An Optimal Allocation Framework of Security Virtual Network Functions in 6G Satellite Deployments

Antonio Petrosino, Giuseppe Piro, Luigi Alfredo Grieco, and Gennaro Boggia,
Department of Electrical and Information Engineering (DEI), Politecnico di Bari, Bari, Italy
CNIT, Consorzio Nazionale Interuniversitario per le Telecomunicazioni
Email: antonio.petrosino@poliba.it, giuseppe.piro@poliba.it, alfredo.grieco@poliba.it, gennaro.boggia@poliba.it

***Abstract*—In the emerging 6G satellite deployments, the interaction between Non-Terrestrial Network terminals and satellite constellations will generate a large surface of attack, which requires the design of novel security architectures. The scientific literature suggests to implement security services as Virtual Network Functions, installed onboard the satellites. The dynamic orchestration of these services, however, still represents a challenging and open research issue. To bridge this gap, this paper presents a novel approach willing to allocate security Virtual Network Functions across satellites, in a dynamic and optimal way. To this end, an optimization problem is formulated by deeply considering the intermittent connectivity between terminals on the Earth and the satellite constellation, the limited computational capabilities of satellites, and the need to provide secure Virtual Network Functions before a given time deadline. Then, the Tabu Search algorithm is used to solve the optimization problem and achieve preliminary results in realistic scenarios. The study illustrates the feasibility of the proposed approach and highlights the issues to address in future research activities.**

***Index Terms*—6G satellite, allocation of security service, optimization problem**

I. Introduction

According to recent standardization activities on 6G, the design and the deployment of Non-Terrestrial Networks (NTNs) represents a fundamental way to foster connectivity wherever the traditional terrestrial infrastructure is not affordable to build [1]. A constellation of satellites can be configured to provide connectivity to NTN terminals, while extending the boundaries of the edge network to the sky [2]. As a result, future 6G satellite infrastructures promise to offer new business opportunities for the massive diffusion of new services (e.g., smart monitoring, smart metering, and smart agriculture) [3].

Behind these benefits, however, there is a critical security concern: the overall 6G satellite infrastructure represents a very large network attack surface, which requires the design and the development of novel schemes and methodologies of protection [4], [5]. The scientific literature already identified some important security services to be deployed onboard the satellite: Intrusion Detection and Prevention as a Service (IDPaas), Authentication as a Service (AaaS), Secure Transmission Channel as a Service (STCaaS). They can be conceived as Virtualized Network Functions (VNFs) and configured through Software-Defined Network (SDN) facilities [6]. Nevertheless, the limited computational capabilities of satellites prevent the offline and fixed deployment of all the required security VNFs (requested by several clusters of NTNs terminals), across the satellite constellation. On the contrary, VNFs must be dynamically deployed throughout the constellation, while carefully taking care of satellite capabilities, visibility time and periodicity offered by the constellation, and the Quality of Service (QoS) requirements associated with the considered security service.

The dynamic deployment of services and applications at the edge of the terrestrial networks is a research topic widely investigated in the current state-of-the-art. Valuable contributions, for example, propose optimal approaches able to minimize latency and energy consumption [7]–[10], or to maximize the user throughput [11]. Unfortunately, these solutions (and many others not mentioned herein because of lack of space) cannot be applied to the considered 6G satellite infrastructure. In fact, they do not consider the movement of satellites and the intermittent connectivity among NTN terminals, satellites, and core network. Also, a recent survey confirms that the optimal provisioning of security services in 6G satellite deployment still represents an unexplored research topic [5].

To bridge this gap, this work presents a novel approach willing to dynamically allocate security VNFs across satellites, so that the requested security service can be provided to clusters of NTN terminals before a given time deadline. To this end, a system model describing the overall 6G satellite infrastructure and the service provisioning delay is firstly derived in Section II. Then, an optimization problem willing to minimize the sum of experienced service provisioning delays, under system and QoS constraints, is formulated in Section III. The optimization problem is solved through the Tabu Search algorithm. Preliminary results, discussed in IV, demonstrate the capability of the proposed approach to successfully allocate the security services before a given time deadline and highlights some complexity issues arose in larger deployments. Finally, Section V concludes the work and draws future research activities.

