Reinforcement Learning-based Security Orchestration for 5G-V2X Network Slicing at Cross-borders

Abdelwahab Boualouache*, Abdelaziz Amara Korba†‡, Sidi-Mohammed Senouci§,
Yacine Ghamri-Doudane† and Thomas Engel*
* FSTM, University of Luxembourg, Luxembourg. † L3I Laboratory, University of La Rochelle, France.
‡ LRS, Badji Mokhtar Annaba University, Algeria. § DRIVE Lab, University of Bourgogne, Nevers, France.

***Abstract*—As part of the 5G, Connected and Automated Vehicles (CAVs) will benefit from Network Slicing (NS) in several tailored 5G-Vehicle-to-Everything (V2X) services running on the same physical infrastructure. However, the use of 5G-NS may also increase the risk of cyber-attacks that could compromise 5G-V2X network slices (5G-V2X-NSs) and cause significant harm to CAV's passengers. This risk is particularly high at cross-borders, where CAVs move from their Home Mobile Network Operator (H-MNO) to a Visited MNO (V-MNO), with similar 5G-V2X-NSs in place. Therefore, deploying security services to neutralize 5G-V2X NS threats in this scenario is mandatory. However, if H-MNO and V-MNO act independently, deploying these security services could be inefficient and may result in increased memory, processing, and network resource consumption. Thus, MNOs should collaborate to orchestrate their security services to neutralize 5G-V2X NS attacks and optimize their costs efficiently. In this context, this paper proposes a novel approach to enhance the security of 5G-V2X NS at cross-borders using Reinforcement Learning (RL) based security orchestration. Specifically, we trained and deployed an RL agent interacting with both H-MNO and V-MNO. The RL agent efficiently deploys security services to effectively remove threats, optimize resource utilization, and minimize the impact on 5G-V2X-NSs. The performance results show that the RL-based security orchestration neutralizes threats with an average success rate of almost 100%. Additionally, resource consumption is minimal at less than 8%, and the acceptable impact on 5G-V2X-NSs is negligible, averaging less than 12%.**

***Index Terms*—5G-V2X; Network Slicing; Security Orchestration; Machine learning; Reinforcement Learning**

I. Introduction

Network slicing (NS) is a key feature of 5G, enabling several verticals and use-cases to co-exist in the same environment with stringent requirements and sharing the same physical infrastructure [1]. By integrating CAVs into the 5G and beyond ecosystem, they can benefit from 5G NS. Specifically, NS enables various isolated 5G-V2X networks with different requirements, such as autonomous driving and platooning, to be created and utilized effectively [2]. Besides, CAVs moving on the road can perform frequent handovers from one 5G cell to another. In open border environments such as the Schengen area, CAVs can cross borders seamlessly. When crossing a border, the roaming procedure is automatically triggered, causing the CAVs to switch their network attachment from their H-MNO to a V-MNO to allow for uninterrupted connectivity. In this case, the V-MNO allocates a 5G-V2X-NS with the same characteristics as 5G-V2X-NS in H-MNO [3].

However, 5G-V2X NS at cross-border is facing several security issues, leading to road hazards and threatening users' lives [4]. Specifically, attackers can exploit the lack of synchronization between the security policies of the H-MNO and the V-MNO, misconfigurations, and information obtained from system infiltration to target the data plane related to CAVs [5–7]. Several intelligent security services have recently been proposed to detect and mitigate attacks in 5G-V2X network slicing [8, 9]. However, deploying these services requires computation, memory, and storage resources to be allocated by the MNOs. Moreover, these security services should be deployed in the 5G-V2X-NSs, while meeting the requirements specified in the Service-Level Agreements (SLAs) between tenants and MNOs. Thus, deploying these security services requires careful attention from MNOs to optimize the allocation of security resources and prevent any adverse impact on the performance of 5G-V2X-NSs. This task is more complicated in cross-border and requires collaboration between H-MNO and V-MNO to ensure efficient neutralization of 5G-V2X NS threats. In addition, both MNOs should optimize their resources and meet SLA applied to 5G-V2X-NSs.

