

# Efficient 5G Edge Caching Over Satellite

Thang X. Vu\*, Nicola Mauro\*, Satyanarayana Vuppala\*, Symeon Chatzinotas\*, Joel Grotz<sup>†</sup>, and Nader Alagha<sup>‡</sup>

\* Interdisciplinary Centre for Security, Reliability and Trust (SnT) – University of Luxembourg

Email: {thang.vu, nicola.mauro, satyanarayana.vuppala, symeon.chatzinotas}@uni.lu.

<sup>†</sup> SES Engineering, 6815 Betzdorf, Luxembourg. Email: joel.grotz@ses.com.

<sup>‡</sup> European Space Agency, Noordwijk, The Netherlands. Email: nader.alagha@esa.int.

**Abstract**—The fifth generation (5G) wireless networks have to deal with the high data rate and stringent latency requirements due to the massive invasion of connected devices and data-hungry applications. Edge caching is a promising technique to overcome these challenges by prefetching the content closer to the end users at the edge node’s local storage. In this paper, we analyze the performance of edge caching 5G networks with the aid of satellite communication systems. Firstly, we investigate the satellite-aided edge caching systems in two promising use cases: a) in dense urban areas, and b) in sparsely populated regions, e.g., rural areas. Secondly, we study the effectiveness of satellite systems via the proposed satellite-aided caching algorithm, which can be used in three configurations: i) mono-beam satellite, ii) multi-beam satellite, and iii) hybrid mode. Thirdly, the proposed caching algorithm is evaluated by using both empirical Zipf-distribution data and the more realistic MovieLens dataset.

## I. INTRODUCTION

During the past 40 years, the Internet has followed an extraordinary evolution and has become an integral part of the modern society. However, this evolution has kept momentum and there are constantly new services and contents distributed through this global communication network. Based on Cisco’s report [1], it is predicted that the mobile data traffic will grow 74% by 2021. Particularly, the mobile video will increase eleven-fold between the mentioned years. The main causes of this traffic growth are the vast availability of mobile devices, e.g., smart phones, tablets, and notebooks, as well as the fast growth of video content on the Internet and their increasing quality. Using these mobile devices, more and more users are immigrating from traditional linear broadcasting services (TV channels) to streaming services, such as YouTube and Netflix. Another factor that contributes to the traffic is the increasing video quality, i.e., 3D, 4K video, Virtual Reality etc., which can be translated to increased bandwidth requirements for both the core and access networks. This perspective seems very promising for content providers, since they can provide improved services and reap the benefits either through subscriptions or advertising. Nevertheless, looking at this from a telecom operator’s point of view, it is obvious that video traffic will become a bottleneck and put excessive strain on current communication infrastructure. On the other hand, video traffic also results in revenue growth. However, telecom operators do not have direct access to the revenue generated through video content delivery and as a result they cannot use it to upgrade their infrastructure. In parallel, spectrum has become scarce and the operators cannot easily access new frequency bands to expand their wireless access and backhaul segments.

These are the reasons why in the last few years, telecom operators started to build their own content delivery networks (CDNs). The aim of a CDN is to serve contents to end-users with high performance by using edge caching. The main benefit of CDNs is represented by the higher degree of cross-optimization between the physical infrastructure and the network service that leads to an improved transmission efficiency. It becomes obvious that closer interaction between operators and content providers will be needed in order to optimize content delivery and overcome the projected bottleneck due to video traffic [3]. Some steps have been made in this direction, for instance Dhiraagu, the Maldives operator, have deployed Google caches successfully in its network [2].

One of the challenges in the edge caching is how to effectively prefetch the popular content to the caches considering the high volume of data [4]. In order to overcome this issue, we propose to use satellite backhauling for cache placement phase to exploit the large coverage of the satellite beams. Satellite systems have the ability to provide high throughput links and to operate in multi/broad-cast modes for immense area coverage.

Due to their multi-hop unicast architecture, the cached content via terrestrial networks has to go through multiple links and has to be transmitted individually towards each base station (BS). On the other hand, with wide area coverage, the satellite backhaul can broadcast content to all BSs or multi-cast contents to multiple groups of BSs. Therefore, bringing these two technologies together can further off-load the network. The main idea is to integrate the satellite and terrestrial telecommunication systems in order to create a hybrid federated content delivery network, which can improve the user experience. The joint deployment of satellite and terrestrial networks can be found in [7], [8]. In this paper, we consider the satellite channel as the only mean for cache placement. The application of satellite communications in feeding several network caches at the same time using broad/multi-cast is investigated in [5], [6], [7]. The work of [6] proposes using the broad/multi-cast ability of the satellite to send the requested contents to the caches located at the user side. Online satellite-assisted caching is studied in [7]. In this work, satellite broadcast is used to help placing the files in the caches located in the proxy servers. Each server uses the local and global file popularity to update the cache.

