Distributed Caching Algorithms for Interconnected Operator CDNs

Valentino Pacifici and György Dán

***Abstract*—Fixed and mobile network operators increasingly deploy managed CDNs with the objective of reducing the traffic on their transit links and to improve their customers' quality of experience. As network operator managed CDNs (nCDNs) become commonplace, operators will likely provide common interfaces to interconnect their nCDNs for mutual benefit, as they do with peering today. In this paper we consider the problem of using distributed algorithms for computing a cache allocation for nCDNs. We show that if every ISP aims to minimize its cost and bilateral payments are not allowed then it may be impossible to compute a cache allocation. For the case when bilateral payments are possible, we propose two distributed algorithms, the aggregate value compensation (AC) and the object value compensation (OC) algorithms, which differ in terms of the level of parallelism they allow and in terms of the amount of information exchanged between nCDNs. We prove that the algorithms converge, and we propose a scheme to ensure ex-post individual rationality. Simulations performed on a real AS-level network topology and synthetic topologies show that the algorithms have geometric rate of convergence, and scale well with the graphs' density and the nCDN capacity.**

***Index Terms*—operator managed CDNs, cooperative caching, content allocation, game theory.**

I. Introduction

Real-time and on-demand video are consumed by a large and ever increasing fraction of mobile and fixed Internet users. To fuel this growth, major over-the-top content providers, such as Netflix, Hulu, etc., try to maintain customer satisfaction through increasing Quality of Experience (QoE): 3D content has become commonplace, and super HD content has become available recently [1]. To face the new challenges caused by increasing demands for digital content and in order to maintain a competitive QoE for their users, most content providers outsource content delivery to commercial content distribution networks (CDNs). For content providers, CDNs offer relatively low delivery costs compared to investing in an own infrastructure, they provide dynamically scaling bandwidth to satisfy sudden surges of demand, and through multiple surrogate servers they provide better quality of experience (QoE) for customers than a system based on a single content delivery server [2], [3].

The increasing demand and the improved QoE result in increased bitrates, which stresses network operators' networks, yet in the traditional CDN-based content distribution model network operators are not part of the revenue chain. At the same time, good QoE may also require control of the network resources between the CDN surrogate and the customers' premises and needs content to be placed closer to the customers.

Many network providers have started to deploy their own CDNs for the above reasons, and recent industry efforts aim to interconnect these network operator managed CDNs (nCDNs), potentially also with traditional commercial CDNs [3], [4]. For content providers, nCDN interconnection provides a transparent solution for bringing content closer to the customers than any single CDN would be able to provide. For network providers, nCDN interconnection can improve CDN availability and customer QoE.

As nCDNs often prefetch content based on predicted demands during periods of low demands (e.g., NetFlix Open Connect), successful nCDN interconnection requires that given predicted demands, the nCDNs be able to agree on a content allocation that serves all service providers' interests. In lack of a central authority the agreement has to be based on a distributed algorithm, the algorithm should not reveal confidential information, and the resulting allocation should be such that no nCDN fares worse due to interconnection, as otherwise nCDNs would have no incentive to interconnect.

In this paper we address the design of distributed algorithms for content allocation among interconnected nCDNs. We propose a model of CDN interconnection assuming that CDNs aim to maximize the QoE of their customers and we show that content allocations that maximize the aggregate QoE are infeasible to compute and to implement. We show that self-enforcing content allocations may not exist if payments are not allowed among nCDNs. We propose two distributed algorithms that use bilateral compensations to guarantee convergence to a content allocation and we propose an opt-out scheme, which combined with the two algorithms ensures that the resulting allocations are individually rational. Thus, participation according to the proposed algorithms is *ex-post individually rational* for all nCDNs. We use simulations on a measured Internet AS-level topology to evaluate the proposed algorithms, and we show that faster convergence can be achieved if nCDNs reveal more private information, such as content demands. To the best of our knowledge ours is the first work to consider the design of ex-post individually rational distributed algorithms for CDN interconnection.

The rest of the paper is organized as follows. In Section II we describe the system model and address the complexity of computing an optimal content allocation. Section III shows

Manuscript submitted for review April 29, 2016; revised August 31, 2016 and September 21, 2016; accepted November 28, 2016.

Valentino Pacifici and György Dán are with the Department of Communication Networks, School of Electrical Engineering, and the ACCESS Linnaeus Centre, KTH Royal Institute of Technology. Osquldas väg 10, 10044 Stockholm, Sweden. E-mail: {pacifici,gyuri}@kth.se.

The work was partly funded by the Swedish Research Council through project 621-2014-6035.

that a satisfactory content allocation may not exist without payments. In Section IV we design two distributed algorithms and we prove their convergence. Section V evaluates the proposed algorithms in terms of convergence rate and achieved cost savings. In Section VI we review the related work. Section VII concludes the paper.

II. System Model

We consider a set of autonomous service providers N. Each service provider manages a CDN; we refer to the CDN managed by service provider $i \in N$ as network CDN (nCDN) i. The customers of service provider i generate requests for content items from the set $\mathcal{O}$ of all content items. We make the common assumption that content is divisible in same-sized chunks, thus every item $o \in \mathcal{O}$ has the same size [5], [6]. The customers of service provider i generate requests for content item o at an average rate of $w_i^o \in \mathbb{R}_+$. We denote the set of content items stored by nCDN i by the set $A_i \in \mathcal{A}_i = \{A \subset \mathcal{O} : |A| = K_i\}$, where $K_i \in \mathbb{N}_+$ is the maximum number of items that nCDN i can store. In what follows, we use the terms *nCDN* and *service provider* interchangeably.

We model the relationships between the nCDNs by an undirected graph $\mathcal{G}(N, E)$, called the *interconnection graph*. There is an edge between nCDN i and nCDN j iff they are connected, and we use $\mathcal{N}(i) = \{j|(i,j) \in E\}$. We denote by $A_{-i} = (A_j)_{j \in N\setminus\{i\}}$ the content item allocations of every nCDN other than nCDN i and by $\mathcal{R}_i(A_{-i})$ the set of items that can be retrieved from the nCDNs connected to nCDN i, $\mathcal{R}_i(A_{-i}) \triangleq \bigcup_{j \in \mathcal{N}(i)} A_j$.

We consider that each service provider aims to improve the quality of experience of its customers through decreasing the average access latency to content items. We denote by α_i the unit cost (i.e., access latency) of service provider i for serving an item stored in nCDN i, i.e., locally. If an item is not stored locally at nCDN i, it can be retrieved from one of the nCDNs $\mathcal{N}(i) \subset N$ connected to nCDN i. We denote by β_i^j the unit cost for serving an item from a connected nCDN $j \in \mathcal{N}(i)$. As β_i^j is a model of the access latency, it depends on the infrastructure connecting nCDN i and nCDN j. If item o is available neither locally nor at a connected nCDN, it needs to be retrieved from the origin content provider in the network. We denote by γ_i the unit cost of retrieving an item from the origin content provider. We make the reasonable assumption that it is faster to access an item stored in the local nCDN than to retrieve it from a connected nCDN, and it is faster to retrieve an item stored in a connected nCDN than retrieving it from the content provider, i.e., $\alpha_i < \beta_i^j < \gamma_i$. This assumption is not restrictive, as if $\beta_i^j \geq \gamma_i$, we can remove (i,j) from E.

A. Average Access Latency Cost

We express the cost in terms of average access latency incurred by service provider i in allocation A as

$$C_i(A) = \sum_{o \in \mathcal{O}} C_i^o(A_i, A_{-i}), \tag{1}$$

where $C_i^o(A_i, A_{-i})$ is the cost for accessing item $o \in \mathcal{O}$,

$$C_i^o(A_i, A_{-i}) = w_i^o \begin{cases} \alpha_i & \text{if } o \in A_i \\ \min\limits_{j \in \mathcal{N}(i)} \{\beta_i^j | o \in A_j\} & \text{if } o \in \mathcal{R}_i(A_{-i}) \setminus A_i \\ \gamma_i & \text{otherwise.} \end{cases} \tag{2}$$

Observe that (i) the content allocations of the nCDNs in $\mathcal{N}(i)$ influence the cost of nCDN i through the set A_{-i}, and (ii) if item o is stored at several connected nCDNs then nCDN i retrieves it from the one with lowest unit cost. The cost incurred by service provider i for serving item o can be rewritten as

$$\begin{aligned} C_i^o(A_i, A_{-i}) &= C_i^o(\varnothing, A_{-i}) - (C_i^o(\varnothing, A_{-i}) - C_i^o(A_i, A_{-i})) \\ &= C_i^o(\varnothing, A_{-i}) - CS_i^o(A_i, A_{-i}), \end{aligned}$$

where $CS_i^o(A_i, A_{-i})$ is the cost saving that service provider i achieves by allocating item o given the content allocation at the nCDNs connected to nCDN i. Since the cost $C_i^o(\varnothing, A_{-i})$ is independent of the allocation A_i of nCDN i, finding the minimum cost is equivalent to finding the maximum aggregated cost saving

$$\begin{aligned} \arg\min_{A_i} C_i(A_i, A_{-i}) &= \arg\min_{A_i} \sum_o C_i^o(A_i, A_{-i}) \\ &= \arg\max_{A_i} \sum_o CS_i^o(A_i, A_{-i}). \end{aligned}$$

If nCDN i allocates item o, i.e., $o \in A_i$, then the cost saving can be rewritten as

$$CS_i^o(\{o\}, A_{-i}) = \begin{cases} w_i^o \left[\gamma_i - \alpha_i\right] & \text{if } o \notin \mathcal{R}_i(A_{-i}) \\ w_i^o \left[\beta_i^o(A_{-i}) - \alpha_i\right] & \text{if } o \in \mathcal{R}_i(A_{-i}) \end{cases} \tag{3}$$

where $\beta_i^o(A_{-i})$ is the lowest unit cost at which nCDN i can retrieve item o from a connected nCDN

$$\beta_i^o(A_{-i}) \triangleq \min_{j \in \mathcal{N}(i)} \{\beta_i^j | o \in A_j\}. \tag{4}$$

If instead $o \notin A_i$, the cost saving $CS_i^o(A_i, A_{-i}) = 0$. Hence, $CS_i(A_i, A_{-i}) = \sum_{o \in \mathcal{O}} CS_i^o(A_i, A_{-i}) = \sum_{o \in A_i} CS_i^o(A_i, A_{-i})$. Observe that finding the minimum cost for service provider $i \in N$ corresponds to solving a knapsack problem where the values of the items are their cost savings given the allocations A_{-i} of the other nCDNs, and where the total weight is K_i.