This paper aims to address the orchestration of security services between H-MNO and V-MNO with a primary objective of quickly and efficiently eliminating 5G-V2X network slicing attacks. Simultaneously, the paper aims to optimize the resources of both MNOs while ensuring minimal impact on 5G-V2X-NSs. Our scheme leverages Reinforcement Learning (RL) to train and deploy an RL-enabled orchestration agent. This agent interacts with both H-MNO and V-MNO to make decisions on the placement of security services over 5G-V2X-NSs of both MNOs. The agent considers varying contextual factors, such as threat level and resource consumption, while respecting SLAs.

The remainder of this paper is organized as follows. Section II describes related works. The proposed architecture is presented in Section III. The RL-based security orchestration designed for the 5G-V2X NS is described in IV. Section V describes the obtained. Finally, Section VI concludes the paper.

II. Related work

This section discusses related works regarding optimal network and security services placement, considering multiple constraints. The authors of [10] proposed a placement scheme of virtual security network functions considering both performance optimization and security aspects. The authors of [11] introduced a potential resolution to the issue of placing virtualized security functions in security service chaining within cloud environments. The authors of [12] proposed a mathematical model for optimally placing virtualized network functions at the edge, considering network security and operator's best practices. The authors in [13] proposed a scheme that combines a mathematical model and heuristic solutions to enable efficient deployment of security function chaining. The authors of [14] proposed an adaptive RL-based scheme to automatically select adequate security function chaining that meets dynamic context requirements. The authors in [15] proposed an RL-based scheme that deploys an RL agent that mitigates attacks by redirecting network traffic flows and reconfiguring security settings based on the observable network state. The authors in [16] proposed an RL-based hierarchical architecture for service function chaining placement on multiple domains while meeting the SLA requirements. The authors in [17] assessed the effectiveness of an RL-based approach for misbehavior detection in V2X scenarios considering the case of sudden-stop attacks. Building upon their previous work, the authors [18] proposed an RL-based detection model that processes V2X data broadcast by vehicles as time series at the roadside units. However, the studies in [10–13] are based on traditional optimization techniques, which have difficulty scaling for NP-Hard problems since their calculation time grows exponentially, rendering them unsuitable for real-time use. In addition, the works in [14–16] did not consider security service orchestration in the context of 5G-V2X and network slicing. Moreover, in [17, 18], RL was used for attack detection, not security orchestration. To this end, this paper addresses security orchestration for 5G-V2X in cross-border scenarios. As far as we know, this work is the first to address this gap.

III. Architecture

In this section, we describe the proposed architecture where we consider a scenario of 5G-V2X NS at a cross-border illustrated in Figure 1. In this scenario, H-MNO and V-MNO comprise a 5G core network and New Radio (NR), which includes Multi-access Edge Computing (MEC) nodes, gNodeBs, CAVs, and VRUs. The N32 standard reference point at the control plan interconnects H-MNO and V-MNO. In addition, the N9 standard reference point at the data plane interconnects H-MNO and V-MNO. Following an SLA between an MNO and a tenant, the network slice manager allocates all the necessary memory, network, and computation resources for creating the 5G-V2X-NSs and deploying all the Virtual Network Function (VNF) belonging to this 5G-V2X-NS. After deploying a 5G-V2X-NS, both MNOs should respect the requirements specified in the SLA. After a CAV or VRU crosses the borders, the V-MNO allocates a 5G-V2X-NS with the same functionality as the one in the H-MNO. In addition, in our scenario the H-MNO is responsible for managing data sessions, and data is only routed from the H-MNO to the V-MNO upon request.

The 3GPP standards propose security solutions for data and control plane exchanges between H-MNO and V-MNO [19]. To protect the core of each MNO and the internetwork packet exchange, a Security Edge Protection Proxy (SEPP) is placed at the edge of H-MNOs and V-MNOs. The SEPPs utilize Transport Layer Security (TLS) to secure their exchange via the N32. Similarly, to protect the data plane, the Inter-PLMN User Plane Security (IPUPS) feature is activated in each H-User Plane Function (UPF) and V-UPF to discard invalid data exchanges via the N9. However, more than 5G security measures are needed to protect 5G-V2X-NSs at cross-borders. Specifically, the roaming of CAVs is time-sensitive to ensure service continuity and road safety. Consequently, V-MNO should promptly allocate for CAVs a 5G-V2X-NS with features similar to that the H-MNO uses. However, this may result in several vulnerabilities, such as non-timely synchronization of security policies and misconfigurations on both the H-MNO and V-MNO sides, which could be exploited to launch attacks on the data plane [5–7]. In addition to 5G security solutions, other security services should be deployed to prevent, detect, and mitigate 5G-V2X NS attacks at cross

