On Edge Caching in Satellite–IoT Networks

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY 4.0

SUBMISSION DATE / POSTED DATE

21-10-2021 / 28-10-2021

CITATION

Ngo, Quynh; Phan, Tran Khoa; Xiang, Wei; Mahmood, Abdun; Slay, Jill (2021): On Edge Caching in Satellite–IoT Networks. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.16842979.v1>

DOI

[10.36227/techrxiv.16842979.v1](https://doi.org/10.36227/techrxiv.16842979.v1)

On Edge Caching in Satellite-IoT Networks

Quynh T. Ngo, Khoa T. Phan, Wei Xiang, Abdun Mahmood, and Jill Slay

Abstract—The implementation of the Internet of Things (IoT) is mostly done through cellular networks which do not cover the whole world. In addition, the explosive growth of global Internet access demand introduces the need for integrating satellites with cellular IoT networks for coverage extension and backhaul offloading. Operating hybrid satellite-IoT (SIoT) networks, however, might incur excessive service latency and high satellite bandwidth consumption. To tackle these issues, edge caching technology has been considered in SIoT. This article reviews existing research on edge caching-based SIoT networks with illustrative performance evaluation. Various caching design criteria with a focus on two-tier cache-enabled SIoT are discussed. In addition, open research problems on edge caching in SIoT are identified as future research directions and opportunities.

I. INTRODUCTION

To offer Internet services across all geographic regions, deploying cellular IoT networks over satellite systems has attracted extensive attention from both industry and academia. For example, the third Generation Partnership Project (3GPP) has initiated a study item for new radio supporting integrated satellite into terrestrial networks [1]. The Satellite and Terrestrial Network for 5G (SaT5G) consortium funded by the European Commission aims to integrate satellite communications into 5G by defining satellite-based backhaul and traffic offloading solutions [2]. SpaceX's Starlink satellite project, Amazon and Blue Origin's project Kuiper, and OneWeb plan to build large satellite constellations to deliver high speed broadband Internet globally in 2021. According to the Allied Business Intelligence research, it is through satellite communications that the global IoT service will be achieved with an estimation of 24 million IoT connections by 2024 [3].

The satellites in a general SIoT architecture can either be low earth orbit (LEO), medium earth orbit (MEO), or geosynchronous earth orbit (GEO). GEO satellites have been traditionally used to provide static network topology, non-interrupting transmission between the satellite transceiver and terrestrial stations, and wide coverage even when only a single satellite is deployed. However, due to the high altitude, GEO satellites are more suitable for IoT applications that offers latency non-sensitive services. LEO satellites have been developed to form large LEO constellations that allow global coverage with high-throughput broadband rate and low service latency. Using LEO satellites in SIoT poses a challenging problem in dynamic network topology since they are only in contact with ground stations for a brief period of time due to their traveling speed. For MEO satellites, a constellation of

twenty satellites (O3b) has been placed to deliver low-latency broadband services. The O3b constellation is unique as it has successfully achieved the trade-off between constellation size and service latency [4]. The O3b constellation also poses the dynamic network topology challenge since it operates in the non-geostationary earth orbit. In a typical satellite-terrestrial network (STN), GEO satellites usually belong to the space-based backbone network. MEO and LEO satellites are part of the space-based access network which is responsible for the realization between IoT space and intersatellite links, data exchange and relaying between the space-based and terrestrial-based backbone networks.

Fig. 1: An example of the satellite-IoT network architecture.

An example of the SIoT network architecture when deploying satellites for IoT in rural areas is depicted in Fig. 1. In this use-case, the ground stations cannot access core networks directly due to their remote locations, i.e., rural area, dispersed terrain, marine offshore bases or vessels. Services are usually served through satellites relaying data from core networks. Deploying satellites for IoT focuses on solving the problems of excessive service delay due to satellite altitude and costly service due to pricey satellite bandwidth consumption. To tackle these issues, edge caching has been introduced to relocate content storage during off-peak hours to edge devices, e.g., satellites, ground stations, that are closer to IoT devices. In terrestrial wireless systems, caching popular contents is used to reduce the network congestion and delivery latency. Migrating edge caching technique to SIoT will not only reduce the service latency but also avoid re-transmission of the same data at satellites, hence, bringing down the satellite bandwidth consumption. A typical caching protocol usually consists of two phases: the placement of files into cache units during

Quynh T. Ngo, Khoa T. Phan, Wei Xiang, and Abdun Mahmood are with the Department of Computer Science and Information Technology, La Trobe University, VIC 3086, Australia. Email: {T.Ngo,K.Phan,W.Xiang,A.Mahmood}@latrobe.edu.au.