B. Utility Model

We consider that bilateral compensations between nCDNs are possible, if they are necessary to implement a content allocation that serves all service providers' interests. The bilateral compensations model monetary transfers between federated nCDNs, that would settle such payments periodically, similar to peering agreements. We model a monetary payment p_i^j from service provider i to service provider j as an additive term in the utility function of nCDNs i and j, thus we can define the utility function of nCDN i in allocation A

$$U_i(A, \boldsymbol{p}) = CS_i(A_i, A_{-i}) + \sum_{j \in N\setminus\{i\}} (p_j^i - p_i^j). \tag{5}$$

Observe that the utility is determined by the content allocation A and by the payment allocation $\boldsymbol{p} = (p_i^j)_{i,j \in N}$.

C. Complexity of Aggregate Utility Maximization

Before we turn to the problem formulation, let us consider the problem of maximizing the aggregate utility of the nCDNs. This optimization problem would have to be solved by a central coordinator, if such a coordinator existed, in order to enforce or to recommend allocations to the individual nCDNs. Solving this optimization problem is also central to cooperative game theoretic solutions, as we discuss later. To formulate the optimization problem, observe that the aggregate utility $U(A) = \sum_{i\in N} U_i(A, \boldsymbol{p})$ does not depend on the payment allocation $\boldsymbol{p}$, and thus maximizing the aggregate utility is equivalent to minimizing the average access latency C_i. An optimal allocation would then be

$$\tilde{A} \in \arg\max_{A} \sum_{i\in N} U_i(A, \boldsymbol{p}) = \arg\max_{A} \sum_{i\in N} \boldsymbol{CS}_i(A_i, A_{-i}). \tag{6}$$

In the following we show that solving (6) is NP-hard.

Theorem 1. *Computing allocation $\tilde{A}$ that minimizes the average access latency is NP-hard in general.*

Proof. We prove the theorem by showing that the problem of finding the chromatic number $\chi(\mathcal{G})$ of a graph $\mathcal{G} = (V, E)$ can be reduced to finding an allocation $\tilde{A}$ that solves (6).

A vertex coloring of a graph $\mathcal{G}(V, E)$ is a mapping $c : V \mapsto S$. We denote by $c^{-1}(s) = \{v \in V | c(v) = s,\ s \in S\}$ the inverse mapping of c. A proper coloring c is such that $c(i) \neq c(j)$ for all $(i, j) \in E$. The chromatic number of graph $\mathcal{G}$, denoted by $\chi(\mathcal{G})$, is the smallest number of colors such that a proper coloring of graph $\mathcal{G}$ exists.

We first show that every proper coloring of $\mathcal{G} = (V, E)$ corresponds to an allocation $A = (A_i)_{i\in V}$ of content items to the nCDNs $i \in V$ in which each nCDN has unit storage capacity, $K_i = 1$ and non-zero cost saving, i.e., $\boldsymbol{CS}^o_i(A_i, A_{-i}) > 0$. For a coloring of $\mathcal{G} = (V, E)$, we let $N = V$, and the interconnection graph $\mathcal{G} = (N, E)$. We set $K_i = 1$ for all $i \in \mathcal{G}$ and use homogeneous costs among nCDNs, i.e., $\beta^j_i = \beta$, $\alpha_i = \alpha$, $\gamma_i = \gamma$ for all $i, j \in V$. In addition, we let $|\mathcal{O}| = |V|$ and $\beta = \alpha$. Observe that $\beta = \alpha$ implies that for all nCDNs $i \in N$

$$\boldsymbol{CS}^o_i(A_i, A_{-i}) > 0 \Rightarrow o \in A_i, o \notin \mathcal{R}_i(A_{-i}). \tag{7}$$

Finally, we let the average rates of customer requests for content items be homogeneous among the nCDNs (i.e. $w^o_i = w^o$ $\forall i \in N$) but different among items, $w^o \neq w^p$ $\forall o \neq p$, $o, p \in \mathcal{O}$.

To every coloring c of $\mathcal{G} = (V, E)$ we define the corresponding allocation A^c as $A^c_i = c(i)$, $S^c \subseteq \mathcal{O}$. Such an allocation exists since $|\mathcal{O}| = |V|$. Analogously, we denote by $c^A : V \mapsto S^A$ the coloring corresponding to allocation A. If c is proper, it follows from the definition of A^c that $A^c_i \cap A^c_j = \varnothing$ $\forall (i, j) \in E$ and therefore $\boldsymbol{CS}^o_i(A^c_i, A^c_{-i}) = w^o[\gamma - \alpha] > 0$ for item $o \in A^c_i$ and $\forall i \in V$. Consequently, the aggregate utility $U(A^c)$ can be expressed in the form

$$U(A^c) = \sum_{i\in V} \sum_{o\in A^c_i} w^o[\gamma - \alpha] = [\gamma - \alpha] \sum_{o\in S} w^o |c^{-1}(o)|. \tag{8}$$

Consider now a solution $\tilde{A}$ of (6). Since $|\mathcal{O}| = |V|$, $\boldsymbol{CS}^o_i(\tilde{A}_i, \tilde{A}_{-i}) > 0$ for item $o \in \tilde{A}_i$ for all nCDNs $i \in N$. Therefore, it follows from (7), that a coloring $c^{\tilde{A}} : V \mapsto S^{\tilde{A}}$ is proper and $U(\tilde{A})$ is in the form (8). Since $\tilde{A}$ is optimal, it follows from (8) that

$$w^o > w^p \Leftrightarrow |c^{\tilde{A}^{-1}}(o)| \geq |c^{\tilde{A}^{-1}}(p)| \ \ \forall o, p \in S^{\tilde{A}}, \tag{9}$$

as otherwise switching item o with item p in $\tilde{A}$ would lead to higher aggregate utility. To conclude the proof, let us define the vector $\tilde{\boldsymbol{c}} = (|c^{\tilde{A}^{-1}}(o)|)_{o\in S^{\tilde{A}}}$, which contains in decreasing order for each item the number of times it is replicated.

Clearly $|S^{\tilde{A}}| \geq \chi(\mathcal{G})$, since $c^{\tilde{A}}$ is proper. In the following we show by contradiction that $|S^{\tilde{A}}| = \chi(\mathcal{G})$. Assume $|S^{\tilde{A}}| > \chi(\mathcal{G})$. Since $\sum_{o\in\mathcal{O}} |c^{-1}(o)| = |V|$, then there exists a proper coloring $\bar{c}$ such that $\bar{\boldsymbol{c}} >_L \tilde{\boldsymbol{c}}$ (i.e., it is lexicographically bigger). Therefore $U(\tilde{A}) < U(A^{\bar{c}})$, which contradicts the assumption and implies $|S^{\tilde{A}}| = \chi(\mathcal{G})$. Therefore, given $\tilde{A}$, it is possible to compute $\chi(\mathcal{G})$ as $|S^{\tilde{A}}|$ in the coloring $c^{\tilde{A}}$. □

Besides being NP-hard, computing an allocation $\tilde{A}$ that maximizes the aggregate utility is also impractical because it requires the service providers to disclose business confidential information such as their content demands.

D. Problem Formulation

If the service providers do not cooperate, i.e., the set of connected nCDNs $\mathcal{N}(i) = \varnothing$ for every service provider i, service provider i would optimize the content allocation in nCDN i in isolation, and would prefetch the K_i items with highest demands. We denote the resulting allocation, which is *optimal in isolation*, by A^I_i. The corresponding cost is $C_i(A^I_i, \varnothing) = \sum_{o\in A^I_i} w^o_i \alpha_i + \sum_{o\in\mathcal{O}\setminus A^I_i} w^o_i \gamma_i$.

Cooperation could allow service providers to decrease their average access latency cost compared to isolation. For an allocation A and a payment allocation $\boldsymbol{p} = (p^j_i)_{i,j\in N}$, we define the cost saving gain as

$$r_i(A) = \frac{C^I_i(\varnothing) - C_i(A) + \sum_{j\in N\setminus\{i\}} (p^i_j - p^j_i)}{C^I_i(\varnothing) - C_i(A^I_i, \varnothing)} \tag{10}$$

where $C^I_i(\varnothing) = \sum_{o\in\mathcal{O}} w^o_i \gamma_i$ is the cost incurred by service provider i with no nCDN. We call an allocation A *individually rational* if $r_i(A) \geq 1$. Observe that service provider i benefits from cooperating only if $r_i(A) > 1$.

Since there is no central authority, cooperation requires a *distributed algorithm* that (i) needs information exchange between connected service providers only, (ii) reveals little private information such as content demands, (iii) in a *finite* number of steps leads to a content allocation A and (iv) is *ex-post individually rational* for all service providers.

III. Stable Allocations Without Payments

Without payments, a content allocation among interconnected nCDNs would have to let every nCDN i allocate content items that minimize its cost C_i, given the allocations of its connected nCDNs $\mathcal{N}(i)$. Such an allocation is self-enforcing, as no nCDN could gain by deviating from it. Modeling the interaction of nCDNs as a strategic game $\Gamma =< N, (\mathcal{A}_i)_{i\in N}, (U_i)_{i\in N} >$, where the utility of player i is the sum of its cost savings $U_i(A_i, A_{-i}) = \boldsymbol{CS}_i(A_i, A_{-i})$,

At time step t:

1) nCDN $i_t \in N$ chooses a content allocation $A_{i_t}(t)$ such that $CS_{i_t}(A_{i_t}(t), A_{-i_t}(t-1)) > CS_{i_t}(A(t-1))$.
2) nCDN i_t communicates to $\forall j \in \mathcal{N}(i_t)$ the sets of evicted and inserted items, $E_{i_t}(t)$ and $I_{i_t}(t)$, respectively.

Fig. 1: Pseudocode of the *Local-Greedy* algorithm.

such a content allocation corresponds to a pure strategy Nash equilibrium A^* of Γ, i.e., a set of allocations $(A_i^*)_{i \in N}$ such that

$$A_i^* \in \arg\max_{A_i} U_i(A_i, A_{-i}^*). \tag{11}$$

It follows from (2) that $C_i(A^*) \leq C_i(A_i^I, A_{-i}^*) \leq C_i(A_i^I, \varnothing)$, therefore A^* is individually rational.

