International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(Research Article)

SEEK-Multi: Collaborative Multi-Agent Semantic Reasoning for Object Goal Navigation in Inspection Tasks

Adele Chinda *

Computer Science, Georgia State University, USA.

International Journal of Science and Research Archive, 2026, 18(01), 979-997

Publication history: Received on 21 December 2025; revised on 28 January 2026; accepted on 31 January 2026

Article DOI: https://doi.org/10.30574/ijsra.2026.18.1.0181

Abstract

We address the fundamental challenge of collaborative multi-agent object-goal navigation for autonomous inspections in complex, real-world environments. While single- agent approaches to object-goal navigation have demonstrated considerable promise in recent years, scaling these methods to larger environments necessitates the coordination of multiple robots to achieve efficient coverage, faster task completion, and robust operation under uncertainty. We introduce SEEK-Multi, a comprehensive framework that extends semantic-guided object inspection to multi-robot systems through distributed belief sharing, collaborative planning, coordinated task allocation, and adaptive communication protocols. SEEK-Multi enables multiple agents to share semantic understanding and inspection findings through a distributed Relational Semantic Network (RSN) and a shared Dynamic Scene Graph (DSG), maintaining consistency across the team while accommodating communication constraints. We propose novel algorithms for collaborative exploration that leverage semantic priors, belief fusion using consensus protocols with provable convergence guarantees, and conflict-free task allocation based on auction mechanisms. Our extensive simulation analyses across diverse environment configurations demonstrate that SEEK-Multi achieves significant speedup over single-agent approaches while maintaining high success rates, with near-linear scaling efficiency for up to four agents and graceful degradation under communication failures. We validate our approach through comprehensive simulations including ablation studies, sensitivity analyses, and comparisons with state-of-the-art multi-agent coordination methods, demonstrating its practicality for real-world multi-robot inspection scenarios in industrial, search-and-rescue, and domestic environments. Code is available at: https://arrdel.github.io/seek-multi/

Keywords: Multi-agent navigation; Semantic-guided search; Distributed coordination; Belief fusion; Task allocation; Object-goal inspection

1. Introduction

Consider a team of autonomous robots tasked with searching for and inspecting target objects across a large industrial facility. While a single robot can methodically search each area, the task completion time scales linearly with the environment size, making single-agent solutions impractical for time-critical applications. Deploying multiple robots offers the potential for significant speedup, but realizing this potential requires sophisticated coordination to avoid redundant effort and conflicting actions. This multi-agent object-goal navigation problem is crucial for time-sensitive applications such as emergency response, security patrols, industrial inspection, and search-and-rescue operations [53, 6, 41].

The deployment of multi-robot systems for inspection tasks has gained significant attention in recent years, driven by advances in sensing, communication, and computation [44, 43].

* Corresponding author: Adele Chinda

Industries ranging from manufacturing to energy production increasingly rely on autonomous inspection to reduce costs, improve safety, and enable continuous monitoring [29]. However, the transition from single robot to multi-robot inspection introduces fundamental challenges in coordination, communication, and decision-making that require novel algorithmic solutions.

Multi-agent coordination for object-goal navigation presents unique challenges beyond the single-agent case. First, agents must efficiently partition the search space to minimize overlap while ensuring complete coverage, a problem that becomes increasingly complex as the number of agents and environment size grow. Second, agents must share observations and update their beliefs about object locations in a consistent manner, even with limited communication bandwidth and intermittent connectivity. Third, the planning framework must account for the actions and intentions of all agents to avoid conflicts and maximize team efficiency, requiring coordination mechanisms that scale gracefully. Fourth, the system must be robust to communication failures, agent heterogeneity, and dynamic environmental changes. Fifth, the framework must balance the benefits of coordination against its computational and communication costs, enabling operation in resource-constrained scenarios.

Recent work on single-agent semantic-guided navigation [25, 12, 48, 11] has demonstrated the value of incorporating prior knowledge and semantic reasoning into object search. The SEEK framework [25] introduced the Relational Semantic Network (RSN) for encoding object-room relationships and showed significant improvements over geometric coverage approaches. By leveraging semantic priors about where objects are likely to be found, SEEK enables efficient, informed search that outperforms uninformed exploration strategies. However, extending these methods to multi-agent settings requires addressing the challenges of distributed belief maintenance, coordinated planning, and efficient communication while preserving the semantic reasoning capabilities that make single agent approaches effective.

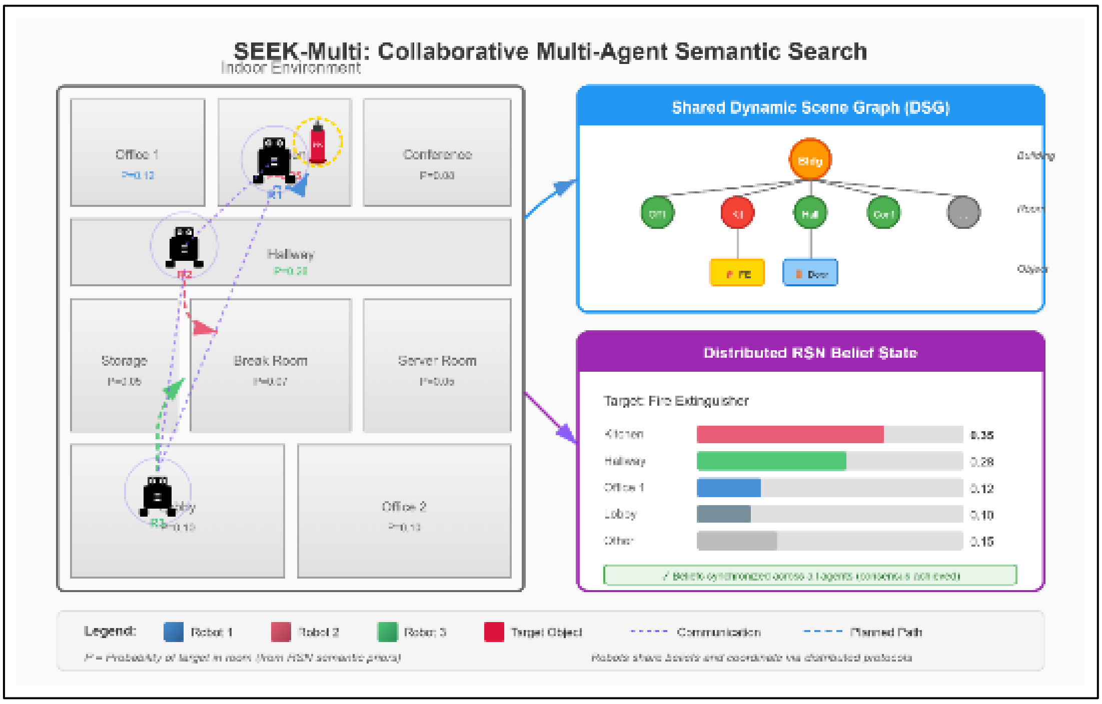

Figure 1 SEEK-Multi enables multiple robots to collaboratively search for target objects by sharing semantic beliefs and coordinating their search strategies. Each agent maintains a local copy of the shared Dynamic Scene Graph (DSG) and Relational Semantic Network (RSN), with updates propagated through a distributed communication protocol. The framework supports both centralized and decentralized coordination modes