In this work, a satellite-aided caching algorithm is proposed. We use off-line caching approaches [8], [9], [10] to off-load the backhaul of the terrestrial network. We focus on the role

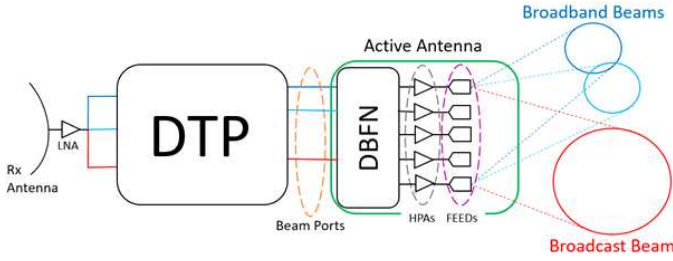


Fig. 1: Hybrid Satellite Architecture

of multimodal satellite backhauling, which provides flexible backhaul's transmission modes, e.g., broadcast and broadband, and its effectiveness on edge caching. The proposed algorithm can be used with three different satellite's configurations: i) broadcast mono-beam, ii) broadband multi-beam where the content of each beam can be selected independently, and iii) hybrid mode that is a combination of the first two modes. Focusing on promising satellite markets, we evaluate the performance of the proposed caching algorithm in two use cases: in dense urban areas and in sparsely populated rural regions. Based on both empirical Zipf-distribution data [11] and the more realistic MovieLens dataset [12], we show via numerical results that the multi-beam satellite will outperform the mono-beam system when the demands are less correlated, and that the hybrid achieves a cache hit ratio between the multi-beam and mono-beam schemes. These observations are very much dependent on the popularity of the content. Despite that, the proposed caching algorithm is capable of adapting to different means of content delivery to optimize the system cost function.

## II. TECHNOLOGICAL ENABLER FOR SATELLITE-ASSISTED EDGE CACHING

### A. Hybrid Satellite

The satellite architecture in general can be classified in two traditionally different cases: satellite with a very wide single beam, mainly used for broadcast services, and satellite with many small beams, mainly used for broadband services. While the distinction between these two types of architectures is well established nowadays, in the future satellite system there may be no real distinction between broadcast and broadband payload. Thanks to the adoption of new technologies, like digital transparent payload (DTP) and (semi-)active on-board antennas, it will be possible to have both services sharing the same hardware power and spectrum resources in reconfigurable hybrid broadcast/broadband payloads. Many satellite manufacturers are currently working towards this new type of payload [13] and in Fig. 1 we provide a pictorial representation of this hybrid architecture.

As an example shown in Fig. 1, the satellite is receiving 3 different streams: 2 broadband streams (the light and the dark blue lines) and 1 broadcast stream (the red line). At a beam ports level (the output of the DTP) we still have 3 different streams that are the input of the active antenna. Thanks to the active antenna technology, it will be possible to create overlapping beams of variable granularity and use a single

broadcasting system to drive a large broadcast beam and at the same time it will be possible to generate on the same payload separated broadband beams. The flexible payload would be able also to shape the broadcasting beams to concentrate the transmitted power where it is needed. It is interesting to note that in this configuration there isn't a direct relation between the beams and the feed, but basically all the feeds cooperate to create all the beams using digital beamforming (BF). In this way is possible to optimize the power consumption of the high power amplifiers (HPAs) serving the feeds.

Because the creation of the beams is driven by the digital beamforming network (DBFN), the beams design can be very easily reconfigured. This type of flexibility is certainly extremely important for satellite operators, so that they can modify their satellite configuration in order to better accommodate the variation of the traffic demands during the satellite lifetime. In this context we refer in particular to the content-related traffic evolution determined by various factors such as population evolution and changes in the content consumption preferences.