Given an initial allocation of content items, $(A_i)_{i \in N}$, a distributed algorithm that might be used to compute A^* and one that reveals little private information is the *Local-Greedy* algorithm shown in Fig. 1. According to the *Local-Greedy* algorithm, at time step t a single nCDN i_t can update its allocation from $A_{i_t}(t-1)$ to an allocation $A_{i_t}(t)$ that increases its cost saving given the allocations of the other nCDNs $A_{-i_t}(t-1)$. The *Local-Greedy* algorithm requires little signaling: upon time step t nCDN i_t has to send the set $E_{i_t}(t) \triangleq A_{i_t}(t-1) \setminus A_{i_t}(t)$ of evicted and the set $I_{i_t}(t) \triangleq A_{i_t}(t) \setminus A_{i_t}(t-1)$ of inserted items to its neighboring nCDNs. The *Local-Greedy* algorithm terminates when no nCDN i can increase its cost saving by updating its allocation. By definition (11), if the *Local-Greedy* algorithm terminates, then the content allocation reached by the nCDNs is a pure strategy Nash equilibrium A^* of Γ.

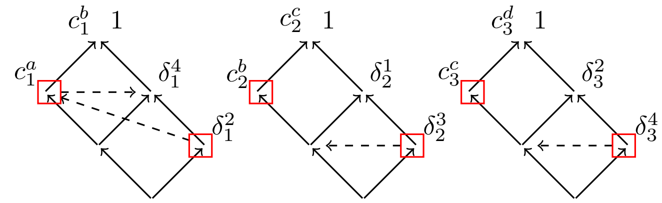
In the particular case that the unit cost β_i^j of serving a request from a connected nCDN is not neighbor specific, it was shown in (Theorem 1 in [7]) that Γ has a pure strategy Nash equilibrium, which allows us to formulate the following.

Proposition 2 ([7])**.** *If the link costs β_i^j are neighbor-homogeneous, i.e., $\beta_i^j = \beta_i$ $\forall i \in N$, $\forall j \in \mathcal{N}(i)$, then the strategic game Γ has a pure strategy Nash equilibrium.*

Proposition 2 is so far the most general sufficient condition for the existence of a Nash equilibrium of Γ, but the assumption of neighbor-homogeneous costs is hard to justify in the case of link costs modeling access latency. For the general case, i.e., if link costs are not neighbor-homogeneous, it is not known whether (i) an equilibrium allocation always exists and whether (ii) the *Local-Greedy* algorithm would lead to an equilibrium even if it exists. In what follows we show that there are instances of the content allocation problem for which an equilibrium allocation A^* that satisfies (11) does *not exist*.

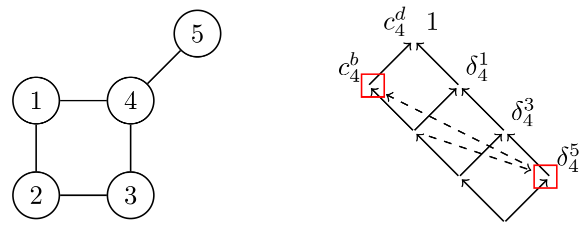
A. Non-Existence of Equilibrium Content Allocations

The strategic game Γ can be interpreted as a resource allocation game where the resources are the items, $c_i^o \triangleq w_i^o[\gamma_i - \alpha_i] \in \mathbb{R}_+$ is the value of resource o for player i and $0 < \delta_i^j \triangleq \frac{\beta_i^j - \alpha_i}{\gamma_i - \alpha_i} < 1$ is the penalty due to sharing the resource with player j. The expression of the cost saving in (3) becomes



(a) (b) (c)

Fig. 2: Cost saving graphs of nCDNs 1 to 3 in Example 1. The squares show the cost savings of each nCDN given the content allocation of its neighbors in the content allocation (a, b, c, d, d).



(a) (b)

Fig. 3: Interconnection graph (a) and cost saving graph (b) of nCDN 4 in Example 1. The squares show the cost savings of nCDN 4 in the content allocation (a, b, c, d, d).

$$CS_i^o(\{o\}, A_{-i}) = \begin{cases} c_i^o & \text{if } o \notin \mathcal{R}_i(A_{-i}) \\ c_i^o \min\limits_{j \in \mathcal{N}(i)} \{\delta_i^j | o \in A_j\} & \text{if } o \in \mathcal{R}_i(A_{-i}). \end{cases} \tag{12}$$

Observe that the cost incurred by player i for retrieving item o depends on which neighboring players store item o, not only on whether any neighboring player stores it as in [8], [9], [7]. As a consequence, results on the existence of Nash Equilibria in player-specific graphical congestion games do not apply.

Consider the following example.

Example 1. *Consider nCDNs $N = \{1, \ldots, 5\}$ and the set $\mathcal{O} = \{a, b, c, d\}$ of content items. The nCDNs are interconnected according to the graph in Figure 3a. For nCDN 5*

$$c_5^d \delta_4^5 > c_5^o \quad \forall o \in \mathcal{O} \setminus \{d\}. \tag{13}$$

For nCDN 1 the demands and the costs satisfy

$$\delta_1^2 < \delta_1^4, \; c_1^a < c_1^b, \tag{14}$$

$$c_1^b \delta_1^2 < c_1^a < c_1^b \delta_1^4. \tag{15}$$

Inequalities (14-15) specify a lattice (a poset with least and greatest element) over the cost savings $CS_1^o(\{o\}, A_{-1})$, which is shown in Figure 2a; we call it the cost saving graph. *An arrow between two cost savings points towards the greater of the two.*

The lattice is on the one hand the product of two totally ordered sets (solid arrows): values $\{c_i^o \,|o \in \mathcal{O}\}$ and link costs $\{\delta_i^j | j \in \mathcal{N}(i)\} \cup \{1\}$. The greatest element of the lattice is the cost saving c_i^o of the item $o \in \mathcal{O}$ with highest rate w_i^o at nCDN i when it is not allocated by any connected nCDN $j \in \mathcal{N}(i)$, i.e. $o \notin \mathcal{R}_i(A_{-i}) \Rightarrow CS_i^o(\{o\}, A_{-i}) = c_i^o$. The least element of the lattice is the cost saving $c_i^p \delta_i^j$ of the item $p \in \mathcal{O}$ with lowest rate w_i^p when it is allocated by the connected nCDN $j \in \mathcal{N}(i)$ such that $j = \arg\min_{k \in \mathcal{N}(i)} \delta_i^k$. On the other hand, the lattice is specified through additional inequalities, *such as (15) for nCDN 1 (dashed lines).*

The cost saving graphs for nCDNs $i \in \{2,3,4\}$ are shown in Figure 2b, 2c and 3b, respectively. The squares in Figures 2 and 3b represent the cost savings $CS_i^o(A_i, A_{-i})$ of the corresponding nCDN $i \in \{1,2,3,4\}$ at content allocation $A = (a,b,c,d,d)$. We omit the relations between cost savings that are not relevant for the example.

We are now ready to prove the following.

Theorem 3. *There are instances of the strategic game Γ that do not possess a pure strategy Nash equilibrium. Consequently, an equilibrium allocation A^* that satisfies (11) does not always exist.*

Proof. We prove the theorem by showing that the game described in Example 1 does not possess a Nash equilibrium. From (13) it follows that no content allocation A such that $A_5 \neq \{d\}$ is a Nash equilibrium. We can furthermore restrict the action set of each nCDN to the items that are included in its cost saving graph, according to the following

$$\mathcal{A}_1 = \big\{\{a\},\{b\}\big\} \quad \mathcal{A}_2 = \big\{\{b\},\{c\}\big\}$$
$$\mathcal{A}_3 = \big\{\{c\},\{d\}\big\} \quad \mathcal{A}_4 = \big\{\{d\},\{b\}\big\}$$

This results in a total of 16 possible content allocations. From the cost saving graphs in Figures 2 and 3b it follows that any content allocation where two interconnected nCDNs store the same item is not a Nash equilibrium; there are 12 such allocations. Therefore, it is enough to focus on the 4 content allocations in which there is no pair of interconnected nCDNs that store the same item.

The following sequence of content allocations starts with these four content allocations (marked with bold) and shows a cycling sequence of updates of nCDNs that follow the *Local-Greedy* algorithm. We omit nCDN 5, which always stores item (d).

$$(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d}) \underset{4}{\rightarrow} (\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{b}) \underset{3}{\rightarrow} (\boldsymbol{a},\boldsymbol{b},\boldsymbol{d},\boldsymbol{b})$$
$$\underset{2}{\rightarrow} (\boldsymbol{a},\boldsymbol{c},\boldsymbol{d},\boldsymbol{b}) \underset{1}{\rightarrow} (b,c,d,b) \underset{4}{\rightarrow} (b,c,d,d)$$
$$\underset{3}{\rightarrow} (b,c,c,d) \underset{2}{\rightarrow} (b,b,c,d) \underset{1}{\rightarrow} (a,b,c,d)$$

There is thus no NE, which proves the theorem. □

In the following we consider two assumptions on the link costs between service providers that can be considered reasonable. We show that, alas, neither of them is sufficient for equilibrium content allocations to exist.

B. Symmetric Link Costs

The first assumption concerns the relationship between the link costs δ_i^j and δ_j^i of interconnected nCDNs. According to recent measurement studies [10], [11], the delay between hosts tend to be symmetric, thus it may be reasonable to assume that link costs are symmetric between interconnected nCDNs, i.e., $\delta_i^j = \delta_j^i\ \forall j \in \mathcal{N}(i)$. In the case of Example 1, the requirement of symmetric link costs implies a feasible total order on the link costs, $1 > \delta_1^4 > \delta_1^2 > \delta_2^3 > \delta_3^4 > \delta_4^5$, which leads to the following.

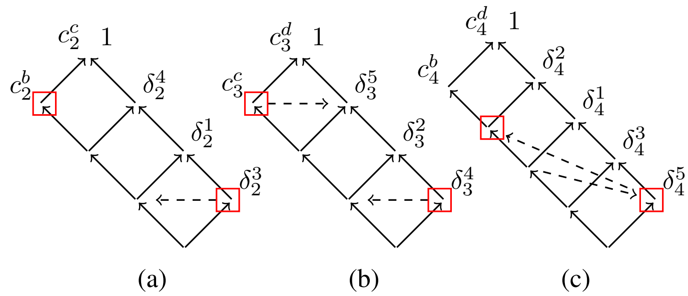


Fig. 4: Cost saving graphs of nCDNs 2 (a), 3 (b) and 4 (c) as described in Section III-C. The squares show the cost savings of each nCDN given the content allocation of its neighbors in the content allocation (a,b,c,d,d).

