Cache-Assisted Hybrid Satellite-Terrestrial Backhauling for 5G Cellular Networks

Ashkan Kalantari, Marilena Fittipaldi, Symeon Chatzinotas, Thang X. Vu, and Björn Ottersten
SnT, University of Luxembourg. Emails: {ashkan.kalantari, symeon.chatzinotas, thang.vu, bjorn.ottersten}@uni.lu, marilena.fittipaldi@gmail.com

***Abstract*—Fast growth of Internet content and availability of electronic devices such as smart phones and laptops has created an explosive content demand. As one of the 5G technology enablers, caching is a promising technique to off-load the network backhaul and reduce the content delivery delay. Satellite communications provides immense area coverage and high data rate, hence, it can be used for large-scale content placement in the caches. In this work, we propose using hybrid mono/multi-beam satellite-terrestrial backhaul network for off-line edge caching of cellular base stations in order to reduce the traffic of terrestrial network. The off-line caching approach is comprised of content placement and content delivery phases. The content placement phase is performed based on local and global content popularities assuming that the content popularity follows Zipf-like distribution. In addition, we propose an approach to generate local content popularities based on a reference Zipf-like distribution to keep the correlation of content popularity. Simulation results show that the hybrid satellite-terrestrial architecture considerably reduces the content placement time while sustaining the cache hit ratio quite close to the upper-bound compared to the satellite-only method.**

Keywords—5G, content delivery, content placement, hybrid satellite-terrestrial, off-line caching, Zipf-like distribution.

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 percent between 2015 to 2020. 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.

As a solution to this tremendous traffic growth, caching [2], [3] has been suggested as a promising solution, particularly as a key technology for 5G networks [4], to bring the content closer to the users so that the core and access networks are off-loaded and the contents are delivered with less delay. In this direction, the works of [5], [6] investigate the users' service quality in terms of outage probability and delay for uniform channels. These works are extended to non-uniform channels in [7]. Caching is a promising approach to reduce the traffic load and can be more effective if the caching and physical layer are designed together. Physical layer precoding by considering the cache content availability are studied in [8]–[10]. On top of storing the popular content in the cache, which is referred to as local gain, network coding can be used to reduce the network traffic load and get a global gain [11].

In addition, satellite systems have the ability to provide wideband backhaul links and to operate in multi/broadcast modes for immense area coverage. Latest advances such as medium earth orbit constellations (e.g. O3B) or the planned medium earth orbit Mega-constellations, e.g., LeoSat, OneWeb, can provide more intricate ways of backhauling due to their dynamic topologies. The terrestrial backhaul is a multi-hop unicast network, hence, the cached content has to go through multiple links and has to be transmitted individually towards each base station (BS). On the other hand, the satellite backhaul can use its wide area coverage and broadcast content to all BSs or multi-cast contents to multiple groups of BSs. The application of satellite communications in feeding several network caches at the same time using broad/multi-cast is investigated in [12], [13]. The work of [12] 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 [13]. 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. Pushing content to the caches using hybrid satellite-terrestrial network is investigated in [14].

As we see, caching is a promising 5G technique to off-load the terrestrial network. Also, the satellite communications offers a vast area coverage with high transmission rate. Hence, bringing these two technologies together can further off-load the network. The main direction is to combine the satellite and terrestrial telecommunication systems in order to create a hybrid federated content delivery network, which can improve the cache-feeding performance and user experience. In this work, we introduce using hybrid satellite-terrestrial network along with off-line caching approach to off-load the backhaul of the terrestrial network. We introduce a transition matrix in order to generate local content popularity distributions.

This work was supported by the National Research Fund (FNR) of Luxembourg under AFR grant for the project SeMIGod and Shared Access Terrestrial-Satellite Backhaul Network enabled by Smart Antennas SANSA H2020 European project.

