Environmental-Aware Scheduler for Trustworthy 6G Communication

Roya Khanzadeh[1,2], Fjolla Ademaj-Berisha[2,3], Andreas Springer[1,2], and Hans-Peter Bernhard[1,3]

[1] Johannes Kepler University Linz, Institute for Communications Engineering and RF-Systems, Linz, Austria
[2] JKU LIT SAL IWS Lab, Linz, Austria [3] Silicon Austria Labs GmbH, Linz, Austria

***Abstract*—Trustworthiness is emerging as a critical design objective for 6G networks, ensuring consistent and reliable performance under dynamic conditions. While several studies have addressed its evaluation, enhancing trustworthiness through intelligent network optimization remains largely unexplored. This paper proposes a deep reinforcement learning-based scheduler that integrates environmental-aware knowledge from channel knowledge maps to optimize reliability, availability, and fairness aspects of trustworthiness. Compared to the round-robin scheduler, the proposed method improves reliability by over 300% with only an 8% drop in availability. Relative to proportional fair, it improves availability and fairness by 61.4% and 40.6%, respectively, with just a 4.5% reduction in reliability. These results demonstrate a balanced and effective environmental-aware scheduling solution for trustworthy 6G.**

***Index Terms*—Trustworthiness, Reliability, Availability, Fairness, Reinforcement Learning, Environmental-aware knowledge.**

I. Introduction

THE vision for 6G networks extends beyond throughput and latency improvements, aiming for consistent trustworthy operation, defined as predictable performance of the network under dynamic conditions [1]. While several studies have addressed the evaluation of trustworthiness in communication and networking systems, enhancing it through intelligent network optimization remains largely unexplored. To this end, recent discussions propose incorporating a dedicated trustworthiness management layer within 6G architectures, guiding key network functions and tasks such as scheduling, resource allocation, and routing [2].

This work focuses on the scheduling task, proposing a novel framework that enhances trustworthiness in terms of reliability, availability, and fairness. Conventional schedulers primarily optimize standard network performance indicators, such as throughput and latency. In contrast, this paper defines measurable trustworthiness indicators and aims to enhance them via an intelligent, environmental-aware network scheduler. The considered trustworthiness indicators in this paper are service reliability, which reflects the system's ability to maintain uninterrupted connectivity, service availability, referring to the timely provision of requested services, and fairness, which ensures equitable service access across nodes [3].

The considered scenario consists of mobile nodes experiencing time-varying channel conditions, significantly influenced by changes in the surrounding physical environment. Recent advances in integrated sensing and communication (ISAC) and multimodal sensing have highlighted the benefits of leveraging environmental-aware knowledge (EaK) for network optimization [4]. The proposed scheduler incorporates the EaK, derived from channel knowledge map (CKM) based on site-specific line-of-sight (LOS) or non-line-of-sight (NLOS) conditions, to enhance reliability, availability, and fairness. These objectives often conflict with each other, i.e., favoring nodes with reliable communication channels may reduce availability for others or lead to unfair access, making the task inherently multi-objective. A deep reinforcement learning (DRL) method is introduced to find a solution for this multi-objective optimization. The agent learns to improve reliability by serving nodes in LOS regions, enhancing availability by prioritizing nodes approaching NLOS, while preserving fairness across the nodes in the network.

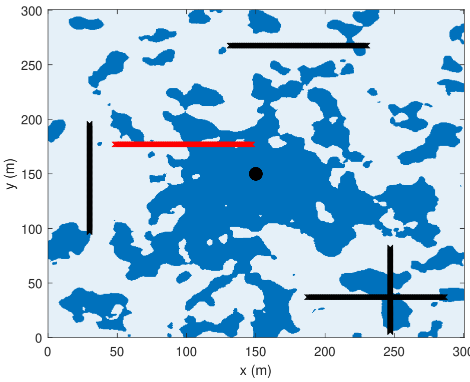
II. System Model

We consider frequency division duplexing (FDD) downlink communication in a cellular network, where the base station (BS) communicates with K moving nodes in a dynamic environment. Our scenario is assumed to be interference-free. The BS has access to the CKM that reflects the channel conditions for the entire environment. Based on this CKM, the BS can determine the current channel state for each node and generate short-term predictions based on their movement trajectories. In many controlled environments, such as factory floors, node mobility follows predictable paths, enabling trajectory estimation from periodic position updates.