Corollary 1. *There are instances of the strategic game Γ with symmetric link costs that do not possess a pure strategy Nash equilibrium.*

A corresponding non-existence result for the case of a linear cost function and a directed interconnection graph was provided in [12]. Observe that, in our model, a link (i,j) is effectively directed if δ_i^j is sufficiently smaller than δ_j^i, such that the content allocated at nCDN i does not affect the allocation that minimizes the cost function of nCDN j. The importance of Corollary 1 is that it extends the non-existence results to undirected interconnection graphs.

C. Link Costs Satisfying the Triangle Inequality

Second, let us consider that the link costs on the interconnection graph $\mathcal{G}$ satisfy the triangle inequality. This assumption might be reasonable both in the case of latency and traffic cost minimization, as measurement studies showed that the triangle inequality is violated by less than 4% of host triplets in the Internet [13]. The triangle inequality implies that for every service provider $i \in N$ the following holds

$$\delta_i^j + \delta_j^k \geq \delta_i^k\ \forall j \in \mathcal{N}(i). \tag{16}$$

Corollary 2. *There are instances of the strategic game Γ with symmetric link costs satisfying the triangle inequality that do not possess a pure strategy Nash equilibrium.*

Proof. Let us impose constraint (16) on the link costs of Example 1. The interconnection graph in Figure 3a contains now 5 new edges, namely $(1,3)$, $(1,5)$, $(2,5)$, $(2,4)$ and $(3,5)$. Note that the items included in the cost saving graphs of service providers 1 and 3 in Figure 2 form disjoint sets, hence the link $(1,3)$ does not change the cost saving graphs. The same holds for the links $(1,5)$ and $(2,5)$. Hence we restrict our analysis to the links $(2,4)$ and $(3,5)$. From (12) it follows that the addition of a link (i,j) between nCDNs i and j can only decrease the cost savings CS_i and CS_j. Choosing δ_2^4 and δ_3^5 large enough, as shown in the cost saving graphs of Figure 4, would therefore not break any of the arguments of the proof of Theorem 3.

Assuming also symmetric link costs, the cost saving graph of nCDN 1 in Figure 2 and the cost saving graphs in Figure 4 imply the following inequalities

$$\delta_2^4 > \delta_1^4 > \delta_1^2 > \delta_2^3 > \delta_3^4 > \delta_4^5, \quad (17)$$

$$\delta_3^5 > \delta_2^3. \quad (18)$$

In addition to the above lower-bounds on δ_2^4 and δ_3^5, the triangle inequality in (16) introduces the following upper-bounds.

$$\delta_2^4 \leq \min(\delta_1^2+\delta_1^4, \delta_2^3+\delta_3^4) = \delta_2^3+\delta_3^4 \quad (19)$$

$$\delta_3^5 \leq \delta_3^4+\delta_4^5. \quad (20)$$

It is easy to see that a set of link costs satisfying (17)-(20) exists.

□

Theorem 3 and Corollary 1 and 2 imply that if payments are not allowed then cost minimizing nCDNs may not be able to compute a content allocation, and no distributed algorithm, including *Local-Greedy*, would ever terminate. Since it is infeasible to determine a-priori whether an equilibrium allocation exists (for reasons of computational complexity and because doing so would require global knowledge), computing a content allocation must involve payments to guarantee finite execution time.

IV. Bilateral Compensation-based Allocation

A natural solution involving payments would be to model the problem as a cooperative game with transferable utility and use an existing solution concept, such as the Shapley value [14], for computing the compensations.

A. Cooperative solution concepts are infeasible

To model the problem as a cooperative game, we define a coalition as a subset $S \subseteq N$ of service providers connecting their nCDNs so as to decrease their average access latency. The nCDNs in coalition S are interconnected according to the subgraph $\mathcal{G}_S = (S, E_S)$ induced on $\mathcal{G}$ by S. The cooperative game (N, v) is defined by the value function $v : 2^N \mapsto \mathbb{R}$ that maps any coalition S to a real value $v(S)$. As the service providers aim at minimizing the average access latency to content items, it is natural to define the value function $v(S)$ of coalition $S \subseteq N$ as the maximum cost saving achievable by the set of players that form the coalition,

$$v(S) = \sum_{i\in S} C_i^I(\varnothing) - \min_{A_S} \sum_{i\in S} C_i(A_S, \varnothing). \quad (21)$$

The value $v(N)$ of the grand coalition N can be shared among the members of the coalition through compensations to form an imputation, i.e., an efficient and individually rational utility vector. As we next show, an imputation exists.

Lemma 1. *The Shapley value for the coalitional game (N, v) is individually rational.*

Proof. To prove this result, we have to show that the value function v is superadditive, i.e., for any two coalitions $S, S' \subset N$ and $S \cap S' = \varnothing$, it holds that $v(S \cup S') - v(S) - v(S') \geq 0$. For a coalition $T \subseteq N$ let us define $\bar{A}^T \triangleq \arg\min_{A_T} \sum_{i\in T} C_i(A_T, \varnothing)$ as the content allocation achieving the minimum aggregate cost for the players in coalition T. Using the expression of the value function (21) we obtain

$$v(S\cup S') - v(S) - v(S') =$$
$$-\min_{A_{S\cup S'}} \sum_{i\in S\cup S'} C_i(A_{S\cup S'}, \varnothing) + \sum_{i\in S} C_i(\bar{A}^S, \varnothing) + \sum_{i\in S'} C_i(\bar{A}^{S'}, \varnothing).$$

Observe that (2) implies that for any two coalitions $S, S' \subset N$ such that $S \cap S' = \varnothing$ and any $A \in \mathcal{A}$,

$$\sum_{i\in S} C_i(A_S, \varnothing) + \sum_{i\in S'} C_i(A_{S'}, \varnothing) \geq \sum_{i\in S\cup S'} C_i(A_{S\cup S'}, \varnothing).$$

Hence, for $\bar{A}^S$, $\bar{A}^{S'}$ and the aggregate allocation $(\bar{A}^S, \bar{A}^{S'}) \triangleq ((\bar{A}_i^S)_{i\in S}, (\bar{A}_i^{S'})_{i\in S'})$ we obtain

$$\sum_{i\in S} C_i(\bar{A}^S, \varnothing) + \sum_{i\in S'} C_i(\bar{A}^{S'}, \varnothing) \geq \sum_{i\in S\cup S'} C_i((\bar{A}^S, \bar{A}^{S'}), \varnothing)$$
$$\geq \min_{A_{S\cup S'}} \sum_{i\in S\cup S'} C_i(A_{S\cup S'}, \varnothing),$$

which implies $v(S \cup S') - v(S) - v(S') \geq 0$ and proves the lemma. □

Nonetheless, computing the Shapley value and other cooperative solutions that rely on the characteristic function is infeasible for two reasons. First, computing the characteristic function of a coalition requires a single entity to know all content demands. Second, observe that computing the value $v(S)$ of coalition S as defined in (21) corresponds to solving (6) for the set S of nCDNs, and is thus NP-hard as shown in Theorem 1.

Motivated by the failure of existing solution concepts, in the following we propose two distributed algorithms that involve *bilateral* compensations for computing an individually rational content allocation. The two algorithms differ in the amount of revealed private information, in the level of parallelism that they allow and, as we will see, in terms of convergence rate.

B. Aggregate-value Compensation Algorithm

Following the aggregate value compensation (AC) algorithm, at every time step t there is a set $N_t \subseteq N$ of nCDNs that is allowed to update its content allocation. Given an allocation of content items $A(t-1)$, an update made by nCDN $i_t \in N_t$ from $A_{i_t}(t-1)$ to $A_{i_t}(t)$ can result in an increase of the cost (1) for one or more connected CDNs $j \in \mathcal{N}(i_t)$. According to the AC algorithm, an nCDN $j \in \mathcal{N}(i_t)$, $i_t \in N_t$, that would suffer an increase of cost $C_j(A_{N_t}(t), A_{-N_t}(t-1)) > C_j(A(t-1))$, offers a compensation $p_j^{i_t}(t)$ to a nCDN $i_t \in \mathcal{N}(j) \cap N_t$ equal to its cost increase

$$p_j^{i_t}(t) = \Delta C_j(t) \triangleq C_j(A_{i_t}(t), A_{-i_t}(t-1)) - C_j(A(t-1)).$$

We use $D_t \subseteq \mathcal{N}(i_t)$ to denote the set of nCDNs that offer a compensation,

$$j \in D_t \Leftrightarrow \Delta C_j(t) > 0. \quad (22)$$

The compensations are used to deter nCDNs from performing updates: nCDN $i_t \in N_t$ performs the update despite the offered compensation if the aggregate compensation offered by all connected nCDNs is lower than the gain it achieves from updating the content allocation from $A_{i_t}(t-1)$ to $A_{i_t}(t)$. We

At time step t:

1) Every nCDN $i_t \in N_t$ computes a content allocation $A^P_{i_t}(t)$ s.t. $\boldsymbol{CS}_{i_t}(A^P_{i_t}(t), A_{-i_t}(t-1)) > \boldsymbol{CS}_{i_t}(A(t-1))$.
2) Every nCDN $i_t \in N_t$ communicates to $\forall j \in \mathcal{N}(i_t)$ the set of items it plans to evict $E_{i_t}(t)$ and insert $I_{i_t}(t)$.
3) Every nCDN $j \in \mathcal{N}(N_t)$ s.t. $\Delta C_j(t) > 0$ offers a compensation $p^{i_t}_j(t) = \Delta C_j(t)$ to one nCDN $i_t \in \mathcal{N}(j) \cap N_t$
4) **If** $\sum_{j \in D_t} p^{i_t}_j(t) \geq -\Delta C_{i_t}(t)$, then nCDN i_t accepts the compensation and it does not make the update, i.e., $A_{i_t}(t) = A_{i_t}(t-1)$.
Otherwise nCDN i_t refuses the compensation and updates its allocation from $A_{i_t}(t-1)$ to $A_{i_t}(t) = A^P_{i_t}(t)$.
5) Every nCDN $i_t \in N_t$ communicates to $\forall j \in \mathcal{N}(i_t)$ its decision, i.e. whether $A_{i_t}(t) = A_{i_t}(t-1)$ or $A_{i_t}(t) = A^P_{i_t}(t)$.

