The Impact of Social-Network Diffusion on Wireless Edge Resource Allocation

Iordanis Koutsopoulos
Department of Informatics
Athens University of Economics and Business
Athens, Greece

***Abstract*—Content providers (CPs) increasingly deploy network infrastructures that oftentimes reach up to the wireless network edge, i.e. base stations or small cells. Hence, they are interested in optimizing resource allocation and relevant performance metrics for that infrastructure. On the other hand, mobile apps featuring streaming content (e.g. video, music) come with social-networking and content-sharing capabilities among users. These need to be taken into account in resource allocation since they decisively shape content demand. In this work, we introduce mathematical optimization problems about resource allocation at the wireless network edge, which obtain interesting twists when social-network diffusion is considered. Specifically, we consider, (*i*) the problem of content caching and user targeting through the recommender system of the app, with the goal to maximize the social diffusion effect of cached content, and (*ii*) the problem of user targeting through the mobile app recommender system, so that the available wireless bandwidth is utilized as efficiently as possible.**

I. Introduction

In the evolving landscape of content distribution, streaming content (e.g. video, music) providers increasingly deploy their own content delivery network infrastructure, aiming at improving system-level aspects of Quality of Experience (QoE) for end-users such as content access latency, streaming quality, and throughput. Netflix OpenConnect, Amazon CloudFront and Google Global Cache are examples of such content delivery infrastructures. Further, major music streaming service providers such as Spotify and Pandora have their own content delivery networks (CDNs), while other content providers (CPs) such as Hulu rely on the infrastructure of the CDN giant Akamai for content delivery, with the latter bringing its server infrastructure closer to the user, namely at the base station premises [1].

The interface of these CPs to end-users is realized through a mobile app, which further offers mechanisms and services for improving QoE in terms of user satisfaction from delivered content, discovery of new content, and content sharing in a user community. Recommender systems are an integral part of these apps and suggest new content to users that presumably matches their interests. Furthermore, content-sharing mechanisms offer the possibility to share videos, songs and playlists and to follow the choices of other users. Spotify is known for its aggressive startup acquisition strategy so as to incorporate extensive social-networking features in its platform [2]. Pandora is known for its AMPcast service which allows musicians to communicate with their fans, and Youtube Community service features video sharing, as well as communication of content creators with consumers [3]. Interestingly, CPs seldom rely on third-party Online Social Networks (OSNs) such as Facebook for their social features; instead, they develop their own social features' services, hence they have direct access to relevant data.

978-1-5386-4725-7/18/$31.00 ©2018 IEEE

Content demand and consumption pattern prediction are important for spatio-temporal resource provisioning across a CDN, while quick demand satisfaction is important for enhancing user QoE. When the CDN extends up to the wireless network edge by featuring caches installed at the premises of base stations or small cells, these issues become challenging due to the spatial locality of demand at different cells. On the other hand, mobile apps featuring streaming content (e.g. video, music) come with *social-networking* and *content-sharing* capabilities among users. These need to be judiciously taken into account in resource allocation decisions since social-diffusion decisively shapes content demand.

The impact of social networking on content delivery has been studied, starting from the seminal social-network diffusion paper [4] and more recent works (e.g. [5]-[7], and refererences therein). In [5], knowledge from OSNs and device encounters is used to place content to device caches close to content requesters. In [7], a hypergraph framework is proposed that exploits user social ties and common interests for content placement in a device-to-device setting. Social diffusion-aware social-media recommender systems are studied in [8], while in [9] a joint consideration of content caching and recommendation to users is proposed for maximizing the cache-hit ratio, where the recommender system shapes content demand.

The novelty of this work in progress is the joint consideration of wireless edge resource allocation and social-network diffusion. This joint view is motivated by the real-life scenarios above in which both the wireless edge network infrastructure and the social-networking aspects of the platform are controlled by a CP. First, we consider the joint problem of content caching and user targeting through the mobile app recommender system, with the goal to maximize the social-network diffusion effect of cached content. This objective pertains to ensuring good QoE for all users that may view the content as a result of diffusion i.e. content sharing. This is also plausible when the CP receives profit through advertisements that are embedded to each content

