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 / 01-02-2022

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.v2>

DOI

[10.36227/techrxiv.16842979.v2](https://doi.org/10.36227/techrxiv.16842979.v2)

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.

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 3rd 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 planned to build large satellite constellations to deliver high-speed broadband Internet globally in 2021. According to Allied Business Intelligence research, it is through satellite communications that 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 offer latency-insensitive 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 20 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 non-GEO. In a typical satellite-terrestrial network (STN), GEO satellites usually belong to the space-based back-

bone 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.

An example of the satellite-IoT (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 to SIoT will not only reduce the service latency but also avoid retransmission 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 off-peak hours and the delivery of files to end devices on demand. Designing edge caching usually involves studying:

- Cache deployment on the location selection for cache units
- Content placement on the file selection to prefetch at each cache unit
- 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 onboard capacities of the satellites in terms of both low processing units and limited power supply. Regardless, recent advances in hardware components, power generation, and energy harvesting technologies have enhanced onboard processing units to enable innovative communication technologies [4].

The remainder of this article is organized as follows. We describe the related work on edge caching in SIoT. We discuss the design criteria of edge caching in SIoT. A performance

Quynh T. Ngo, Khoa T. Phan, Wei Xiang, and Abdun Mahmood are with La Trobe University, Australia.

Jill Slay is with the University of South Australia, Australia.

Digital Object Identifier: 10.1109/IOTM.001.2100062

FIGURE 1. An example of the satellite-IoT network architecture.

evaluation of different cache-enabled SIoT systems is presented. Directions and opportunities for future research are considered. Finally, conclusions are made.

THE STATE OF THE ART OF EDGE CACHING IN SATELLITE TERRESTRIAL NETWORKS

Although three types of satellites can be employed in STNs, as introduced later, research work on applying edge caching in MEO-based STNs 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 STNs. A classification of research works on cache-enabled STNs is summarized in Table 1, in which research works are classified based on satellite types deployed (LEO or GEO), network architecture (single-tier vs. two-tier), cache deployment strategies (where to cache), performance evaluation criteria, content exploration (content query and placement problems), and content delivery.

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 an in-network caching mechanism, [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 of assembling 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 decreasing 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.

CACHE-ENABLED STN FOR CELLULAR TRAFFIC OFFLOADING

Edge caching and satellite deployment can help offload cellular traffic and enhance 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 networking with a cooperative caching scheme. The proposed scheme allows terrestrial access points and satellite to share the cached content and 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-enabled system where caches are placed in both base stations and LEO satellite. To offload cellular traffic, [7] proposes a cooperative multicast-unicast transmission mechanism for satellite, where popular requests are served in the multicast phase and unpopular ones in the 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 back-up 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 disasters. Proposing offline caching approaches to offload terrestrial backhaul networks, [9] designs a cache-aided GEO STN focusing on the multimode satellite backhaul in two use cases: dense urban areas and sparsely populated regions. Reference [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-complexity approach and achieves better performance over the most popular caching policy.

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, and delivery delay. Caching 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 in 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 the traditional terrestrial network with a 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

Related work	Satellite		Network architecture		Cache deployment	Traffic off-loading	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 bandwidth	✓	✓

TABLE 1. Summary of research works on edge caching in STN.

caching. To offload the backhaul of terrestrial networks, [12] proposes using GEO STN in combination with an offline 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 locally popular files, while the other part is filled with the most globally popular files. The proposed offline 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 the case of two-tier cache systems [8, 13], globally popular content can be cached at satellites to leverage satellites' broadcast nature. This placement strategy leads to improvement of satellite spectral efficiency and content retrieval delay.

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 from which the content should be transmitted, the power consumption, and the satellite bandwidth consumption for transmitting the content. The performance metrics (i.e., throughput, power consumption, bandwidth consumption, and delay) are usually considered when designing the content delivery strategies. Combining terrestrial cooperation and satellite broadcast functionality to relay requested content from core networks, [9–11] design 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:

- Reduce service latency since more content is stored closer to the end users.
- 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, due to the fact that satellites are in contact with ground stations

for only a period of time, the dynamic network topology also creates a great challenge in designing caching policy and transmission scheme compared to conventional terrestrial networks.

