Reinforcement Learning with 3D UAV Network Digital Twin in Silico

NC STATE UNIVERSITY

NC State College of Engineering
Department of Computer Science

Everett Tucker[1], Jiayuan Huang[1], Yuchen Liu, PhD[1]

[1] Networking and IntelligenCE (NICE) Lab @ North Carolina State University, Raleigh

Abstract

As wireless communication technology quickly advances to 5G and 6G in urban environments, the network infrastructure must adapt to suit new needs. The higher frequency signals of 5G and 6G networks decrease the service radius of cell towers while making signals more sensitive to physical obstructions. To meet this need, an emerging approach uses Uncrewed Aerial Vehicles (UAVs) to act as uplinks between Ground Users (GUs), such as pedestrians and cell towers. To best serve Ground Users, we want to optimize the UAVs' flight patterns, network structure, and signal power. This Digital Twin (DT), essentially a simulation of the UAVs' urban environment, seeks to answer this optimization problem. Then, we can create deterministic or reinforcement learning (RL) models using the Digital Twin to approximate optimal UAV behavior. Previous research has constructed similar Digital Twins to model UAV communications alongside various models. However, these simulations are often simplified by removing GU movement, removing physical obstructions like buildings, or ignoring UAV power consumption from movement. This research seeks to design a comprehensive digital twin for aerial base stations in an urban environment that could train a more thorough reinforcement learning model. It incorporates a Sionna Ray Tracing for path quality computation, Simulation of Urban Mobility (SUMO) for realistic traffic simulation, and a custom UAV control framework. This simulation allows users to model UAV and GU movement, communication, and power consumption in a time-discrete, space-continuous environment. The simulation also uses deterministic algorithms to optimize network structure and UAV transmitter power. In the future, this Digital Twin could train a more advanced model to optimize the positions of aerial base stations in an urban setting.

Background

Previous research has considered similar digital twins but emphasized reinforcement learning techniques at the expense of a more thorough simulation. A 2024 paper demonstrated a comprehensive communication-focused digital twin with SUMO, Sionna, and OpenStreetMaps, which used Deep Deterministic Policy Gradient (DDPG) to optimize UAV positions. However, this simulation assumed a single UAV device and omitted power consumption (Lee, Jeongyoon, et al.). Another 2023 work implemented a simple space-discrete digital twin combined with Multi-Agent Deep Reinforcement Learning (MADRL) to solve a target search problem (Shen, Gaoquin, et al.). While this simulation lacks many of the features of a communications simulation, the reinforcement learning approaches are applicable more broadly. Research has also considered multi-UAV cooperation, such as a 2023 paper that addressed task scheduling over time in a physically representative digital twin simulation (Tang, Xin, et al.). This paper also considers energy consumption and uses Deep-Q Learning to optimize task scheduling. In general, prior research features more specific digital twins and more powerful machine-learning models compared to this simulation environment. By fusing ideas and techniques from these previous studies, a complete digital twin can be constructed to yield sounder models.