Jill Slay is with the School of Computer and Information Science, University of South Australia, Adelaide, Australia. Email: {Jill.Slay}@unisa.edu.au.

off-peak hours and the delivery of files to end-devices upon demand. Designing edge caching usually involves studying (i) cache deployment on the location selection for cache units; (ii) content placement on the file selection to prefetch at each cache unit; and (iii) the impact of edge caching on the network performance. In SIoT, cache deployment takes place at ground stations or satellites forming a single-tier cache-enabled network, and at both ground stations and satellites forming a two-tier cache-enabled network. Moreover, when incorporating edge caching into practical SIoT, one concern regards on-board capacities of the satellites in terms of both low processing units and limited power supply. Regardless, recent advances in hardware components, power generating, and energy harvesting technologies have enhanced on-board processing units to enable innovative communication technologies [4].

The remainder of this article is organized as follows. Section II describes the related work on edge caching in satellite-IoT. Section III discusses the design criteria of edge caching in SIoT. A performance evaluation between different cache-enabled SIoT systems is presented in Section IV. Section V is on the directions and opportunities for future research. Finally, conclusions are made in Section VI.

II. THE STATE-OF-THE-ART OF EDGE CACHING IN SATELLITE TERRESTRIAL NETWORKS

Although there are three types of satellite can be employed in STN as introduced in Section I, research work on applying edge caching technique into MEO-based STN is very limited since O3b is the only commercially successful MEO constellation [4]. It is more common to employ LEO and GEO satellites in cache-enabled STN. A classification of research works on cache-enabled STN is summarized in Table I. In Table I, research works are classified based on satellite types deployed (LEO or GEO), network architecture (single-tier versus two-tier), cache deployment strategies (where to cache), performance evaluation criteria, content exploration (content query and placement problems), and content delivery.

A. Cache-enabled STN with dynamic network topology

In LEO-based STN, due to the dynamic network topology, determining the end-to-end paths for content delivery to end-users is one of the most challenging problems. The content delivery delay is prolonged not only by the transmission distance but also by the excessive repeated transmissions due to the intermittent connectivity between satellites and ground stations. Traditional routing strategies are not effective enough since repeated transmissions are still required [5]. By enabling in-network caching mechanism, the works [5], [6] propose on-path caching file distribution strategies that can capture the time-varying network topology of LEO STN. Both network models in [5], [6] deploy cache units at ground stations/relays forming single-tier cache-aided systems. With the goal to assemble the fractional transmission chances of inter-satellite links as cross-slot contacts, [5] proposes an event updating graph model that can capture the real-time topology state. Based on the event updating graph model, [5]

selects the caching satellite candidates following a minimum time-evolving covering set (MTCS) algorithm. The file access procedure is designed with a query-based minimum delay distribution path algorithm. The proposed strategy in [5] is also proved to lower the cache overhead and file access delay compared to the network central location (NCL), a current typical method for file distribution in dynamic network topology. Also capturing the time-varying network topology, [6] proposes a cross-time-slot graph file distribution and a back-tracing partition directed on-path caching distribution mechanism to reduce redundant transmissions of the target files, thereby decreases the caching overhead. The cache deployment in [6] follows a modified Dijkstra routing algorithm, and the cache content placement follows a popularity-aware multiple region cooperative strategy. File access delay of the designed mechanism in [6] is proved to be better than NCL.