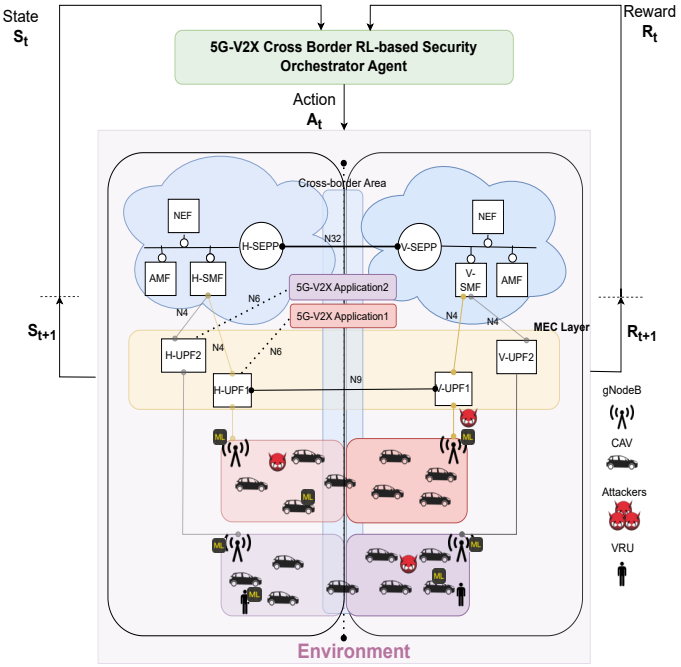

Fig. 1: RL-based Security Orchestration for 5G-V2X Network Slicing at Cross Borders

AMF: Access and Mobility Management Function, UPF: User Plane Function, SMF: Session Management Function, NEF: Network Exposure Function, SEPP: Security Edge Protection Proxy

borders. These services may include firewalls, intrusion detection, and prevention systems. However, these security services may be inefficient without collaboration between H-MNO and V-MNO. Moreover, non-cooperation between MNOs can result in a wastage of resources, impede the performance of 5G-V2X-NS, and even violate the SLA. To address this issue, we propose the deployment of a federated RL-based orchestrator on top of both MNOs. This orchestrator will ensure all necessary security services are deployed and effectively eliminate potential security threats. Our proposed approach optimizes resource utilization while respecting the SLA constraints.

IV. RL-based Security Orchestration for 5G-V2X Network Slicing at Cross Borders

This section presents our RL-based model that leverages Q-learning to orchestrate security in 5G-V2X NS at cross borders [20]. The RL model, depicted in Figure 1, consists of an agent interacting with the environment to gather information and determine the best actions to get the highest possible rewards. The RL agent aims to find an optimal policy that maximizes commutative rewards. In the following, we provide a detailed description of each element involved in the RL process.

A. Agent

The RL agent is a security orchestrator deployed at the top of the infrastructure of the two MNOs (H-MNO and V-MNO). The agent collects information from the environment (the state s) regarding the status of available resources, the threat level, and the negative impact on the 5G-V2X-NS after taking the previous decision (the action a). The agent gets a reward r as feedback for each taken action a. This reward helps the agent continuously improve its model, hence the decision-making process. The agent uses Equation (1) to iteratively update its RL model using Q-learning [20]. The agent stores all Q-values in the Q-table, which updates it in the training phase and uses it to pick optimal actions in the deployment.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha(r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)) \quad (1)$$

Where α and γ denote the learning rate and discount factor, respectively.

B. Environment

The RL model's environment controls the agent's training. It takes the agent's action a as input to produce a reward r and the subsequent environment state s for the agent. As illustrated in Figure 1, the environment is the 5G-V2X NS cross-border infrastructures of H-MNO and V-MNO. Specifically, the two MNOs provide interfaces to the RL agent for interacting with the environment, such as collecting information from the environment and sending the actions to execute securely. These interfaces could be the Network Exposure Functions (NEFs) deployed at 5G MNOs cores to facilitate interaction with third-party applications like the RL agent.