DESIGN CRITERIA OF EDGE CACHING IN SLoT

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

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 served by the cached devices, a higher cache hit ratio is 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 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. Another 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 patterns can be monitored and applied in cache placement design. Deep learning can also be applied in this process as in [10]. Since the cache placement design phase is offline and is usually done at the gateways/servers with high computational capacity, machine learning is a potential means to achieve a precise requested content pattern in content placement design.

SERVICE DELAY

Service delay is related to 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 SLoT from a single tier 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 a two-tier cache-enabled GEO-based SLoT system was 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 denoting the percentage of the file cached at satellite. It is observed that using the two-tier caching model can help achieve better

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

FIGURE 3. Content distribution comparison for [5, 6].

SDP. With more than 70 percent of the file cached, service delivery for both caching models is guaranteed. Relaying protocol also plays an important role in leveraging the network SDP. The delivery time is shorter for non-cached content with a suitable relaying protocol. Some common protocols that can be used in STN are fast-forward/non-fast-forward relaying protocols, and amplify-and-forward relaying protocol.

BANDWIDTH CONSUMPTION

One main concern when deploying satellites for IoT is 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 lower uplink bandwidth consumption since less content is retrieved directly from the gateway. Downlink bandwidth is saved when duplicated content is broadcast once if it is cached at satellite. Reference [13] formulates a joint caching problem for a 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 nonlinear integer programming problem that can be solved effectively using genetic algorithm.

RESOURCE ALLOCATION

Caching in SIoT can help improve energy efficiency by reducing

redundant transmissions and spectrum efficiency by decreasing network congestion. The dynamic resource allocation strategies for SIoT are designed to balance energy efficiency and spectral efficiency [15]. The resource allocation problem for a cache-enabled SIoT system was investigated in [7] for traffic offloading. The problem is formulated as a nonlinear fractional programming problem that jointly considers content placement and power allocation, which 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. The number of end users the network can support is sometimes indicated. 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 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, and maximum power constraints.

PERFORMANCE EVALUATION

To illustrate the performance evaluation of research works reviewed earlier, 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 a comparison of the network performance in terms of average SDP and satellite bandwidth consumption for single-tier vs. two-tier cache-enabled GEO-based SIoT systems is conducted.

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 the Earth's surface with 45° 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 distributed 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. 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 to more than 700 kB. The delay growth rate of the MTSC algorithm is mostly stable through different file sizes.

SINGLE-TIER VS. TWO-TIER CACHE-ENABLED GEO-BASED SIoT SYSTEMS

A performance comparison between single-tier (adopted from the system models in [9, 11, 12]) vs. 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

FIGURE 4. Successful delivery probability of two-tier vs. single-tier caching.

FIGURE 5. Satellite bandwidth consumption of two-tier vs. single-tier caching under uniform caching policy.

constructed with two cache placement policies, namely uniform caching (caching the same portion of all the files) and ground station most popular caching (caching the most popular files at the ground station). The request probability for a file hosted at the gateway follows Zipf distribution. In a two-tier cache system, under uniform caching policy, the same portion μ_1 and μ_2 of all files are prefetched at the ground station and satellite, respectively; under the ground station most popular caching policy, the most popular files are prefetched at the ground station, and the most popular of the remaining files are cached at a satellite. In a single-tier cache system, under uniform caching policy, the same portion of all files are cached at the ground station, and under the ground station most popular caching policy, the most popular files are prefetched at the 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.

Figure 4 shows the performance comparison in terms of SDP under total caching capacity. The total caching capacity in a single-tier cache system refers to the cache capacity at the ground station, whereas in two-tier cache system, it refers to the total cache capacity at the two tiers in which the 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 the uniform and ground station most popular caching policies because this setup only has one ground station. In a practical system with multiple ground stations, the performance of the latter caching policy will excel. With 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 shortened compared to a single ground station system. Furthermore, to achieve higher SDP with multiple ground stations in the two-tier cache scheme, most popular content can be prefetched at the satellite to be broadcast once to multiple ground stations.

Figure 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 fewer files are cached at the ground station, the non-cached files are served directly from the gateway, which leads to more bandwidth consumption in the single-tier caching model. When more than 50 percent of the content are cached at the 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 a single-tier cache system under appropriate caching policy can achieve almost the same SDP while consuming the same amount of satellite bandwidth as a two-tier cache scheme.

FUTURE RESEARCH DIRECTIONS AND OPPORTUNITIES

Although there is tremendous potential combining edge caching with SLoT 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 SLoT 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 an offline caching algorithm but with a periodically content caching updating algorithm according to the lifetime of IoT data.