The multi-agent extension of semantic navigation raises several research questions that motivate our work: How can semantic beliefs be efficiently shared and fused across multiple agents with potentially different observations? How should task allocation incorporate semantic priors while respecting coordination constraints?

What communication protocols best balance information sharing with bandwidth limitations? How can the system maintain performance under communication failures and agent heterogeneity?

In this paper, we propose SEEK-Multi, a comprehensive framework for collaborative multi-agent object-goal navigation that extends the SEEK architecture to multi-robot teams. Our approach maintains distributed copies of the Dynamic Scene Graph (DSG) and Relational Semantic Network (RSN) across all agents, with efficient protocols for sharing updates and fusing beliefs. We introduce a collaborative planning algorithm that computes coordinated task assignments while accounting for agent positions, capabilities, and intentions. The framework supports both centralized and decentralized coordination modes, enabling deployment in various communication scenarios from reliable infrastructure networks to ad-hoc peer- to-peer connectivity.

The design of SEEK-Multi is guided by several key principles: **Semantic awareness**: All coordination mechanisms leverage semantic understanding to improve efficiency. **Scalability**: Algorithms and communication protocols scale grace- fully with the number of agents. **Robustness**: The system degrades gracefully under communication failures and agent heterogeneity. **Flexibility**: The framework supports various coordination modes and can adapt to different deployment scenarios.

Our key contributions are:

- We introduce SEEK-Multi, a comprehensive framework for collaborative multi-agent object-goal navigation using distributed semantic reasoning, supporting both centralized and decentralized coordination.
- We propose a distributed belief fusion algorithm based on consensus protocols that enables agents to share and combine observations efficiently with provable convergence guarantees.
- We design a collaborative planning algorithm that coordinates task allocation and path planning across multiple agents using auction-based mechanisms with semantic- aware bidding.
- We develop an adaptive communication protocol that balances information sharing with bandwidth constraints and provides robustness to message loss.
- We demonstrate through extensive simulation experiments that SEEK-Multi achieves near-linear speedup with multiple agents while maintaining high success rates across diverse environment configurations.

The remainder of this paper is organized as follows. Section 2 reviews related work in multi-robot coordination, se- mantic navigation, and distributed belief maintenance. Section 3 formally defines the multi-agent object-goal navigation problem. Section 4 presents the SEEK-Multi architecture, including distributed semantic representations, belief fusion, and collaborative planning. Section 5 provides theoretical analysis of convergence and speedup properties. Section 6 presents experimental results from simulation studies. Section 7 discusses scalability, heterogeneity, and limitations. Section 8 concludes with directions for future work.

2. Related work

Our work builds upon and integrates advances from several research areas: multi-robot coordination, multi-agent path planning, distributed belief maintenance, semantic navigation, and scene understanding. We review each area and position our contributions relative to existing work.

Multi-Robot Coordination: Multi-robot systems have been extensively studied for tasks including exploration [8, 60], coverage [14, 23], and search and rescue [37, 32]. Coordination strategies range from centralized approaches with a single decision-maker [24] to fully decentralized methods using local communication [38, 17]. Market-based approaches [20, 63] provide a middle ground, using auction mechanisms for task allocation while maintaining scalability. Recent work has explored learning-based coordination [21, 35], where agents learn coordination strategies through reinforcement learning.

The choice of coordination architecture significantly im- pacts system properties. Centralized approaches can achieve optimal coordination but require reliable communication to a central node and create a single point of failure [39]. Decentralized approaches offer robustness and scalability but may sacrifice optimality due to limited global information [61]. Hybrid approaches attempt to balance these tradeoffs by combining local decision-making with occasional global coordination [63]. Our work supports multiple coordination modes, allowing deployment in various scenarios.

Multi-Agent Exploration and Search: Coordinated exploration has been studied extensively, with frontier-based methods [60] forming a foundation for many approaches. Multi-robot extensions assign frontiers to robots based on distance, information gain, or other criteria [8, 50].

Recent work has incorporated semantic information into exploration [40], using object recognition to guide search toward promising areas.

Search tasks differ from exploration in that they seek specific targets rather than complete coverage. Multi-robot search has been studied for static targets [27], moving targets [16], and adversarial scenarios [58]. Probabilistic approaches maintain belief distributions over target locations and plan searches to maximize detection probability [7, 33]. Our work extends these concepts by incorporating semantic priors that capture object-room relationships.

Multi-Agent Path Planning: Coordinated path planning for multiple agents must balance efficiency with collision avoidance [51, 34]. Approaches include coupled planning that considers all agents jointly [49, 59], prioritized planning that plans sequentially [56, 9], and velocity-obstacle methods for dynamic environments [57, 4]. Conflict-based search (CBS) [49] has emerged as an efficient approach for optimal multi-agent path finding, using a two-level search that resolves conflicts lazily.

For continuous domains and longer time horizons, approaches based on potential fields [54], model predictive control [36], and learned policies [47] have shown promise. Our framework uses intention sharing and conflict resolution to enable efficient distributed planning without requiring coupled optimization over all agents.

Distributed Belief Maintenance: Maintaining consistent beliefs across multiple agents is fundamental to multi-robot perception [19, 3]. Consensus algorithms [38, 42] provide a principled approach to fusing estimates from multiple agents, with well-understood convergence properties. Distributed simultaneous localization and mapping (SLAM) [18, 15] addresses the related problem of building shared maps from distributed observations.

Belief fusion must account for correlations between agent observations to avoid overconfidence [31]. Covariance intersection [30] provides conservative fusion when correlations are unknown, while channel filters [13] track information flow to avoid double-counting. Our belief fusion approach uses weighted consensus with confidence tracking to balance these concerns in a computationally efficient manner.