B. Cache-enabled STN for cellular traffic offloading

Edge caching and satellite deployment can help offload cellular traffic and enhance the cellular network capacity. Optimizing traffic offloading can improve caching performance. Leveraging the broadcast characteristic of satellite to efficiently provide traffic offloading for terrestrial networks, [7] integrates LEO satellite into terrestrial network with a cooperative caching scheme. The proposed scheme allows terrestrial access points and satellite sharing the cached content, hence offload the traffic from base stations. The offloading problem in [7] is designed to maximize the network energy efficiency in terms of content placement, power allocation, and cache sharing decisions. Leveraging the potential of edge caching, [8] proposes a two-tier cache-enable system where caches are placed in both base stations and LEO satellite. To offload cellular traffic, [8] proposes a cooperative multicast-unicast transmission mechanism for satellite, where the popular requests are served in multicast phase and the unpopular ones in unicast phase. An extra terrestrial link between the core networks and base stations is designed to maintain services when the satellite connection is absent. The intention of enabling caching capacity at LEO satellite in [8] is to offer one service simultaneously. However, [8] does not consider the handover problem of the link between satellite and base stations. In addition, having a backup terrestrial link to account for the dynamic network is not practical enough because the terrestrial link to core networks is usually missing in remote areas or during natural disaster. Proposing off-line caching approaches to offload terrestrial backhaul networks, [9] designs a cache-aided GEO STN focusing on the multimode satellite backhaul in two use cases, i.e., dense urban areas and sparsely populated regions. [9] introduces a file delivery over unidirectional transport protocol for implementing caching over satellite channels. The system in [9] is proved to be effective by virtue of the cache hit ratio and cost per bit for satellite transmission. Using deep reinforcement learning in optimizing the long-term average network delay, [10] proposes a transmission protocol for non-cached files at one ground station to be fetched from the nearest station in the network. The deep reinforcement learning algorithm proposed is a low-

TABLE I: Summary of research works on edge caching in STN

Related work	Satellite		Network architecture		Cache deployment	Traffic offloading	Criteria	Content exploration	Content delivery
	LEO	GEO	Single-tier	Two-tier					
Z. Yang <i>et al.</i> [5]	✓		✓		✓		access delay		
Y. Li <i>et al.</i> [6]	✓		✓		✓		access delay		
J. Li <i>et al.</i> [7]	✓		✓			✓	energy efficiency	✓	
L. Liu <i>et al.</i> [8]	✓			✓		✓	access delay		✓
T. Vu <i>et al.</i> [9]		✓	✓			✓	cache hit ratio		✓
N. Garg <i>et al.</i> [10]		✓	✓			✓	service delay		✓
K. An <i>et al.</i> [11]		✓	✓				outage probability	✓	✓
A. Kalantari <i>et al.</i> [12]		✓	✓				cache hit ratio	✓	
H. Wu <i>et al.</i> [13]		✓		✓			satellite BW	✓	✓

complexity approach and achieves better performance over the most popular caching policy.

C. Content exploration

In STN, the content exploration refers to choosing which content should be cached in which satellite or ground stations. This process is the first phase of operation in cache-enabled STN and is usually executed during off-peak hours. The cache placement strategies are designed to optimize some network performance metrics such as cache hit ratio, satellite's spectral efficiency, energy efficiency, delivery delay etc. Caching technique in general has two types of placement strategies: coded and uncoded placement. The coded placement strategy [7] is employed to save the cache space consumption by segmenting each file into small parts then having them encoded. The uncoded placement [11], [12], on the other hand, directly stores the whole file or part of it into the cache. Using block placement strategy in addition to a cache sharing scheme, [7] optimizes the energy efficiency and traffic offloading for LEO STN. The proposed caching scheme provides significant improvement in terms of energy efficiency under the same traffic offloading performance compared to traditional terrestrial network with cooperative transmission mechanism. Considering the most popular and uniform content-based cache placement schemes for GEO STN, the proposal in [11] shows a substantial improvement in outage probability over the traditional approach without caching. To offload the backhaul of terrestrial networks, [12] proposes using GEO STN in combination with an off-line edge caching algorithm. Caches are placed at base stations, in which each cache is divided into two parts: the first part is filled with the most local popular files whilst the other part is filled with the most global popular files. The proposed off-line caching algorithm is designed to improve cache hit ratio. Note that the cache placement strategies proposed in [7], [11], [12] are for single-tier cache systems where the caching capacity is enabled at ground stations. In case of two-tier cache systems [13], global popular content can be cached at satellites to leverage satellites' broadcast nature. This placement strategy leads to the improvement of satellite spectral efficiency and content retrieving delay.

