Learning to Cover: Imitation Learning from a Geometric Expert for MANET Deployment

Edwin Meriaux   Shuo Wen   Antonio Loría   Gregory Dudek

***Abstract*—We present an analysis of geometric and data-driven approaches to Mobile Ad Hoc Network (MANET) deployment in unknown environments. Our geometric algorithm, CADENCE (Centralized Agent-Directed Exploration for Network Coverage), grounds the deployment problem in the Art Gallery Problem (AGP) and its cooperative extension (CGAGP), achieving provably correct coverage and connectivity by placing agents at reflex vertices of orthogonal environment polygons. Across 1,500 test environments of varying complexity, CADENCE achieves 100% coverage while Multi-Agent Reinforcement Learning (MARL) baselines trained from scratch achieve below 11%. We argue that this gap reveals two bottlenecks in current MARL methods: a behavioral one, as random exploration cannot discover the coordination structure the problem requires, and a representational one, as grid-based observations fail to capture the communication and coverage graph that CADENCE reasons over directly. To bridge these paradigms, we combine Graph Attention Networks (GATs) [1], [2] to encode multi-agent state as a communication graph with imitation learning from CADENCE demonstrations to provide a behavioral prior. Preliminary results confirm that the GAT-based pipeline successfully learns from CADENCE demonstrations. This positions geometric algorithms not as competitors to learning but as structured scaffolding that makes learning tractable on hard coordination problems.**

I. Introduction

Geometric methods have long provided robotics with provable guarantees, interpretable structure, and worst-case bounds that learning-based approaches cannot yet match. At the same time, data-driven methods promise generality beyond the assumptions that make geometric analysis tractable. This paper analyzes how geometric and data-driven algorithms fare in the setting of Mobile-Ad Hoc Networks (MANETs), and proposes a path forward based on ongoing research that can be generalized to various other robotics settings.

The Art Gallery Problem (AGP) [3] asks how many observers suffice to see every point of a polygon. The Cooperative Guard Art Gallery Problem (CGAGP) [4] extends the AGP by requiring that the placed guards also form a connected communication graph, with a known worst-case bound of $M_{\text{CGAGP}} = \frac{n+2h-4}{2}$ guards for orthogonal polygons with n corners and h holes. In prior work [5], we expanded CGAGP into an online setting called the Partially Observable Cooperative Guard Art Gallery Problem (POCGAGP), where agents must incrementally discover an initially unknown environment while maintaining coverage and connectivity throughout deployment. This presents a mathematically rooted representation of what MANETs need to do, as the visibility of guards can represent the communication coverage from a given node in the MANET. To solve this problem with orthogonal polygons, we introduced CADENCE (Centralized Agent-Directed Exploration for Network Coverage), a geometric deployment algorithm that provably solves POCGAGP within the CGAGP worst-case agent bound. Its key insight is that, in orthogonal polygons, placing observers at a subset of reflexive (270) vertices guarantees full visibility, a result derived from visibility graph analysis, and the number of such vertices matches the CGAGP bound. A sample deployment illustrating agent coverage and connectivity is shown in Fig. 1.

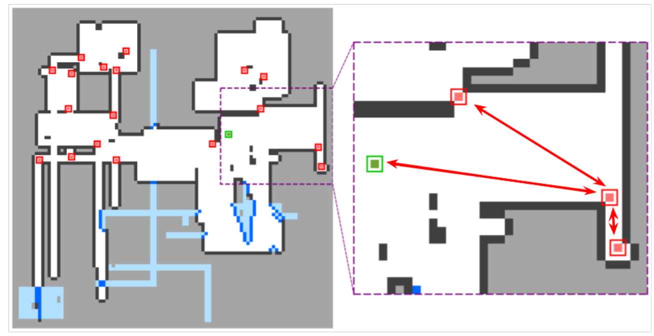

Fig. 1. Sample deployment of agents in a given world on the left. Areas covered by agents are shown in white, with the boundary of vision in dark blue and unseen portions in light blue. A zoomed-in view of the connectivity between agents is shown on the right.