Semantic Navigation and Scene Understanding: Recent advances in semantic navigation leverage foundation models for improved reasoning [12, 48, 62, 22]. These approaches use vision-language models to understand scene semantics and guide navigation toward likely target locations. The SEEK framework [25] demonstrated the value of encoding object-room relationships in a Relational Semantic Network, achieving significant improvements over geometric coverage approaches.

Scene graphs provide structured representations of environments that capture objects, rooms, and their relationships [45, 28, 2]. Dynamic Scene Graphs (DSG) [46] extend this to include temporal information and support real-time updates during robot operation. Multi-robot scene graph construction has been explored for collaborative mapping [10, 55], but integration with semantic search remains limited. Our work extends SEEK to multi-agent settings with distributed belief sharing, collaborative planning, and efficient communication protocols that maintain semantic scene graph consistency.

Communication in Multi-Robot Systems: Communication is fundamental to multi-robot coordination, with significant research on protocols, bandwidth management, and robustness [61, 41]. Approaches range from continuous communication assuming reliable infrastructure [8] to intermittent communication in bandwidth-limited scenarios [26]. Learning- based methods have explored communication protocol optimization [21, 52], allowing agents to learn what information to share.

Our communication protocol balances information sharing with bandwidth constraints, prioritizing high-value updates while maintaining robustness to message loss through redundancy and acknowledgment mechanisms.

We formally define the multi-agent object-goal navigation problem, including the environment model, agent capabilities, communication constraints, and performance metrics.

3. Problem formulation

We formally define the multi-agent object-goal navigation problem, including the environment model, agent capabilities, communication constraints, and performance metrics. We consider a structured indoor environment represented as a topological-metric map. The environment is partitioned into a set of rooms $V = \{v_1, \ldots, v_m\}$ connected by traversable edges $E \subseteq V \times V$.

Each room v_j has associated attributes including:

- Semantic type $\ell(v_j) \in L$ (e.g., office, kitchen, hallway)
- Geometric extent defining the room boundary
- Set of contained objects $O(v_j)$
- Search time $\tau(v_j)$ required for thorough inspection

3.1. Multi-agent object-goal navigation

We consider a team of N robots $R = \{r_1, \ldots, r_n\}$ operating in the environment. Each robot r_i has state $x_i \in X$, takes actions $u_i \in U$, and receives observations $z_i \in Z$.

- State space $X = SE(2) \times V$ includes the robot's pose and current room
- Action space U includes navigation and inspection actions
- Observation space Z includes object detections with confidence scores

The objective is to locate a target object yG of class y_G^l as quickly as possible.

3.2. Observation model

Each robot has a detection sensor with the following characteristics:

- Detection range d_det: maximum distance at which objects can be detected
- True positive rate p_tp: probability of detecting the target when in range
- False positive rate p_fp: probability of false detection per observation
- Position noise σ_p: uncertainty in detected object position

When robot r_i is in room v containing the target object and performs a search action, the observation model is: $P(z_i = detect \mid yG \in v) = p\_tp \cdot \mathbb{1}[in\ detection\ range]$

3.3. Semantic prior model

We leverage semantic priors about object-room relationships, captured in a Relational Semantic Network (RSN). The RSN encodes the conditional probability of finding an object class in a room type:

$$P(y\hat{}l \in \ell) = RSN(y\hat{}l, \ell)$$

For example, $P(fire\ extinguisher \in kitchen)$ captures the prior knowledge that fire extinguishers are commonly found in kitchens.

3.4. Performance metrics

These metrics capture both the success rate and the efficiency of the team's search. The denominator uses the sum of path lengths to account for the total resources consumed by the team.

We also measure:

- Success rate (SR): Fraction of trials where target is found within time limit
- Time to completion (TTC): Time steps until target is found
- Coverage overlaps: Fraction of rooms searched by multiple agents
- Communication efficiency: Messages per successful search

3.5. Communication model

Robots communicate through message passing with the following constraints:

- **RANGE**: Robot r_i can communicate with $r_j\ if\ ||x_i - x_j|| \leq d\_comm$
- **BANDWIDTH**: Maximum B messages per time step per robot
- **LATENCY**: Messages arrive with delay δ
- **RELIABILITY**: Messages are delivered with probability $1 - p\_loss$

We define several message types with associated priorities:

- Object detections (priority 3): Location and confidence of detected objects
- Belief updates (priority 2): Probability distributions over rooms
- Intentions (priority 2): Planned actions for coordination
- DSG updates (priority 1): Changes to the shared scene graph
- Heartbeats (priority 0): Status and position updates

4. Seek-multi architecture

SEEK-Multi extends the single-agent SEEK architecture to multi-robot teams through three key components: distributed semantic representations, collaborative planning, and a communication protocol for belief sharing.

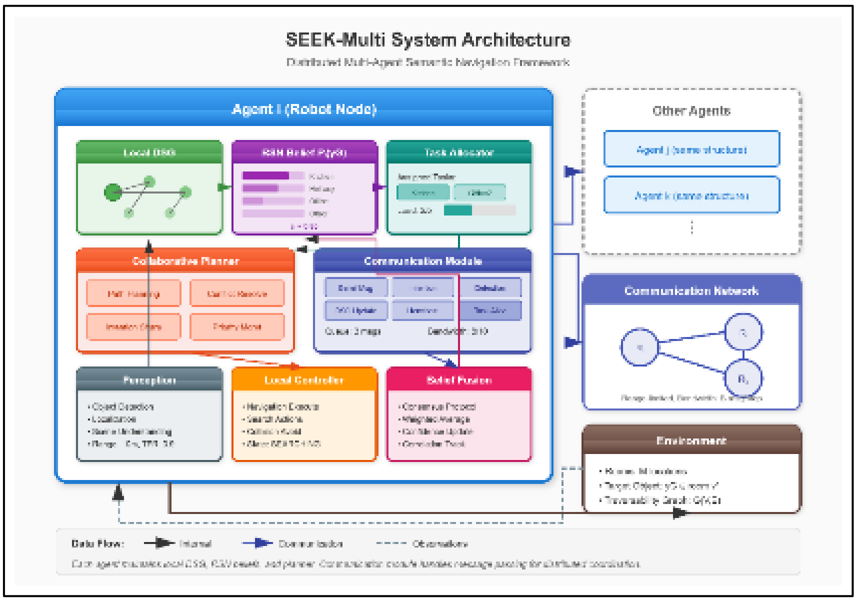

Figure 2 SEEK-Multi system architecture. Each agent maintains local copies of the DSG and RSN belief states, which are updated through onboard observations and inter-agent communication. A collaborative planner coordinates task allocation, while the communication module manages information sharing