### B. Integrated Satellite-Terrestrial System

The classic cellular backhaul comprises of fibre, copper lines and wireless links, which can efficiently transfer the backhaul traffic from point to point. The multimodal architecture enhances the conventional backhaul by overlaying wide satellite beams and narrower broadcasting cells (e.g. HPHT), which are capable of P2MP (point to multi-point) multi/broadcasting. Studies in [14] have shown that a small percentage of video files can be extremely popular posing a huge load on current networks that do not support broadcasting on the physical layer. The proposed satellite overlay caching solution can relieve this load by simultaneously and efficiently distributing popular content the edge caches by integrating traditional P2P backhauling with broad/multicasting P2MP backhauling. It should be noted that these multi/broadcast systems can inherently reach a large number of BSs with a single PHY-layer transmission in contrast to the NET-layer multicasting which implies packet replication and multiple individual PHY connections. In addition to above mentioned benefit, this converged solution brings several distinct benefits: 1) additional backhaul capacity based on existing infrastructure, 2) spectrally-efficient physical-layer multi/broadcasting, 3) variable cell sizes for broadcast backhauling (wide coverage for satellite, narrow coverage for high broadcasting towers).

## III. SYSTEM MODEL AND CACHING ALGORITHM

We consider a satellite system serving the users within its coverage via a number of edge nodes, e.g., base stations, which are equipped with a satellite receiver. The considered system can find application in villages, in rural area or small business customers such as hotels and residents where only one satellite receiver/decoder is needed for each base station (BS). Each BS has a local storage which can store up to  $M$  content files. For the sake of simplicity, all files are assumed to have same size of  $Q$  bits, however, the extension to arbitrary file sizes is



TABLE I: CACHING ALGORITHM VIA SATELLITE

<b>Input:</b> $C$ - caching capacity, $M$ - cache size at the base station, $K_b$ - number of BSs, $K_u$ - number of users per BS, $L$ - number of beams.	
1.	The BSs estimate $\mathbf{p}_{k,u}$ - interested content from user $u$ at BS $k$ , then forward it to the gateway
2.	<b>if</b> (mono - beam) 2.1. The global cache manager (GCM) at the gateway computes the global popularity $\mathbf{p}_G = \frac{1}{K_b K_u} \sum_{k,u} \mathbf{p}_{k,u}$ 2.2. The GCM sorts $\mathbf{p}_G$ in the decreasing order 2.3. The GCM chooses the first $C$ files in $\mathbf{p}_G$ to broadcast for caching 2.4. Each BS $k$ computes its local popularity $\mathbf{p}_k = \frac{1}{K_u} \sum_{k=1}^{K_u} \mathbf{p}_{k,u}$ . It then sorts $\mathbf{p}_k$ in the decreasing order, and chooses the first $M$ files in the sorted local popularity for caching. A file is cached at the local cache of BS $k$ if it is at both in the list and is sent by the GCM of $M$ files
3.	<b>else_if</b> (multi - beam) 3.1. The GCM constructs $L$ beams. $\mathcal{K}_l$ denotes the set of users in the $l$ -th beam. 3.2. For each beam, the GCM applies the caching policy in step 2.
4.	<b>else_if</b> (hybrid - mode) 4.1. Determine $C_{\text{mono}}, C_{\text{multi}}$ as the caching capacity for mono-beam and multi-beam, respectively. 4.2. The GCM applies step 2 to decide $C_{\text{mono}}$ files to be broadcasted via mono-beam satellite, denoted by $\mathcal{F}_{\text{mono}}$ 4.3. The GCM removes files in $\mathcal{F}_{\text{mono}}$ from the requests $\mathbf{p}_{k,u}$ , then applies step 3 with the caching capacity $C_{\text{multi}}$ .

is a comparison on the basis of the total satellite cost per transmitted information bit.

We assume multi-cast link budget assessment results indicating the required  $E_S/N_0$  values for both the considered transmission modes. From this assessment, we derive the transmission efficiency, taking into account all overhead parameters required.

The power and bandwidth used in the link budgets to assess the transmission efficiency have to be considered. If it is multiplexed with other traffic, the proportional DC power resources and bandwidth utilization resources is taken into account. The DC power equivalent is deduced from the assessment of the power assessment taking into account the amplifier efficiency and output backoff (OBO) of the amplifiers required. The total bandwidth available over all beams is considered as bandwidth reference for the satellite. Here the total bandwidth of the simultaneously transmitted traffic has to be considered for all possible beams which are active at the same time.

The total cost of the transmission can be estimated based on the reference of same satellite services, which takes into account CAPEX (and OPEX in a refined model, but not required in first order approximation) costs with related assumptions on fill rates per lifetime. This is deduced as cost per Mbps assessment for the computed traffic load. We may consider different levels of costs to take into account the fact that an underutilized time period could be used for the data transfer. The total is a  $COST_{LINK}$  result as  $RESOURCE_{PERCENT} * COST_{SAT}$ .