CADENCE achieves 100% coverage across 1,500 test environments of varying complexity. Multi-agent reinforcement learning (MARL) baselines trained from scratch on the same benchmark achieve below 11%. This gap is striking, but we argue it is not a verdict against learning: it is evidence that geometric structure is the missing behavioral prior that makes learning tractable on hard coordination problems. MARL policy training is unable to replicate the understanding of spatial geometry required to solve problems like POCGAGP, along with AGP and CGAGP. CADENCE specifically takes into account the geometry of the space in how agents move, making this knowledge encoded by design.

We propose imitation learning from a geometric expert along with Graph Attention Networks (GATs) [1], [2] as the bridge between these paradigms. CADENCE demonstrations provide the behavioral initialization that MARL currently

Edwin Meriaux, Shuo Wen and Gregory Dudek are affiliated with the McGill University Computer Science Department and MILA Institute. (e-mail: edwin.meriaux@mail.mcgill.ca)

Edwin Meriaux is also affiliated with L2S at the University Paris-Saclay, CentraleSupélec.

Antonio Loría is affiliated with CNRS L2S.

lacks, while learned policies can in turn relax the orthogonality assumption, reduce agent counts below the worst-case bound, and generalize to environments where the geometric guarantees no longer hold. Rather than positioning geometry and learning as competitors, this work treats geometric algorithms as scaffolding that makes learning possible and learning as the mechanism that extends geometric insight beyond its formal domain of validity.

Let us first begin by presenting our geometric-based algorithm followed by our data-driven approach.

II. CADENCE

CADENCE (Centralized Agent-Directed Exploration for Network Coverage) is a centralized deployment algorithm that provably solves POCGAGP within the CGAGP worst-case agent bound for orthogonal polygons.

The algorithm is built on two concepts: visibility graphs and valid corners. In an orthogonal polygon, every interior angle is either 90 or 270. The 270 vertices, called reflex vertices, govern the geometry of visibility. It has been shown that placing guards at all reflex vertices produces a visibility graph in which the shortest path between any two mutually non-visible points passes through a sequence of reflex vertices [**?**], [**?**]. This means that guards placed at reflex corners alone simultaneously solve both the AGP and the CGAGP: every point in the polygon is seen by at least one guard, and the visibility graph connecting those guards is connected.

However, not all reflex vertices are necessary. CADENCE refines the set of reflex vertices into *valid corners*: all 270 vertices along the outer boundary, plus all 270 vertices on holes except one per hole. The excluded vertex, the top-left corner of each hole, is omitted to maintain coverage and connectivity in the world. This exclusion reduces the worst-case agent count from $M_{\text{refl}} = \frac{n+4h-4}{2}$ (the total number of reflex vertices) to

$$M_{\text{valid}} = \frac{n + 2h - 4}{2}, \tag{1}$$

matching the CGAGP bound exactly.

To solve POCGAGP, CADENCE operates incrementally. It maintains a frontier of valid corners that have been observed but not yet occupied, and dispatches agents to these corners sequentially, preserving connectivity at every deployment step. This design illustrates both the strengths and limitations of a purely geometric approach. On one hand, CADENCE provably matches the CGAGP worst-case bound and guarantees that no agent ever loses communication or coverage, handling the complexities of line-of-sight connectivity by construction. On the other hand, the algorithm is conservative: it deploys an agent to every valid corner it discovers, and in many environments the number of valid corners far exceeds the minimum needed for full coverage. Naive post-hoc deallocation of redundant agents can partially mitigate this over-deployment, but only to a limited extent [5].

In our previous work, we evaluated CADENCE and competing methods by benchmarking on 1,500 orthogonal dungeon environments spanning three size classes (50×50, 100×100, and 250×250 cells) and 60 complexity levels measured by quadtree decomposition node count. This complexity measure captures obstacle density and geometric fragmentation more faithfully than raw world size: a 100×100 dungeon can be structurally more complex than a 250×250 one, and the quadtree rank reflects this.

III. MARL Failure and What It Reveals

We evaluate two MARL algorithms trained from scratch on POCGAGP. IQL (Independent Q-Learning [6]) is a fully decentralized method in which each agent trains an independent Deep Q-Network using local rewards. VDN (Value Decomposition Networks [7]) is a centrally trained, decentrally executed method that decomposes the joint action-value function additively, enabling implicit coordination during training. Both methods receive a reward that encourages coverage, penalizes disconnection, and penalizes collisions. Observations are stacked 2D grid frames encoding each agent's field of view, processed by convolutional layers followed by fully connected value heads.