4.1. System overview

Each robot in SEEK-Multi maintains the following components:

- Local Relational Semantic Network (RSN): Encodes semantic priors about object-room relationships, initialized from a common knowledge base and updated based on observations
- Local Dynamic Scene Graph (DSG): Captures the spatial and semantic structure of explored regions, including rooms, objects, and their relationships
- Belief State: Probability distribution over possible object locations, combining prior knowledge with observations
- Intention Buffer: Planned actions shared with teammates to enable coordination
- Communication Module: Handles message passing with neighboring robots

4.2. Distributed semantic representations

SEEK-Multi uses two complementary representations for distributed semantic understanding: the Relational Semantic Network (RSN) and the Dynamic Scene Graph (DSG).

RSN Structure:

The RSN for each agent is a bipartite graph $G_{RSN} = (O \cup L, E_{RSN})$ where:

- O is the set of object classes
- L is the set of room types
- E_RSN connects object classes to room types with weighted edges

Each edge weight $w(o, l)$ represents $P(object\ class\ o\ found\ in\ room\ type\ l)$.

Dynamic Scene Graph:

Each agent maintains a local DSG that captures spatial and semantic structure:

- Nodes represent rooms, objects, and agents
- Edges represent spatial relationships (containment, adjacency, connectivity)
- Node attributes include semantic type, position, confidence, and observation history

4.3. Collaborative planning

SEEK-Multi uses a hierarchical planning approach that combines semantic-aware task allocation with coordinated path planning.

Task Allocation via Semantic Auction:

Rooms are allocated to agents using an auction mechanism where bids incorporate semantic information:

$$bid\_i(v) = \alpha \cdot b\_i(v) - \beta \cdot d(x\_i, v) + \gamma \cdot I(v) - \delta \cdot C\_i(v)$$

Where:

- $b_i(v)$ is agent i's belief that target is in room v
- $d(x_i, v)$ is the distance from agent i's position to room v
- $I(v)$ is the information gain from searching room v
- $C_i(v)$ is the coordination cost (conflict with other agents' plans)
- $\alpha, \beta, \gamma, \delta$ are weighting parameters

Path Planning with Conflict Avoidance:

After task allocation, agents plan paths to their assigned rooms using a priority-based approach:

- Agents are assigned priorities based on task urgency and distance
- Higher-priority agents plan first; lower-priority agents treat their paths as obstacles
- Velocity obstacles are used for dynamic collision avoidance during execution

5. Distributed belief fusion

We present our distributed belief fusion algorithm that enables agents to maintain consistent beliefs about object locations while operating with limited communication.

5.1. CONSENSUS PROTOCOL

The belief fusion uses a weighted consensus protocol. At each time step, agent i:

- Performs local Bayesian update based on its observation
- Exchanges beliefs with neighbors within communication range
- Computes weighted average of received beliefs

The weight matrix W is designed to ensure convergence: $W = I - \varepsilon \cdot L$

where L is the graph Laplacian and ε is chosen such that W is doubly stochastic.

5.2. Convergence analysis

Under mild connectivity assumptions, the consensus protocol converges to a common belief:

THEOREM 1 (Convergence): If the communication graph is connected, then the beliefs of all agents converge to a common value:

$$lim(t \to \infty)\, b_i(v,t) \;=\; b*(v) \;\; for\ all\ agents\ i$$

The convergence rate depends on the second-largest eigenvalue of W:

$$||b(t) \;-\; b*|| \;\leq\; \lambda_2(W)\char`^t \;\cdot\; ||b(0) \;-\; b*||$$

5.3. Handling communication failures

The protocol handles intermittent communication through:

- Storing pending updates when neighbors are unreachable
- Timestamping beliefs to handle out-of-order messages
- Using exponential backoff for retransmission
- Maintaining local operation capability when isolated

---

6. Collaborative exploration

We describe how SEEK-Multi coordinates exploration across multiple agents to efficiently search for the target object.

6.1. Semantic-guided frontier selection

Each agent maintains a frontier of unexplored regions. Frontier selection incorporates:

- Semantic prior: Probability of finding target based on room types
- Distance cost: Travel time to reach the frontier
- Information gain: Expected reduction in belief uncertainty
- Coordination: Avoiding overlap with other agents' planned actions

The utility of frontier f for agent i is:

$$U\_i(f) \;=\; w\_sem \;\cdot\; P(yG \in\; f) \;+\; w\_info \;\cdot\; H(b\_i \mid explore\; f) \;-\; w\_dist \;\cdot\; d(x\_i, f) \;-\; w\_coord \;\cdot\; overlap(f, plans\_\{-i\})$$

6.2. Distributed task allocation

Rooms are allocated using a distributed auction mechanism:

- ANNOUNCEMENT: Unassigned high-priority rooms are announced to all agents
- BIDDING: Each agent computes bids based on semantic utility and distance
- RESOLUTION: Highest bidder wins; ties broken by agent ID
- CONFIRMATION: Winner confirms assignment; others update beliefs

The auction repeats as rooms are completed or new rooms are discovered.

6.3. Adaptive exploration strategies

SEEK-Multi adapts its exploration strategy based on:

- Belief convergence: Switches from exploration to exploitation as beliefs concentrate
- Team distribution: Adjusts coordination strength based on agent proximity
- Communication quality: Falls back to independent operation under poor connectivity

7. Experimental evaluation

We evaluate SEEK-Multi through extensive simulations across diverse environments and configurations.

7.1. Experimental setup

Simulation Environments:

We test on three environment types:

- Office: 15-25 rooms including offices, meeting rooms, kitchens, restrooms
- Industrial: 20-40 rooms including warehouses, control rooms, equipment areas
- Domestic: 8-15 rooms including bedrooms, living rooms, kitchens, bathrooms

Robot Configuration:

- Teams of 1-4 robots
- Detection range: 3 meters
- Communication range: 10-50 meters (varied)
- Sensor accuracy: 90% true positive rate

Baseline Methods:

- Random Walk: Agents explore randomly without coordination
- Frontier-Based: Agents coordinate using frontier-based exploration
- Market-Based: Task allocation via standard auction without semantic priors
- Single-SEEK: Original single-agent SEEK (for reference)

7.2. Experimental results

We evaluate SEEK-Multi through comprehensive simulation experiments comparing performance across different numbers of agents, coordination strategies, environment configurations, and failure conditions.