Then the total file transfer cost results in:

$$COST_{file} = Size * COST_{LINK} / Efficiency * Bw$$

The relative cost comparison between mono-beam and multi-beam used in the simulation is given in Table II.

TABLE II: RELATIVE COMPARISON FOR COST PER BIT

	Widebeam	Multibeam	Unit
	Multicarrier	Singlecarrier	
$E_S/N_0$ (worst case)	8.0	7.5	dB
Efficiency	2.4	2.1	bps/Hz
Total Sat RF Power	8556	5398	W
Total Sat Bandwidth	4320	80000	MHz
Beams per coverage	1	10	#
Power per beam	120	60	W
Bandwidth per beam	60	500	MHz
Power Ratio	1.4%	11.1%	
Bandwidth Ratio	1.4%	6.3%	
$RESOURCE_{PERCENT}$	1.4%	11.1%	
$COST_{LINK}$	0.014	0.111	relative
Size	5000	5000	Mbyte
$COST_{file}$	3.89	4.19	
Relative Cost	1	1.09	

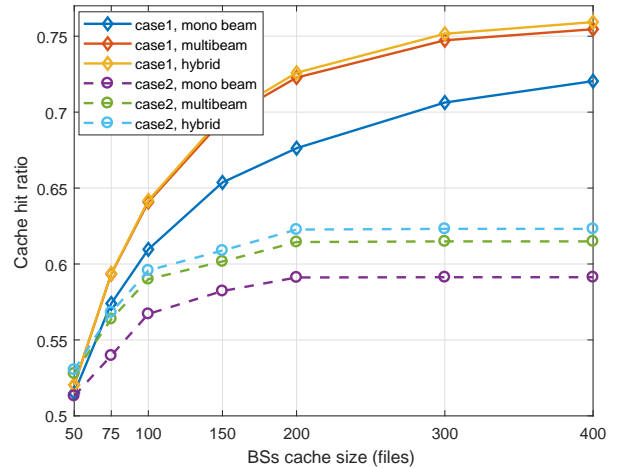


Fig. 3: CHR performance of the proposed caching algorithm in two use cases. For each case, the caching algorithm is designed for three different beam settings: mono beam only, multi-beam only, and hybrid mono beam and multi beam.

## V. NUMERICAL RESULTS

In this section, we present the cache hit ratio (CHR) performance of the proposed caching algorithm via numerical results. The user's requests follow Zipf distribution with the skewness factor equal to 0.8. The library consists of  $F = 500$  files. The cache size  $M$  indicates that it can store  $M$  movies in its storage. To make a fair comparison, the caching capacity in mono-beam and multi-beam settings are chosen such that both schemes have same total placement cost. This means that we fix the caching capacity of the monobeam mode and we scale the caching capacity of the multibeam mode in accordance with the results of Table II. In particular, let  $C_{\text{mono}}$  and  $C_{\text{mul}}$  denote the caching capacity in the mono-beam and multi-beam modes, respectively. The caching capacity of the hybrid mode is calculated in accordance with the usage percentage of the monobeam,  $C_{\text{hybrid},1}$ , and the multibeam mode,  $C_{\text{hybrid},2}$ . In order to meet the same placement cost, we have  $C_{\text{mono}} = 1.09C_{\text{mul}} = C_{\text{hybrid},1} + 1.09C_{\text{hybrid},2}$ . The satellite provides service for an area consisting of 1000 BSs, each is serving 2000 users in use case 1 and 400 users in use case 2.



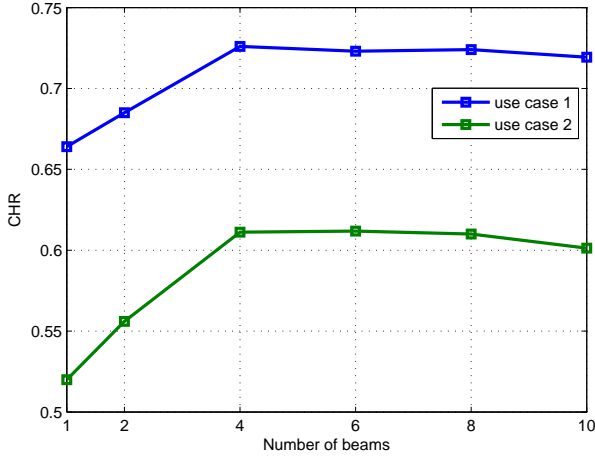


Fig. 4: CHR performance of the multi-beam scheme v.s. number of beams. The user requests form 4 weakly correlated regions.