D. Content delivery

Content delivery strategies in STN are transmission mechanisms when serving the requested content to users. The content

delivery design usually relates to the location that the content should be transmitted from, the power consumption, and the satellite bandwidth consumption for transmitting the content. The performance metrics, i.e., throughput, power consumption, bandwidth consumption, delay are usually considered when designing the content delivery strategies. Combining terrestrial cooperative manner and satellite broadcast functionality to relay requested content from core networks, [9]–[11] design the content delivery strategies to improve the cost per bit for satellite transmission, the long-term average network delay and the network outage probability, respectively. Enabling cache capacity at satellite in addition to at ground stations, [8], [13] focus on boosting the network average wait time and satellite bandwidth consumption, respectively. The two-tier cache-enabled system models used in [8], [13] can: (i) reduce service latency since more content is stored closer to the end-users; (ii) efficiently utilize satellite bandwidth as the content is prefetched at the satellite during off-peak hours; Furthermore, when the same content is requested by multiple users, having the content cached at satellite and broadcasting at once can save satellite bandwidth.

The main differences when migrating edge caching to STN lie in the channel modeling and the nature of non-geostationary satellites. In conventional terrestrial networks, the channels are commonly modeled with Rayleigh distribution, while satellite channels are modeled with a more complicated combination of Rayleigh distribution for scatter components and Nakagami distribution for line of sight components [14]. It is more challenging when designing and evaluating edge caching schemes for STN. When employing non-geostationary satellites in STN, the dynamic network topology due to the fact that satellites are in contact with ground stations for only a period of time also creates a great challenge in designing caching policy and transmission scheme comparing to conventional terrestrial networks.

III. DESIGN CRITERIA OF EDGE CACHING IN SIoT

When designing a caching scheme in SIoT, several aspects can be considered to improve both the caching and the network performances. In this section, the design criteria on cache content placement, service latency, satellite uplink/downlink bandwidth consumption and resource allocation will be discussed.

A. Cache content placement

Content placement is the process of choosing which files to cache at the edge devices. An effective content placement strategy is reflected through a high cache hit ratio. Cache hit ratio is defined as the ratio between the number of requested files prefetched at cache units over the total number of files cached [12]. With more requested files are served by the cached devices, a higher cache hit ratio will be achieved. This ratio plays a crucial role in enhancing the network performance, especially for two-tier cache-enabled systems since having a better cache hit ratio will lower the file access delay and the satellite bandwidth consumption. A trivial way to enhance cache hit ratio is to enlarge the caching capacity of edge devices. However, there is a trade-off between the cache size and the infrastructure cost since the cache units can be placed at satellites in STN. Other direction to improve cache hit ratio is through cache placement design. Rather than using general popular file caching policy as in [9], [12], requested content pattern can be monitored and applied into cache placement design. Deep learning can also be applied in this process as in [10]. Since the cache placement design phase is off-line and is usually done at the gateways/servers with high computational capacity, machine learning is a potential mean to achieve a precise requested content pattern in content placement design.

B. Service delay

Service delay is related with the successful delivery probability (SDP) concept. SDP is defined as the probability a requested file is served within the user active time [14]. A high average SDP implies low service delays in a network. Expanding the caching capacity of SIoT from single to two tiers in addition to having cache placement strategies with a high cache hit ratio will significantly improve the SDP since more requested content is moved closer to end-users. A closed-form expression for SDP of two-tier cache-enabled GEO based SIoT system has been derived in [14]. An example of SDP of a file cached in single-tier and two-tier models is illustrated in Fig. 2 with μ_2 denotes the percentage of the file cached at satellite. It is observed that using two-tier caching model can help achieve better SDP. With more than seventy percent of the file cached, it is a guaranteed service delivery for both caching models. Relaying protocol also plays an important role in leveraging the network SDP. The delivery time will be shorter for non-cached content with suitable relaying protocol. Some common protocols can be used in STN are fast-forward/non-fast-forward relaying protocols, and amplify-and-forward relaying protocol.

C. Bandwidth consumption

One main concern when deploying satellites for IoT is the satellite bandwidth consumption since the satellite bandwidth is pricey and often limited. To reduce satellite bandwidth consumption, caching capacity must be enabled at satellite in combination with content placement strategy [13]. During peak hours, more requested content cached at satellite will

Fig. 2: An example on the successful delivery probability of a file cached in single-tier and two-tier models.

lower uplink bandwidth consumption since less content is retrieved directly from gateway. Downlink bandwidth is saved when duplicated content will be broadcast once if it is cached at satellite. [13] formulates a joint caching problem for two-tier cache GEO-based SIoT system to optimize the satellite uplink and downlink bandwidth consumption simultaneously. The problem in [13] is formulated as a non-linear integer programming problem which can be solved effectively using genetic algorithm.