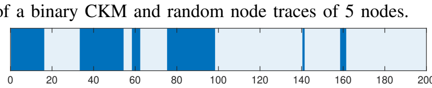
To model EaK, we adopt a simulation-based framework using the 3rd generation partnership project (3GPP) 2D geometry-based stochastic channel model (GSCM). This model captures realistic multipath propagation through location-dependent statistical distributions. These statistical distributions are parameterized by site-specific measurements, reflecting a realistic behavior of the propagation environment. Spatial consistency is maintained in the simulation, reflecting how neighboring positions share similar propagation effects. Following the probabilistic sampling method in [5], a binary CKM is generated to represent LOS/NLOS states across the environment (see an example of a binary CKM in Fig. 1). The LOS/NLOS state of each node k at transmission time interval (TTI) t, denoted by $q(t,k)$, serves as the primary input to the proposed environmental-aware scheduler.

III. Proposed Method

The environmental-aware scheduler aims to find a policy to select nodes to be served in each TTI t, in a way that enhances reliability, availability, and fairness. We propose a DRL method to find the solution for this multi-objective optimization task, wherein the reward function is a weighted sum



(a) an example of a binary CKM and random node traces of 5 nodes.



(b) LOS or NLOS states for a random node trace marked in red from Fig. 1a.

Fig. 1: A binary CKM for a 300×300 m^2 area with a resolution of 1 meter, where the black circle at the center shows the BS, and dark blue and light blue represent LOS and NLOS states, respectively.

of these three objectives. The proposed DRL model consists of an environment, a DRL agent, and a reward function. The agent interacts with the environment and learns how to select nodes at each environment condition to maximize rewards by improving communication reliability and availability while ensuring fairness. Fig. 2 shows the architecture for the proposed method.

We define matrix $\mathbf{S}(t)$, with dimensions $(T+2) \times K$, as the state of the environment at TTI t, where the kth column correspond to the kth node and is defined as

$$\mathbf{s}(t,k) = [h(t,k), s(t,k), s(t+1,k), \ldots, s(t+T,k)], \quad (1)$$

with $s(t,k) = 1$ or 0, if $q(t,k)$ is equivalent to LOS or NLOS, respectively. Similarly, $s(t+i,k), i = \{1,\ldots,T\}$ are the predicted LOS or NLOS states of the kth node in the T upcoming TTIs. $h(k,t) \in \{0,\ldots,L\}$ denotes the number of TTIs the kth node has been served over the past L transmissions. For simplicity, we assume only a single node is served at each TTI. Thus, the RL agent observes at each TTI the state $\mathbf{S}(t)$ and selects the current action $a(t) \in \{1,\ldots,K\}$ that maximizes the expected reward, i.e., it chooses a single node to be served. A feedforward neural network (FNN) approximates the action-value function by mapping the vectorized state $\mathbf{S}(t)$ to Q-values for all possible actions. At each TTI, the agent selects an action following the ε-greedy policy.

To encourage the agent to take an action that leads to maximum reliability, availability, and fairness, the reward function at TTI t is defined as

$$r(t) = c_1 \mathrm{Rel}(t) + c_2 \mathrm{Avl}(t) + c_3 \mathrm{Fair}(t). \quad (2)$$

The parameters c_1, c_2, c_3 are positive scalar weights that align the reward function with the objectives. At each TTI t, the reliability score, i.e. $\mathrm{Rel}(t)$, is one if the selected node to be served is located in LOS state, i.e $q(t,a(t)) \equiv \mathrm{LOS}$. The agent does not receive a reliability-related reward if $q(t,a(t)) \equiv \mathrm{NLOS}$ but there are no other nodes in the LOS state in time t.

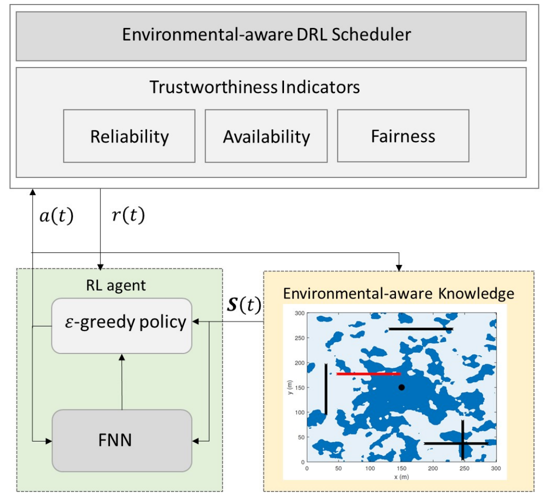


Fig. 2: The proposed environmental-aware scheduler for trustworthy 6G networks based on DRL.