References

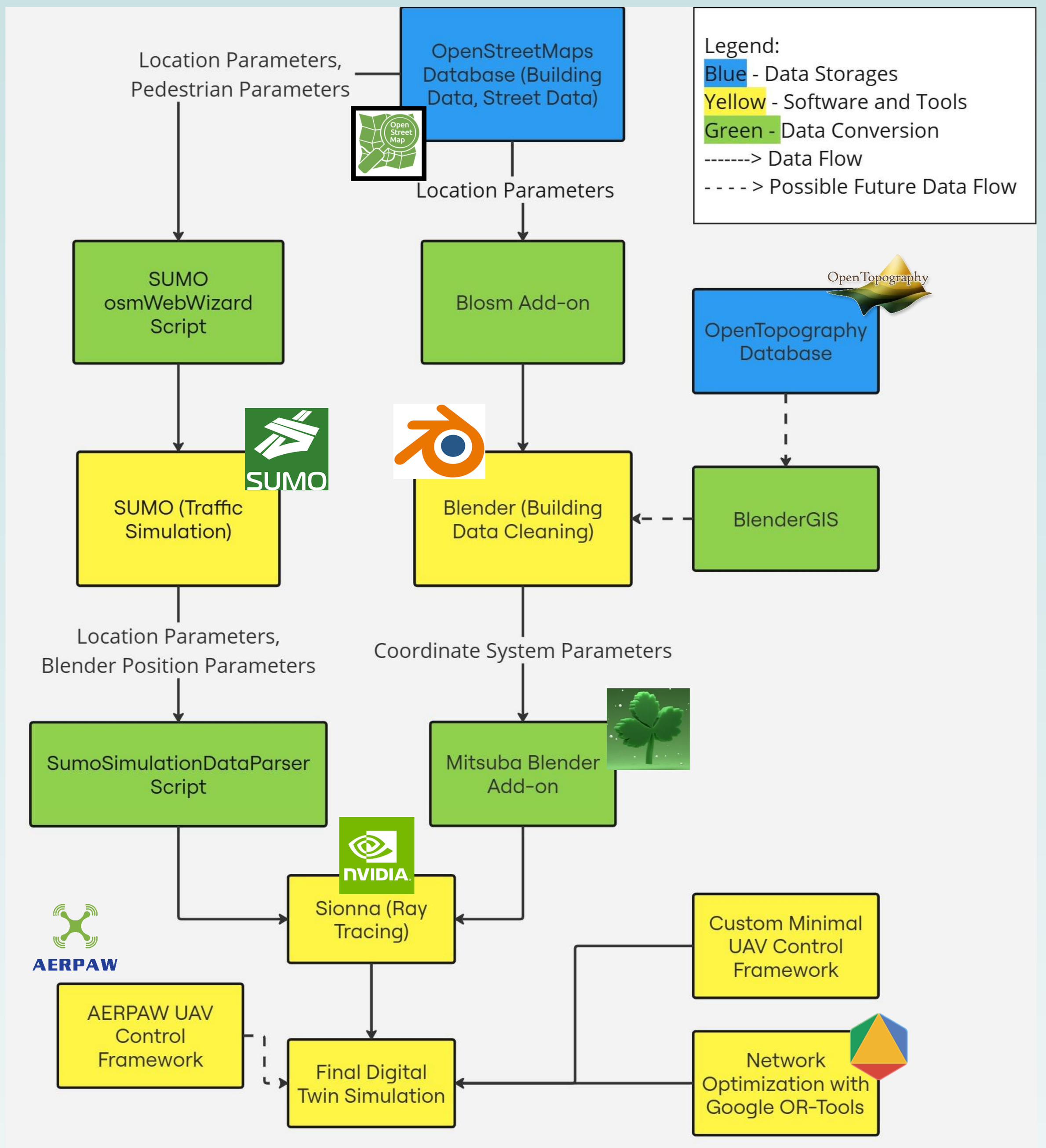
Lee, Jeongyoon et al. Digital twin based DDPG reinforcement learning for sum-rate maximization of AI-UAV communications. J Wireless Com Network 2024, no. 57 , 18 July 2024.

Shen, Gaoquin et al. "Multi-UAV Cooperative Search Based on Reinforcement Learning With a Digital Twin Driven Training Framework," in *IEEE Transactions on Vehicular Technology*, vol. 72, no. 7, pp. 8354-8368, July 2023.

Tang, Xin *et al*. "Digital-Twin-Assisted Task Assignment in Multi-UAV Systems: A Deep Reinforcement Learning Approach," in *IEEE Internet of Things Journal*, vol. 10, no. 17, pp. 15362-15375, Sept.1, 2023.

Methodology

Digital Twin Data Pipeline



Tools Used

Simulation of Urban Mobility (SUMO)
SUMO is a time-discrete, space-continuous traffic simulation framework that can generate pedestrian, bike, or vehicle traffic over a 2D road network (Figure 1). SUMO generates randomized traffic flow that is then parsed to create realistic urban pedestrian routes.

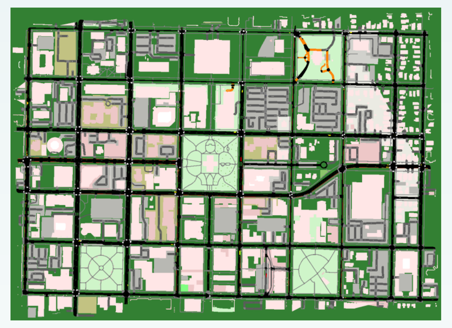


Figure 1: SUMO Road Network of Downtown Raleigh

Sionna
Sionna is a Python ray tracing library for calculating signal paths and coverage maps (Figure 2) with respect to building geometries and materials. Once initialized, Sionna is utilized to calculate path qualities between UAVs and GUs.

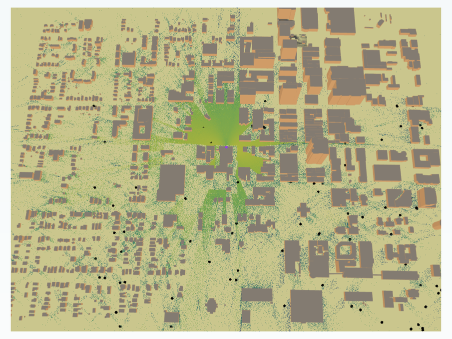


Figure 2: Sionna Coverage Map Visualization in Downtown Raleigh

Custom UAV Control Framework
This framework is used to control UAV flight, energy usage, and interface with Sionna methods. 3D cubic Bezier curves model UAV movement (Figure 3) and minimize power usage. Consumption considers movement, hovering, and transmitter power costs.