D. Resource allocation

Caching in SIoT can help improve the energy efficiency by reducing redundant transmissions and the spectrum efficiency by decreasing the network congestion. The dynamic resource allocation strategies for SIoT are to balance energy efficiency and spectral efficiency [15]. The resource allocation problem for cache-enabled SIoT system has been investigated in [7] for traffic offloading. The problem is formulated as a nonlinear fractional programming problem which jointly considers content placement and power allocation, that can be solved by adopting iteration and sub-problem decomposition. When adapting the resource allocation strategies in [15] with caching, the dynamic coordination of resource (power) allocation, caching scheme, and traffic distribution should be thoroughly designed. Allocating transmit power should also be reasonable between the content placement phase and the serving phase. In addition, the energy resource at satellites is limited and hard to renew since they are in space. This factor should be taken into account not only for effective energy allocation but also for prolonging the energy source of satellites. The spectrum efficiency reflects the network capacity over a given bandwidth. It is sometimes indicated the number of end-users the network can support. With caching, the network capacity is increased, hence, it can accommodate more users. The deployment of cache units at satellites in addition to multiple cache-enabled ground stations also helps enhance the spectrum efficiency. Taking all design criteria into consideration, a typical resource allocation problem for

caching design is maximizing the system SDP subject to various system constraints such as minimum satellite bandwidth consumption, total cache capacity, or maximum power constraints.

IV. PERFORMANCE EVALUATION

To illustrate the performance evaluation of research works reviewed in Section II, this section presents the numerical results for average access delay to compare the performance of the proposals in [5], [6] for single-tier cache-enabled LEO-based SIoT systems. Then the comparison on the network performance in terms of average SDP and satellite bandwidth consumption for single-tier versus two-tier cache-enabled GEO-based SIoT systems is conducted.

A. Single-tier cache-enabled LEO-based SIoT systems

A system is set up with 16 LEO satellites in the constellation with scale of 32. The satellites orbit at 780 km above earth surface with 45° of inclination. The observing time is 20 minutes with 20 time slots. A performance comparison in terms of average access delay (a.k.a. the average file distribution delay time) to distribute the content to the chosen cache nodes following algorithms in [5], [6] and the conventional central nodes selected for caching (NCL metric) is constructed in Fig. 3. Two caching policies are used for the content dis-

Fig. 3: Content distribution comparison for [5] and [6].

tributed to the chosen cache nodes: (i) most popular caching, where the most likely requested files are distributed to the cache satellites; and (ii) random caching, where the files with random popularity are distributed to the cache satellites. It can be observed that the file access delay shows no significant gap for the two aforementioned caching policies. The MTSC [5] gives less access delay than that of Dijkstra [6] and NCL algorithms when the file size is less than 700 KB. Increasing the file size also means increasing the transmission rate, the delay growth rate in NCL is higher than that of Dijkstra and MTSC when the file size increases for more than 700 KB. The delay growth rate of MTSC algorithm is mostly stable through different file sizes.

B. Single-tier versus two-tier cache-enabled GEO-based SIoT systems

A performance comparison between single-tier (adopted from the system models in [9], [11], [12]) versus two-tier (adopted from [13]) caching models is constructed in Figs. 4 and 5. A network includes a GEO satellite, a gateway hosting 1000 files of equal size of 100 Mb, and a ground station serving 20 end-users is constructed with two cache placement policies, namely the uniform caching (caching same portion of all the files) and the ground station most popular caching (caching most popular files at ground station). The request probability for a file hosted at gateway follows Zipf distribution. In two-tier cache system, under uniform caching policy, the same portion μ_1 and μ_2 of all files are prefetched at ground station and satellite, respectively; under ground station most popular caching policy, the most popular files are prefetched at ground station and the most popular of the remained files are cached at satellite. In single-tier cache system, under uniform caching policy, the same portion of all files are cached at ground station, and under ground station most popular caching policy, the most popular files are prefetched at ground station. Since the cache capacity and spectral resource are limited to satellite, we compare the performance of the two-tier and single-tier cache systems in terms of SDP under the system total caching capacity and satellite bandwidth consumption.