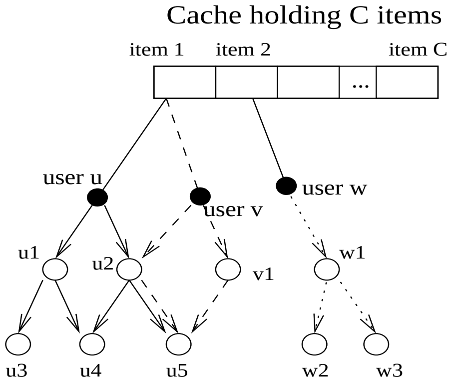

Fig. 1. A cache with C items, where content item 1 is initially recommended to users u and v, i.e. $y_{1u} = 1$, $y_{1v} = 1$ and item 2 to user w, i.e. $y_{2w} = 1$. Arrows in solid, dashed and dotted lines show the diffusion process for each of these 3 initial recommendations. The resulting reachable sets are, $\mathcal{R}_{1u} = \{u_1, u_2, u_3, u_4, u_5\}$, $\mathcal{R}_{1v} = \{u_2, v_1, u_5\}$ and $\mathcal{R}_{2w} = \{w_1, w_2, w_3\}$.

view, so that the CP would aim at maximizing content reach.

Next, we introduce the problem of user targeting through the app recommender system so that the available wireless bandwidth is utilized as efficiently as possible for content consumed as a result of the social-network diffusion effect. This objective is meaningful when the CP leases a fixed amount of bandwidth from a network operator and wishes to neither exceed this amount nor underutilize it. We also discuss future directions on these issues.

II. Model and problem statement

A. Model

There exists a catalogue of a set $\mathcal{I}$ of content items e.g. videos that are available at the CP back-end server. We consider a single cache with capacity C items, where each item is of unit size. The cache is installed at a base station or small cell, and a set of users $\mathcal{U}$ are served by the cache. The cache is updated overnight, when content demand is low, with content items that are predicted to be of high demand in the next day. During the next day, items viewed by users possibly become *viral* through consecutive shares with their friends. The caching policy is a binary vector $\mathbf{x} = (x_i : i \in \mathcal{I})$ where $x_i = 1$ if item i is cached, and 0 otherwise.

For the first view of an item i, say by a user u, there exist several possibilities. In one of them, the item may be recommended to the user through the mobile app. A recommended item may be viewed by a user with certain probability which can be estimated through machine-learning techniques on historical data. Without loss of generality, we assume here that a recommended item will be viewed by the user. The *recommendation* or *user targeting* or *seeding policy* is defined through binary variables y_{iu}, for $i \in \mathcal{I}$ and $u \in \mathcal{U}$, with $y_{iu} = 1$ if item i is recommended to user u, and 0 otherwise, and let $\mathbf{y} = (y_{iu} : i \in \mathcal{I}, u \in \mathcal{U})$. Another possibility is that an item i is not recommended to a user u, but u may anyway request to view the item with probability p_{iu} based on the match of item attributes with the user profile. This probability may be estimated with machine-learning and information-retrieval techniques.

The mobile app is equipped with social-networking features. Users are organized in a directed social graph $G = (V, E)$, where $V = \mathcal{U}$ and E is the set of edges. A directed edge (u, v) between users u and v signifies that user v "follows" u, with the understanding that v has visibility on all content that is shared or re-posted by u. For item i and user u, let w_{iu} be the probability that item i is shared by u; this depends on how often a user tends to share items, and on the relevance of item attributes to the user profile.

1) Diffusion model: When content item i is initially recommended and viewed by a user u, user u may share it, and thus the item may be viewed by her followers as well. In turn, these followers may share the item so it will reach users that follow them, and so on. It can be seen that the initial recommendation of item i to user u gives rise to a *diffusion process* in the social-network graph. For each item i, the diffusion model with sharing probabilities $\{w_{iu}\}$, $u \in \mathcal{U}$ gives rise to a probabilistic graph $\widetilde{G}_i$, whose edges are assigned an existence probability. Graph $\widetilde{G}^i$ is defined as a triple (V, E, p^i), where V and E are the node set and edge set of graph G respectively, and $p^i : E \to (0, 1]$ is a function that assigns a probability of existence to each edge $(u, v) \in E$ so that $p^i(u, v) = w_{iu}$.