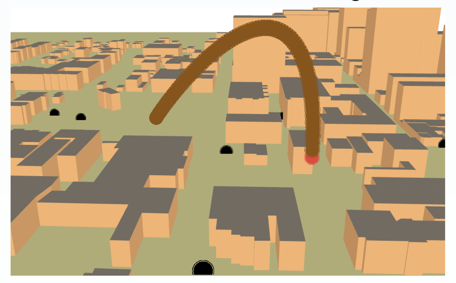


Figure 3: Bezier Curve Routing for High-G Maneuver in Sionna

Network Optimization & Google OR-Tools
To optimize GU-UAV assignments, we formulate an integer constraint problem and solve it using OR-Tools' CP-Solver. This algorithm allows compensations for load balancing, power minimization, and throughput maximization.

$$r_{max} = B\log_2(1+\frac{\alpha^2 P_t}{kTB})$$

Equation 1: Theoretical Maximum Data Rate formula based on path coefficients a, bandwidth B, temperature T, transmitter power P, and Boltzmann Constant k.

Results & Discussion

Simulation Evolution & User Assignment

To validate results, test data is generated for an environment in downtown Raleigh, North Carolina. UAV movement patterns are preset to disperse around the environment, and the pedestrian data is randomly generated to simulate typical traffic flow.

Figures 4-5 provide a comprehensive view of the simulation evolution:
- Top – Coverage map with pedestrians color-coded according to UAV assignment
- Far Left – The number of GUs assigned to each UAV
- Middle Left – Combined theoretical maximum throughput for each UAV
- Middle Right – Actual combined throughput for each UAV
- Far Right – UAV power consumption from flight and transmission

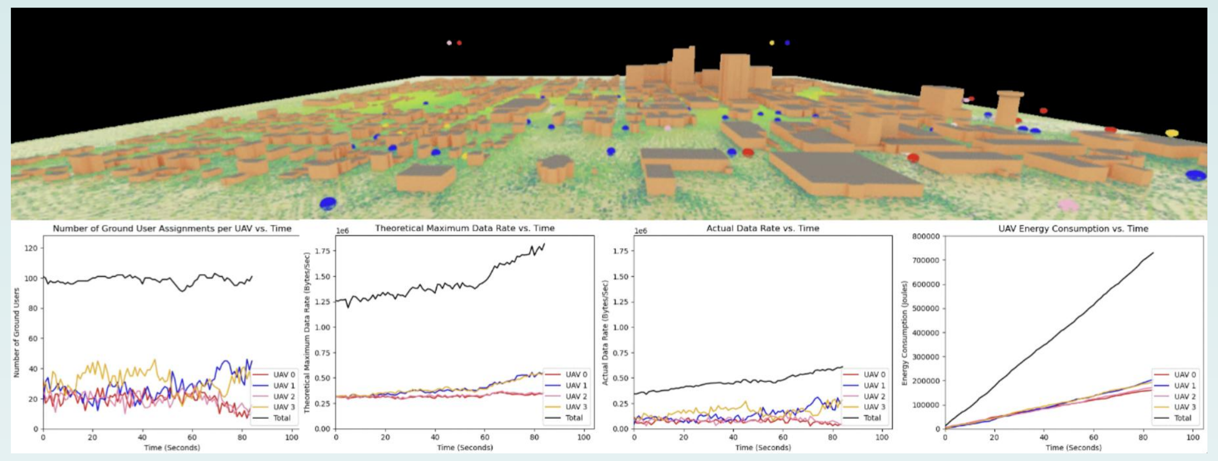


Figure 4: Digital Twin evolution with a throughput-focused Ground User assignment algorithm

The assignments in Figure 4 are optimized to maximize the actual data rate at each time step. This evolution shows that:
- As the UAVs become more dispersed, both the actual data rate and the theoretical data rates increase expectedly. However, the actual data rate is bounded by UAV data capacity and thus increases more modestly.
- There is a high variability in the GU assignments, which indicates a high degree of variability in the path qualities over position.

Figure 5 shows a similar evolution, but instead GU assignments are optimized with a bi-objective that considers both actual data rate and load-balancing by maximizing the minimum number of GUs assigned to any UAV.