Fig. 4: Successful delivery probability of two-tier versus single-tier caching.

Fig. 4 shows the performance comparison in terms of SDP under total caching capacity. The total caching capacity in single-tier cache system refers to the cache capacity at ground station, whereas in two-tier cache system refers to the total cache capacity at the two tiers in which satellite's cache is fixed at 20 GB. It is observed that the two-tier model outperforms its single-tier counterpart, and applying the ground station most popular caching policy achieves higher SDP than that of the uniform caching policy. The two-tier cache system achieves approximately the same SDP under uniform and ground station most popular caching policies because this set up only has one ground station. With a practical system having multiple ground stations, the performance of the later caching policy will excel. When having popular files cached at multiple

ground stations, a cache hit even is more likely to happen at the ground station tier; Hence, the serving time is further shorten comparing to a single ground station system. Besides, to achieve higher SDP with multiple ground stations in the two-tier cache scheme, most popular content can be prefetched at satellite to be broadcast once to multiple ground stations.

Fig. 5: Satellite bandwidth consumption of two-tier versus single-tier caching under uniform caching policy.

Fig. 5 displays the performance comparison in terms of SDP and satellite bandwidth consumption under uniform caching policy. As expected, the single-tier caching model consumes more bandwidth and achieves lower SDP than the two-tier caching one. When less files are cached at the ground station, the non-cached files are served directly from the gateway, which leads to more bandwidth consumption in single-tier caching model. When more than fifty percent of the content are cached at ground station, the deviations in SDP and bandwidth between the two systems are very small. This behavior should be noted when designing cache deployment as single-tier cache system under appropriate caching policy can achieve almost the same SDP while consuming the same amount of satellite bandwidth as two-tier cache scheme.

V. FUTURE RESEARCH DIRECTIONS AND OPPORTUNITIES

Although there is tremendous potential combining edge caching with satellite IoT networks, research works in this specific area are limited with many open issues still in need of further investigation. With the need for global connectivity and the growth of massive device connection, caching in SIoT has been applied to reduce the load of terrestrial IoT networks. Research works in this category usually do not consider the IoT data lifetime in their designs. Therefore, the cached IoT data might not represent the actual physical status. Content placement in this case cannot be done with off-line caching algorithm but a periodically content caching update algorithm according to the lifetime of IoT data.

With the increasing number of LEO satellites launched recently, the networks might suffer from inter-satellite co-channel interference and interference with the existing GEO satellites when they share the same spectrum resource. The interference is fail to be considered by most caching policies

in current literature. Hence, the optimized policy to improve spectrum efficiency and energy efficiency might not perform well in practice as well as the achieved network capacity. To reflect the practical networks accurately, the investigation on satellites' interference and share spectrum is quite urgent.

Beside solving the fundamental problems in designing edge caching in SIoT, securing the transmissions has also raised growing concerns as the broadcast nature of satellites is susceptible to be wiretapped by unauthorized parties. Physical layer security (PLS) has been considered as an alternative solution over upper-layer encryption since it does not require any modification at upper-layer protocol. PLS only exploits features of wireless channels to safeguard the transmissions. There has been a wide range of research works on PLS in space information networks that makes this research area quite mature. Research on PLS in cache-enabled SIoT, however, is still in its infancy with very few studies. Securing transmissions in cache-enabled SIoT by PLS requires a protocol design for file access scheme. The cached content can be used to create interference to eavesdroppers when encoding the confidential information. The secrecy communication can also be ensured by allocating part of the transmit power to prefetched files in order to restrict the signal-to-interference-plus-noise ratio of the eavesdroppers. When designing PLS, the data rate is often sacrificed. Hence, the trade-off between having secure transmissions and network performances such as access delay and bandwidth consumption should be taken into consideration. In addition, eavesdroppers have evolved with unlimited resource (ie., hardware, power, signal processing capacity). In SIoT, eavesdroppers can also be other satellites within the transmission range of a gateway when data is purposely transmitted to a specific satellite. Those eavesdroppers are referred to as smart eavesdroppers with the ability to gain full channel state information and void current PLS solutions. Caching policies and file access protocols designed in this case need to address the enhanced version of eavesdroppers.