This paper is accepted in the IEEE GLOBECOM 2017, Dec. 2017, Singapore. Personal use is permitted, but republication/redistribution requires IEEE permission

Compared to the previous works, first, we show that there is a trade-off between the cache hit ratio and the required time for off-line file placement, which depends on the distribution of local and global content popularities. Second, we propose using multi-beam satellites to place the contents in caches of BSs where the placed contents in the cache of a specific BS match more the most popular content of the local popularity of that BS compared to the case when mono-beam satellite is used for content placement. In this work, we use the *cache hit ratio* as the performance metric to evaluate the selection efficiency of cached content. Cache hit ratio can be measured by the ratio of requests served through the edge cache over the entire requests. In the case of multiple caches, this metric can be derived by averaging over multiple cache hit ratios. We measure the cache hit ratio for mono/multi-beam hybrid satellite-terrestrial networks with respect to cache memory size. We also quantify the cache hit ratio of the hybrid satellite-terrestrial scheme for different content popularity distributions and compare it versus the benchmark schemes.

The remainder of this paper is organized as follows. In Section II, the mono/multi-beam hybrid satellite-terrestrial architectures and the related assumptions are mentioned. The model of content popularity distribution and the process to generate local and global popularities are discussed in Section IV. In Section V, we present simulation results by comparing the proposed methods with the benchmark schemes. Finally, we draw the conclusions in Section VI.

Notation: Upper-case and lower-case bold-faced letters are used to denote matrices and column vectors, respectively. The operator $\lfloor\cdot\rfloor$ rounds to the nearest integer less than or equal to the number inside the bracket and $|\cdot|$ represents the absolute value of a scalar.

II. HYBRID SATELLITE-TERRESTRIAL ARCHITECTURE

In this section, we introduce the hybrid satellite-terrestrial architecture. We propose telecommunication content delivery network architectures that include mono/multi-beam satellites and terrestrial components as shown in Figures 1 and 2. Such architectures can benefit wide area coverage of satellite broad/multi-casting. In these architectures, each edge BS is equipped with a limited memory cache. In addition to the terrestrial backhaul fiber, the edge caches are equipped with two interfaces to receive the content, one for terrestrial backhauling and the other for satellite backhauling. In these architectures, both satellite and terrestrial networks are used for placing the files in the caches of each cache-enabled edge cellular base station (BS). The mono-beam architecture uses wide area broadcast based on global content popularity for content placement in the caches. The multi-beam architecture places the content in the caches according to the content popularity of each beam, which results in more accurate content placement compared with the architecture with mono-beam satellite. In the multi-beam architecture, it is possible to adapt the content placement to the popularity of smaller regions and refine the content placement compared to the mono-beam architecture. Hence, the multi-beam architecture results in higher cache hit ratio compared to the mono-beam architecture.

In the following part, we describe the designed caching algorithms for the proposed mono/multi-beam hybrid satellite-terrestrial architectures.

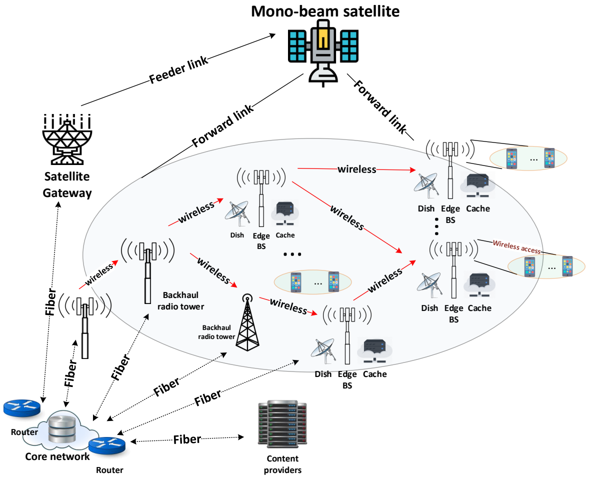

Fig. 1: Hybrid satellite-terrestrial network using a mono-beam satellite.