II. Reference Architecture and System Model

This work assumes to exploit a 6G satellite architecture to serve NTN terminals spread on the Earth, that sporadically

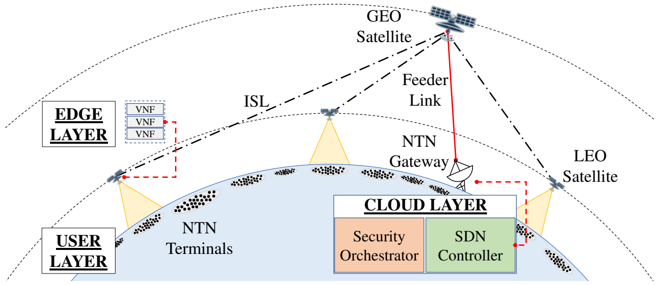


Fig. 1. The proposed architecture.

generate small data packets. Indeed, the reference scenario well fits the description of a large set of use cases, such as smart monitoring, smart metering, and smart agriculture. The 6G satellite architecture is configured to receive and process data generated by NTN terminals, while offering security services, such as firewall, intrusion detection, and prevention system.

Fig. 1 depicts the three-layer infrastructure considered in this work, embracing user, edge, and cloud layers.

At the user layer, groups of NTN terminals deployed within a delimited geographical area form different clusters. Without loss of generality, it is assumed that terminals belonging to the same cluster request the same set of security services.

The edge layer hosts a constellation of Low Earth Orbit (LEO) and Geostationary Orbit (GEO) satellites and the NTN gateways. It is important to remark that LEO CubeSat is one of the most adopted cheap satellite platform. Therefore, the deployment cost of a CubeSat constellation is lower than the other type of satellites. Additionally, it is the best candidate to foster connectivity in remote areas on the Earth, thanks to its low propagation loss and low propagation time, exploiting radio access technologies such as Narrow-Band IoT (NB-IoT) and Long Range Wide Area Network (LoRaWAN) [12], [13]. At the same time, however, LEO satellites cannot be always connected with NTN gateways [12]. To cope with this issue, GEO satellites are used to guarantee persistent feeder links: LEO satellites communicate with NTN gateways, and in turn with the core network, via GEO satellites. Security services are installed onboard LEO satellites as VNFs.

Finally, Security Orchestrator and SDN controller are parts of the cloud layer. The Security Orchestrator defines, during the time, which VNFs should be deployed throughout LEO satellites. This objective is achieved by solving the optimization problem discussed hereafter. Then, VNFs will be installed and/or configured on satellites by using SDN facilities. Specifically, the SDN controller, contacted by the Security Orchestrator, will forward instructions and related data to satellites by using the persistent feeder link. OpenFlow and RESTCONF are possible control plane technologies to be used for this purpose.

Definitively, the interaction between these network elements is summarized as in what follows:

- NTN terminals communicate to a visible satellite the activation of a service, requesting the deployment of a given security VNFs;
- the request is delivered to the Security Orchestrator through the feeder link;
- the Security Orchestrator collects all the requests and decides to deploy security VNFs in the satellite constellation, over a specific time horizon (dynamically computed, as discussed later);
- the SDN controller configures the constellation according to the decision taken by the Security Orchestrator;
- NTN terminals are served by security VNFs, installed in specific LEO satellite.

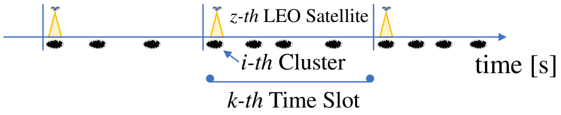


Fig. 2. Intermittent connectivity between terminals on the Earth and the satellite constellation.

Since the orbits of the satellite constellation work autonomously, the system model considered herein directly focuses on a set of satellites belonging to a specific orbit and the set of clusters served by that satellites.

Let $\mathbf{\Lambda} = \{\lambda_1, ..., \lambda_L\}$ and $\mathbf{\Sigma} = \{\sigma_1, ..., \sigma_S\}$ be the list of clusters served by LEO satellites of the considered orbit and the list of satellites of that orbit, respectively. The total number of available clusters is $L = \|\mathbf{\Lambda}\|$. The number of LEO satellites is equal to $S = \|\mathbf{\Sigma}\|$. The computational capability of the z-th LEO satellite is denoted with $c(\sigma_z)$.

T_o is the time needed by a satellite to complete one revolution around the Earth. It means that each satellite can periodically serve each cluster every T_o. Moreover, T_p denotes the elapsed time between two subsequent satellite visibility for a given cluster on the ground. It holds that $T_p = T_{orbit}/S$.

The proposed system model assumes to divide the time into slots, lasting T_p. During a slot, a cluster can communicate with only one LEO satellite, but just for a short visibility time (as depicted in Fig. 2). In this context, $\mathbf{V}(k)$ represents the visibility matrix for the k-th time slot, which reports the reciprocal visibility between the i-th cluster and the z-th LEO satellite. In other words, $v_{i,z}(k) = 1$, with $v_{i,z}(k) \in \mathbf{V}(k)$, if the z-th LEO satellite can communicate with the i-th cluster in the k-th time slot. Otherwise, $v_{i,z}(k) = 0$.