Finally, $\mathrm{Rel}(t) = -1$ if $q(t,a(t)) \equiv \mathrm{NLOS}$, even though other nodes in LOS states are available.

Enhancing service availability requires a reduction of time between access (TBA) for all nodes evolved in the network, where $\mathrm{TBA}(t,k)$ represents the number of TTIs that have elapsed from the latest access of the kth node to the network, measured at TTI t. It means if the kth node is served at TTI t, then $\mathrm{TBA}(t,k) = 0$; otherwise, it grows as $\mathrm{TBA}(t,k) = \mathrm{TBA}(t-1,k) + 1$. We define ζ as the threshold for the acceptable number of TTIs between accesses. If the TBA for a node exceeds this threshold, the service for that node is considered unavailable. To enhance availability, the agent prioritizes nodes transitioning from LOS to NLOS, taking their movement trajectories into account. This prioritization is essential because delaying service for these nodes reduces their chances of being served in future TTIs, as their service reliability scores decrease over time due to their movement. Thus, the agent earns an extra corresponding reward for selecting a node in LOS that is transiting to NLOS state within the next T TTIs, with

$$\mathrm{Avl}(t) = \begin{cases} \frac{1}{\sum_{i=1}^{T} s(t+i,a(t))}, & \text{if } q(t,a(t)) \equiv \mathrm{LOS} \\ 0, & \text{Otherwise} \end{cases}. \quad (3)$$

Finally, actions that improve network fairness result in greater reward, following Jain's fairness index [6]:

$$\mathrm{Fair}(t) = \frac{\left(\sum_{k=1}^{K} \sum_{\tau=t-L}^{t} \left(a(\tau) \equiv k\right)\right)^2}{K \sum_{k=1}^{K} \left(\sum_{\tau=t-L}^{t} \left(a(\tau) \equiv k\right)\right)^2}, \quad (4)$$

where $a(\tau) \equiv k$ evaluates to one if true, and zero otherwise, and L denotes the number of TTIs considered as the recorded history at the BS.

IV. Simulation Results

The CKM simulation is based on the urban micro (UMi) scenario specified in 3GPP TR 38.901 and the implementation from [5]. The simulation setup includes $K = 5$ nodes, and the history and prediction lengths are both set to 10 TTIs,

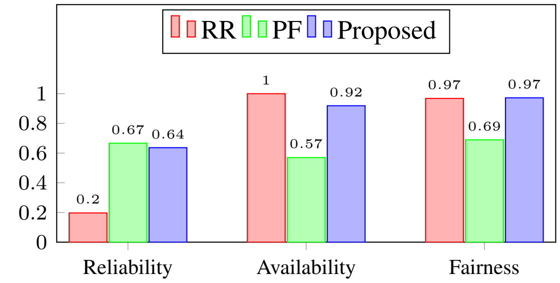


Fig. 3: Average performance comparison between the proposed algorithm and RR and PF methods.

i.e. $T = L = 10$. The DRL agent uses a discount factor $\gamma = 0.95$, a learning rate of 0.1, and a target network update every 100 steps. Following a trial-and-error approach, the reward weights are set as $c_1 = 5$, $c_2 = 10$, and $c_3 = 5$, with $\zeta = 40$ controlling the sensitivity to availability gaps. The DRL scheduler undergoes training for 1000 iterations, with a new $300 \times 300\,\mathrm{m}^2$ 2D map randomly generated for each iteration. Node initial positions and movement angles are sampled from a uniform distribution, while the BS remains fixed at the map centre. For each iteration, nodes move for 200 steps (200 TTIs), with their movement trajectories assumed to be known for T future TTIs. All nodes are configured with identical traffic and service requirements.