C. States

At each time step, the RL agent observes the state of the 5G-V2X-NS, which includes information on several key factors. Firstly, it considers the consumption state of the resources (CPU, memory, network) allocated for security services in all 5G-V2X-NSs. To simplify the notation, we use Γ_i to represent the overall percentage of remaining resources in a given 5G-V2X-NS i, where Γ_{max} is the maximum percentage of available resources (100%). Secondly, the agent considers the threat level of each 5G-V2X-NS, denoted as Ψ_i on a scale from 0 to 10, with Ψ_{max} being the maximum threat level of 10. Finally, the agent considers the remaining tolerable impacts of security services specified in the SLA for each 5G-V2X-NS. We denote Λ_i as the remaining tolerable impact percentage (from 0% to 100%) in a given 5G-V2X-NS i, where Λ_{max} is the maximum tolerable impact percentage (100%).

D. Actions

The action set is the candidate security services $SR = \{sr_j\}_{j=1}^{m}$ that the agent can select to neutralize the threats from the 5G-V2X-NSs. Each agent's action a determines the domain (the MNO) and the 5G-V2X-NS where the security service should be placed. Each security service sr_j is characterized by its power Υ (sr_j) to neutralize the threat, the resources Γ (sr_j) required to deploy sr_j, and Λ (sr_j) the impact on 5G-V2X-NS after deploying sr_j. Given the threat level, resource limitation, and tolerable impact on the 5G-V2X-NSs, selecting sr_j should meet the following objectives:

$$max\ \Upsilon = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_{ij} * \Upsilon(sr_j)) \quad (2)$$

$$min\ \Gamma = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_{ij} * \Gamma(sr_j)) \quad (3)$$

$$min\ \Lambda = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_{ij} * \Lambda(sr_j)) \quad (4)$$

Where (i) x_{ij} denotes the number of times that sr_j is selected, (ii) in Equation (2), Υ denotes the total security power over all the 5G-V2X-NSs, (iii) in Equation (3), Γ denotes the total of resources consumed over all the 5G-V2X-NSs, and (iv) in Equation (4) Λ denotes the negative impact caused by the selected security services over all the 5G-V2X-NSs. The selection of security services should also satisfy a set of requirements. First, in Equation (5), the sum of powers of selected security services for a given 5G-V2X-NS i must be greater or equal to the i's threat level Ψ_i.

$$\sum_{j=1}^{m}(x_j * \Upsilon(sr_j)) \geq \Psi_i,\ x_j = 0, 1, 2, 3, ..l; \quad (5)$$

In addition, as 5G-V2X-NSs have limited resources and tolerable impact, the selection of security services must meet certain criteria. According to Equation (6), the total sum of selected security services' resources for a given 5G-V2X-NS i should not exceed the maximum available resources Γ_{max} in i. Also, Equation (7) indicates that the total sum of tolerable

impact of selected security services for 5G-V2X-NS i should not exceed the maximum tolerable impact Λ_{max} on i.

$$\sum_{j=1}^{m}(x_j * \Gamma(sr_j)) \leq \Gamma_{max},\ x_j = 0, 1, 2, 3, ..l; \tag{6}$$

$$\sum_{j=1}^{m}(x_j * \Lambda(sr_j)) \leq \Lambda_{max},\ x_j = 0, 1, 2, 3, ...l; \tag{7}$$

E. Reward

The RL agent aims to neutralize the threats on 5G-NS-V2Xs while minimizing resource usage and negative impacts. Thus, the agent receives a reward r for each action a taken in a given state s.