Fig. 5: Aggregate-value Compensation (AC) Algorithm

call this the *Aggregate-value Compensation* (AC) algorithm, and we show its pseudo-code in Figure 5.

Observe that the AC algorithm does not specify how N_t is chosen at each time step t. Before proving convergence for specific choices of N_t, we make the following definition.

Definition 1. A sequence $N_t \subseteq N$, $t = 1, \dots$ of sets of nCDNs is a complete sequence, if for all t and each nCDN $i \in N$ there exists a time step $t' > t$ such that $i \in N_{t'}$.

1) **Asynchronous operation:** Let us first consider that only one nCDN $i_t \in N_t$ is allowed to update its allocation at each time slot t. Thus, the sets N_{t_1}, N_{t_2}.. are singletons and nCDN i_t is the only recipient of the compensation of each nCDN $j \in D_t$. The following result shows that the AC algorithm converges if used asynchronously.

Theorem 4. *Let N_t be a complete sequence of singleton sets and every nCDN use the* AC *algorithm. We refer to this as the* 1-AC *algorithm. The* 1-AC *algorithm converges to an allocation $\boldsymbol{A}$ of content items to interconnected nCDNs in a finite number of time steps.*

Proof. We prove the theorem by showing that the aggregate cost $C(A) \triangleq \sum_{i \in N} C_i(A_i, A_{-i})$ strictly decreases at every update made by any nCDN following the 1-AC algorithm.

Consider a nCDN i_t that updates its content allocation from $A_{i_t}(t-1)$ to $A_{i_t}(t)$ at time step t. It follows from (2) that $C_k(A(t)) = C_k(A(t-1))$ for any nCDN $k \notin \mathcal{N}(i_t)$. We can calculate the aggregate cost function $C(A(t))$ after the update of nCDN i_t as

$$C(A(t)) = C(A(t-1)) + \Delta C_{i_t}(t) + \sum_{j \in D_t} \Delta C_j(t) + \sum_{j \in \mathcal{N}(i_t) \setminus D_t} \Delta C_j(t).$$

From (22) it follows that $\sum_{j \in \mathcal{N}(i_t) \setminus D_t} \Delta C_j(t) \leq 0$. Moreover, since nCDN i_t refused the compensation offered by the connected nCDNs in D_t, it follows that $\Delta C_{i_t}(t) + \sum_{j \in D_t} \Delta C_j(t) < 0$. Hence, at every update of the 1-AC algorithm $C(A(t)) < C(A(t-1))$. Since the set of all content allocations is finite and the sequence N_t is complete, this proves the theorem. □

A significant shortcoming of the 1-AC algorithm is that it requires global synchronization. Furthermore, if nCDN i_t is chosen uniformly at random at every time step t, the probability that nCDN i_t can decrease its cost C_{i_t} by updating its content allocation $A_{i_t}(t-1)$ at time step t decreases as the 1-AC approaches allocation $\boldsymbol{A}$. As a consequence, the convergence of the 1-AC algorithm may be slow.

2) **Plesiochronous operation:** In the following we show that convergence can be guaranteed even if the sets N_t are not singletons. Before we formulate our result, we recall the following definition from graph theory.

Definition 2. A *k-independent set* $\mathcal{I}^k$ of a graph $\mathcal{G} = (N, E)$ is a subset $\mathcal{I}^k \subseteq N$ of the vertexes of $\mathcal{G}$ such that the distance between any two vertexes of $\mathcal{I}^k$ in $\mathcal{G}$ is at least $k+1$. We denote by $\mathfrak{I}^k$ the set of all the k-independent sets of the interconnection graph $\mathcal{G}$.

We can now prove the following.

Theorem 5. *Let N_t be a complete sequence of* 2*-independent sets and every nCDN use the AC algorithm. We refer to this as the $\mathcal{I}^2$-*AC *algorithm. The $\mathcal{I}^2$-*AC *algorithm converges to an allocation $\boldsymbol{A}$ of content items to interconnected nCDNs in a finite number of time steps.*

Proof. Consider a nCDN $j \in \mathcal{N}(i_t)$, connected to $i_t \in \mathcal{I}^2_t$. From the definition of 2-independent set follows that i_t is the only nCDN in $\mathcal{N}(j)$ that is allowed to update its content allocation A_{i_t} at time step t. Hence, it is possible to compute the aggregate cost function $C(A(t))$ from $C(A(t-1))$ as follows

$$C(A(t)) = C(A(t-1)) \; + \\ + \sum_{i_t \in U_t} \Big(\Delta C_{i_t}(t) + \sum_{j \in D_t} \Delta C_j(t) + \sum_{j \in \mathcal{N}(i_t) \setminus D_t} \Delta C_j(t)\Big),$$

where $U_t \subseteq \mathcal{I}^2_t$ is the set of nCDNs i_t such that $A_{i_t}(t) \neq A_{i_t}(t-1)$. From the same argument in the proof of Theorem 4 it follows that $C(A(t)) < C(A(t-1))$ at every update of the $\mathcal{I}^2$-AC algorithm. □

3) **Equilibria under the AC algorithm:** So far we showed that the AC algorithm converges to a content allocation $\boldsymbol{A}$ in a finite number of time steps if N_t is a complete sequence of singleton sets or 2-independent sets. In the following we provide a characterization of the set of content allocations that it converges to.

Proposition 6. *Let us define the utility function $G_i(A) \triangleq CS_i(A) + \sum_{j \in \mathcal{N}(i)} CS_j(A)$ and the strategic game $\Gamma' \triangleq < N, (\mathcal{A}_i)_{i \in N}, (G_i)_{i \in N} >$. A content allocation $\boldsymbol{A}$ computed by the* AC *algorithm is a Nash equilibrium of the strategic game Γ'.*

Proof. The proof follows from the condition at stage 4) of the AC algorithm in Figure 5, which implies that an update of nCDN i_t at time step t decreases the aggregate cost of nCDN i_t and its connected nCDNs $\mathcal{N}(i_t)$. □

We can provide an alternative characterization of the set of content allocations reached by the AC algorithm based on the utility function (5). Let us define the set of possible

At time step t:

1) Every nCDN $i_t \in N_t$ computes a content allocation $A^P_{i_t}(t)$ such that $CS_{i_t}(A^P_{i_t}(t), A_{-i_t}(t-1)) > CS_{i_t}(A(t-1))$.
2) Every nCDN $i_t \in N_t$ communicates to $\forall j \in \mathcal{N}(i_t)$ the set of items it plans to evict $E_{i_t}(t)$ and insert $I_{i_t}(t)$.
3) Every nCDN $j \in \cup_{i_t \in N_t}\mathcal{N}(i_t)$ calculates
$$\Delta\widetilde{C}^o_j(t) = C^o_j(A^P_{N_t}(t), A_{-N_t}(t-1)) - C^o_j(A(t-1))$$
for all $o \in \cup_{N_t} E_{i_t}$. If $\Delta\widetilde{C}^o_j(t) > 0$, nCDN j offers to a nCDN $k \in N_t$ such that $o \in E_k$ and $\beta^k_j = \beta^o_j(A_{-k}(t))$ a compensation $p^o_{j,k}(t) \triangleq \Delta\widetilde{C}^o_j(t)$.
4) **If** $\sum_{j \in D_{i_t}} \sum_{o \in E_{i_t}} p^o_{j,i_t}(t) \geq -\Delta C_{i_t}(t)$, then nCDN i_t accepts the offer and it does not make the update, i.e., $A_{i_t}(t) = A_{i_t}(t-1)$.
Otherwise nCDN i_t refuses the offer and updates its allocation from $A_{i_t}(t-1)$ to $A_{i_t}(t) = A^P_{i_t}(t)$.
5) Every nCDN $i_t \in N_t$ communicates to $\forall j \in \mathcal{N}(i_t)$ its decision, i.e. whether $A_{i_t}(t) = A_{i_t}(t-1)$ or $A_{i_t}(t) = A^P_{i_t}(t)$.

Fig. 6: Object-value Compensation (OC) Algorithm

updates of nCDN i in allocation A as $PU_i(A) \triangleq \{A'_i \in \mathcal{A}_i | CS_i(A'_i, A_{-i}) > CS_i(A)\}$.

Proposition 7. *A content allocation $\boldsymbol{A}$ computed by the* AC *algorithm is a Nash equilibrium of the strategic game $\Gamma =< N, (\mathcal{A}_i)_{i \in N}, (U_i)_{i \in N} >$ where the payment allocation $\boldsymbol{p} = (p^j_i)_{i,j \in N}$ is defined as*
$$p^i_j \triangleq \max\{0, \max_{A'_i \in PU_i(A)} C_j(A'_i, A_{-i}) - C_j(A)\}.$$

Proof. The proof follows directly from the condition at stage 4) of the AC algorithm in Figure 5. □

The set of content allocations reached by the *AC* algorithm is not a singleton in general. To see this, observe that a Nash equilibrium of the strategic game Γ with null payment allocation (i.e., $p^i_j = 0\ \forall i, j \in N$), if it exists, is a stable allocation under the AC algorithm. Furthermore, it is easy to construct instances of the strategic game Γ with null payment allocation that possess more than one Nash equilibrium.

C. Object-value Compensation Algorithm

Since the 2-independent sets of $\mathcal{G}$ are typically small, the number of nodes that can make updates simultaneously in the $\mathcal{I}^2$-AC algorithm is small, and thus the convergence rate of $\mathcal{I}^2$-AC may be only marginally faster than that of 1-AC. The number of simultaneous updates could be increased by using 1-independent sets, i.e., $\mathcal{I}^1$-AC, but the convergence of the $\mathcal{I}^1$-AC algorithm can not be guaranteed. We therefore propose an alternative to the AC algorithm.

The object value compensation (OC) algorithm, shown in Figure 6, is similar to the AC algorithm; the difference is that nCDNs offer a compensation for each individual object that is to be evicted, instead of offering a compensation for the set of objects to be evicted. As we will see this difference allows for significantly faster convergence, but at the price of revealing more information about content item popularities.