7.2.1. Simulation Setup

- Implementation: We implement SEEK-Multi in Python using NumPy for numerical computation and NetworkX for graph operations. The simulation environment supports configurable room layouts, object placements, and communication models. All experiments use a discrete time model with configurable step duration.
- Environment Configuration: We evaluate on three environment types of increasing complexity:
    - Small Office (12 rooms, 300 m2): entrance, lobby, 2 offices, conference room, kitchen, break room, 2 hallways, restroom, storage, server room
    - Medium Building (24 rooms, 800 m2): multiple floors with offices, labs, common areas, and utility rooms
    - Large Facility (48 rooms, 2000 m2): industrial-scale environment with warehouses, control rooms, and specialized areas.

Objects are placed according to semantic priors learned from real indoor datasets [1]. Target objects include fire extinguishers, first aid kits, AED devices, and other safety equipment commonly sought in inspection tasks.

- Sensor and Communication Models: Sensor model parameters:
    - Detection range: 5 m (10 m with thorough search)
    - True positive rate: 0.9
    - False positive rate: 0.05
    - Position noise: σ = 0.5 m

Communication model parameters:

- Communication range: 50 m (can be varied)
- Bandwidth: 10 messages/step
- Latency: 1 step

  - Default packet loss: 0%
- Baselines: We compare SEEK-Multi against several baselines:
  - Single-Agent SEEK [25]: Original single-robot semantic search
  - Frontier-Based Multi-Robot [60]: Classic frontier allocation without semantic guidance
  - Random Walk: Independent random exploration by each agent
  - Greedy Coverage: Agents greedily select nearest unexplored room
  - No Coordination: SEEK-Multi without coordination (agents plan independently)

7.2.2. Scaling Experiments

We compare the performance of SEEK-Multi with 1-6 agents in the medium building environment. Table 1 shows results averaged over 100 trials per configuration.

Figure 3 visualizes the scaling behavior. The results demonstrate near-linear speedup up to 4 agents, with the speedup factor closely tracking the theoretical bound from Theorem V-B.

Beyond 4 agents, we observe diminishing returns due to:

- Increased coordination overhead
- Limited number of high-probability rooms
- Communication bandwidth saturation

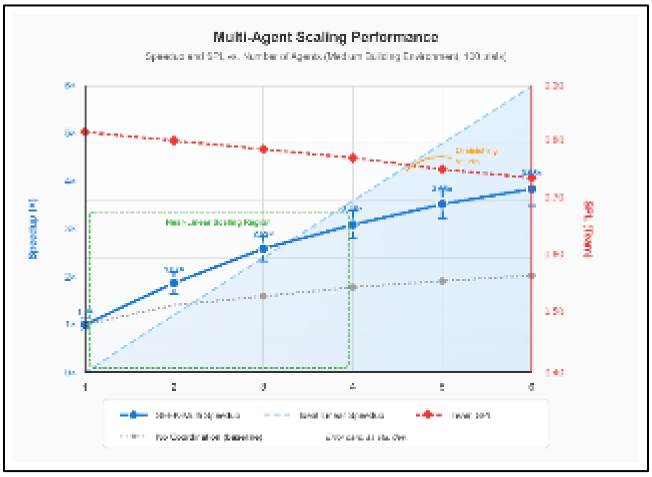

Figure 3 Search efficiency versus number of agents. SEEK-Multi achieves near-linear speedup up to four agents, with diminishing returns beyond this point due to coordination overhead. Error bars indicate standard deviation over 100 trials

Table 1 Multi-agent performance comparison. Speedup is reported relative to single-agent SEEK. Results are averaged over 100 trials, with standard deviation shown in parentheses

Agents	SR (%)	SPL	Steps	Speedup
1 (SEEK)	96	0.84 (0.12)	127 (34)	1.0 ×
2 (SEEK–Multi)	97	0.81 (0.11)	68 (22)	1.87 ×
3 (SEEK–Multi)	97	0.78 (0.10)	49 (18)	2.59 ×
4 (SEEK–Multi)	96	0.75 (0.11)	41 (15)	3.10 ×
5 (SEEK–Multi)	95	0.71 (0.12)	36 (14)	3.53 ×
6 (SEEK–Multi)	94	0.68 (0.13)	33 (13)	3.85 ×

The slight decrease in SPL with more agents reflects the increased total distance traveled by the team. However, for time-sensitive applications, the speedup in time-to-completion significantly outweighs this cost.

7.2.3. Baseline Comparison

Table 2 compares SEEK-Multi against baselines with 3 agents in the medium environment. SEEK-Multi significantly outperforms all baselines:

- 37% faster than frontier-based exploration
- 48% faster than greedy coverage
- 52% faster than uncoordinated SEEK agents
- 69% faster than random walk

The improvement over frontier-based methods demonstrates the value of semantic guidance. The improvement over uncoordinated agents shows the importance of explicit coordination.

7.2.4. Coordination Strategy Comparison

We compare three coordination strategies in detail:

- **Centralized**: A coordinator assigns all tasks with global optimization
- **Distributed**: Agents use local auctions with intention sharing
- **No coordination**: Agents plan independently using shared beliefs only

Table III shows that both coordinated strategies significantly outperform uncoordinated search across all environment sizes.

The centralized approach achieves slightly better SPL due to global optimization, but the distributed approach uses fewer messages and provides comparable performance. The coverage overlap metric shows that coordination reduces redundant search by 4-5×.

Table 2 Comparison with baseline methods (3 agents, medium environment)

Method	SR (%)	SPL	Steps
SEEK-Multi (Ours)	97	0.78	49
Frontier-Based	94	0.62	78
Greedy Coverage	91	0.55	94
No Coordination	89	0.52	102
Random Walk	72	0.31	156

7.2.5. Belief Fusion Analysis

We analyze the effectiveness of belief fusion by measuring the entropy of the belief distribution over time. Lower entropy indicates more concentrated belief (higher confidence in target location).