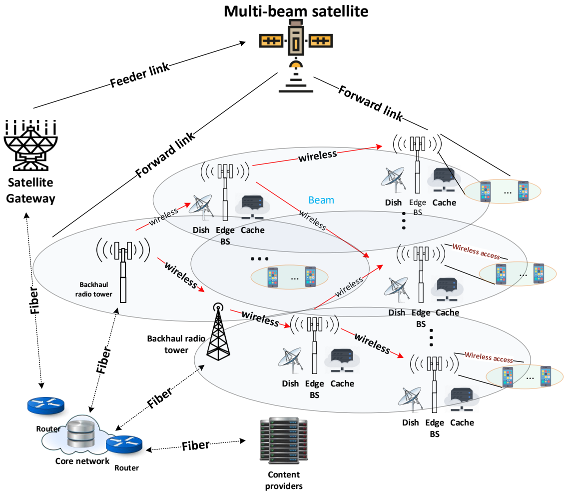

Fig. 2: Hybrid satellite-terrestrial network using a multi-beam satellite for more accurate content placement phase.

III. OFF-LINE CACHING ALGORITHMS

In this part, we describe the suggested caching algorithms for the architectures mentioned in Section II. Here, we focus on off-line, non-real time, hybrid satellite-terrestrial caching, which is comprised of content placement and content delivery phases. The content placement is carried out in off-peak hours and is divided into two rounds where the satellite broadcast and the terrestrial unicast are used for content placement in the first and second round, respectively. In the first round, the satellite performs the content placement based on the global content popularity, mono-beam satellite case, or global content popularities, multi-beam satellite case. In the second round,

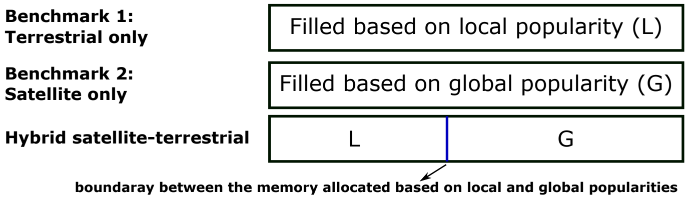

Fig. 3: Schematic of the proposed hybrid satellite-terrestrial and the benchmark caching algorithms.

the terrestrial network accomplishes the content placement based on local content popularity. In the proposed off-line caching algorithm, the cached content remain intact during the content delivery period and are updated during the next content placement period. For example, putting videos of YouTube or Netflix on the caches in advance during the night, i.e., content placement, and serving the user requests through the cached video files or the terrestrial backhaul during the day, i.e., content delivery.

In this work, we consider three different caching schemes as illustrated in Fig. 3. The hybrid satellite-terrestrial scheme is the proposed scheme as well as the terrestrial-only and satellite-only benchmark schemes. We assume that there exists a file library with arbitrary amount of files and the users' request falls within this file library. Furthermore, it is presumed that the edge caches keep track of the requested content by the users. The recorded requests by each cache corresponds to the local content popularity of the corresponding edge BS. The local content popularities are transmitted to the satellite gateway to calculate the global content popularity. In the mono-beam architecture of Fig. 1, the global file popularity is derived by averaging over all the local content popularities. On the other hand, in the multi-beam architecture, the global content popularity for each beam is derived using the local content popularities of the edge BSs located in the corresponding beam. In the hybrid satellite-terrestrial scheme of Fig. 3, each cache is filled with the most local popular related files till the threshold, then the rest of the cache is filled with the most global popular files. In the case of multi-beam architecture, the global popularity of a beam is used for content placement in the caches of the BSs covered by that beam. Note that the files placed based on the local popularity are removed from the global content popularity profile(s) before content placement based on global popularity.

In the terrestrial-only method of Fig. 3, the files are placed in the caches based on local content popularities which results in the highest cache hit ratio. Hence, we use it as the upper bound to evaluate the cache hit ratio of the hybrid mono/multi-beam architecture. On the other hand, in the satellite-only method, the content placement is carried out only based on global content popularity profile(s), which results in the lowest cache hit ratio, hence, it is used as a lower bound for the cache hit ratio.

Since the multi-beam hybrid satellite-terrestrial architecture considers the global popularity of each beam for the content placement phase, it results in placing the content in the edge caches that are more similar to the most popular files of the local content popularity of the corresponding edge BSs. Therefore, using a multi-beam satellite for the content placement phase results in a higher cache hit ratio compared to using a mono-beam satellite.

In the following section, we mention the content popularity model.

IV. CONTENT POPULARITY MODEL

In this section, we define the content popularity distribution model and the proposed approach to generate the local content popularity distribution. To generate the local and global content popularities, first, we define the model for a reference content popularity distribution. Here, we assume that the reference content popularity follows the Zipf-like distribution as [15]

$$p\left(i\right) = \frac{\Omega}{i^{\alpha}}, \tag{1}$$

where $\Omega = \left(\sum\limits_{i=1}^{N} \frac{1}{i^{\alpha}} \right)^{-1}$ is the normalization factor to have $\sum\limits_{i=1}^{N} p\left(i\right) = 1$. The parameter α controls the distribution tail. For illustration, a lower α means that more files have similar popularity and a higher α means that a few files have very high popularity and a large number of files have very low popularity.

The local content popularity generation process for each BS is as follows. Assume that $\mathbf{p}$ is an $N \times 1$ vector which accommodates the elements of the reference popularity $p\left(i\right)$ introduced in (1) for $i = 1, ..., N$. Let us assume that this is the first local content popularity distribution that we have generated. To generate the second local content popularity, we swap two elements of the vector $\mathbf{p}$. To swap these two elements of $p\left(i\right)$, first, we randomly select an element of $\mathbf{p}$, which is a file with a specific popularity. Second, we use the probability matrix $\mathbf{P}$ defined in (2) to find the second file which is going to the swapped with the first file.

$$\mathbf{P} = \begin{bmatrix} 0 & p_{1,2} & \cdot & \cdot & \cdot & p_{1,N} \\ p_{1,2} & 0 & \cdot & p_{2,l} & \cdot & p_{2,N} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & p_{m,2} & \cdot & 0 & \cdot & p_{m,N} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ p_{N,1} & p_{N,2} & \cdot & p_{N,l} & \cdot & 0 \end{bmatrix}. \tag{2}$$

The m-th row of the matrix $\mathbf{P}$ shows the probability of swapping the m-th file, the first randomly selected file, with other files, the second file. In the defined swapping matrix of (2), $\mathbf{P}$, the entries closet to the diagonal entries have the highest value and then these values decrease as we move away toward further off-diagonal entries. This decrement can have arbitrary intensity. In other words, given the m-th and l-th element of $\mathbf{P}$, $p_{m,l}$ becomes lower as $|m - l|$ increases. Therefore, the files with closer popularity have a higher chance to be swapped for local content popularity generation. This defined structure for matrix $\mathbf{P}$ is less likely to generate consecutive local content popularity distributions that have drastic difference in content popularity. This results in correlation between the consecutive generated local content popularities since the swapped files are more likely to be closer to each other.

To generate the next local content popularity distribution, we repeat a similar procedure as before with the difference

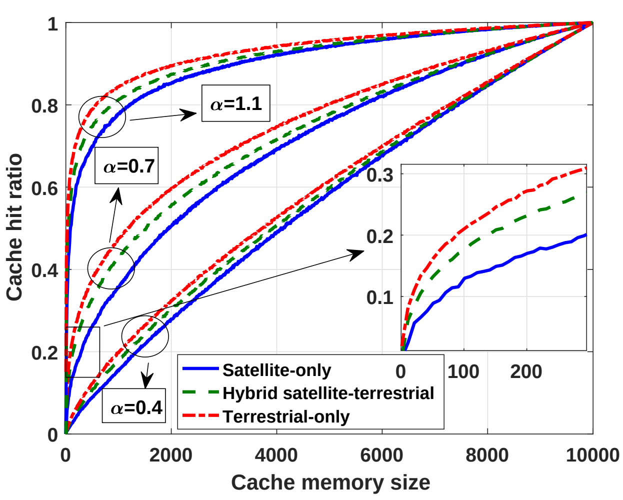

Fig. 4: Comparison among the cache hit ratio of satellite-only, hybrid satellite-terrestrial, and terrestrial-only backhaul architectures versus the cache memory size for different values of α.

that we exclude all previously swapped files from the files which are going to be swapped. In total, we can generate $[N/2]$ possible local popularity distributions by swapping the elements of $p(i)$ in (1) using the transition matrix defined in (2).

To generate the global content popularity in the mono-beam architecture, we average over all the local content popularities. On the other hand, to generate the global content popularity of each beam in the multi-beam architecture, we average over the local content popularities of the corresponding beam. Note that the generated local and global content popularities do not necessarily follow the Zipf distrconten popularities do not follow the Zipf distribution since

V. SIMULATION RESULTS

In this section, we simulate the performance of the proposed off-line caching algorithm in the mono/multi-beam hybrid satellite-terrestrial architectures and compare them with the satellite-only and terrestrial-only benchmark architectures. We measure the cache hit ratio and the required time for content placement to evaluate the proposed caching algorithms versus the benchmark ones. Here, we assume that half of the cache size of a BS is filled with the most popular files based on the local content popularity distribution of that BS and the other half of the cache is filled with the most popular content based on the global content popularity distribution. In all simulations, we consider a library with 10^4 files and the satellite and terrestrial backhaul (for each BS) throughputs are considered to be 50 M byte/s. during the content delivery phase, we consider 10^5 of user request. In simulations, we consider one period, e.g., a day or weak, consisting of content placement and content delivery. The content placement is carried out during the night, off-peak hours, and the content delivery is performed during day, peak hours. The number of requests can be increased to simulate longer periods. The simulations are carried out in MATLAB.

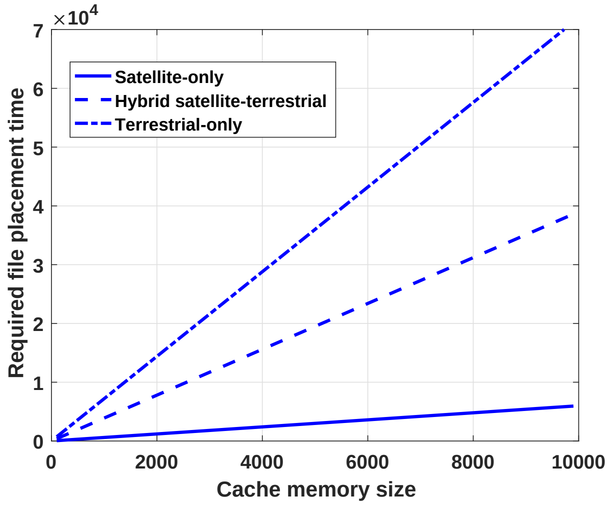

Fig. 5: Comparison among the required time for content placement by satellite-only, mono-beam hybrid satellite-terrestrial, and terrestrial-only backhaul architectures for different cache memory sizes when each file is 30 M bytes in size.

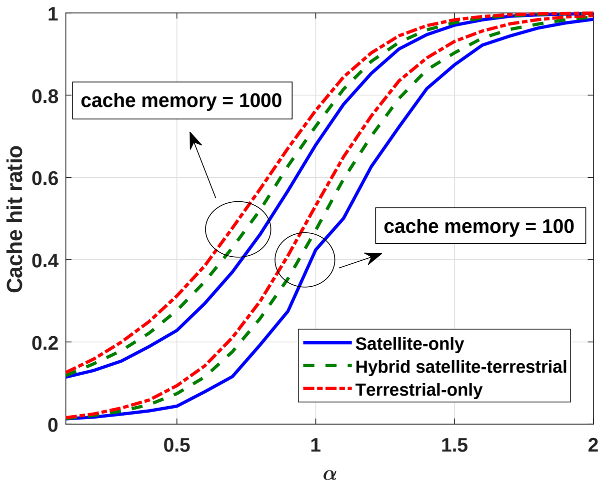

Fig. 6: Cache hit ratio with respect to α for different cache memory sizes of mono-beam hybrid satellite-terrestrial scheme.

In the first scenario, we evaluate the cache hit ratio and the required time for file placement, 10^4 files, with respect to the cache memory size. These simulations are carried out for the proposed mono-beam hybrid satellite-terrestrial architecture and the benchmark schemes when there are 8 BSs in the satellite beam and all the library files have the same size equal to 30 M byte. As wee see in Fig. 5, the satellite-only backhauling requires the least amount of time for content placement. This is due to the fact that a single satellite broadcast can send the content to a vast number of BSs thanks to immense area coverage. On the other hand, the terrestrial-only method takes the most time for content placement since the content need to be sent to each BS through the fiber separately. The hybrid satellite-terrestrial required time for

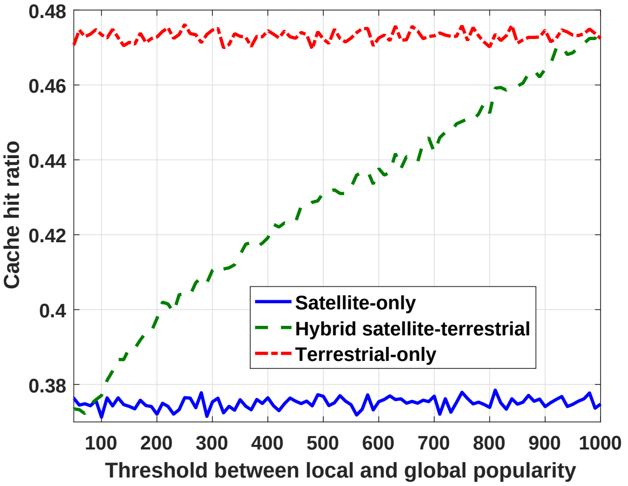

Fig. 7: Cache hit ratio with respect to the threshold of memory allocation between local and global popularities when $\alpha = 0.7$ and total cache memory is 1000.

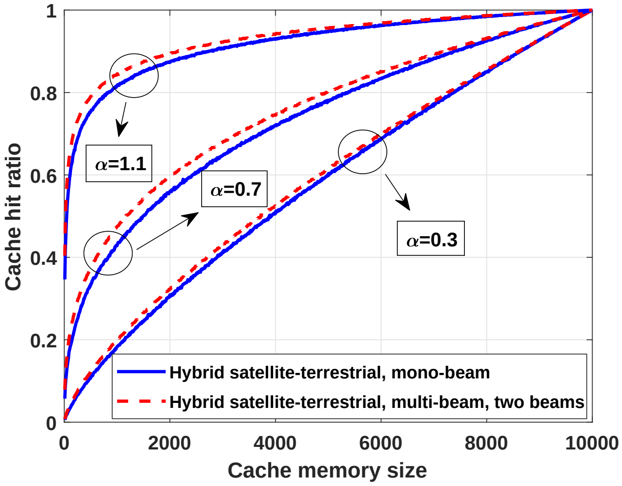

Fig. 8: Comparison of the cache hit ratio between hybrid satellite-terrestrial architectures with mono-beam and multi-beam satellites for different values of α.

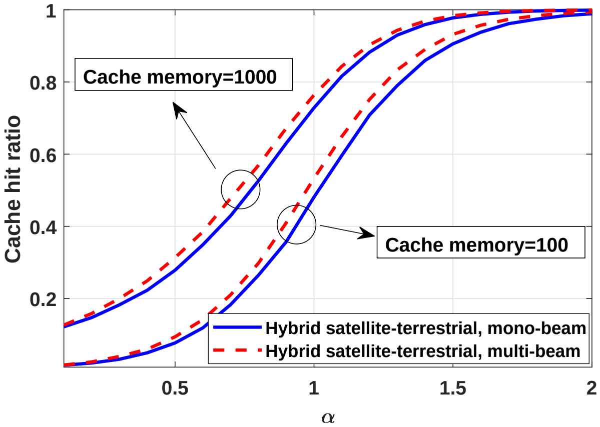

Fig. 9: Cache hit ratio versus α for various cache memory sizes for mono/multi-beam hybrid satellite-terrestrial architectures.

content placement is between the required time of satellite-only and terrestrial-only methods since it uses both satellite backhaul and terrestrial backhauls. In the hybrid method, the satellite backhaul is used to place the content based on global popularity. On the other hand, the terrestrial fiber is used for content placement based on the local popularity of each BS, which requires using the fiber.

We present the cache hit ratio for the 8-th BS with respect to the cache memory size in Fig. 4 for different values of α and the required file placement time for all BSs with respect to the cache memory size in Fig. 5. As we see, the cache hit ratio of the proposed mono-beam hybrid satellite-terrestrial architecture is closer to the terrestrial-only benchmark scheme compared to the satellite-only benchmark for specific ranges of the cache memory. For example, for a cache memory with the capacity of 200 files and $\alpha = 0.7$, the cache hit ratio of the hybrid satellite-terrestrial method is 0.04 lower than the terrestrial-only method while the cache hit ratio of the satellite-only method is 0.1 lower than the terrestrial-only scheme. On the other hand, we see that the hybrid satellite-terrestrial method considerably cuts on the required file placement time. Both hybrid satellite-terrestrial and satellite-only methods cut on the required time for content placement considerably, however, the cache hit ratio of the hybrid approach is much closer to the terrestrial-only approach, upper bound, compared to the satellite-only method. This is due to the fact that the user requests around each BS are based on the local content popularities of that BS, however, the global popularity may not be an accurate representation of the local popularities. Hence, the hybrid scheme results in a higher cache hit ratio compared to the satellite-only method. In addition, we observe that for relatively low or high cache memory sizes, the cache hit ratio of all approaches are close. For relatively low value of popularity distribution parameter, $\alpha = 0.4$, we see that the difference between the proposed and the benchmark schemes is lower since files do not have much difference in popularity. For relatively high values of popularity distribution parameter, $\alpha = 1.1$, the cache hit ratio difference between the proposed and the benchmark schemes is noticeable for relatively low values of cache memory.

Next, we show the cache hit ratio of the 8-th BS with respect to the α in Fig. 6. We observe that for relatively low or high values of α, the cache hit ratio of all approaches are almost the same. For low values of α, the files have close popularity values and for high values of α, very few files have high popularity and the rest of the files have close and very low amount of popularity. Hence, based on the aforementioned information, the cache memory does not influence the cache hit ratio for rather low and high values of α. For values of $0.4 \leq \alpha \leq 1.6$, the difference among the three caching schemes increases as the cache memory decreases. As the last part of the first scenario, we investigate the effect of the cache memory threshold between content placement based on local and global content popularity distributions. The cache hit ratio of the 8-th BS with respect to the cache memory allocation to content placement based on local and global content popularities is shown in Fig. 7. As we see, the cache hit ratio of the hybrid satellite-terrestrial scheme falls between that of the satellite-only and terrestrial-only schemes. Also, as the memory allocated to file placement based on local popularity increases, the cache hit ratio consistently advances toward the terrestrial-only approach.

In the second scenario, we investigate the effect of multi-beam hybrid satellite-terrestrial architecture on the cache hit ratio and compare it versus the mono-beam hybrid satellite-terrestrial architecture. In this scenario, we consider two beam of a multi-beam satellite where each beam covers a specific region. There are 9 and 3 cache-enabled BSs in the first and second beam, respectively. Fig. 8 shows the average cache hit ratio of the region covered by the second beam with respect to the cache memory size for different values of α. As we see, the multi-beam architecture can provide up to 0.05 better cache hit ratio over a long range of cache memory sizes compared to the mono-beam architecture. This is due to the fact that a multi-beam satellite can perform the content placement based on the global popularity of each region. Compared to the global content popularity of a mono-beam satellite, the global content popularity of a specific beam of a multi-beam satellite is closer to the local content popularities of the BSs covered by that beam. This results in higher cache hit ratio for multi-beam hybrid satellite-terrestrial architecture compared to the mono-beam hybrid satellite-terrestrial architecture.