With the increasing number of LEO satellites launched recently, 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 has failed to be considered by most caching policies in the 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 shared spectrum is quite urgent.

Besides solving the fundamental problems in designing edge caching in SLoT, securing the transmissions has also raised growing concerns as the broadcast nature of satellites is susceptible to wiretapping 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 of the upper-layer protocol. PLS only exploits features of wireless channels to safeguard transmissions. There have 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 SLoT, however, is still in its infancy with very few studies. Securing transmissions in cache-enabled SLoT by PLS requires a protocol design for a file access scheme. The cached content can be used to create interference to eavesdroppers when encoding confidential information. The secrecy of 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 eavesdroppers.

When designing PLS, the data rate is often sacrificed. Hence, the trade-off between having secure transmissions and network performance such as access delay and bandwidth consumption should be taken into consideration. In addition, eavesdroppers have evolved with unlimited resources (i.e., hardware, power, signal processing capacity). In SLoT, 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. In this case, caching policies and file access protocol designs need to address the enhanced version of eavesdroppers.

CONCLUSIONS

In this article, a comprehensive review on edge caching in satellite-IoT with illustrative performance evaluation has been conducted. First, the utilization of a satellite system as a promising solution complementing terrestrial networks is discussed. Then several satellite architectures deployed in SLoT and the edge caching techniques are introduced. The state-of-the-art 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.

ACKNOWLEDGMENTS

This work was supported by the La Trobe University Postgraduate Research Scholarship, the La Trobe University Full-Fee Research Scholarship, and the Net Zero Scholarship sponsored by Sonepar/SLS.

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"; <https://www.sat5g-project.eu/wp-content/uploads/2019/04/Sat5gproject-document-2019-v31.pdf>, accessed Aug. 15, 2021.
- [3] "Satellite in IoY: New Satellite Networks for New Wave of IoT Investment," *Allied Business Intelligence Research*, 2019, pp. 1–20.
- [4] O. Kodheli et al., "Satellite Communications in the New Space Era: A Survey and Future Challenges," *IEEE Commun. Surveys & Tutorials*, vol. 23, no. 1, Feb. 2021, pp. 70–109.
- [5] Z. Yang et al., "TCSC: A Novel File Distribution Strategy in Integrated Leo Satellite-Terrestrial Networks," *IEEE Trans. Vehic. Tech.*, vol. 69, no. 5, May 2020, pp. 5426–41.
- [6] Y. Li et al., "Popularity-Aware Back-Tracing Partition Cooperative Cache Distribution for Space-Terrestrial Integrated Networks," *IET Commun.*, vol. 13, no. 17, Oct. 2019, pp. 2786–96.
- [7] J. Li et al., "Energy Efficiency and Traffic Offloading Optimization in Integrated Satellite Terrestrial Radio Access Networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, Apr. 2020, pp. 2367–81.
- [8] L. Liu et al., "Design and Analysis of Cooperative Multicast-Unicast Transmission Scheme in Hybrid Satellite-Terrestrial Networks," *2018 IEEE Int'l. Conf. Commun. Systems, Wuxi, China*, 2018, pp. 309–14.
- [9] T. Vu et al., "Modeling and Implementation of 5G Edge Caching over Satellite," *Int'l. J. Satellite Commun. and Networking*, vol. 38, no. 5, Mar. 2020, pp. 395–406.
- [10] N. Garg, M. Sellathurai, and T. Ratnarajah, "In-Network Caching for Hybrid Satellite-Terrestrial Networks Using Deep Reinforcement Learning," *ICASSP 2020, Barcelona, Spain*, 2020, pp. 8797–8801.
- [11] K. An et al., "On the Performance of Cache-Enabled Hybrid Satellite-Terrestrial Relay Networks," *IEEE Wireless Commun. Letters*, vol. 8, no. 5, June 2019, pp. 1506–09.
- [12] A. Kalantari et al., "Cache-Assisted Hybrid Satellite-Terrestrial Backhauling for 5G Cellular Networks," *2017 IEEE GLOBECOM, Singapore*, 2017.
- [13] H. Wu et al., "A Two-Layer Caching Model for Content Delivery Services in Satellite-Terrestrial Networks," *2016 IEEE GLOBECOM, Washington, DC*, 2016.
- [14] Q. Ngo et al., "Two-Tier Cache-Aided Full-Duplex Hybrid Satellite-Terrestrial Communication Networks," *IEEE Trans. Aerospace and Electronic Systems*. DOI: 10.1109/TAES.2021.3121657.
- [15] Y. Peng et al., "A Review of Dynamic Resource Allocation in Integrated Satellite and Terrestrial Networks," *2018 Int'l. Conf. Networking and Network Applications, Xi'an, China*, 2018, pp. 127–32.

BIOGRAPHIES

QUYNH T. NGO (T.Ngo@latrobe.edu.au) is currently a Ph.D. candidate in the Department of Computer Science and Information Technology, La Trobe University, Melbourne, Australia. She received her B.Sc. in electrical engineering (Magna

Cum Laude) from California State University in 2013, and her M.Sc. in telecommunications from the Vietnam National University — University of Sciences, HCM, in 2016. Her research interests are medium access control for vehicular ad hoc networks, edge caching, and physical layer security for satellite-IoT. She has been awarded the Jet Propulsion Laboratory Undergraduate Scholarship (JPLUS), the visiting research award funded by the Norwegian - Vietnamese Industrial & Infrastructure Safety Consortium, the La Trobe University Post Graduate Research Scholarship, and the Net Zero Scholarship.

KHOA T. PHAN [M] (K.Phan@latrobe.edu.au) received his B.Eng. degree in telecommunications (First Class Honors) from the University of New South Wales, Sydney, Australia, in 2006, M.Sc. degrees in electrical engineering from the University of Alberta, Edmonton, Canada, in 2008 and California Institute of Technology, Pasadena, in 2009, respectively, and his Ph.D. degree in electrical engineering from McGill University, Montreal, Quebec, Canada, in 2017. He is currently an Australia Research Council Discovery Early Career Researcher Award Fellow. His current research interests are broadly design, control, optimization, and operation of 5G mobile communications networks with applications in the Internet of Things, satellite communications, machine-type communications, smart grids, and cloud computing.

WEI XIANG [S'00, M'04, SM'10] (W.Xiang@latrobe.edu.au) is Cisco Research Chair of AI and IoT and director of the Cisco-La Trobe Centre for AI and IoT at La Trobe University. Previously, he was Foundation Chair and Head of Discipline of IoT Engineering at James Cook University, Cairns, Australia. He is an elected Fellow of the IET in the United Kingdom and Engineers Australia. He received the TNQ Innovation Award in 2016, the Pearcey Entrepreneurship Award in 2017, and Engineers Australia Cairns Engineer of the Year in 2017. He has published over 250 peer-reviewed papers including 3 books and 200 journal articles. He has served in a large number of international conferences in the capacity of General Co-Chair, TPC Co-Chair, Symposium Chair, and others. His research interests include the Internet of Things, wireless communications, machine learning for IoT data analytics, and computer vision.

ABDUN MAHMOOD [SM] (A.Mahmood@latrobe.edu.au) received his Ph.D. from the University of Melbourne, Australia, in 2008, his M.Sc. (Research) degree in computer science, and his B.Sc. degree in applied physics and electronics from the University of Dhaka, Bangladesh, in 1999 and 1997, respectively. He has had an academic career since 2000, working at the University of Dhaka, RMIT University, UNSW Canberra, and currently at La Trobe University as an associate professor (Reader). He leads a group of researchers focusing on machine learning and cybersecurity, including anomaly detection in smart grid, SCADA security, memory forensics, and false data injection. He has had success in attracting over \$1 million in grant funding as a CI, including two ARC Linkage Projects.

JILL SLAY [M] (Jill.Slay@unisa.edu.au) was Optus Chair of Cyber Security with La Trobe University. Previously, she was director of the Australian Centre for Cyber Security, University of New South Wales Canberra at ADFA. She was made a member of the Order of Australia (AM) for service to the information technology industry through contributions in the areas of forensic computer science, security, protection of infrastructure, and cyber-terrorism. She is a Fellow of ACS and the International Information Systems Security Certification Consortium. She has published one book and more than 120 refereed book chapters, journal articles, and research papers in information assurance, critical infrastructure protection, security, and forensic computing in the last 10 years. She has been awarded over AU\$2 million in Australian Government Category 1 research income, including a Future Fellowship.