VI. CONCLUSIONS

In this article, a comprehensive review on edge caching in satellite-IoT with illustrative performance evaluation has been conducted. Firstly, the utilization of satellite system as a promising solution complimenting terrestrial networks is discussed. Then several satellite architectures deployed in SIoT and the edge caching techniques are introduced. The state-of-the-art for caching-enabled satellite-based terrestrial system is reviewed with the performance comparison of current research works. Finally, design criteria and future research directions and opportunities are addressed.

REFERENCES

- [1] T. 22.822, "Study on using satellite access in 5g," V16.0.0, Aug. 2018.
- [2] Sat5G, "Satellite and terrestrial network for 5g," Available at <https://www.sat5g-project.eu/wp-content/uploads/2019/04/Sat5g-project-document-2019-v31.pdf> (2021/08/15).
- [3] "Satellite in iot: New satellite networks for new wave of iot investment," *Allied Business Intelligence (ABI) Research*, pp. 1–20, 2019.
- [4] O. Kodheli, E. Lagunas, N. Maturo, S. K. Sharma, B. Shankar, J. F. M. Montoya, J. C. M. Duncan, D. Spano, S. Chatzinotas, S. Kisseleff, J. Querol, L. Lei, T. X. Vu, and G. Goussetis, "Satellite communications in the new space era: A survey and future challenges," *IEEE Communications Surveys Tutorials*, vol. 23, no. 1, pp. 70–109, Feb. 2021.

- [5] Z. Yang, Y. Li, P. Yuan, and Q. Zhang, "Tcsc: A novel file distribution strategy in integrated leo satellite-terrestrial networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5426–5441, May 2020.
- [6] Y. Li, Y. Wang, P. Yuan, Q. Zhang, and Z. Yang, "Popularity-aware back-tracing partition cooperative cache distribution for space-terrestrial integrated networks," *IET Communications*, vol. 13, no. 17, pp. 2786–2796, Oct. 2019.
- [7] J. Li, K. Xue, D. S. L. Wei, J. Liu, and Y. Zhang, "Energy efficiency and traffic offloading optimization in integrated satellite terrestrial radio access networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2367–2381, Apr. 2020.
- [8] L. Liu, J. Zhang, X. Zhang, P. Wang, Y. Wang, and L. Ouyang, "Design and analysis of cooperative multicast-unicast transmission scheme in hybrid satellite-terrestrial networks," in *2018 IEEE International Conference on Communication Systems (ICCS)*, Wuxi, China, 2018, pp. 309–314.
- [9] T. Vu, Y. Poirier, S. Chatzinotas, and N. Maturo, "Modeling and implementation of 5g edge caching over satellite," *International Journal of Satellite Communications and Networking*, vol. 38, no. 5, pp. 395–406, Mar. 2020.
- [10] N. Garg, M. Sellathurai, and T. Ratnarajah, "In-network caching for hybrid satellite-terrestrial networks using deep reinforcement learning," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 2020, pp. 8797–8801.
- [11] K. An, Y. Li, X. Yan, and T. Liang, "On the performance of cache-enabled hybrid satellite-terrestrial relay networks," *IEEE Wireless Communications Letters*, vol. 8, no. 5, pp. 1506–1509, Jun. 2019.
- [12] A. Kalantari, M. Fittipaldi, S. Chatzinotas, T. X. Vu, and B. Ottersten, "Cache-assisted hybrid satellite-terrestrial backhauling for 5g cellular networks," in *2017 IEEE Global Communications Conference (GLOBECOM)*, Singapore, 2017, pp. 1–6.
- [13] H. Wu, J. Li, H. Lu, and P. Hong, "A two-layer caching model for content delivery services in satellite-terrestrial networks," in *2016 IEEE Global Communications Conference (GLOBECOM)*, DC, USA, 2016, pp. 1–6.
- [14] Q. Ngo, K. Phan, W. Xiang, A. Mahmood, and J. Slay, "Two-tier cache-aided full-duplex hybrid satellite-terrestrial communication networks," *TechRxiv. Preprint.*, 2021, <https://doi.org/10.36227/techrxiv.15172038.v1>.
- [15] Y. Peng, T. Dong, R. Gu, Q. Guo, J. Yin, Z. Liu, T. Zhang, and Y. Ji, "A review of dynamic resource allocation in integrated satellite and terrestrial networks," in *2018 International Conference on Networking and Network Applications (NaNA)*, Xi'an, China, 2018, pp. 127–132.