The cache hit ratio for the region covered by the second beam with respect to α is presented in Fig. 9. As we see, the multi-beam architecture outperforms the mono-beam over a long range of α. For relatively low and high values of α, the mono-beam and multi-beam architectures have close values of cache hit ratio. This can be justified similar as the explanation for Fig. 6.

VI. CONCLUSIONS

In this paper, we proposed mono/multi-beam hybrid satellite-terrestrial architectures along with an off-line caching algorithm to reduce the required time for content placement and preserving the cache hit ratio. We performed extensive simulations with respect to various parameters to investigate the performance of the proposed off-line caching algorithm for the mono/multi-beam hybrid satellite-terrestrial architectures. The results showed that the proposed hybrid satellite-terrestrial architecture is able to considerably reduce the required time for file placement while keeping the cache hit ratio very close to the terrestrial-only upper bound. Based on the simulations, we figured out that the hybrid satellite-terrestrial caching method can be beneficial for content popularity distribution with the range of $0.4 \leq \alpha \leq 1.6$. In addition, in order to have improvement in cache hit ratio, the required cache memory for the proposed off-line hybrid satellite-terrestrial scheme varies depending on the content popularity distribution parameter α.

It was demonstrated that the proposed multi-beam hybrid satellite-terrestrial off-line caching scheme can further improve the cache hit ratio compared to the mono-beam hybrid architecture. This is done by considering the global popularities of each beam which covers a smaller region compared to the mono-beam hybrid satellite-terrestrial architecture and can improve the accuracy of content placement. We also revealed that for a long range of values for α, the multi-beam architecture outperforms the mono-beam architecture in terms of the cache hit ratio.

REFERENCES

[1] “Cisco visual networking index: Global mobile data traffic forecast update, 2015–2020,” CISCO White paper. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.pdf

[2] E. Bastug, M. Bennis, and M. Debbah, “Living on the edge: The role of proactive caching in 5G wireless networks,” *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 82–89, Aug. 2014.

[3] Y. Kawamoto, Z. M. Fadlullah, H. Nishiyama, N. Kato, and M. Toyoshima, “Prospects and challenges of context-aware multimedia content delivery in cooperative satellite and terrestrial networks,” *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 55–61, Jun. 2014.

[4] 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.

[5] 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.

[6] N. Abedini and S. Shakkottai, “Content caching and scheduling in wireless networks with elastic and inelastic traffic,” *IEEE/ACM Trans. Netw.*, vol. 22, no. 3, pp. 86–874, Jun. 2014.

[7] B. Zhou, Y. Cui, and M. Tao, “Optimal dynamic multicast scheduling for cache-enabled content-centric wireless networks,” in *IEEE Int. Symp. on Inf. Theory (ISIT)*, Jun. 2015, pp. 1412–1416.

[8] X. Peng, J. C. Shen, J. Zhang, and K. B. Letaief, “Joint data assignment and beamforming for backhaul limited caching networks,” in *IEEE Int. Symp. on Personal, Indoor, and Mobile Radio Commun. (PIMRC)*, Sep. 2014, pp. 1370–1374.

[9] H. Zhou, M. Tao, E. Chen, and W. Yu, “Content-centric multicast beamforming in cache-enabled cloud radio access networks,” in *IEEE Global Commun. Conf. (GLOBECOM)*, San Diego, CA, Dec. 2015.

[10] D. Christopoulos, S. Chatzinotas, and B. Ottersten, “Cellular-broadcast service convergence through caching for CoMP cloud RANs,” in *IEEE Symp. on Commun. and Veh. Tech. in the Benelux (SCVT)*, Luxembourg, Nov. 2015.

[11] M. A. Maddah-Ali and U. Niesen, “Fundamental limits of caching,” *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.

[12] 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.

[13] C. 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 *AIAA Int. Commun. Satellite Syst. Conf.*, Florence, Italy, Oct. 2013.

[14] 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.

[15] 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.