Given the probabilistic graph $\widetilde{G}^i$, the goal is to specify the set of users that are reached through sharing when the diffusion process is initiated by the recommendation of an item i to a user u. Namely for each user $u \in V$, we seek the set of nodes that are reachable from u with high probability when item i is initially recommended and viewed by u. This corresponds to finding the so-called *reliable node set* of user u in graph $\widetilde{G}^i$ [10]. That is, given the probabilistic graph $\widetilde{G}^i$, a source user u, and a probability threshold $\gamma \in [0, 1]$, we need to find the set $\mathcal{R}_{iu}$ of all nodes (users) that are *reachable* through a (multihop) path from u with probability larger than γ.

Computing the exact reliable node set of a node is a hard problem as it relies on the two-terminal reliability problem which is about computing the probability that a node is reachable from another node in a probabilistic graph. However, the reachable set $\mathcal{R}_{iu}$ can be accurately and efficiently estimated via Monte Carlo sampling. The idea is to sample a set $\mathcal{D}$ of deterministic graphs from the probabilistic graph according to edge probabilities and, for each of these deterministic graphs, compute the nodes reachable from u. The reachable set $\mathcal{R}_{iu}$ is computed as the set of all nodes that are reachable from u in the majority of sampled graphs.

B. Problem statement

Given the social-network diffusion model above, we would like to find the joint item-to-user recommendation policy and item caching policy so as to maximize the total diffusion effect of cached content, i.e. maximize the number of users included in the union of all reachable sets resulting from diffusion. In that case, the cache hit ratio would be maximized, which further implies that the positive impact of the cache towards fast retrieval of content demand emanating through diffusion would be maximized, while the amount of content demand traffic, and thus the amount of bandwidth consumed, at the backend network would be minimized.

For each item i, the overlap between different reachable sets initiated by users u, v should be small so that the cumulative social-network diffusion effect of item i is higher. Indeed, for a pair of users (u, v) and item i, the cumulative social diffusion effect is quantified as $|\mathcal{R}_{iu} \cup \mathcal{R}_{iv}| \leq |\mathcal{R}_{iu}| + |\mathcal{R}_{iv}|$, thus the effect is maximized when $\mathcal{R}_{iu} \cap \mathcal{R}_{iv} = \emptyset$. Let $e^i_{uv} = |\mathcal{R}_{iu} \cap \mathcal{R}_{iv}|$ be the size of the overlap between reachable sets $\mathcal{R}_{iu}$ and $\mathcal{R}_{iv}$. Also, let $e^i_u = |\mathcal{R}_{iu}|$. An upper bound on the number of users that receive item i when it is initially recommended to user u is, $\left(e^i_u - \sum_{v \in \mathcal{U}: v \neq u} e^i_{uv}\right)$. Also define the indicator function $I^i_{uv} = 1$ if user v is in the reachable set $\mathcal{R}_{iu}$, and 0 otherwise. The problem of maximizing content diffusion outreach may be formulated as follows:

$$\max_{\mathbf{x},\mathbf{y}} \sum_{i \in \mathcal{I}} x_i \sum_{u \in \mathcal{U}} \left((1 + e^i_u) y_{iu} - \frac{1}{2} \sum_{v \in \mathcal{U}: v \neq u} e^i_{uv} y_{iu} y_{iv} \right) \quad (1)$$

subject to the contraints:

$$\sum_{i \in \mathcal{I}} x_i = C \, , \quad (2)$$

$$\sum_{i \in \mathcal{I}} y_{iu} \leq M, \ \forall u \in \mathcal{U} \, , \quad (3)$$

$$\sum_{i \in \mathcal{I}} \sum_{v \in \mathcal{U}} I^i_{vu} y_{iu} \leq \, , \ \forall u \in \mathcal{U} \, . \quad (4)$$

The first term in (1) counts the users to which an item is initially recommended and the users in the ensuing reachable sets. The second term in (1) is due to the overlap of reachable sets $\mathcal{R}_{iu}, \mathcal{R}_{iv}$. Constraint (2) is the cache capacity constraint. Constraint (3) says that at most M items can be initially recommended to each user, while (4) says that each user should be in at most K reachable sets; these capture content consumption, attention span or time constraints.

III. Extensions and future work

The model above calls for an evaluation process with real data, and it can also be extended in various ways. Here we outline some of these future directions.