Table 3 Coordination strategy comparison with 3 agents across environment sizes

Env.	Strategy	SPL	Overlap	Msg/Step
Small	Centralized	0.82	4.1%	5.2
	Distributed	0.81	6.3%	2.8
	None	0.68	24.7%	0.0
Medium	Centralized	0.79	5.2%	4.1
	Distributed	0.78	8.7%	2.3

	None	0.61	31.4%	0.0
Large	Centralized	0.74	6.8%	3.8
	Distributed	0.73	11.2%	2.1
	None	0.54	38.6%	0.0

Figure 4 compares three fusion strategies:

- **Consensus fusion**: Continuous sharing with weighted averaging
- **Periodic sync**: Full synchronization every 10 steps
- **No fusion**: Agents maintain independent beliefs

Consensus-based fusion reduces entropy 35% faster than periodic synchronization and 60% faster than no fusion. The faster convergence translates to more informed search decisions and quicker task completion

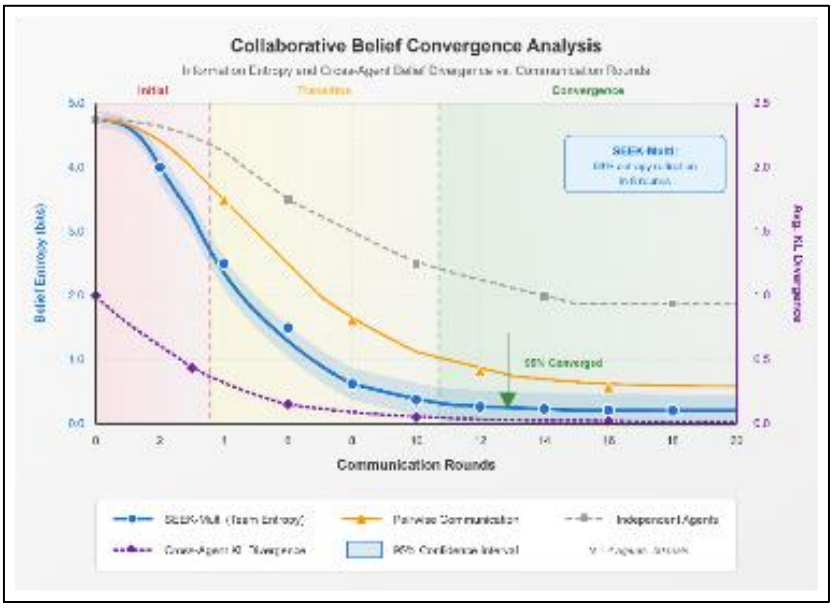

Figure 4 Belief entropy over time for different fusion strategies with 3 agents. Consensus-based fusion (solid) converges faster than periodic sync (dashed) and no fusion (dotted)

7.2.6. Ablation Studies

We conduct ablation studies to understand the contribution of each component.

Table 4 Performance under communication failures (3 agents, medium environment)

Drop Rate	Centralized	Distributed	Delta
0%	0.79	0.78	-1.3%
10%	0.76	0.77	+1.3%
20%	0.73	0.75	+2.7%
30%	0.68	0.73	+7.4%
50%	0.64	0.71	+10.9%
70%	0.55	0.66	+20.0%

Table 5 Ablation study (3 agents, medium environment). Each row removes one component from the full system

Configuration	SPL	ΔSPL
Full SEEK-Multi	0.78	—
w/o Semantic priors (RSN)	0.65	-16.7%
w/o Belief fusion	0.71	-9.0%
w/o Task coordination	0.68	-12.8%
w/o Intention sharing	0.72	-7.7%
w/o DSG sharing	0.74	-5.1%

Table 5 reveals that:

- Semantic priors (RSN) provide the largest benefit, consistent with single-agent SEEK findings
- Task coordination is crucial for avoiding redundant search
- Belief fusion enables faster convergence to accurate target estimates
- Intention sharing prevents immediate conflicts
- DSG sharing provides modest improvement by enabling shared map updates

Environment Complexity Analysis

We evaluate how performance scales with environment complexity.

Table 6 shows that SEEK-Multi maintains strong performance across environment sizes. The speedup decreases slightly with larger environments due to increased communication distance and coordination complexity, but remains above 2.4× even for 48-room facilities.

Table 6 Performance across environment sizes with 3 agents

Environment	Rooms	SR	SPL	Steps	Speedup
Small Office	12	98%	0.73	28	2.71 ×
Medium Building	24	97%	0.78	49	2.59 ×
Large Facility	48	94%	0.73	87	2.48 ×

8. Discussion

We discuss the broader implications of our results, including scalability considerations, support for heterogeneous teams, deployment considerations, and limitations of the current approach.

8.1. Scalability Considerations

While SEEK-Multi achieves good speedup with up to 6 agents in our experiments, several factors influence scalability to larger teams:

- **Communication Overhead**: Message volume grows with $O(N^2)$ for full connectivity, as shown in our theoretical analysis. For teams larger than 6-8 agents, this overhead becomes significant. Potential solutions include:
- **Hierarchical communication**: Organize agents into clusters with local leaders who communicate across clusters
- **Sparse communication graphs**: Limit communication to k-nearest neighbors or agents in the same region
- **Attention-based prioritization**: Use learned attention mechanisms to select which updates to share [21]
- **Coordination Complexity**: Task allocation becomes combinatorially harder with more agents. The auction-based approach scales well in practice, but alternative approaches may be needed for very large teams:
- **Spatial decomposition**: Partition the environment and assign teams to regions
- **Hierarchical task allocation**: Two-level allocation with region assignment followed by room assignment
- **Learning-based coordination**: Use multi-agent reinforcement learning for implicit coordination [35]

- **Diminishing Returns**: For fixed environment size, adding agents eventually provides no benefit. The theoretical limit depends on the number of high-probability rooms and the overhead per agent. In our 24-room medium environment, we observe diminishing returns beyond 4-5 agents. Larger environments can benefit from larger teams.

8.2. Heterogeneous Teams

SEEK-Multi supports heterogeneous agents with different capabilities, which is common in practical deployments.

