in V \setminus \mathcal{D}^i : s \in p^{(v,i)}\}| \leq \max_{v \in V, i \in \mathcal{I}} |p^{(v,i)}| - 1$. Hence for $g_s(\cdot)$, we have

$$\begin{aligned} g_s(\mathbf{x}_s, \mathbf{x}_{-s}) &= \sum_{i \in Z_s} \lambda_{(s,i)} \sum_{k=1}^{|p^{(s,i)}|-1} w_{p_{k+1}p_k} \\ &+ \sum_{i \in \mathcal{I} \setminus Z_s} \lambda_{(s,i)} \sum_{k=1}^{|p^{(s,i)}|-1} w_{p_{k+1}p_k} \left(1 - \prod_{k'=2}^k (1 - x'_{p_{k'}i})\right) \\ \stackrel{(d)}{\geq} & \sum_{i \in Z_s} \lambda_{(s,i)} \sum_{k=1}^{|p^{(s,i)}|-1} w_{p_{k+1}p_k} \\ \stackrel{(e)}{\geq} & \sum_{i \in Z_s} \lambda_{(s,i)} \sum_{k=1}^{|p^{(s,i)}|-1} w_{p_{k+1}p_k} \prod_{k'=2}^k (1 - x_{p_{k'}i}) \\ \stackrel{(f)}{\geq} & \frac{1}{\alpha} \cdot (G(\mathbf{x}_s, \mathbf{x}_{-s}) - G(\mathbf{0}, \mathbf{x}_{-s})) \end{aligned}$$

Step (d) is due to $1 - \prod_{k'=2}^k (1 - x'_{p_{k'}i}) \geq 0$. Step (e) is due to $\prod_{k'=2}^k (1 - x_{p_{k'}i}) \leq 1$. Step (f) is due to step (c). This completes our proof.

APPENDIX D PROOF OF THEOREM 5

To characterize the PoA, the key step is to find the relationship between the Nash equilibria and the socially optimal solution. Let \mathbf{x}^{NE} and \mathbf{x}^* denote the caching strategy under any Nash equilibria and socially optimal solution, respectively. For any Nash equilibria \mathbf{x}^{NE} , we have

$$\begin{aligned} G(\mathbf{x}^{\text{NE}}) &\stackrel{(a)}{=} \sum_{s \in V} g_s(\mathbf{x}_s^{\text{NE}}, \mathbf{x}_{-s}^{\text{NE}}) \\ &\stackrel{(b)}{\geq} \sum_{s \in V} g_s(\mathbf{x}_s^*, \mathbf{x}_{-s}^{\text{NE}}) \\ &\stackrel{(c)}{\geq} \sum_{s \in V} \frac{1}{\alpha} (G(\mathbf{x}_s^*, \mathbf{x}_{-s}^{\text{NE}}) - G(\mathbf{0}, \mathbf{x}_{-s}^{\text{NE}})) \\ &\stackrel{(d)}{=} \frac{1}{\alpha} \sum_{s \in V} (G(Z_s^*, Z_{-s}^{\text{NE}}) - G(\emptyset, Z_{-s}^{\text{NE}})) \\ &\stackrel{(e)}{\geq} \frac{1}{\alpha} \sum_{s \in V} (G(Z_1^*, \dots, Z_s^*, Z_{-s}^{\text{NE}}) - G(Z_1^*, \dots, Z_{s-1}^*, Z_{-s}^{\text{NE}})) \\ &\stackrel{(f)}{=} \frac{1}{\alpha} (G(Z_1^*, \dots, Z_{|V|}^*, Z^{\text{NE}}) - G(Z^{\text{NE}})) \\ &\stackrel{(g)}{\geq} \frac{1}{\alpha} (G(Z^*) - G(Z^{\text{NE}})) \\ &\stackrel{(h)}{=} \frac{1}{\alpha} (G(\mathbf{x}^*) - G(\mathbf{x}^{\text{NE}})) \end{aligned}$$

Step (a) is from the definition of $G(\cdot)$. Step (b) is from the definition of Nash equilibrium. Step (c) is due to (9). Step

(d) is due to the one-to-one correspondence between \mathbf{x}_s and Z_s . Step (e) is due to the submodularity of $G(\cdot)$. Step (f) is derived by calculating the summation over all $s \in V$. Step (g) is because $G(\cdot)$ is non-decreasing. Step (h) is due to the one-to-one correspondence between \mathbf{x} and Z . The relationship between $G(\mathbf{x}^{\text{NE}})$ and $G(\mathbf{x}^*)$ leads to

$$\frac{G(\mathbf{x}^{\text{NE}})}{G(\mathbf{x}^*)} \geq \frac{1}{1 + \alpha} = \frac{1}{\max_{v \in V, i \in \mathcal{I}} |p^{(v,i)}|}.$$

This completes our proof.