1) Set $\ell \leftarrow 1$ and $N^\ell_c \leftarrow N$
2) At round ℓ:
   - The nCDNs in N^ℓ_c run algorithm AC or OC until it terminates, in allocation $\boldsymbol{A}$.
   - Set $A^\ell \leftarrow \boldsymbol{A}$ and $N^{\ell+1}_c \leftarrow \{i \in N^\ell_c | r_i(A^\ell) \geq 1\}$.
3) If $|N^\ell_c \setminus N^{\ell+1}_c| > 0$:
   - Set $\ell \leftarrow \ell + 1$ and go to step 2).

Fig. 7: OPT OUT scheme

For the OC algorithm we can prove the following.

Theorem 8. *Let N_t be a complete sequence of 1-independent sets and every nCDN use the* OC *algorithm. We refer to this as the $\mathcal{I}^1$-*OC *algorithm. The $\mathcal{I}^1$-*OC *algorithm converges to an allocation $\boldsymbol{A}$ of content items to interconnected nCDNs in a finite number of time steps.*

Proof. Consider the compensation $p^o_{j,k}(t)$ offered by nCDN j to nCDN k for the eviction of item $o \in E_k$ at time step t. Substituting (3) in the expression of $\Delta\widetilde{C}^o_j(t)$ we obtain
$$p^o_{j,k}(t) = w^o_j \left[\beta^o_j\left((A^P_{\mathcal{I}^1_t}(t), A_{-\mathcal{I}^1_t}(t-1))\right) - \beta^o_j\left(A(t-1)\right)\right].$$
We call U_t the set of nCDNs that update their content allocation at time step t of the algorithm, i.e. $U_t = \{i_t \in \mathcal{I}^1_t | A_{i_t}(t) \neq A_{i_t}(t-1)\}$. Since $U_t \subseteq \mathcal{I}^1_t$, it follows from (4) that $\beta^o_j\left((A^P_{\mathcal{I}^1_t}(t), A_{-\mathcal{I}^1_t}(t-1))\right) \geq \beta^o_j\left((A^P_{U_t}(t), A_{-U_t}(t-1))\right)$ and thus
$$p^o_{j,k}(t) \geq C^o_j(A(t)) - C^o_j(A(t-1)). \tag{23}$$
In the following we use (23) to prove that $C(A(t)) < C(A(t-1))$ at every update of the $\mathcal{I}^1$-OC algorithm. We can express the aggregate cost change $\Delta C(t) = C(A(t)) - C(A(t-1))$ as
$$\Delta C(t) = \sum_{i_t \in U_t} \Delta C_{i_t}(t) + \sum_{j \in D_t} \Delta C_j(t) + \sum_{j \in \mathcal{N}(i_t) \setminus D_t} \Delta C_j(t). \tag{24}$$
From (23) it follows that the second term
$$\sum_{j \in D_t} \Delta C_j(t) = \sum_{j \in D_t} \sum_{o \in \mathcal{O}} \Delta C^o_j(t) \leq \sum_{j \in D_t} \sum_{o \in \mathcal{O}} p^o_{j,k}(t). \tag{25}$$
Substituting (25) into (24) we obtain
$$\begin{aligned}\Delta C(t) &\leq \sum_{i_t \in U_t} \Delta C_{i_t}(t) + \sum_{j \in D_t} \sum_{o \in \mathcal{O}} p^o_{j,k}(t) \\ &= \sum_{i_t \in U_t} \Big(\Delta C_{i_t}(t) + \sum_{j \in D_{i_t}} \sum_{o \in \mathcal{O}} p^o_{j,i_t}(t)\Big).\end{aligned}$$
Since every nCDN $i_t \in U_t$ refused the offer and updated its allocation, it holds that $\Delta C_{i_t}(t) + \sum_{j \in D_{i_t}} \sum_{o \in \mathcal{O}} p^o_{j,i_t}(t) < 0$ for all $i_t \in U_t$. Since the set of all content allocations is finite and the sequence N_t is complete, this proves the theorem. □

D. Achieving Individual Rationality

The proposed algorithms terminate in a finite number of time steps in a content allocation $\boldsymbol{A}$ from which no nCDN adhering to the compensation algorithm would like to deviate. However, the resulting content allocation $\boldsymbol{A}$ may not be individually rational, i.e., there may be some nCDNs i for which $r_i(\boldsymbol{A}) < 1$. The nCDNs $i \in \{i \in N | r_i(\boldsymbol{A}) < 1\}$

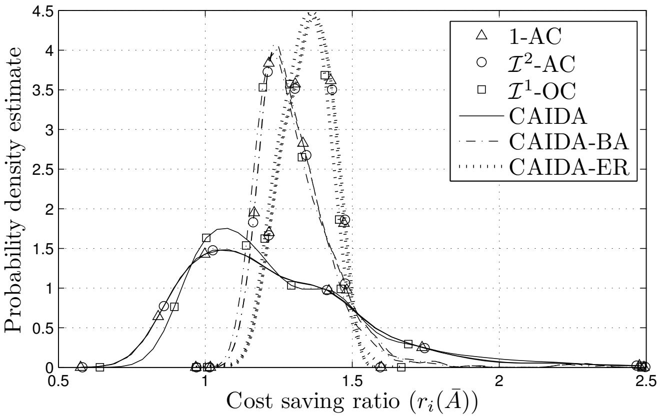


Fig. 8: Probability density estimate of $r_i(\bar{A})$ for the three algorithms 1-AC, $\mathcal{I}^2$-AC and $\mathcal{I}^1$-OC on the CAIDA, CAIDA-BA and CAIDA-ER graphs. Results from 400 simulations.

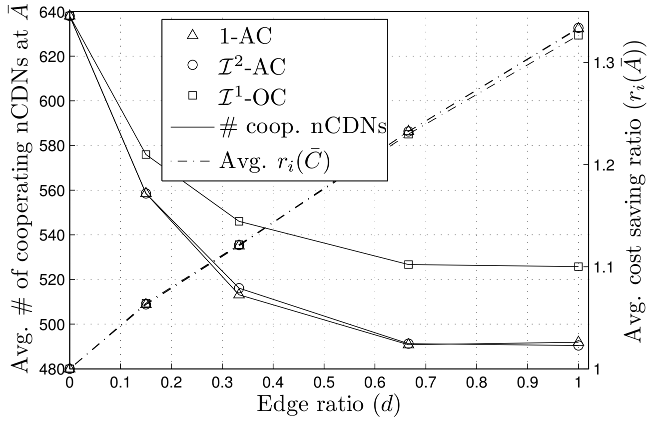


Fig. 9: Average number of nCDNs choosing to cooperate and average cost saving ratio $r_i(\bar{A})$ at allocation $\bar{A}$, as a function of the edge ratio d for the three algorithms on the CAIDA graph.

would not have an incentive to implement $\boldsymbol{A}_i$, and would instead implement A_i^I. The OPT OUT scheme, shown in Figure 7, allows these nCDNs to implement A_i^I instead of $\boldsymbol{A}_i$ and iteratively re-executes the distributed algorithm with the remaining nCDNs; hence the final allocation is individually rational.

Corollary 3. *The* OPT OUT *scheme reaches an individually rational content allocation $\bar{A}$ in a finite number of iterations.*

Proof. Observe that the OPT OUT scheme terminates in allocation $\bar{A} = (A^{\ell}_{N_c^{\ell}}, A^I_{N \setminus N_c^{\ell}})$ only if $N_c^{\ell} = N_c^{\ell+1}$ at step 3). This is true if either $|N_c^{\ell}| = 0$ or $r_i(A^{\ell}) < 1 \ \forall i \in N_c^{\ell}$. In both cases $\bar{A}$ is individually rational. □

Thus the $\mathcal{I}^2$-AC and the $\mathcal{I}^1$-OC algorithms combined with the OPT-OUT scheme are ex-post individually rational distributed algorithms for computing content allocations in a finite number of time steps, without prior global knowledge of the item popularities.

V. EVALUATION

We use simulations to validate the results in Section IV and to evaluate the convergence rate and the achieved gains for the cooperating nCDNs.

We consider three network topologies for the evaluation. The first topology is based on the Internet's AS-level topology reported in the CAIDA dataset [15] as of 1 Nov. 2013. In order to have a fairly large interconnection graph, we consider the ASes in the CAIDA dataset that are in Europe. As very small ASs are unlikely to deploy their own CDNs, we only consider ASs that have more than 2^{16} IP addresses allocated. We consider two ASes connected if they have a business relationship (peering or transit) reported in the CAIDA dataset. We call *CAIDA graph* the largest connected component of the resulting topology, which consists of 638 ASes with an average node degree of 10.77. The other two topologies are Erdős-Rényi (CAIDA-ER) and Barabási-Albert (CAIDA-BA) random graphs that have same number of vertexes, average node degree and node degree ranking as the CAIDA graph. The node degree distributions of the three topologies do, however, differ in terms of their skeweness. We computed distance-1 and distance-2 colorings of all graph topologies by using the Welsh-Powell [16] and the Lloyd-Ramanathan [17] algorithms, respectively. We used $\alpha_i = 0.5$, $\gamma_i = 20$ at every nCDN and we computed the β_i^j as the propagation delay between nCDNs i and j assuming a signal propagation speed of $2 \cdot 10^5$ km/s. We considered $|\mathcal{O}| = 3000$ objects and the demands w_i^o for the content items at the various nCDNs follow Zipf's law with exponent 1. To simulate the algorithms, at each time step we choose an nCDN or a k-independent set uniformly at random, thus the sequence is complete. If not otherwise specified, the results shown are the averages of 200 simulations and $K_i = 20$ for every nCDN $i \in N$. We omit the confidence intervals as the results are within 5% of the average at a 0.95 confidence level.

A. Individual Rationality

We start with considering the gain of cooperation and the necessity of the OPT OUT scheme. Figure 8 shows the probability density estimate of the cost saving ratios $r_i(A^{\ell=1})$ at the end of the first round of the OPT OUT scheme ($\ell = 1$) for all nCDNs for the three interconnection graphs and the three algorithms. The results show that the share of nCDNs for which the allocation is individually rational after the first round is determined by the graph topology. For the CAIDA-ER and the CAIDA-BA graphs, the content allocation is individually rational, $r_i(A^{\ell=1}) \geq 1$, for all nCDNs, and thus the OPT OUT scheme terminates after the first round, i.e. $\bar{A} = A^{\ell=1}$. On the contrary, for the CAIDA graph for many nCDNs $r_i(A^{\ell=1}) < 1$ after the first round. The difference is due to that the degree distribution of the CAIDA-BA graph is the most right-skewed among all the interconnection graphs, while the degree distribution of the CAIDA-ER graph is not skewed.