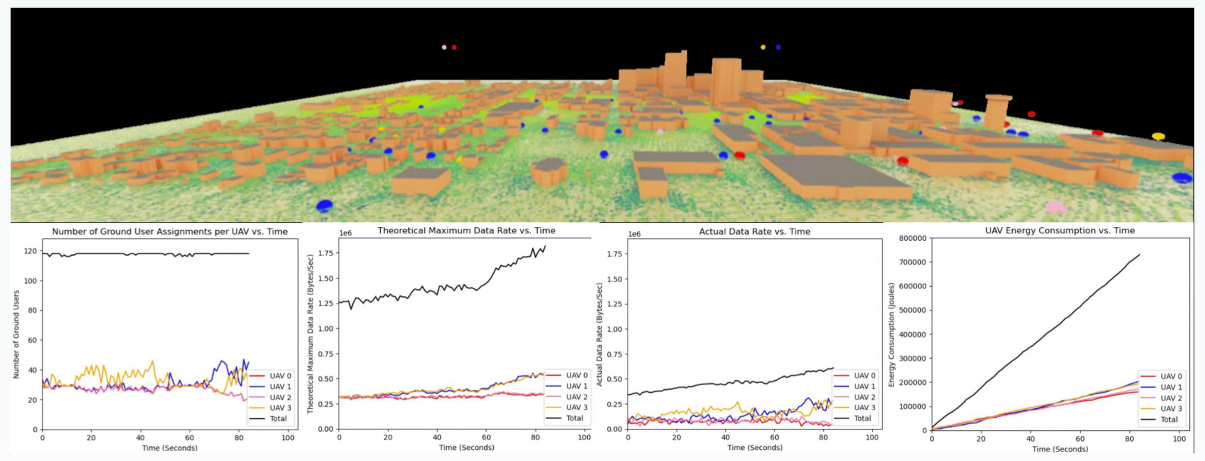


Figure 5: Digital Twin evolution with a throughput and load-balancing bi-objective assignment algorithm

Figure 5 shows that the load-balancing bi-objective increases the GU coverage to 100% in most instances while simultaneously limiting the variability in the number of GUs assigned to each UAV. Finally, this algorithm doesn't significantly affect the actual data rate progression, so this optimization increases coverage without sacrificing the data rate.

K-Means Proof of Concept

To measure the Digital Twin's machine learning potential, we use a basic K-Means algorithm that moves the UAVs to the cluster centers based on the positions of the ground users, minimizing the total horizontal distance between UAVs and GUs. Figures 6-7 compare the actual data rate of the K-Means approach with random UAV movement.

Figure 6: Actual Data Rates for Random Positioning

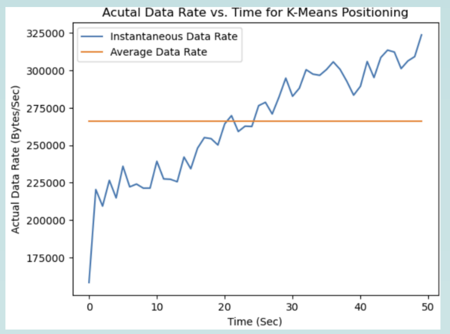


Figure 7: Actual Data Rates for K-Means Positioning

Figures 6-7 demonstrate that the K-Means positioning protocol both outperforms the random positioning on average and is less variable overall, which is better for guaranteeing signal quality for most GUs.

Conclusion

This poster highlights a comprehensive Digital Twin simulation for implementing RL models that maximize the communication output between UAVs and Ground Users. The simulation combines tools for traffic simulation, path quality computation with ray tracing, efficient UAV routing, and multiple algorithms for GU assignment. The demonstrations show the various data metrics available to a machine learning pipeline, including Ground User assignments, theoretical and actual data rate, and power consumption. These quantities form an invaluable data package for reinforcement learning in simulation. Furthermore, the demonstration of assignment optimizations shows that the simulation can assist reinforcement learning algorithms with the throughput maximization objective by performing deterministic lower dimensional advances. These algorithms also decrease the action space of the reinforcement learning model, helping it converge to accurate UAV movements faster. Finally, the K-Means proof-of-concept demonstration shows that this model can help a model converge to optimal flight patterns, which is promising for its future use in reinforcement learning.

Future Work

The next step for this Digital Twin is to train a UAV-position optimizing reinforcement learning model. A good starting point is a form of Q-Learning based on position with compensation to reduce the continuous state space to a discrete one. Additionally, algorithms such as Deep Deterministic Policy Gradient or Actor-Critic have likewise been used in previous research and are promising areas for future exploration.

Acknowledgements

Thank you to PhD candidate Jiayuan Huang for his invaluable assistance and passionate advice during the research process. I am also incredibly grateful to the head of NICE Lab, Dr. Yuchen Liu, for his crucial support and technical advice. Finally, I thank the NC State-based Aerial Experimentation and Research and Platform for Advanced Wireless (AERPAW) for providing references and technical assistance during the development of the Digital Twin.
Disclosure of Funding: This research is sponsored by an REU Award from the Department of Computer Science in the NC State College of Engineering. None of the authors have any financial incentive regarding the development of this research.

NC STATE UNIVERSITY

AERPAW
Aerial Experimentation and Research Platform for Advanced Wireless