A. Caching in a multiple-cache setting

An interesting extension would be to consider a set $\mathcal{C}$ of several caches, where each cache c is attached to a small cell. Each user $u \in \mathcal{U}$ has a subset $\mathcal{C}_u$ of caches within range from which it can retrieve content, while each cache c covers a subset $\mathcal{U}_c$ of users. The caching policy, $\mathbf{x} = (x_{ic} : i \in \mathcal{I}, c \in \mathcal{C})$ now concerns the cache at which to store each item, if at all. There is also a small-cell user association policy $\mathbf{z} = (z_{uc} : u \in \mathcal{U}, c \in \mathcal{C})$. The objective is again to maximize the total diffusion effect of cached content, therefore a desirable joint caching and association policy $\mathbf{x}, \mathbf{z}$ should aim to allocate users to caches such that the overall overlap between reachable sets is small. A challenge now is that different users in each reachable set $\mathcal{R}_{iu}$ may receive the item from different caches, hence a residence probability of each user in each cache needs to be assumed.

B. Small-cell bandwidth utlization

Consider a single small cell, and let B be the total bandwidth available in that cell for content transmission. The total consumed bandwidth is a function of the seeding policy $\mathbf{y}$ and it is equal to

$$W(\mathbf{y}) = \sum_{i \in \mathcal{I}} \sum_{u \in \mathcal{U}} e^i_u y_{iu} \, . \quad (5)$$

Uncontrolled content virality may lead to excessive bandwidth consumption and deterioration of user QoE. Namely, if $W(\mathbf{y}) > B$, then some users may experience the negative effects of bandwidth overconsumption on content streaming quality. On the other hand, if $W(\mathbf{y}) < B$, a certain portion of the available bandwidth would remain unutilized. If the CP has leased the amount B of bandwidth from an access network operator, it makes sense that the CP would like to utilize all of it or as much as possible. Therefore, a valid objective would be to control content virality by deciding on the initial user seeding policy $\mathbf{y}$ with the objective,

$$\min_{\mathbf{y}} \left| W(\mathbf{y}) - B \right| . \quad (6)$$

In the case of multiple small cells, the association policy $\mathbf{z}$ can also be included in the formulation, since it affects bandwidth consumption in a cell.

Acknowledgment

This work was supported by the European Commission H2020 Research Program under Grant number 688768 netCommons (Network Infrastructure as Commons) and by AUEB-RC under the internal project “Original Scientific Publications”.

References

[1] https://www.akamai.com/us/en/resources/mobile-optimization.jsp

[2] https://www.billboard.com/articles/business/6850996/spotify-buys-cord-project-soundwave-startups

[3] https://techcrunch.com/2016/09/13/youtube-gets-its-own-social-network-with-the-launch-of-youtube-community/

[4] D. Kempe, J.M. Kleinberg, and E. Tardos, “Maximizing the spread of influence through a social network”, *In Proc. KDD,* 2003.

[5] F. da S. Moraes, K. V. Cardoso and V. C. M. Borges, “Improving Video Content Access with Proactive D2D Caching and Online Social Networking”, *in Proc. IEEE ISCC,* 2017.

[6] H. Hu, Y. Wen, T.-S. Chua, J. Huang, W. Zhu and X. Li, “Joint Content Replication and Request Routing for Social Video Distribution Over Cloud CDN: A Community Clustering Method”, *IEEE Trans. Cir. Syst. for Video Tech.,* vol.26, no.7, pp.1320-1333, July 2017.

[7] B. Bai, L. Wang, Z. Han, W. Chen and T. Svensson, “Caching-based Socially-aware D2D Communications in Wireless Content Delivery Networks: A hypergraph framework”, *IEEE Wireless Comm. Mag.,* vol.23, no.4, pp.74-81, Aug. 2016.

[8] P. Vahabi, I. Koutsopoulos, F. Gullo and M. Halkidi, “DifRec: A social diffusion-aware recommender system”, *In Proc. ACM Conf. Inf. Knowl. Manag. (CIKM),* 2015.

[9] L. Chatzieleftheriou, M. Karaliopoulos and I. Koutsopoulos, “Caching-aware recommendations: Nudging user preferences towards better caching performance”, *in Proc. IEEE INFOCOM,* 2017.

[10] A. Khan, F. Bonchi, A. Gionis, F. Gullo, Fast Reliability Search in Uncertain Graphs. *In Proc. EDBT,* 2014.