Observe that the probability densities for the 1-AC and $\mathcal{I}^2$-AC algorithms overlap, and are similar to that for the $\mathcal{I}^1$-OC algorithm. This suggests that the choice of the algorithm seems to have little impact on the gain from cooperation achieved by the nCDNs.

We evaluate the sensitivity of the results on synthetic topologies based on the CAIDA graph. The synthetic topologies

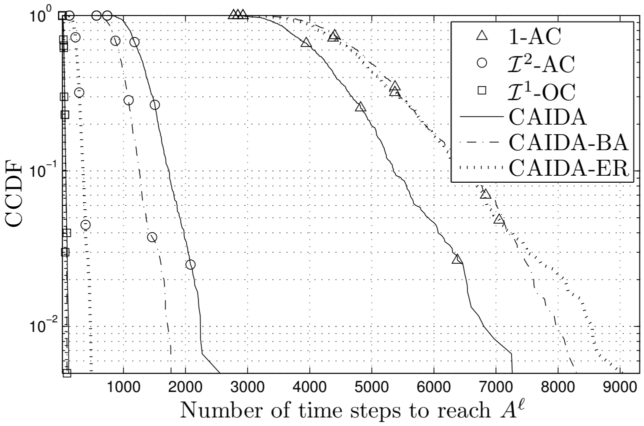


Fig. 10: Complementary CDF of the number of time steps needed to reach allocation A^ℓ for the algorithms 1-AC, $\mathcal{I}^2$-AC and $\mathcal{I}^1$-OC on the CAIDA, CAIDA-BA and CAIDA-ER graphs.

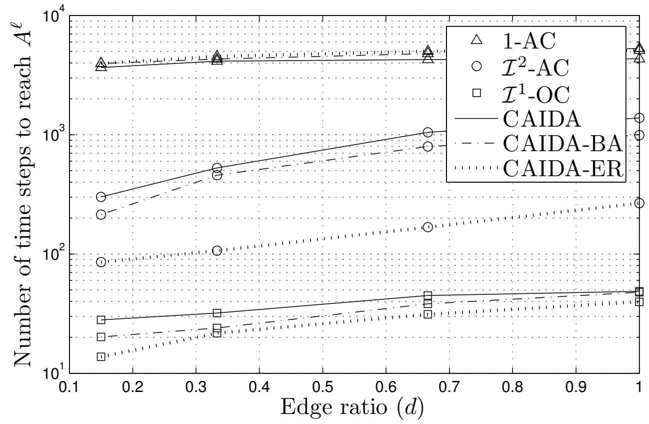


Fig. 11: Average number of time steps needed to reach allocation A^ℓ as a function of the edge ratio d for the three algorithms on the CAIDA, CAIDA-BA and CAIDA-ER graphs.

were created by removing all edges from the CAIDA graph, and then reintroducing d fraction of the edges at random; the probability of reintroducing an edge between ASes i and j was proportional to the product of the number of IP addresses allocated to AS i and j.

Figure 9 shows the number of nCDNs choosing to cooperate and the average cost saving ratio $r_i(\bar{A})$ for the allocation $\bar{A}$ reached by the OPT OUT scheme, for the algorithms 1-AC, $\mathcal{I}^2$-AC and $\mathcal{I}^1$-OC on the CAIDA graph. The figure shows that the number of cooperating nCDNs is a decreasing convex function of the edge ratio d, suggesting that the majority of the nCDNs would not opt out from cooperation even if the graph was denser. Furthermore, the number of nCDN that would not opt out from cooperation is about 6% higher for the $\mathcal{I}^1$-OC algorithm compared to the $\mathcal{I}^2$-AC algorithm. At the same time the average cost saving ratio increases linearly, which is due to that the nCDNs have access to a linearly increasing amount of storage at neighbors.

B. Convergence Rate

We characterize the rate of convergence of the 1-AC, $\mathcal{I}^2$-AC and $\mathcal{I}^1$-OC algorithms by comparing the number of time steps needed to reach allocation A^ℓ during one round ℓ of the OPT OUT scheme. The number of time steps needed to reach A^ℓ is proportional to the time required by the algorithms to converge, as it also captures the parallelism embedded in the plesiochronous $\mathcal{I}^2$-AC and $\mathcal{I}^1$-OC algorithms.

Figure 10 shows the complementary CDF of the number of time steps needed to reach allocation A^ℓ based on 400 simulations for the three algorithms on the CAIDA, CAIDA-BA and CAIDA-ER graphs. The tail of each distribution decreases exponentially or faster as the number of time steps increases, which suggests that the rate of convergence is geometric. As expected, the 1-AC algorithm performs worst in terms of convergence rate, as it does not allow the nCDNs to update their allocations simultaneously. $\mathcal{I}^2$-AC and $\mathcal{I}^1$-OC are up to two orders of magnitude faster than 1-AC. Note that the fast convergence of the $\mathcal{I}^1$-OC algorithm is achieved at the price of increased information exchange between connected nCDNs compared to the 1-AC and $\mathcal{I}^2$-AC algorithms. In practice, the object-wise information exchange between ASs may be problematic due to privacy concerns.

Figure 11 shows that the average number of time steps needed to reach allocation A^ℓ is an increasing concave function of the edge ratio d for all algorithms and interconnection graphs. Observe that the three interconnection graphs rank analogously for algorithms $\mathcal{I}^2$-AC and $\mathcal{I}^1$-OC but not for algorithm 1-AC. The reason lies in the average sizes of the k-independent sets used by the algorithms, which are reported in Table I. The higher the average size of the k-independent sets, the higher the parallelism achieved by the $\mathcal{I}^k$-COMP algorithm, and the faster the convergence. As the coloring of the interconnection graph does not affect the performance of the 1-AC algorithm, the rankings of the three curves in both Figures 10 and 11 reflect other characteristics of the different network topologies.

Graph	#1-ind. sets	avg. size	#2-ind. sets	avg. size
CAIDA	16	39.8	219	2.9
CAIDA-BA	10	63.8	131	4.9
CAIDA-ER	8	79.8	36	17.8

TABLE I: Number of k-independent sets and corresponding average size for the CAIDA, CAIDA-BA and CAIDA-ER interconnection graphs.

C. Scaling for Storage Capacity

In the following we investigate the effect of increasing the storage capacity K_i on the convergence rate of the proposed algorithms. Figure 12 shows the average number of time steps to reach allocation A^ℓ during one round ℓ of the OPT OUT scheme. We plot one curve for each algorithm on each of the CAIDA and CAIDA-BA graphs, as a function of the storage capacity K_i. The convergence rate is surprisingly insensitive to the storage capacity and the algorithms rank analogously to Figure 11. To explain this insensitivity we plot the average number of content item updates performed by the nCDNs for the same algorithms and graphs in Figure 13. We make two observations. First, the number of content item updates is the same for 1-AC and $\mathcal{I}^2$-AC, as they are both

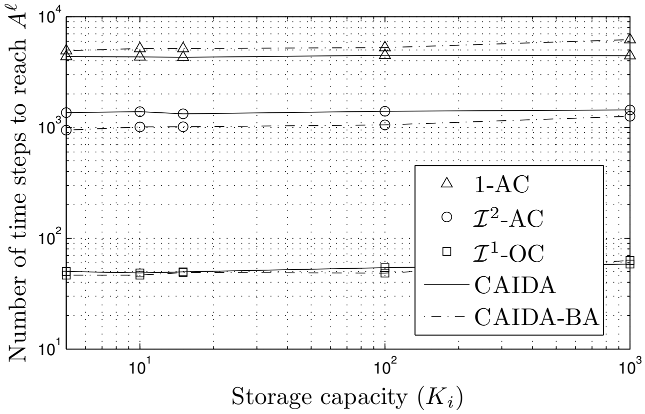


Fig. 12: Average number of time steps needed to reach allocation A^ℓ as a function of the storage capacity K_i for the algorithms 1-AC, $\mathcal{I}^2$-AC and $\mathcal{I}^1$-OC on the CAIDA and CAIDA-BA graphs.

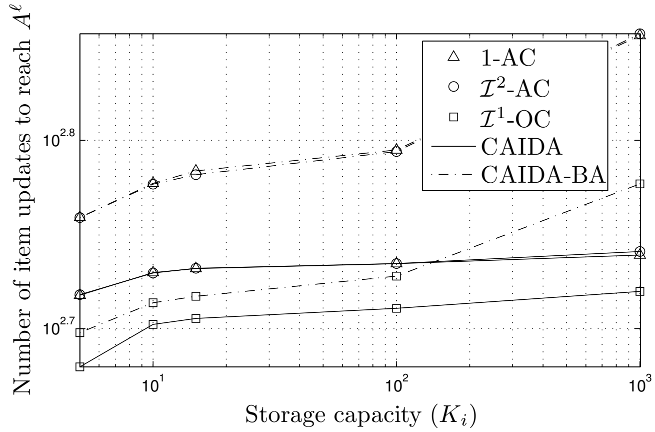


Fig. 13: Average number of content item updates needed to reach allocation A^ℓ as a function of the storage capacity K_i for the three algorithms on the CAIDA and CAIDA-BA graphs.

based on aggregate value compensation. The nCDNs perform less updates using the $\mathcal{I}^1$-OC algorithm, as they exchange object-wise compensations. The nCDNs perform less updates using the $\mathcal{I}^1$-OC algorithm, as they exchange object-wise compensations. Second, an nCDN can do an arbitrary number of content item updates during one time step, thus although the number of content items increases for larger storage sizes, this does not result in slower convergence in Figure 12.

VI. Related Work

Our work is related to recent works on content placement in the context of CDNs [2], [18], [19], [20]. The majority of these works assume a single CDN operator and optimize content placement given a single performance objective. The authors in [18], [19] considered centralized algorithms for content placement and compared the retrieval cost for the different algorithms. Recently, [20], [21] considered distributed algorithms that optimize for a single performance objective and provided analytical results for tree networks. The authors in [22], [23] considered a hybrid network with in-network caching and they proposed centralized algorithms for the joint problem of request routing and content replication. [23] considered strict bandwidth constraints at the storage sites. A more generic cost model was studied in [24], where the authors developed a centralized algorithm with approximation guarantee by rounding the optimal solution of the LP-relaxation of the problem. In contrast to these works, in this paper we consider distributed algorithms for operator-managed CDNs, and thus the allocations need to be individually rational.