Fig. 3 presents the CHR for the two considered use cases. The total caching capacity is equal to 200 files in use case 1, and 100 files in use case 2. We assume that the user requests form four geographical regions which are weakly correlated. More specifically, 10% of the files is globally popular across all regions. The popularity of other files are randomly assigned for each region. The multi-beam mode sends the data via 4 beams. It is observed that the multi-beam satellite achieves a CHR significantly larger than the mono-beam scheme. This is because the multi-beam satellite can better match the local popularity with the user demands. More importantly, the proposed hybrid mode, which prefetches the cache via both the mono-beam and the multi-beam satellite, can further improve the CHR performance. The usage percentage of the mono-beam mode in the hybrid scheme is optimized and equal to the most popular files in all regions, e.g.,  $C_{\text{hybrid},1} = 50$  and  $C_{\text{hybrid},2} = 138$  in use case 1. It is also shown that usecase 1 outperforms usecase 2 due to larger caching capacity. Indeed, this initial result shows one of the major advantages of the hybrid and multi-beam caching with respect to the mono-beam caching, i.e. the access to the geographical diversity of the content. This way, the hybrid and multibeam modes can cache the less correlated content popularity in a more efficient manner.

Fig. 4 shows the CHR of the two use cases as a function of number of beams or cluster of beams according to the fact that we are considering the new hybrid satellite architecture, presented in Section II-A, or a more conventional and less flexible payload. It is shown that the number of beams equal to 4 gives the largest CHR for both scenarios. This is because there the requests form 4 geographical areas which are weakly correlated. In this case, using four beams is sufficiently efficient. It is noted that using more beams than the regions may degrade the CHR since the cost per bit of multi beam mode is larger than the that in the mono beam mode.

Fig. 5 presents the CHR of usecase 1 with Movielens demand [12]. The system parameters are similar in the previous

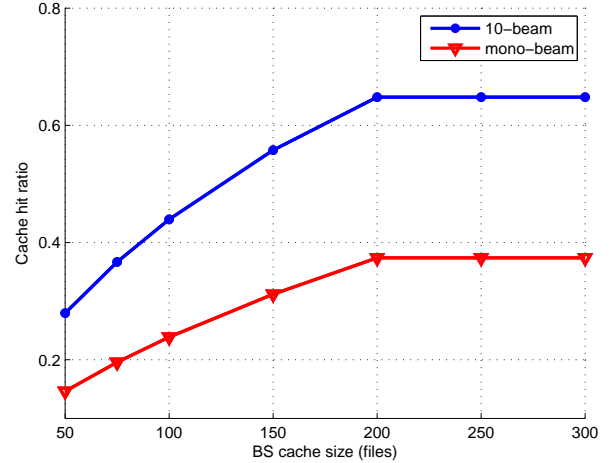


Fig. 5: CHR performance based on Movielens dataset.

section. A significant gain is observed for multi beam satellite (4 beams) compared with the mono beam counterpart. In particular, a CHR gain of 30% is achieved by the 10-beam setup compared with the mono beam. This is because the user demands across the beams are weakly correlated. Therefore, using the multi-beam satellite can serve the local demands better than the mono beam. Indeed, multi-beam satellites bring the CHR in this case to a level which is much more beneficial. It is also observed that the CHR in both setups saturates due to the caching capacity limit. This is due to the fact that the caching capacity of the transmission scheme is lower than the storage capacity of the terminals. This would be not a limitation if we had run the simulation for a longer period of time, where the storage capacity would have been filled in the caching period of several consecutive days.

## VI. CONCLUSIONS AND FUTURE WORKS

We have demonstrated the effectiveness of the multimodal satellite backhauling on edge caching systems in the presence of highly uncorrelated content which is likely to be the trend of the future content consumption. The proposed offline caching algorithm is shown to be capable for flexible deployments of the satellite channels: mono beam, multi beam and hybrid. We have shown that the multi-beam and the hybrid modes become useful with respect to widebeam in caching as the geographical distribution of content popularities becomes uncorrelated. In particular, the future flexible multi-beam payloads with their adjustable coverage will be able to better distinguish between the clusters of beams with less popularity correlation and hence further improve the CHR. As future development, a cost based optimization of the rates of the different modes of the hybrid setup, in order to better fit the traffic distribution, can be realized. In addition, these results suggest a promising extension to on-line caching strategies [15], [16], where the user preference periodically changes time to time.