Let $\mathbf{B}(k)$ be the services allocation matrix. Also in this case, $b_{i,z}(k) \in \mathbf{B}(k)$ is a boolean flag that denotes if the z-th satellite hosts the security VNF, serving the i-th cluster. The difference between $\mathbf{V}(k)$ e $\mathbf{B}(k)$ is that the former depends on the position of both clusters and satellites into the 6G architecture, while the latter is the outcome of the optimization problem. Each request delivered to the Security Orchestrator contains the following information: the cluster that generated the request, λ_i; the time slot in which the request has been generated, $t(\lambda_i)$; the acceptable upper bound delay for the provisioning of the requested security VNFs, $\tau(\lambda_i)$; the computational requirement associated with the request, $\xi(\lambda_i)$.

III. Optimization problem

The allocation of security VNFs is performed every time the Security Orchestrator receives a new request. Let $\mathbf{R}(k)$ the set of clusters with pending requests at the k-th time slot. Given $\mathbf{R}(k)$, the allocation considers a time horizon equal to $T = \max_{\mathbf{R}(k)} \tau(\lambda_i)$. As anticipated in the Introduction, the goal of the optimization problem formulated herein is to allocate security VNFs across satellites of the considered orbit, while ensuring that (i) the system constraint of the satellite is never violated, (ii) the cluster is served by the requested security VNF satisfying the QoS constraints, and (iii) sum of experienced service provisioning delays is minimized during the entire time horizon T. Assuming that the cluster λ_i has not been served by the requested security VNF at the k-th time slot, the accumulated service provisioning delay is equal to $k - t(\lambda_i)$ time slots. This delay will increase slot by slot, until a given satellite will host the requested security VNF in one of the consecutive time slots. Therefore, considering the allocation horizon T, the service provisioning delay experienced by the cluster λ_i, that is $\delta(\lambda_i, k)$, can be formally described as:

$$\delta(\lambda_i, k) = \sum_{\nu\,=\,1}^{T} \sum_{\sigma_z \in \Sigma} \Big(b_{i,z}(\nu)\,[k + \nu - t(\lambda_i)] \Big). \tag{1}$$

The (1) assumes that $\sum_{\nu\,=\,1}^{T} \sum_{\sigma_z \in \mathbf{\Sigma}} b_{i,z}(\nu) = 1, \forall \lambda_i \in \mathbf{R}(k)$. In this case, in fact, it is possible to ensure that the cluster will be served by a single satellite during T.

Based on these premises, the objective function to be minimized is formally defined as:

$$\begin{aligned} U(k) &= \sum_{\lambda_i \in\, \mathbf{R}(k)} \delta(\lambda_i, k) = \\ &= \sum_{\lambda_i \in\, \mathbf{R}(k)} \left[\sum_{\nu\,=\,1}^{T} \sum_{\sigma_z \in \mathbf{\Sigma}} \Big(b_{i,z}(\nu)\,[k + \nu - t(\lambda_i)] \Big) \right] \end{aligned} \tag{2}$$

On the other hand, the optimization problem is formalized in (3a). Note that the (3b) takes into account the limited computational capability of LEO satellites. The (3c) considers the deadline constraint to force the system to allocate the requested security VNF before a time deadline. Finally, the (3d) denotes the visibility constraint that checks if the cluster to serve and the LEO satellite equipped with the security services are in reciprocal visibility.

To conclude, it is important to remark that the (3a) represents an Integer Linear Programming (ILP).

$$\min_{\mathbf{R}\,\mathbf{\Sigma}} \sum_{\lambda_i \in\, \mathbf{R}(k)} \left[\sum_{\nu\,=\,1}^{T} \sum_{\sigma_z \in \mathbf{\Sigma}} \Big(b_{i,z}(\nu)\,[k + \nu - t(\lambda_i)] \Big) \right] \tag{3a}$$

$$\text{s.t.} \sum_{\lambda_i \in\, \mathbf{R}(k)} b_{i,z}(\nu)\,\xi(\lambda_i) \le c(\sigma_z), \forall\, z, \nu \tag{3b}$$

$$\sum_{\nu\,=\,t(\lambda_i)}^{T} b_{i,z}(\nu)\, v_{i,z}(\nu) = 1, \forall i, z \tag{3c}$$