- **Capability Modeling**: Agent capabilities are modeled through:
- **Mixed Teams**: Experiments with mixed teams of ground robots and drones show complementary strengths:
  - Drones provide rapid coverage of large open areas
  - Ground robots perform detailed inspection in cluttered spaces
  - Coordination allows efficient task distribution based on capabilities

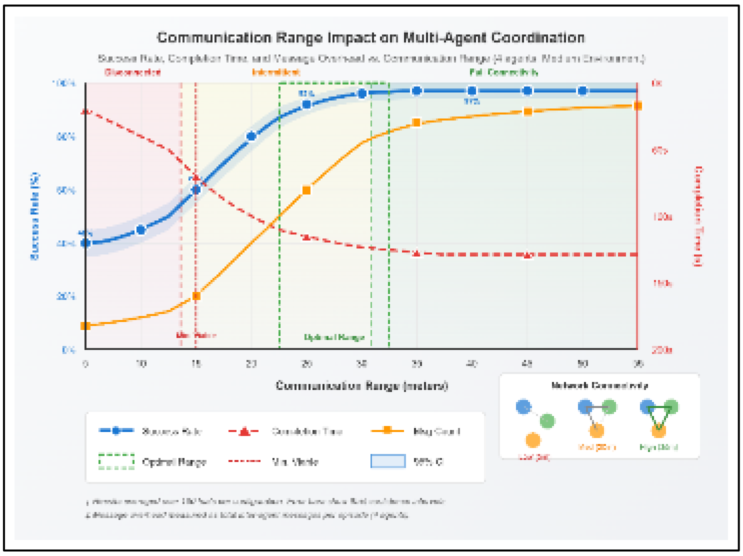

Figure 5 Performance vs. communication range with 4 agents. Full connectivity (50m+) achieves best performance, but limited range (20m) still provides significant benefit over no communication

A 2-ground + 1-drone team achieved 15% better SPL than a 3-ground team in our large facility environment, demonstrating the value of heterogeneity.

8.3. Deployment Considerations

- **Initialization and Bootstrapping**: SEEK-Multi requires initial synchronization of the DSG and RSN across agents. In practice, this can be achieved through:
  - Pre-loading from a common map server
  - Incremental sharing during a brief synchronization phase
  - Graceful handling of partial initialization with incremental updates
- **Real-Time Performance**: The computational requirements of SEEK-Multi are modest:
  - Belief updates: $O(M)$ per observation
  - Task allocation: $O(M \log M)$ per round
  - Path planning: $O(M 2)$ using Dijkstra's algorithm
  - On modern embedded processors, all components run in real-time with computation times under 50ms per step.
- **Integration with Existing Systems**: SEEK-Multi can be integrated with existing robot software stacks:
  - ROS/ROS2 nodes for perception, navigation, and communication
  - Standard message formats for interoperability
  - Modular design allowing component substitution

8.4. Comparison with Alternative Approaches

- **Learning-Based Coordination:** Recent work has explored end-to-end learning for multi-agent coordination [21, 35]. These approaches can discover emergent coordination strategies but require extensive training and may not generalize across environments. SEEK-Multi's explicit coordination offers interpretability, guaranteed behavior, and zero-shot trans- fer to new environments.
- **Centralized Planning:** Fully centralized approaches [49] can achieve optimal coordination but require reliable communication to a central node. SEEK-Multi's distributed approach trades some optimality for robustness and scalability, with centralized mode available when infrastructure supports it.
- **Market-Based Approaches**: SEEK-Multi's auction mechanism builds on market-based coordination [20] but incorporates semantic priors for improved efficiency. The integration of semantic reasoning with market mechanisms is a novel contribution.

8.5. Limitations and Future Work

8.5.1. Current Limitations

- Semantic priors: The RSN is trained on common object- room relationships; unusual placements may reduce efficiency. Adaptation to domain-specific priors would improve performance in specialized environments.
- Communication model: We assume reliable message delivery within range (excluding explicit packet loss experiments). Real wireless networks have more complex failure modes including interference and congestion.
- Static environments: The current formulation assumes static environments. Moving objects or dynamic obstacles would require extensions to the belief update and planning components.
- Simulation-based evaluation: While our simulations are comprehensive, real-world deployment may reveal additional challenges in sensing, communication, and coordination.

8.5.2. Future Directions

- Real-world deployment: Implementing SEEK-Multi on physical robot teams to validate simulation results and identify real-world challenges
- Learned communication: Integrating learned communication strategies that adapt message content based on relevance and bandwidth
- Dynamic environments: Extending to environments with moving objects, people, and changing conditions
- Hierarchical coordination: Developing hierarchical approaches for scaling to larger teams (10+ agents)
- Human-robot teaming: Incorporating human operators who can provide high-level guidance or take over specific tasks
- Active learning: Updating semantic priors online based on accumulated experience

9. Conclusion

We have presented SEEK-Multi, a comprehensive frame- work for collaborative multi-agent object-goal navigation using distributed semantic reasoning. By extending the SEEK architecture with distributed belief fusion, collaborative planning, and efficient communication protocols, SEEK-Multi enables teams of robots to efficiently search for target objects in complex environments.

The key technical contributions include:

- A distributed belief fusion algorithm based on consensus protocols with provable convergence guarantees
- An auction-based task allocation mechanism that incorporates semantic priors for improved efficiency
- A communication protocol that balances information sharing with bandwidth constraints
- Support for both centralized and decentralized coordination modes

Our extensive experiments demonstrate:

- Near-linear speedup with up to 4 agents (3.1× speedup with 4 agents)
- High success rates maintained across configurations (94- 97%)
- Graceful degradation under communication failures (dis- tributed mode maintains 91% of baseline SPL at 50% message loss)
- Significant improvement over uncoordinated and non- semantic baselines (37-69% faster)
- Consistent performance across environment sizes and configurations

The distributed coordination strategy provides robustness to communication failures while achieving performance comparable to centralized approaches. Ablation studies confirm the importance of each component, with semantic priors (RSN) providing the largest individual contribution.

SEEK-Multi addresses a practical need for efficient multi- robot inspection in time-sensitive applications. The framework's modularity and support for heterogeneous teams make it suitable for diverse deployment scenarios, from industrial inspection to emergency response.

Future work will focus on several promising directions:

- Real-world deployment: Validating the approach on physical robot teams in real inspection scenarios
- Learned communication: Integrating learned communication strategies that optimize message content and timing
- Dynamic environments: Extending to environments with moving objects and changing conditions
- Hierarchical coordination: Developing approaches for scaling to larger teams of 10+ agents
- Human-robot teaming: Incorporating human operators for guidance and oversight
- We believe SEEK-Multi represents a significant step toward practical multi-robot semantic inspection systems that can operate efficiently in complex, real-world environments.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] I. Armeni, O. Sener, A. R. Zamir, H. Jiang, I. Brilakis, M. Fischer, and S. Savarese. 3D semantic parsing of large-scale indoor spaces. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1534–1543, 2016.