Algorithm 1: The rewarding function

```
1  Input: f_c = 0, f_i = 0, threat = t, μ, ν, ξ
2  Output: terminated, reward
3  if Γ(sr_j) > Γ_i then
4  |   f_Γ = 1;
5  end
6  if Λ(sr_j) > Λ_i then
7  |   f_Λ = 1;
8  end
9  if ((f_Γ==1) or (f_Λ==1)) then
10 |   terminated = True;
11 |   reward = f_Γ * -Γ(sr_j) + f_Λ * -Λ(sr_j) ;
12 else
13 |   if Ψ_i > 0 then
14 |   |   if Ψ_i > ϒ(sr_j) then
15 |   |   |   reward = Ψ_i - ϒ(sr_j);
16 |   |   else if Ψ_i == ϒ(sr_j) then
17 |   |   |   reward= ϒ(sr_j) ;
18 |   |   |   threat = threat -1 ;
19 |   |   else
20 |   |   |   reward = (μ * Ψ_i) - (ν * Γ(sr_j)) - (ξ * Λ(sr_j));
21 |   |   |   threat = threat -1;
22 |   |   end
23 |   |   if (threat==0) then
24 |   |   |   terminated = True;
25 |   |   |   reward = 100;
26 |   else
27 |   |   reward = -10
28 |   end
29 end
```

The pseudo-code for the rewarding function is presented in Algorithm 1. In line 4, the flag f_Γ is raised if the agent tries to place a security service that consumes more than the resource available in the 5G-V2N-NS. Similarly, in line 7, the flag f_Λ is raised if the agent tries to place a security service that exceeds the remaining capacity of the 5G-V2N-NS to tolerate impacts. Thus, if either f_Γ or f_Λ is equal to one, the agent's action results in failure, and it receives a negative reward that depends on the specific flag that was raised, as well as the available resources and the impact of the security service it attempted to place. If neither of the flags are raised, the agent can apply the security service. In lines 14-15, if the agent applies security service with a power less than the threat level of the 5G-V2X-NS, it receives a reward equal to the remaining threat level on that 5G-V2X-NS. Otherwise, if the agent applies a security service with a power equivalent to the threat level of that 5G-V2X-NS, the agent thus successfully removes the threat from that 5G-V2X-NS and receives a reward equal to the power of the applied security service. In addition, if the power of the security service is more than the threat level of that 5G-V2X-NS, the agent successfully removes the threats from that 5G-V2X-NS. Still, its reward depends on the threats, the resources consumed, and the impact on that 5G-V2X-NS. Here, we define the factors μ, ν, and ξ to establish a trade-off among the previous parameters in the reward calculation. Note that the agent receives a reward of -10 if it tiers to apply a security service in a free-threat 5G-V2X-NS. Finally, if the agent successfully neutralizes all the threats presented in all 5G-V2X-NSs, it succeeds and receives a reward of 100.

V. Performance Evaluation

In this section, we conduct experiments to evaluate the performance of the RL-based security orchestrator and compare it with uncoordinated decision-making by MNOs. This section first presents the RL agent's simulation environment setting and training process. Then, it shows and discusses the results obtained from the testing process.

A. Environment Setting and Training

We consider that the H-MNO manages five 5G-V2X-NSs. To enable service continuity for CAVs and VRUs crossing borders, the V-MNO allocates five 5G-V2X-NSs with similar functionalities. The task of the RL agent is thus to strategically place security services to neutralize attacks from these ten 5G-V2X-NSs. Initially, the percentage of available resources for these security services and tolerable impact in these 5G-V2X-NSs is set to 100 %. Moreover, we offer a range of twenty security services with different combination settings (Υ_i, Γ_i, Λ_i) to enable the RL agent to neutralize the threats from the 5G-V2X-NSs. Each security service has a power level (Υ_i) that ranges from 1 to 10, a resource requirement value (Γ_i) that ranges from 1 to 20, and an impact level (Λ_i) that can be classified as weak (10%), medium (20%), or high (30%). Additionally, we have set the factors μ, ν, and ξ in the rewarding functions to 0.9, 0.09, and 0.01, respectively.

We run 6000 episodes to train the RL agent. The environment is reset at the beginning of each episode, and the number of threats is randomly distributed over 5G-V2X-NS. In our training process, we use the ϵ-greedy exploration strategy. Specifically, a random action is selected with probability ϵ, and the action with the highest Q-value is selected with probability 1-ϵ. After each episode, we update the probability ϵ by applying the exponential decay formula outlined in Equation (8), where ϵ_{min} is set to 0.01, decay is set to 0.001, and e represents the running episode number. The Q-learning value is updated using a learning rate (α) of 0.1 and a discounted factor (γ) of 0.99.

$$\epsilon = max(\epsilon_{min}, \exp^{decay*e}) \tag{8}$$

Figure 2 shows the average reward per action during the training. The average reward per action rises quickly with the increase of episodes. After about 4000 episodes, the increasing

trend stops. One of the reasons why the reward fluctuates continuously is that threats are randomly distributed over the 5G-V2X-NSs at the beginning of each episode. Figure 3 shows the average number of actions per episode during the training. As we can see, the average number of actions decreases quickly with the increase in episodes. Initially, it takes an average of almost 20 security services to be selected to neutralize the threats from the 5G-V2X-NS. Ultimately, it has been observed that mitigating the threats requires the selection of fewer than five security services.

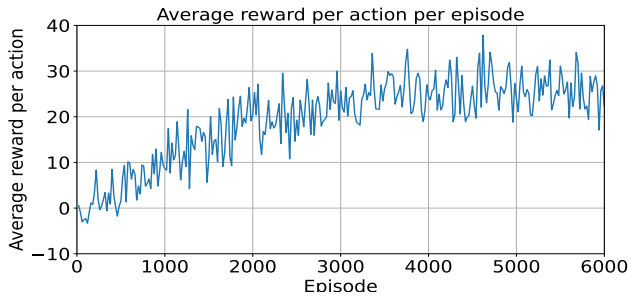

Fig. 2: Average reward per action during the training

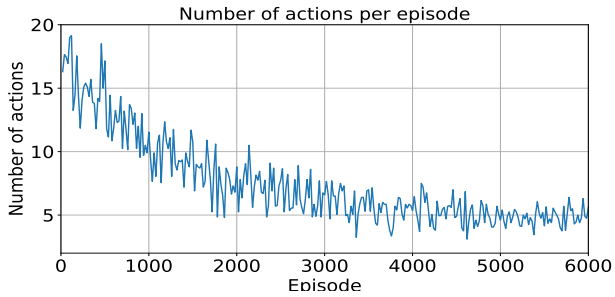

Fig. 3: Average number of actions per episode

B. Performance evaluation

After the training, we use the Q-table for online decisions based on Algorithm 2. Suppose the agent detects threats in the environment (env) consisting of 5G-V2X-NS associated with H-MNO and V-MNO. In that case, it uses the Q-table to efficiently determine the best action (selection of security service) at each step to remove the threats from the 5G-V2X-NS. After applying a security service, the agent gets a reward, a new state, and flags that indicate if an endpoint is reached. If the endpoint is reached (done equal to true), the agent checks the last reward to check if it succeeded in neutralizing the threats or not. The agent collects results for both successful and failed cases, gathering information on success rates, resource consumption, and impacts in the former, and wasted resources and unnecessary impacts, as well as success rates in the latter. In our experiments, we tested our RL agent 6000 times, using the same security services as in the training phase.

The average results are presented in Figures 4, 5, and 6. To evaluate the effectiveness of our RL-based security

Algorithm 2: The selection of security services using Q-table

```
1  Input: env, security services
2  Output: Results on success rate, consumed resources, and impacts
3  while Threat == true do
4      for step <= max_steps) do
5          action = argmax(Q_table[state,:]) ;
6          new_state, reward, done←Apply (env, action);
7          if done == true then
8              if reward==100 then
9                  Succeed;
10             else
11                 Failed;
12             end
13             Break;
14         end
15         state = new_state ;
16         Collect_results()
17     end
18 end
```

orchestration, we compare it with a random orchestration. Specifically, random orchestration here refers to individual actions taken by the H-MNO and V-MNO to neutralize threats without any coordinated effort between them. Figure 4 shows the average success rate of neutralizing the threats from 5G-V2X-NSs. As we can see, our RL agent obtained almost 100% of the success rate. On the other hand, the random strategy only achieved an average success rate of 40%. These results demonstrate the superior effectiveness of our RL-based in neutralizing the threats from 5G-V2X in most cases.

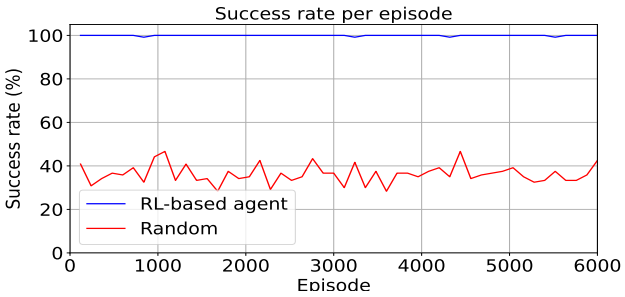

Fig. 4: Success rate for both random and RL-based orchestration strategies

Figure 5 shows the impact of both RL-based and random strategies on the 5G-V2X-NS. Notably, the random strategy appears to have a greater impact on 5G-V2X-NS compared to the RL-based approach, as it fails to take into account the consequences of security service selection. In contrast, the RL-based orchestration considers factors such as power consumption, resource utilization, and overall impact when selecting security services. Thus, the RL-based approach enables a more efficient and effective orchestration of security services for 5G-V2X-NS than the random strategy.

Figure 6 compares the consumed resources in the case of RL-based orchestration and random orchestration. The results show that random orchestration consumes fewer resources on average compared to RL-based orchestration. This is because,

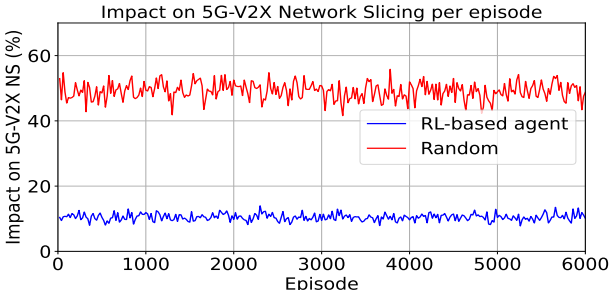

Fig. 5: Impact on 5G-V2X-NSs for both random and RL-based orchestration strategies

in the random selection of security services, the placement failure ratio is primarily due to impact constraint that stops random strategy from consuming more resources. Moreover, the majority of consumed resources in the case of random strategy is wasted due to the weak success rate. In contrast, RL-based orchestration efficiently utilizes resources to neutralize attacks, as demonstrated by its high success rate.

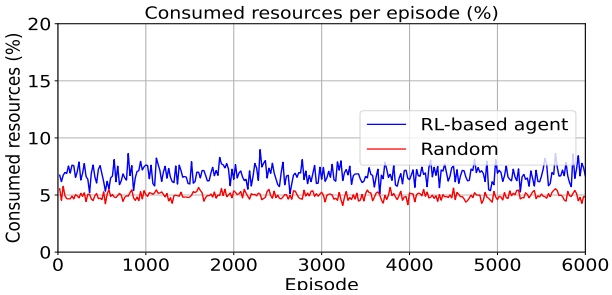

Fig. 6: Consumed resources for both random and RL-based orchestration strategies

VI. Conclusion

The failure to neutralize threats from 5G-V2X-NSs in cross-border scenarios can have catastrophic consequences. Therefore, the collaboration between MNOs is crucial to act efficiently without wasting resources and disrupting services. This paper proposed an RL-based security orchestration for effectively removing threats in 5G-V2X NS at cross borders. The results demonstrate its efficiency compared to random strategy orchestration, achieving a 100% of success rate with optimal resource consumption and tolerable impact of 5G-V2X slice services. In future work, we plan to investigate an extensive scenario involving multiple MNOs and leverage deep RL to handle the large potential state space. Additionally, we intend to implement and validate these results through a proof-of-concept.

Acknowledgment

This work was also supported by the 5G-INSIGHT bilateral project (ID: 14891397) / (ANR-20-CE25-0015-16), funded by the Luxembourg National Research Fund (FNR), and by the French National Research Agency (ANR).

References

[1] I. Afolabi, T. Taleb, K. Samdanis, A. Ksentini, and H. Flinck, "Network Slicing and Softwarization: A Survey on Principles, Enabling Technologies & Solutions," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2429–2453, 2018.

[2] C. Campolo, A. Molinaro, A. Iera, and F. Menichella, "5G network Slicing for Vehicle-to-Everything Services," *IEEE Wireless Communications*, vol. 24, no. 6, pp. 38–45, 2017.