We compare the performance of the proposed method with two well-established scheduling techniques in the literature: round robin (RR) and proportional fair (PF) [7]. In this scenario, RR aims to maximize service availability and fairness by serving nodes in a fixed, regular order, while PF prioritizes service reliability by favoring nodes with better channel conditions. In contrast, the proposed method strikes a balanced trade-off between reliability, availability, and fairness, positioning itself between RR and PF. Fig. 3 illustrates the average performance over 1000 randomly generated environment conditions, each with varying initial node positions and movement trajectories. Compared to RR, the proposed method maintains the same level of fairness while boosting reliability by over 300%, at the cost of a slight 8% reduction in availability. When compared to PF, it improves availability and fairness by 61.4% and 40.6%, respectively, with only a minor drop in reliability of approximately 4.5%. Note that the maximum achievable value for each criterion is one. The radar chart in Fig. 4 visually demonstrates the performance balance among the schedulers. It shows that the proposed method occupies a significantly larger triangular area compared to the other methods, confirming its ability to maintain a more balanced performance across all criteria. Additionally, Fig. 5 compares availability across all methods as parameter ζ varies. As expected, availability improves for all methods as the threshold ζ increases, with the proposed method approaching the availability level of RR. However, the choice of ζ should be application-specific, as it defines what constitutes acceptable service delay and directly impacts how availability is measured.

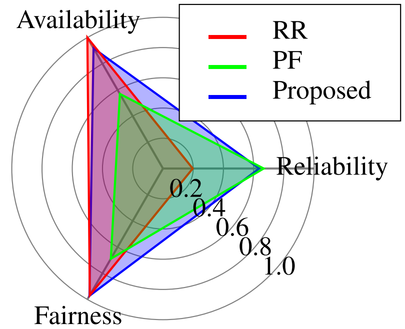


Fig. 4: Radar chart for reliability, availability, and fairness.

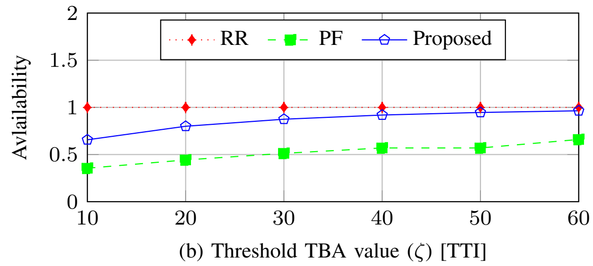


Fig. 5: Availability versus threshold TBA value ζ.

V. Conclusion

This paper proposed a DRL-based scheduling strategy that leverages current and near-future environmental-aware knowledge to optimize reliability, availability, and fairness. The proposed scheduler can be integrated into future network management frameworks, supporting the vision of trustworthy communication in 6G. Simulation results confirm its effectiveness in balancing reliability, availability, and fairness. As future work, we plan to extend the approach to multi-node scheduling and evaluate its performance in more diverse node requirements, while also incorporating additional trustworthiness criteria such as security and privacy.

Acknowledgment

This work has been in part supported by the "University SAL Labs" initiative of Silicon Austria Labs (SAL) and its Austrian partner universities for applied fundamental research for electronic-based system, and also by the European Unions Horizon Europe project DETERMINISTIC6G under grant agreement No. 101096504.

References

[1] D. P. M. Osorio et al., "Trustworthy 6G: Misconceptions, Attributes, and the Labeling Approach," *Authorea Preprints*, 2024.

[2] G. P. Fettweis and H. Boche, "On 6G and trustworthiness," *Communications of the ACM*, vol. 65, no. 4, pp. 48–49, 2022.

[3] J.-H. Cho et al., "Stram: Measuring the trustworthiness of computer-based systems," *ACM Computing Surveys (CSUR)*, vol. 51, no. 6, pp. 1–47, 2019.

[4] Y. Zeng et al., "Toward Environment-Aware 6G Communications via Channel Knowledge Map," *IEEE Wireless Communications*, vol. 28, no. 3, pp. 84–91, 2021.

[5] F. Ademaj, M. K. Mueller, S. Schwarz, and M. Rupp, "Modeling of spatially correlated geometry-based stochastic channels," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2017, pp. 1–6.

[6] D. Perdana et al., "Performance evaluation and comparison of scheduling algorithms on 5G networks using network simulator," *International Journal of Computers Communications & Control*, vol. 14, no. 4, pp. 530–539, 2019.

[7] A. Mamane et al., "Scheduling algorithms for 5G networks and beyond: Classification and survey," *IEEE Access*, vol. 10, pp. 51 643–51 661, 2022.