The results across 1,500 test environments are unambiguous: IQL achieves 10.6% and VDN achieves 10.7% mean coverage [5]. A characteristic failure mode is agent stalling at the deployment point, blocking all subsequent spawns, resulting in near-zero coverage. Even in isolated single-environment training, both methods solve POCGAGP only on instances requiring no more than five to ten agents. Both methods were also limited to 50×50 worlds; 100×100 environments were computationally infeasible at the network sizes evaluated.

This failure is informative in a precise sense. The problem is that random exploration cannot discover the coordination structure POCGAGP requires: expand outward, maintain a connected frontier, and target reflex corners. This structure is not apparent from local reward signals alone. It is, however, exactly what CADENCE encodes as a consequence of its geometric design.

IV. Toward Graph-Based Imitation Learning

The comparison above identifies two distinct bottlenecks for data-driven models such as MARL. The first is *behavioral*: MARL lacks a prior for coverage, and random exploration cannot discover the coordination structure that POCGAGP requires. The second is *representational*: the 2D grid observations used by IQL and VDN encode raw spatial data but fail to capture the communication and coverage structure that is central to the problem. CADENCE succeeds precisely because it reasons about this structure directly, operating on visibility graphs and connectivity constraints rather than pixel-level observations. A learning-based approach must do the same while learning to outperform CADENCE.

We address both bottlenecks simultaneously through the hybrid pipeline illustrated in Fig. 2. For representation, we introduce a GAT [1], [2] that encodes the multi-agent state as a graph: agents are nodes, communication and visibility links are edges, and coverage information is captured as node or edge

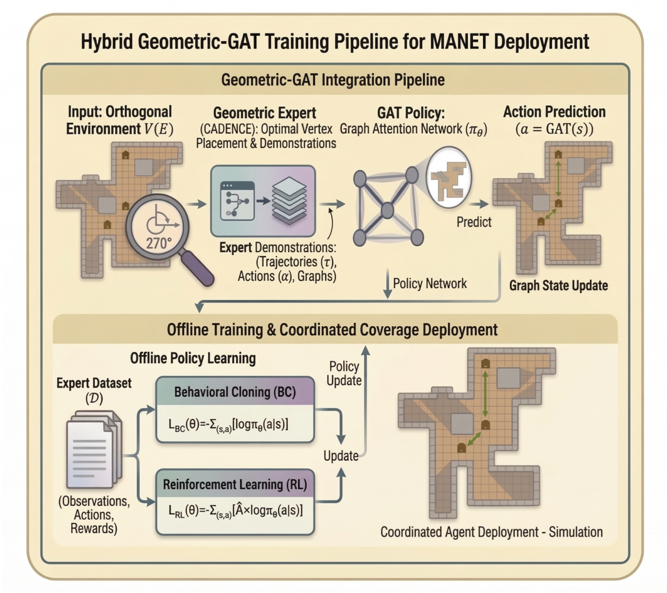

Fig. 2. Proposed hybrid Geometric-GAT training pipeline for MANET deployment. CADENCE generates expert demonstrations from orthogonal environments, which are used for offline policy learning combining behavioral cloning and reinforcement learning losses to train a GAT-based policy network.

features. This formulation gives the policy direct access to the relational structure that grid observations obscure, namely which agents can see each other, which regions are covered, and how the communication graph is connected. Visual observations from individual agents can additionally be embedded as node features, allowing the policy to incorporate local perceptual information alongside the global graph structure. This architecture is designed to work in 2D and be expandable to 3D spaces.