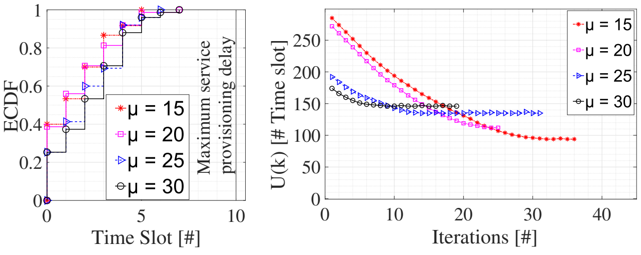
$$b_{i,z}(\nu) \le v_{i,z}(\nu), \forall i, z, \nu \tag{3d}$$

IV. Performance Evaluation

The optimization problem formulated in Section III is solved by using a well-known meta-heuristic solution method, namely Tabu Search [14]. The system model, the optimization problem, and the aforementioned solution method have been implemented in Python. The conducted study considers a scenario where the computational capability of each satellite is set to $c(\sigma_z) = 3, \forall z$, and the computational requirement for each service request is set to $\xi(\lambda_i) = 1$, $\forall i$. It also assumes that security VNFs must be provided within 10 time slots from the generation of the request. On the other hand, the number of the clusters served by the constellation is set to $L = 60$.

First of all, the impact of the number of pending requests on system performance is investigated. To this end, it is assumed that each cluster may have only 1 pending request. Moreover, service requests are generated in order to ensure an average number of pending requests, that is μ, ranging from 15 to 30. Instead, the number of satellites in the orbit is set to $S = 3$. Fig. 3a shows the Empirical Cumulative Distribution Function (ECDF) of the service provisioning delays experienced by all the clusters. As expected, the delays increase with the network load. A high number of requests in the same time slot, in fact, overloads the constellation. Consequently, the clusters may wait for more visibility times before being served by a satellite hosting the requested security VNF on board. In any case, however, the targeted upper bound (equal to 10 time slots) is always satisfied. As well known, the Tabu Search is able to find the optimal solution after a number of iterations, which strictly depends from the initial solution randomly chosen by the algorithm. Indeed, there is not a specific relation between the amount of iterations needed to solve the optimization problem and the traffic load. Results reported in Fig. 3b just show that Tabu Search employs from 19 to 36 iterations to minimize the objective function reported in (3a).

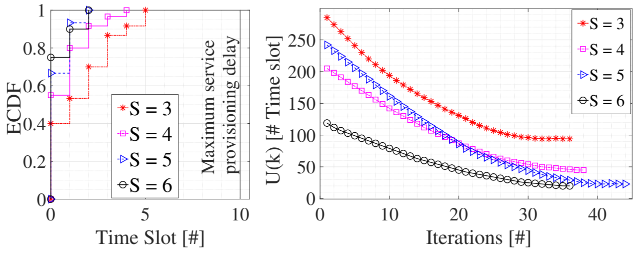
The second study discussed herein investigates the impact of the number of satellites on system performance, while setting the average number of pending request to $\mu = 15$. Fig. 4a confirms (once again) the ability of the optimization problem to satisfy the expected upper bound delay. At the same time, Fig. 4b highlights that the number of iterations needed to find the optimal solution generally increases with the number of satellites in the orbit. This is due to the higher size of matrices managed by the algorithm. The only exception is registered when $S = 6$. In that case, however, even if the number of iterations needed to solve the problem is lower, their duration is higher (as discussed below). Indeed, the higher number of satellites per orbit brings to an increment of the problem complexity. The complexity of the proposed optimization problem is measured on the host machine with 4-core 3.5 GHz CPU and 16 GB of RAM. The amount of time required by the Tabu Search algorithm to find the optimal



(a) Service provisioning delay.

(b) Objective function $U(k)$.

Fig. 3. Impact of μ on network performance.



(a) Service provisioning delay.

(b) Objective function $U(k)$.

Fig. 4. Impact of S on network performance.

solution is reported in Table I, for all the scenarios discussed before. It is possible to observe that a simple machine, like the one adopted in the conducted study, is able to solve the optimization problem before the end of the current time slot only for low number of satellites per orbit. On the contrary, in the case the number of satellites per orbit is higher than 4, the optimal solution is obtained after a higher amount of time. Definitively, it is possible to conclude that the computational complexity required to solve optimization problem increases with the number of satellites per orbit. At the same time, however, it is very important to remark that the feasibility of the proposed approach can be still reached by using machines with higher computing capabilities, as well as by exploiting other meta-heuristic solution methods (as indicated in our future works).