[3] GSMA Association, "An Introduction to Network Slicing," 2017. [Online]. Available: https://www.gsma.com/futurenetworks/wp-content/uploads/2017/11/GSMA-An-Introduction-to-Network-Slicing.pdf

[4] A. Boualouache, B. Brik, Q. Tang, A. A. Korba, S. Cherrier, S.-M. Senouci, E. Pardo, Y. Ghamri-Doudane, R. Langar, and T. Engel, "5G Vehicle-to-Everything at the Cross-Borders: Security Challenges and Opportunities," *IEEE Internet of Things Magazine*, vol. 6, no. 1, pp. 114–119, 2023.

[5] S. Park, S. Kwon, Y. Park, D. Kim, and I. You, "Session management for security systems in 5g standalone network," *IEEE Access*, vol. 10, pp. 73 421–73 436, 2022.

[6] G. Amponis, P. Radoglou-Grammatikis, T. Lagkas, W. Mallouli, A. Cavalli, D. Klonidis, E. Markakis, and P. Sarigiannidis, "Threatening the 5G core via PFCP DoS attacks: the case of blocking UAV communications," *EURASIP Journal on Wireless Communications and Networking*, vol. 2022, no. 1, pp. 1–27, 2022.

[7] S. Kukliński, K. Szczypiorski, K. Wrona, and J. Bieniasz, "5G-Enabled Defence-in-Depth for Multi-domain Operations," in *MILCOM 2022-2022 IEEE Military Communications Conference (MILCOM)*. IEEE, 2022, pp. 1024–1029.

[8] T. Wichary, J. Mongay Batalla, C. X. Mavromoustakis, J. Żurek, and G. Mastorakis, "Network Slicing Security Controls and Assurance for Verticals ," *Electronics*, vol. 11, no. 2, p. 222, 2022.

[9] A. Boualouache and T. Engel, "A Survey on Machine Learning-based Misbehavior Detection Systems for 5G and Beyond Vehicular Networks," *IEEE Communications Surveys & Tutorials*, 2023.

[10] R. Doriguzzi-Corin, S. Scott-Hayward, D. Siracusa, and E. Salvadori, "Application-centric provisioning of virtual security network functions," in *2017 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN)*. IEEE, 2017, pp. 276–279.

[11] H. Wu, Y. Zhang, H. Yang, G. Yu, and J. Cao, "Virtualized Security Function Placement for Security Service Chaining in Cloud," in *2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS)*. IEEE, 2018, pp. 628–637.

[12] M. M. Iordache-Sica, C. Anagnostopoulos, and D. P. Pezaros, "Towards QoS-aware Provisioning of Chained Virtual Security Services in Edge Networks," in *2021 IFIP/IEEE International Symposium on Integrated Network Management (IM)*. IEEE, 2021, pp. 178–186.

[13] R. Doriguzzi-Corin, S. Scott-Hayward, D. Siracusa, M. Savi, and E. Salvadori, "Dynamic and Application-Aware Provisioning of Chained Virtual Security Network Functions," *IEEE Transactions on Network and Service Management*, vol. 17, no. 1, pp. 294–307, 2019.

[14] G. Li, H. Zhou, B. Feng, G. Li, and S. Yu, "Automatic Selection of Security Service Function Chaining Using Reinforcement Learning," in *2018 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2018, pp. 1–6.

[15] M. Zolotukhin, P. Kotilainen, and T. Hämäläinen, "Intelligent IDS Chaining for Network Attack Mitigation in SDN," in *2021 17th International Conference on Mobility, Sensing and Networking (MSN)*. IEEE, 2021, pp. 786–791.

[16] N. Toumi, M. Bagaa, and A. Ksentini, "Hierarchical multi-agent deep reinforcement learning for SFC placement on multiple domains," in *2021 IEEE 46th Conference on Local Computer Networks (LCN)*. IEEE, 2021, pp. 299–304.

[17] R. Sedar, C. Kalalas, F. Vázquez-Gallego, and J. Alonso-Zarate, "Reinforcement Learning-based Misbehaviour Detection in V2X Scenarios," in *2021 IEEE International Mediterranean Conference on Communications and Networking (MeditCom)*. IEEE, 2021, pp. 109–111.

[18] ——, "Reinforcement Learning Based Misbehavior Detection in Vehicular Networks," in *ICC 2022-IEEE International Conference on Communications*. IEEE, 2022, pp. 3550–3555.

[19] 3GPP TS 33.501, "Security architecture and procedures for 5G system (Release 17)," Sep 2022.

[20] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.