Orthogonal to the problem we consider are the recent works in [25], [26], which focus on the commercial interactions between a network operator aiming at minimizing its content delivery costs and the content providers serving content to the operator's subscribers. [25], [26] develop incentive mechanisms, e.g. payment schemes, to compute individually rational allocations that jointly maximize the profit of the content providers and the network operator. In our work we consider content providers outsourcing the content delivery to several nCDNs, and we focus on the commercial interactions among interconnected nCDNs.

Closely related to ours are recent works on distributed selfish replication. Game theoretical analyses of equilibria and convergence for distributed selfish replication were considered in [8], [27], [28], [29], [30], [7]. The authors in [8] showed the existence of equilibria when the access costs are homogeneous and nodes form a complete graph. Similarly, [27], [28] assumed homogeneous costs and calculated the social cost of equilibria. The latter works considered that the nodes had no restriction on where to retrieve the content from. Other works [29], [30], [7] relax this assumption and introduce an interconnection graph to restrict the interaction between nodes. [29] assumed unit storage capacity and an infinite number of objects, showed the existence of equilibria and analyzed the price of anarchy for some special cases. [30] considered a variant of the problem where nodes can replicate a fraction of objects, and showed the existence of equilibria. The authors in [7] showed results in terms of convergence to equilibria in the case of homogeneous neighbor costs. In this paper we show that equilibrium existence results cannot be extended to the general problem of content replication on graphs and propose compensation-based algorithms that are guaranteed to converge.

Individually rational allocation of costs and revenues is the subject of cooperative game theory. Solution concepts such as the Shapley value and the core have found application in Internet routing [31] and in resource allocation [32], but these solution concepts require complete information and global enforcement, which make them impractical in the considered scenario. To the best of our knowledge this is the first work that proposes ex-post individually rational distributed algorithms for interconnected CDNs.

VII. Conclusion

We considered the problem of computing a content allocation among interconnected network CDNs. We showed that finding an allocation that minimizes the aggregate cost of the CDNs is computationally prohibitive and that such an allocation would need a central authority to be enforced. Moreover, we showed that without payments there may be no self-enforcing allocation that minimizes the cost of every CDN, but bilateral payments are sufficient to guarantee the existence of an equilibrium allocation. For the case that payments are possible, we proposed two bilateral compensation-based distributed algorithms that converge to an equilibrium allocation and that are ex-post individually rational. The two algorithms require different amounts of information to be revealed by the CDNs, and allow different levels of parallelism. Numerical results show that the algorithms have geometric convergence, and that if CDNs reveal more private information about their content demands, the convergence of the algorithms becomes faster. Our results also show that the convergence times are fairly insensitive to the graph density and the amount of CDN storage.

References

[1] "Netflix." [Online]. Available: http://www.netflix.com

[2] A.-M. K. Pathan and R. Buyya, "A Taxonomy and Survey of Content Delivery Networks," The University of Melbourne, Tech. Rep., 2007.

[3] G. Bertrand, E. Stephan, T. Burbridge, P. Eardley, K. J. Ma, and G. Watson, "Use Cases for Content Delivery Network Interconnection," Internet Engineering Task Force (IETF), RFC, 2012. [Online]. Available: http://tools.ietf.org/html/rfc6770

[4] L. Peterson, B. Davie, and R. van Brandenburg, "Framework for Content Distribution Network Interconnection (CDNI)," Internet Engineering Task Force (IETF), RFC, 2014.

[5] E. J. Rosensweig, J. Kurose, and D. Towsley, "Approximate Models for General Cache Networks," in *Proc. of IEEE INFOCOM*, 2010, pp. 1–9.

[6] C. Fricker, P. Robert, and J. Roberts, "A versatile and accurate approximation for LRU cache performance," in *Proc. of the 24th International Teletraffic Congress (ITC)*, 2012, pp. 1–8.

[7] V. Pacifici and G. Dán, "Convergence in Player-Specific Graphical Resource Allocation Games," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 11, pp. 2190–2199, 2012.

[8] N. Laoutaris, O. Telelis, V. Zissimopoulos, and I. Stavrakakis, "Distributed selfish replication," *IEEE Trans. Parallel Distrib. Syst.*, vol. 17, no. 12, pp. 1401–1413, 2006.

[9] V. Pacifici and G. Dán, "Content-peering Dynamics of Autonomous Caches in a Content-centric Network," in *Proc. of IEEE INFOCOM*, 2013, pp. 1079 – 1087.

[10] Y. He, M. Faloutsos, and S. Krishnamurthy, "Quantifying routing asymmetry in the internet at the AS level," in *Proc. of IEEE GLOBECOM*, vol. 3, 2004, pp. 1474–1479.

[11] A. Pathak, H. Pucha, Y. Zhang, Y. C. Hu, and Z. M. Mao, "A measurement study of internet delay asymmetry," in *Proc. of Passive and Active Measurement Conference (PAM)*, 2008, pp. 182–191.

[12] V. Bilò, A. Fanelli, M. Flammini, and L. Moscardelli, "Graphical Congestion Games." in *Proc. of WINE*, vol. 5385, 2008, pp. 70–81.

[13] C. Lumezanu, R. Baden, N. Spring, and B. Bhattacharjee, "Triangle inequality variations in the internet," in *Proc. of ACM SIGCOMM IMC*, 2009, pp. 177–183.

[14] L. S. Shapley, "A Value for n-Person Games," in *Contributions to the Theory of Games (AM-28)*, 1953, pp. 307–317.

[15] "CAIDA. Automated Autonomous System (AS) ranking." [Online]. Available: http://as-rank.caida.org/data/

[16] D. J. A. Welsh and M. B. Powell, "An upper bound for the chromatic number of a graph and its application to timetabling problems," *The Computer Journal*, vol. 10, no. 1, pp. 85–86, 1967.

[17] E. Lloyd and S. Ramanathan, "On the complexity of distance-2 coloring," in *Proc. of International Conference on Computing and Information (ICCI)*, 1992, pp. 71–74.

[18] M. R. Korupolu and M. Dahlin, "Coordinated placement and replacement for large-scale distributed caches," *IEEE Trans. Knowl. Data Eng.*, vol. 14, no. 6, pp. 1317–1329, 2002.

[19] Y. Chen, L. Qiu, W. Chen, L. Nguyen, and R. H. Katz, "Efficient and Adaptive Web Replication Using Content Clustering," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 6, pp. 979–994, 2003.

[20] S. Borst, V. Gupta, and A. Walid, "Distributed Caching Algorithms for Content Distribution Networks," in *Proc. of IEEE INFOCOM*, 2010, pp. 1478–1486.

[21] V. Pacifici, S. Josilo, and G. Dán, "Distributed Algorithms for Content Caching in Mobile Backhaul Networks," in *Proc. of International Teletraffic Congress (ITC)*, 2016, pp. 1–9.

[22] 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 *Proc. of IEEE INFOCOM*, 2015, pp. 936–944.

[23] K. Poularakis, G. Iosifidis, and L. Tassiulas, "Approximation Algorithms for Mobile Data Caching in Small Cell Networks," *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3665–3677, 2014.

[24] I. D. Baev and R. Rajaraman, "Approximation algorithms for data placement in arbitrary networks," in *Proc. of ACM SODA*, 2001.

[25] J. Li, H. Chen, Y. Chen, Z. Lin, B. Vucetic, and L. Hanzo, "Pricing and resource allocation via game theory for a small-cell video caching system," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 8, pp. 2115–2129, 2016.

[26] F. Shen, K. Hamidouche, E. Batu, and M. Debbah, "A Stackelberg Game for Incentive Proactive Caching Mechanisms in Wireless Networks," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM)*, 2016, pp. 1–6.

[27] G. Pollatos, O. Telelis, and V. Zissimopoulos, "On the social cost of distributed selfish content replication," in *Proc. of IFIP Networking*, 2008, pp. 195–206.

[28] E. Jaho, M. Karaliopoulos, and I. Stavrakakis, "Social similarity favors cooperation: the distributed content replication case," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 3, pp. 601–613, 2013.

[29] B.-G. Chun, K. Chaudhuri, H. Wee, M. Barreno, C. H. Papadimitriou, and J. Kubiatowicz, "Selfish caching in distributed systems: a game-theoretic analysis," in *Proc. of ACM symposium on Principles of Distributed Computing (PODC)*, 2004, pp. 21–30.

[30] G. Dán, "Cache-to-cache: Could ISPs cooperate to decrease peer-to-peer content distribution costs?" *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 9, pp. 1469–1482, 2011.

[31] R. T. B. Ma, J. C. S. Lui, V. Misra, and D. Rubenstein, "Internet Economics: The Use of Shapley Value for ISP Settlement," *IEEE/ACM Trans. Netw.*, vol. 18, no. 3, pp. 775–787, 2010.

[32] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic Game Theory*. Cambridge University Press, 2007.

Valentino Pacifici is a postdoctoral researcher at the Department of Communication Networks in KTH Royal Institute of Technology, Stockholm, Sweden. He studied computer engineering at Politecnico di Milano in Milan, Italy. In October 2010, he completed a joint M.Sc. degree in computer engineering, between Politecnico di Milano and KTH Royal Institute of Technology. He received the Ph.D. in Electrical Engineering from KTH in 2016. His research interests include the design and analysis of content management systems, and game theoretical and stochastic modeling of distributed systems.

György Dán is an associate professor at KTH Royal Institute of Technology, Stockholm, Sweden. He received the M.Sc. in computer engineering from the Budapest University of Technology and Economics, Hungary in 1999, the M.Sc. in business administration from the Corvinus University of Budapest, Hungary in 2003, and the Ph.D. in Telecommunications from KTH in 2006. He worked as a consultant in the field of access networks, streaming media and videoconferencing 1999-2001. He was a visiting researcher at the Swedish Institute of Computer Science in 2008, a Fulbright research scholar at University of Illinois at Urbana-Champaign in 2012-2013, and an invited professor at EPFL in 2014-2015. His research interests include the design and analysis of content management and computing systems, game theoretical models of networked systems, and cyber-physical system security in power systems.