## VII. ACKNOWLEDGMENTS

We would like to thank Nader Alagha, European Space Agency, The Netherlands, for his useful discussions and sug-

gestions.

This work is partially supported by the Luxembourg National Research Fund under the project FNR CORE ProCAST, and by the European Space Agency, ESA ESTEC, Noordwijk, The Netherlands, under the project title "High Throughput Digital Broadcasting Satellite Systems" (Contract No. 4000120692/17/NL/EM). Opinions, interpretations, recommendations and conclusions expressed herein are those of the authors and are not necessarily endorsed by the European Space Agency.

#### REFERENCES

- [1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update 2016-2021," 2017, white paper.
- [2] "Content delivery networks 3.0," CTOiC White paper. Online: <https://www.slideshare.net/BenSchwarz1/content-deliverynetworks30>.
- [3] W. Han, A. Liu, and V. K. N. Lau, "PHY-caching in 5G wireless networks: design and analysis," *IEEE Commun. Mag.*, vol. 54, no. 8, pp. 30–36, Aug. 2016.
- [4] S. Ramesh, I. Rhee, and K. Guo, "Multicast with cache (Mcache): an adaptive zero-delay video-on-demand service," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 3, pp. 440–456, Mar. 2001.
- [5] B. Evans, M. Werner, E. Lutz, M. Bousquet, G. E. Corazza, G. Maral, and R. Rumeau, "Integration of satellite and terrestrial systems in future multimedia communications," *IEEE Wireless Commun.*, vol. 12, no. 5, pp. 72–80, Oct. 2005.
- [6] H. Linder, H. D. Clausen, and B. Collini-Nocker, "Satellite internet services using DVB/MPEG-2 and multicast web caching," *IEEE Commun. Mag.*, vol. 38, no. 6, pp. 156–161, Jun. 2000.
- [7] C. G. Brinton, E. Aryafar, S. Corda, S. Russo, R. Reinoso, and M. Chiang, "An Intelligent Satellite Multicast and Caching Overlay for CDNs to Improve Performance in Video Applications", in *Proc. 31st AIAA Int. Commun. Satellite Systems Conf.*, 2013-5664.
- [8] A. Kalantari, M. Fittipaldi, S. Chatzinotas, T. X. Vu, and B. Ottersten, "Cache-Assisted Hybrid Satellite-Terrestrial Backhauling for 5G Cellular Networks," in *Proc. IEEE Global Commun. Conf.*, Singapore, 2017, pp. 1-6.
- [9] T. X. Vu, S. Chatzinotas, B. Ottersten, and T. Q. Duong, "Energy Minimization for Cache-Assisted Content Delivery Networks With Wireless Backhaul," *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 332-335, Jun. 2018.
- [10] T. X. Vu, S. Chatzinotas, and B. Ottersten, "Edge-Caching Wireless Networks: Performance Analysis and Optimization," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2827-2839, Apr. 2018.
- [11] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "Web caching and Zipf-like distributions: evidence and implications," in *Annual Joint Conf. of the IEEE Computer and Commun. Societies*, vol. 1, Mar. 1999, pp. 126–134.
- [12] <https://grouplens.org/datasets/movielens/1m/>
- [13] Glyn Thomas. "Enabling Technologies for flexible HTS payloads", 33rd AIAA International Communications Satellite Systems Conference and Exhibition, International Communications Satellite Systems Conferences (ICSSC), (AIAA 2015-4348)
- [14] D. Ciullo, V. Martina, M. Garetto and E. Leonardi, "How Much Can Large-Scale Video-on-Demand Benefit From Users' Cooperation?," in *IEEE/ACM Transactions on Networking*, vol. 23, no. 6, pp. 1846-1861, Dec. 2015.
- [15] L. Lei, L. You, G. Dai, T. X. Vu, D. Yuan and S. Chatzinotas, "A deep learning approach for optimizing content delivering in cache-enabled HetNet," in *Proc. Int. Symp. on Wireless Commun. Syst. (ISWCS)*, Bologna, 2017, pp. 449-453.
- [16] B. N. Bharath, K. G. Nagananda, and H. V. Poor, "A learning-based approach to caching in heterogenous small cell networks," *IEEE Transactions on Communications*, vol. 64, Apr. 2016.
- [17] Christopoulos D., Chatzinotas S., Ottersten B., "Cellular-Broadcast Service Convergence through Caching for CoMP Cloud RANs," in *IEEE Symposium on Communications and Vehicular Technology in the Benelux*, SCVT 2015, Luxembourg, November 2015.