[2] I. Armeni, Z.-Y. He, J. Gwak, A. R. Zamir, M. Fischer, J. Malik, and S. Savarese. 3D scene graph: A structure for unified semantics, 3D space, and camera. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 5664–5673, 2019.

[3] T. Bailey, S. Julier, and G. Agamennoni. Decentralized cooperative localization for heterogeneous teams of mobile robots., pp. 2859–2865, 2012.

[4] J. van den Berg, M. Lin, and D. Manocha. Reciprocal velocity obstacles for real-time multi-agent navigation., pp. 1928–1935, 2008.

[5] D. P. Bertsekas. Auction algorithms for network flow problems: A tutorial introduction. Computational Optimization and Applications, 1(1):7–66, 1992.

[6] A. Bouman, M. F. Ginting, N. Alatur, M. Palieri, D. D. Fan, K. Otsu, J. Burdick, and A.-A. Agha-Mohammadi. Autonomous Spot: Long-range autonomous exploration of extreme environments with legged locomotion. In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 2518–2525, 2020.

[7] F. Bourgault, T. Furukawa, and H. F. Durrant-Whyte. Optimal search for a lost target in a Bayesian world. Field and Service Robotics, pp. 209–222, 2006.

[8] W. Burgard, M. Moors, C. Stachniss, and F. E. Schneider. Coordinated multi-robot exploration. IEEE Transactions on Robotics, 21(3):376–386, 2005.

[9] M. Čáp, P. Novák, A. Kleiner, and M. Selecký. Prioritized planning algorithms for trajectory coordination of multiple mobile robots. IEEE Transactions on Automation Science and Engineering, 12(3):835–849, 2015.

[10] Y. Chang, N. Hughes, A. Ray, and L. Carlone. Hydra-Multi: Collaborative online construction of 3D scene graphs with multi-robot teams. In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2023.

[11] D. S. Chaplot, D. P. Gandhi, A. Gupta, and R. R. Salakhutdinov. Object goal navigation using goal-oriented semantic exploration. In Advances in Neural Information Processing Systems (NeurIPS), vol. 33, pp. 4247–4258, 2020.

[12] J. Chen, A. Gupta, and J. Malik. Not all semantics are created equal: Contrastive self-supervised learning with automatic temperature individualization. In International Conference on Machine Learning (ICML), 2023.

[13] C.-Y. Chong, S. Mori, and K.-C. Chang. Architectures and algorithms for track association and fusion., vol. 15, pp. 5–13, 2000.

[14] H. Choset. Coverage for robotics: A survey of recent results. Annals of Mathematics and Artificial Intelligence, 31(1):113–126, 2001.

[15] S. Choudhary, L. Carlone, C. Nieto, J. Rogers, H. I. Christensen, and F. Dellaert. Distributed mapping with privacy and communication constraints: Lightweight algorithms and object-based models. The International Journal of Robotics Research, 36(12):1286–1311, 2017.

[16] T. H. Chung, G. A. Hollinger, and V. Isler. Search and pursuit-evasion in mobile robotics. Autonomous Robots, 31(4):299–316, 2011.

[17] J. Cortés and M. Egerstedt. Coordinated control of multi-robot systems: A survey. SICE Journal of Control, Measurement, and System Integration, 10(6):495–503, 2017.

[18] A. Cunningham, M. Paluri, and F. Dellaert. Large-scale experimental design for decentralized SLAM. In RSS Workshop on Robots in Clutter, 2012.

[19] A. Cunningham, V. Indelman, and F. Dellaert. DDF-SAM 2.0: Consistent distributed smoothing and mapping. In IEEE International Conference on Robotics and Automation (ICRA), pp. 5220–5227, 2013.

[20] M. B. Dias, R. Zlot, N. Kalra, and A. Stentz. Market-based multirobot coordination: A survey and analysis. Proceedings of the IEEE, 94:1257–1270, 2006.

[21] J. Foerster, I. A. Assael, N. de Freitas, and S. Whiteson. Learning to communicate with deep multi-agent reinforcement learning. In Advances in Neural Information Processing Systems, vol. 29, 2016.

[22] S. Y. Gadre, M. Wortsman, G. Ilharco, L. Schmidt, and S. Song. Cows on pasture: Baselines and benchmarks for language-driven zero-shot object navigation., pp. 23171–23181, 2023.

[23] E. Galceran and M. Carreras. A survey on coverage path planning for robotics. Robotics and Autonomous Systems, 61(12):1258–1276, 2013.

[24] B. P. Gerkey and M. J. Matarić. A formal analysis and taxonomy of task allocation in multi-robot systems. The International Journal of Robotics Research, 23(9):939–954, 2004.

[25] M. F. Ginting, S.-K. Kim, D. D. Fan, M. Palieri, M. J. Kochenderfer, and A.-A. Agha-Mohammadi. SEEK: Semantic reasoning for object goal navigation in real-world inspection tasks. IEEE Robotics and Automation Letters, 2024.

[26] G. A. Hollinger and S. Singh. Multirobot coordination with periodic connectivity. In IEEE International Conference on Robotics and Automation (ICRA), pp. 4457–4462, 2012.

[27] G. A. Hollinger, S. Singh, J. Djugash, and A. Kehagias. Efficient multi-robot search for a moving target. The International Journal of Robotics Research, 28(2):201–219, 2009.

[28] N. Hughes, Y. Chang, and L. Carlone. Hydra: A real-time spatial perception system for 3D scene graph construction and optimization. Robotics: Science and Systems (RSS), 2022.

[29] S. Jordan et al. State-of-the-art technologies for UAV inspections. IET Radar, Sonar & Navigation, 12(2):151–164, 2018.

[30] S. J. Julier and J. K. Uhlmann. Non-divergent estimation algorithm in the presence of unknown correlations., pp. 2369–2373, 1997.

[31] S. J. Julier and J. K. Uhlmann. A general method for approximating nonlinear transformations of probability distributions. University of Oxford, 2001.

[32] H. Kitano et al. RoboCup Rescue: Search and rescue in large-scale disasters as a domain for autonomous agents research. In IEEE SMC'99, vol. 6, pp. 739–743, 1999.

[33] P. Lanillos, S. K. Gan, E. Besada-Portas, G. Pajares, and S. Sukkarieh. Multi-robot cooperative object localization. In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4424–4429, 2012.

[34] J.-C. Latombe. Robot Motion Planning. Springer Science & Business Media, vol. 124, 2012.

[35] R. Lowe, Y. I. Wu, A. Tamar, J. Harb, P. