Multifaceted UAV Fleet Automation via Asynchronous Deep RL

From Mechanized Farming to Hybrid RANs

Bharath Keshavamurthy and Nicolò Michelusi

School of Electrical, Computer and Energy Engineering

Arizona State University, Tempe, AZ

1 Abstract

We envision a future in which UAV-aided hybrid networks encompass every aspect of the modern communication infrastructure: from beyond LoS connectivity & traffic offloading via UAV relays in dense urban neighborhoods to broadband augmentation for mechanized farming & ranching in rural towns. To manifest this vision, we leverage the tools available at our disposal from the A.I. toolbox – namely, optimization, statistical decision making, Dynamic Programming (DP), and Reinforcement Learning (RL) – to tackle communication request scheduling, path planning, data & energy harvesting, and other problems constituent in the implementation of 3D HetNets - now conceived to be integral to the *Internet of Drones* ecosystem in 6G deployment architectures.

Core Research: Our preliminary research on these problems revolves around adaptive communication request scheduling and 2D path planning for a single rotary-wing UAV serving as a cellular BS relay for distributed GNs (UEs, IoT devices, Q-UGVs, etc), with a formulation that incorporates a multi-scale Semi-Markov Decision Process structure: outer decisions on UAV radial velocity (in waiting states) and end position (in communication states) optimize the average long-term delay-power trade-off; and consequently, inner decisions on angular velocity (in waiting states) and scheduling decisions & trajectories (via Hierarchical Competitive Swarm Optimization [HCSO] in communication states) greedily minimize instantaneous delay-power costs[1].

With a need to alleviate the associated computational complexity and to extend this structure to multifaceted autonomous UAV fleet management, our enhanced system model envelops prioritized request scheduling actions; deployment restrictions in terms of NFZ constraints imposed by the FAA and/or structural roadblocks; and drone service pit-stops for recharging and data upload/policy download. Employing an Asynchronous Advantage Actor Critic (A3C) framework – with a Deep Deterministic Policy Gradient (DDPG) training algorithm in the actor and a nested pair of neural networks trained via gradient descent in the critic – we derive the single-agent optimal waiting and communication state policies, which are then mapped to a distributed deployment of several UAVs through *conflict resolution* and *spread maximization* heuristics.

[1] M. Bliss and N. Michelusi, "Power-Constrained Trajectory Optimization for Wireless UAV Relays with Random Requests", IEEE ICC 2020

Subsequent improvements here include 3D trajectory optimization; heterogeneous fleets with both rotary- & fixed-wing UAVs with advanced power models that consider acceleration, directional mobility nuances, and vertical lift; and a DP approach for path planning instead of HCSO for improved computational feasibility.

AERPAW Implementations: Given the heterogeneity of radio frequency equipment available on the NSF AERPAW[2] testbed; UAV piloting provisions; and our experience working with cloud-provisioned SDRs & compute nodes on POWDER (OAI E2E SDR-based LTE network, OTA BRS srsLTE, OTA BRS GNURadio OFDM Tx/Rx, 28GHz V2X measurement campaign, DARPA SC2 CRN tests, and RENEW mMIMO experiments), we are optimally positioned to start testing our multifaceted autonomous UAV fleet orchestration algorithms on AERPAW.

Starting with the straightforward F2F1, F2F2, F2AM, and F2AM-PT scenarios[3] to get a feel for the testbed's capabilities, we plan to scale our experiments on AERPAW: first, we intend to deploy our trajectory optimization algorithms (HCSO, DP) via the F2AM-AT scenario vis-à-vis delay-power cost minimization; and next, integrate this path planning optimization with prioritized communication request scheduling (with commensurate rewards), data harvesting (from IoT-sensor/IoT-gateway PAWs), and drone service pit-stops (recharging and data upload/policy download) – in systems with multiple BSs, multiple GNs, and multiple UAVs – via the MFMM scenario.

Additionally, given the research overlap, we intend to collaborate with researchers from Purdue University's School of Electrical & Computer Engineering and College of Agriculture to augment their research on system-level coverage analyses for cellular networks with UAV data relays[4] through AERPAW implementations of mmWave network deployments for agricultural fleets (harvesters + UAV relays) and UAV-assisted data & energy harvesting for soil sensors.

[2] https://sites.google.com/ncsu.edu/aerpaw-wiki/aerpaw-user-manual

[3] https://sites.google.com/ncsu.edu/aerpaw-wiki/aerpaw-user-manual/3-experiment-structure-overview

[4] Y. Zhang et al., "Large-Scale Cellular Coverage Analyses for UAV Data Relay via Channel Modeling", IEEE ICC 2020

2 Appendix

Companion Research: Recognizing the role of the mmWave spectrum and of mMIMO technologies in next-generation wireless networks, we have stepped-up our efforts to understand optimal realizations of these systems. Firstly, through a 28GHz V2X measurement campaign[5] on the NSF PAWR POWDER[6] testbed in Salt Lake City, UT (equipped with a sliding-correlator channel sounder atop a fully autonomous robotic beam-steering platform), we have gathered ∼400GB of geographically diverse datasets which are being used – in addition to ray-tracing simulations on Wireless InSite[7] – to glean crucial results vis-à-vis mmWave signal propagation (attenuations, reflections, scattering, antenna gains, path loss, coverage, etc) in both urban and suburban radio environments. Secondly, our experiments on the RENEW[8] mMIMO equipment (Skylark/Faros BS with Iris SDRs and an aggregation hub) on the POWDER testbed – namely, a TDD framer, beacon beam-sweeping algorithms, and a sounder with throughput, beam-forming gains, and channel estimation evaluations – have facilitated extensions on beam-management in these systems via different deep learning models and codebook designs.

The electromagnetic midband has turned out to be an extremely valuable sweet-spot (*the goldilocks zone*) for 5G service providers: considering FCC regulations, market dynamics, and network & equipment design constraints, wireless carriers feel that this sub-6GHz spectrum strikes the right balance between data rates and propagation distance. Recognizing the need for dynamic spectrum sharing – owing to this spectrum being a coveted yet scare resource – some of our research has also focused on intelligent channel sensing & access[9]: E2E radio design for the DARPA Spectrum Collaboration Challenge (SC2) with an approximately optimal channel sensing & access strategy in the MAC layer derived using a randomized point-based POMDP formulation and its low-complexity approximations via fragmentation, Hamming distance state filters, and Monte-Carlo methods[10]. A subsequent development to these aspects of our research involves deploying our cognitive radio network on the POWDER testbed at the University of Utah (using USRP X310 and USRP B210 SDRs) with OTA communication in CBRS spectrum.

[5] B. Keshavamurthy, et al., "A Robotic Antenna Alignment and Tracking System for Millimeter Wave Propagation Modeling", USNC-URSI NRSM 2022

[6] https://powderwireless.net/

[7] https://www.remcom.com/wireless-propagation

[8] https://wiki.renew-wireless.org/en/home

[9] B. Keshavamurthy and N. Michelusi, "Learning-based Cognitive Radio Access via Randomized Point-Based Approximate POMDPs", IEEE ICC 2021

[10] B. Keshavamurthy and N. Michelusi, "Learning-based Spectrum Sensing and Access in Cognitive Radios via Approximate POMDPs", IEEE TCCN