V. Conclusion

This paper presented an optimal allocation framework of security Virtual Network Functions in 6G satellite deployments. Specifically, the conceived optimization problem dynamically allocates the security services throughout the satellite constellation, with the aim of minimizing the sum of service provisioning delays and ensuring system and QoS requirements. Preliminary results demonstrated the ability of the proposed approach to reach optimal solutions, while highlighting the complexity issues that arose in scenarios with a high number of satellites. Further research activities will investigate different meta-heuristic methods able to reduce the problem complexity, even in larger deployments. Different application domains and the impact on energy consumption will be investigated as well.

Impact of μ with S = 3			**Impact of S with μ = 15**		
μ	T_p [min]	Solving Time	S	T_p [min]	Solving Time
15	31.53	0.25 T_p	3	31.53	0.26 T_p
20	31.53	0.17 T_p	4	23.65	0.65 T_p
25	31.53	0.21 T_p	5	18.91	1.40 T_p
30	31.53	0.11 T_p	6	15.76	1.89 T_p

TABLE I
Solving time.

Acknowledgment

This work was supported by the PRIN project no. 2017NS9FEY entitled "Realtime Control of 5G Wireless Networks: Taming the Complexity of Future Transmission and Computation Challenges" funded by the Italian MIUR. It has been also partially supported by the Italian MIUR PON projects Pico&Pro (ARS01_01061), AGREED (ARS01_00254), FURTHER (ARS01_01283), RAFAEL (ARS01_00305).

References

[1] M. Giordani and M. Zorzi, "Non-terrestrial networks in the 6g era: Challenges and opportunities," *IEEE Network*, vol. 35, no. 2, pp. 244–251, 2021.

[2] Z. Zhang, W. Zhang, and F.-H. Tseng, "Satellite mobile edge computing: Improving qos of high-speed satellite-terrestrial networks using edge computing techniques," *IEEE Network*, vol. 33, no. 1, pp. 70–76, 2019.

[3] A. Chaoub, M. Giordani, B. Lall, V. Bhatia, A. Kliks, L. Mendes, K. Rabie, H. Saarnisaari, A. Singhal, N. Zhang, S. Dixit, and M. Zorzi, "6g for bridging the digital divide: Wireless connectivity to remote areas," *IEEE Wireless Communications*, pp. 1–9, 2021.

[4] M. Cui, Y. Fei, and Y. Liu, "A survey on secure deployment of mobile services in edge computing," *Security and Communication Networks*, vol. 2021, 2021.

[5] P. Ranaweera, A. D. Jurcut, and M. Liyanage, "Survey on multi-access edge computing security and privacy," *IEEE Communications Surveys Tutorials*, pp. 1–1, 2021.

[6] P. Ranaweera, V. N. Imrith, M. Liyanag, and A. D. Jurcut, "Security as a service platform leveraging multi-access edge computing infrastructure provisions," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[7] W. Sun, H. Zhang, R. Wang, and Y. Zhang, "Reducing offloading latency for digital twin edge networks in 6g," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 12 240–12 251, 2020.

[8] H. A. Almashhadani, X. Deng, S. N. A. Latif, M. M. Ibrahim, and A. H. Alshammari, "An edge-computing based task-unloading technique with privacy protection for internet of connected vehicles," *Wireless Personal Communications*, pp. 1–22, 2021.

[9] ZhenQin, Y. Liao, and D. Shen, "Multi-MEC server multi-user resource allocation in heterogeneous network," *Journal of Physics: Conference Series*, vol. 1792, p. 012005, feb 2021.

[10] J. Plachy, Z. Becvar, E. C. Strinati, and N. d. Pietro, "Dynamic allocation of computing and communication resources in multi-access edge computing for mobile users," *IEEE Transactions on Network and Service Management*, vol. 18, no. 2, pp. 2089–2106, 2021.

[11] Y. Deng, Z. Chen, X. Chen, and Y. Fang, "Throughput maximization for multi-edge multi-user edge computing systems," *IEEE Internet of Things Journal*, pp. 1–1, 2021.

[12] R. Brandborg Sørensen, H. Krogh Møller, and P. Koch, "5g nb-iot via low density leo constellations," *arXiv e-prints*, pp. arXiv–2108, 2021.

[13] M. A. Ullah, K. Mikhaylov, and H. Alves, "Massive machine-type communication and satellite integration for remote areas," *IEEE Wireless Communications*, vol. 28, no. 4, pp. 74–80, 2021.

[14] F. Glover, "Tabu search: A tutorial," *Interfaces*, vol. 20, no. 4, pp. 74–94, 1990.