APPENDIX E PROOF OF THEOREM 6

We first introduce two notations. Given a strategy profile $Z = \{Z_1, \dots, Z_{|V|}\}$, let $Z \oplus Z'_s$ denote the new strategy profile where player s changes its strategy from Z_s to Z'_s . Mathematically, $Z \oplus Z'_s = \{Z_1, \dots, Z_{s-1}, Z'_s, Z_{s+1}, \dots, Z_{|V|}\}$. Given $Z = \{Z_1, \dots, Z_{|V|}\}$, we define $Z^s = \{Z_1, \dots, Z_s, \emptyset_{s+1}, \dots, \emptyset_{|V|}\}$.

The social function $G(\cdot)$ satisfies the following two lemmas.

Lemma 4. [15] *For any feasible strategy profile Z , the socially optimal solution Z^* satisfies:*

$$G(Z^*) \leq G(Z) + \sum_{s: Z_s^* \in Z^* - Z} G'_{Z_s^*}(Z \oplus \emptyset_s) - \sum_{s: Z_s \in Z - Z^*} G'_{Z_s}(Z^* \cup Z^{s-1}).$$

Proof. By the submodularity of G , we have

$$\begin{aligned} & G(Z^* \cup Z) - G(Z) \\ &= [G(Z \cup Z_1^*) - G(Z)] + \dots \\ & \quad + [G(Z \cup Z_1^* \cup \dots \cup Z_j^*) - G(Z \cup \dots \cup Z_{j-1}^*)] \\ &\leq [G(Z \cup Z_1^*) - G(Z)] + \dots + [G(Z \cup Z_j^*) - G(Z)] \\ &= \sum_{s: Z_s^* \in Z^* - Z} G'_{Z_s^*}(Z) \\ &\leq \sum_{s: Z_s^* \in Z^* - Z} G'_{Z_s^*}(Z \oplus \emptyset_s) \end{aligned}$$

Furthermore,

$$G(Z^* \cup Z) = G(Z^*) + \sum_{s: Z_s \in Z - Z^*} G'_{Z_s}(Z^* \cup Z^{s-1}).$$

This completes our proof. \square

Lemma 5. *For the selfish caching game with the homogeneous request pattern and path overlap properties on caching graphs with no mixed request loop, any Nash equilibrium Z and the socially optimal solution Z^* satisfy*

$$G(Z^*) \leq (1 + \alpha)G(Z) - \sum_{s \in V} G'_{Z_s}(Z^* \cup Z - Z_s).$$

Proof. From Lemma 4, we know that

$$\begin{aligned} & G(Z^*) \\ &\leq G(Z) + \sum_{s: Z_s^* \in Z^* - Z} G'_{Z_s^*}(Z \oplus \emptyset_s) \\ &\quad - \sum_{s: Z_s \in Z - Z^*} G'_{Z_s}(Z^* \cup Z^{s-1}) \\ &\stackrel{(a)}{\leq} G(Z) + \alpha \cdot \sum_{s: Z_s^* \in Z^* - Z} g_s(Z) \\ &\quad - \sum_{s: Z_s \in Z - Z^*} G'_{Z_s}(Z^* \cup Z^{s-1}) \\ &\stackrel{(b)}{=} G(Z) + \alpha \cdot \sum_{s: Z_s \in Z - Z^*} g_s(Z) \\ &\quad - \sum_{s: Z_s \in Z - Z^*} G'_{Z_s}(Z^* \cup Z^{s-1}) \\ &\stackrel{(c)}{=} G(Z) + \alpha \cdot [G(Z) - \sum_{s: Z_s \in Z \cap Z^*} g_s(Z)] \\ &\quad - \sum_{s: Z_s \in Z - Z^*} G'_{Z_s}(Z^* \cup Z^{s-1}) \\ &\stackrel{(d)}{\leq} G(Z) + \alpha \cdot [G(Z) - \frac{1}{\alpha} \cdot \sum_{s: Z_s \in Z \cap Z^*} G'_{Z_s}(Z \oplus \emptyset_s)] \\ &\quad - \sum_{s: Z_s \in Z - Z^*} G'_{Z_s}(Z^* \cup Z^{s-1}) \\ &\stackrel{(e)}{\leq} (1 + \alpha)G(Z) - \sum_{s: Z_s \in Z \cap Z^*} G'_{Z_s}(Z^* \cup Z - Z_s) \\ &\quad - \sum_{s: Z_s \in Z - Z^*} G'_{Z_s}(Z^* \cup Z - Z_s) \\ &= (1 + \alpha)G(Z) - \sum_{s \in V} G'_{Z_s}(Z^* \cup Z - Z_s) \end{aligned}$$

Step (a) is due to (9). Step (b) is due to the fact that the set of indices where $Z_s^* \neq Z_s$ is the same as the set of indices where $Z_s \neq Z_s^*$. Step (c) is due to $G(Z) = \sum_{s \in V} g_s(Z)$. Step (d) is due to (9). Step (e) is due to the submodularity of $G(\cdot)$. This completes our proof. \square

Now we are ready to prove Theorem 6. By the submodularity of G , any Nash equilibrium Z and the socially optimal solution Z^* satisfy

$$\begin{aligned} & \sum_{s \in V} G'_{Z_s}(Z^* \cup Z - Z_s) \geq \sum_{s \in V} G'_{Z_s}(\mathcal{I}^{|V|} - Z_s) \\ & \geq \sum_{s \in V} G'_{Z_s}(\mathcal{I}^{|V|} - Z_s) \frac{G'_{Z_s}(Z^{s-1})}{G'_{Z_s}(\emptyset)} \\ & \geq \min_{s \in V} \frac{G'_{Z_s}(\mathcal{I}^{|V|} - Z_s)}{G'_{Z_s}(\emptyset)} \sum_{s \in V} G'_{Z_s}(Z^{s-1}) = (1 - \delta(G))G(Z). \end{aligned}$$

From Lemma 5, we know that

$$\begin{aligned} G(Z^*) &\leq (1 + \alpha)G(Z) - \sum_{s \in V} G'_{Z_s}(Z^* \cup Z - Z_s) \\ &\leq (1 + \alpha)G(Z) - (1 - \delta(G))G(Z) \\ &= (\alpha + \delta(G))G(Z) \end{aligned}$$

This completes our proof.

APPENDIX F PROOF OF THEOREM 7

We show a constructive proof. Since Algorithm 2 can find a β -approximate Nash equilibrium, we can prove the existence of the β -approximate Nash equilibria for Game 2.

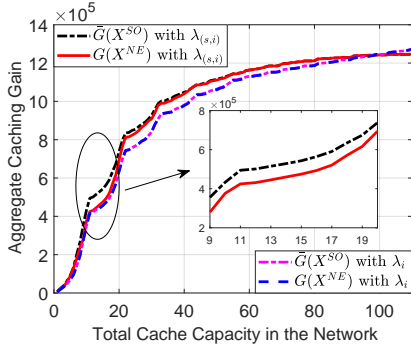


Figure 26. $G(\cdot)$ vs. $\sum_{v \in V} c_v$, under both heterogeneous and homogeneous request patterns.

APPENDIX G PROOF OF THEOREM 8

The procedure to find a β -approximate Nash equilibrium in Algorithm 2 takes at most $|V| - 1$ steps. In each step, the equilibrium strategy of the particular player can be found by solving problem (18) with complexity $\mathcal{O}(|\mathcal{I}|)$ (see Section 9.4.2 of [53]). Hence, the β -approximate Nash equilibrium can be found in $\mathcal{O}(|V||\mathcal{I}|)$ time.

APPENDIX H PROOF OF THEOREM 9

Let $\mathbf{x}^{\beta-NE}$ and \mathbf{x}^* denote the caching strategy under any β -approximate Nash equilibria and the socially optimal solution, respectively. According to the definition of the β -approximate Nash equilibrium, we have

$$g_s(\mathbf{x}_s^{\beta-NE}, \mathbf{x}_{-s}^{\beta-NE}) \geq \frac{1}{\beta} \cdot g_s(\mathbf{x}_s^*, \mathbf{x}_{-s}^{\beta-NE}), \forall s \in V.$$

Summing over all $s \in V$, we have

$$G(\mathbf{x}_s^{\beta-NE}, \mathbf{x}_{-s}^{\beta-NE}) \geq \frac{1}{\beta} \sum_{s \in V} g_s(\mathbf{x}_s^*, \mathbf{x}_{-s}^{\beta-NE}), \forall s \in V.$$

According to (9) and Theorem 5, we know that

$$\begin{aligned} g_s(\mathbf{x}_s^*, \mathbf{x}_{-s}^{\beta-NE}) &\geq \frac{1}{\alpha} \left(G(\mathbf{x}_s^*, \mathbf{x}_{-s}^{\beta-NE}) - G(\mathbf{0}, \mathbf{x}_{-s}^{\beta-NE}) \right) \\ &\geq \frac{1}{\alpha} \left(G(\mathbf{x}^*) - G(\mathbf{x}^{\beta-NE}) \right). \end{aligned}$$

Hence,

$$\begin{aligned} &G(\mathbf{x}_s^{\beta-NE}, \mathbf{x}_{-s}^{\beta-NE}) \\ &\geq \frac{1}{\beta} \sum_{s \in V} g_s(\mathbf{x}_s^*, \mathbf{x}_{-s}^{\beta-NE}) \\ &\geq \sum_{s \in V} \frac{1}{\alpha \cdot \beta} \left[G(\mathbf{x}_s^*, \mathbf{x}_{-s}^{\beta-NE}) - G(\mathbf{0}, \mathbf{x}_{-s}^{\beta-NE}) \right] \\ &\geq \frac{1}{\alpha \cdot \beta} \left[G(\mathbf{x}^*) - G(\mathbf{x}^{\beta-NE}) \right], \end{aligned}$$

which leads to

$$\frac{G(\mathbf{x}^{\beta-NE})}{G(\mathbf{x}^*)} \geq \frac{1}{1 + \alpha \cdot \beta} = \frac{1}{1 + \beta \cdot (\max_{v \in V, i \in \mathcal{I}} |p^{(v,i)}| - 1)}.$$

This completes our proof.

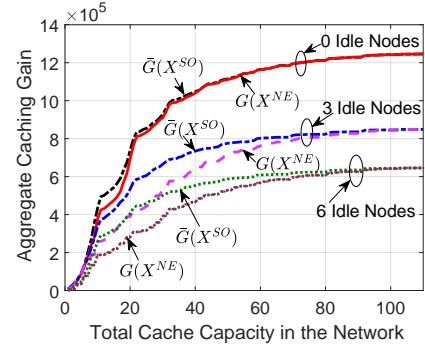


Figure 27. $G(\cdot)$ vs. $\sum_{v \in V} c_v$, under different number of idle nodes.

APPENDIX I PROOF OF LEMMA 3

Since problem (23) maximizes the same function as problem (4) over a larger domain, we have:

$$G(\phi^*) \geq G(\mathbf{x}^*).$$

By Goemans-Williamson inequality, we have that: for any sequence of $z_i \in [0, 1], i \in \{1, \dots, n\}$,

$$\left(1 - \frac{1}{e}\right) \min \left\{ 1, \sum_{i=1}^n z_i \right\} \leq 1 - \prod_{i=1}^n (1 - z_i) \leq \min \left\{ 1, \sum_{i=1}^n z_i \right\}$$

So for any feasible ϕ , we have:

$$\left(1 - \frac{1}{e}\right) L(\phi) \leq G(\phi) \leq L(\phi).$$

Hence, we have

$$L(\phi^*) \geq G(\phi^*).$$

By the optimality of ϕ^{**} to problem (24), we have

$$L(\phi^{**}) \geq L(\phi^*).$$

This completes our proof.

APPENDIX J SIMULATIONS FOR DIFFERENT CACHE CAPACITIES

We perform simulations on how the aggregate caching gains $G(\mathbf{x}^{NE})$ and $\bar{G}(\mathbf{x}^{SO})$ change with the total cache capacity in the Abilene network. We start from the state where the cache capacity of each node is zero. In each trial of the simulation, we add one unit of cache capacity to one node (starting from node 1 to node 11). For example, when the total cache capacity is 3, we have $c_v = 1$ for $v = 1, 2, 3$, and $c_v = 0$ for $v = 4, 5, \dots, 11$. When the total cache capacity is 28, we have $c_v = 3$ for $v = 1, \dots, 6$, and $c_v = 2$ for $v = 7, \dots, 11$.

Figure 26 shows the aggregate caching gains $G(\mathbf{x}^{NE})$ and $\bar{G}(\mathbf{x}^{SO})$ under different total cache capacities in the network, for the case with heterogeneous request patterns $\lambda_{(s,i)}$ (the upper two curves) and for the case with homogeneous request patterns $\lambda_{(s,i)} = \lambda_i, \forall s \in V, i \in \mathcal{I}$ (the lower two curves), respectively. We can see that the gap between $G(\mathbf{x}^{NE})$ and $\bar{G}(\mathbf{x}^{SO})$ under homogeneous request patterns λ_i is smaller than the gap under heterogeneous request patterns $\lambda_{(s,i)}$. This

is consistent with the observation that we obtained from Figure 10 in the paper.

In practice, some cache nodes are intermediate routers which do not request for any content items, i.e., idle nodes. We show the impact of the number of idle nodes in Figure 27. We can see that the gap between $G(x^{\text{NE}})$ and $\bar{G}(x^{\text{SO}})$ decreases with the total cache capacity in the network, while the gap increases with the number of idles nodes. This is consistent with the observations that we obtained from Figure 11 in the paper. This implies that the impact of the selfish behaviors is mitigated when the cache resource increases, and the selfish behaviors of idle nodes degrade the (relative) performance of Nash equilibria (since the selfish idle nodes will not cache content items at equilibrium).