For behavior, we use imitation learning from CADENCE. Given an orthogonal environment $V(E)$ as input, CADENCE produces optimal vertex placements at reflex corners, yielding expert demonstrations consisting of trajectories τ^*, actions a, and the corresponding communication graphs. The GAT policy network π_θ consumes these graph-structured states and produces action predictions $a = \text{GAT}(s)$, with the graph state updating after each deployment step to reflect newly occupied positions and expanded visibility. Demonstration trajectories generated across a large distribution of orthogonal environments are then used for offline policy learning over the collected expert dataset $\mathcal{D}$. Training combines a behavioral cloning loss

$$L_{\text{BC}}(\theta) = -\mathbb{E}_{(s,a)\sim\mathcal{D}}\left[\log \pi_\theta(a|s)\right] \tag{2}$$

which directly supervises the policy to reproduce CADENCE's placement decisions, with a reinforcement learning loss

$$L_{\text{RL}}(\theta) = -\mathbb{E}_{(s,a)\sim\mathcal{D}}\left[A \log \pi_\theta(a|s)\right] \tag{3}$$

weighted by advantage A, which allows the policy to improve beyond the expert by discovering solutions that use fewer agents or shorter deployment paths than the worst-case bound requires. The behavioral cloning objective ensures that the resulting policies inherit CADENCE's coverage-preserving behavior as a prior, while the reinforcement learning objective frees them to deviate where improvements exist. In environments where fewer agents suffice or where shorter deployment paths exist, a learned policy can discover these improvements. The CADENCE demonstrations provide a floor, not a ceiling.

Preliminary results confirm that the GAT-based training pipeline successfully learns from CADENCE demonstrations, validating the feasibility of this hybrid approach.

This combination is qualitatively different from deploying agents with CADENCE directly. A GAT-based policy is not restricted to 270 corners, is not centralized, and is not bound to M_{valid}. The graph representation generalizes naturally to non-orthogonal and 3D environments, where grid-based observations become impractical but the communication graph retains its structure. Furthermore, the graph formulation is not tied to square grid cells or regular voxels. By discretizing the environment into arbitrary convex cells rather than axis-aligned grids, the same architecture can operate over irregular tessellations that conform more naturally to curved or non-orthogonal boundaries, reducing the representational mismatch between the environment geometry and the agent's state space. This represents a further advancement from CADENCE, whose geometric reasoning is rooted in axis-aligned grid discretizations.

V. Conclusion

CADENCE demonstrates that geometric reasoning produces reliable, provably correct multi-robot deployment policies in the POCGAGP setting, while MARL methods trained from scratch fail by a wide margin on the same benchmark. The gap is not a verdict against learning: it is evidence that geometric structure is the missing ingredient that makes learning tractable on this problem. The hybrid approach we propose uses geometry to solve the learning problem, providing the behavioral initialization and sample efficiency that MARL currently lacks, and uses learning to solve the geometry problem, relaxing the orthogonality assumption and improving beyond the worst-case bound. Preliminary results with the GAT-based pipeline validate this direction. Geometric methods are not a historical baseline to be displaced by data-driven approaches. They are a source of structured demonstrations, architectural priors, and correctness certificates that remain indispensable precisely where learning struggles most.

References

[1] S. Brody, U. Alon, and E. Yahav, "How attentive are graph attention networks?" in *International Conference on Learning Representations (ICLR)*, 2022.

[2] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *International Conference on Learning Representations (ICLR)*, 2018.

[3] J. O'rourke, *Art gallery theorems and algorithms*. Oxford University Press, Inc., 1987.

[4] P. Żyliński, "Cooperative guards: combinatorial bounds and complexity," Ph.D. dissertation, University of Gdańsk, Gdańsk, Poland, 2004.

[5] E. Meriaux, S. Wen, L.-R. Langevin, D. Precup, A. Loría, and G. Dudek, "On mobile ad hoc networks for coverage of partially observable worlds," *arXiv preprint arXiv:2512.09495*, 2025.
[6] A. Tampuu, T. Matiisen, D. Kodelja, I. Kuzovkin, K. Korjus, J. Aru, J. Aru, and R. Vicente, "Multiagent cooperation and competition with deep reinforcement learning," *PloS one*, vol. 12, no. 4, p. e0172395, 2017.
[7] P. Sunehag, G. Lever, A. Gruslys, W. M. Czarnecki, V. Zambaldi, M. Jaderberg, M. Lanctot, N. Sonnerat, J. Z. Leibo, K. Tuyls *et al.*, "Value-decomposition networks for cooperative multi-agent learning," *arXiv preprint arXiv:1706.05296*, 2017.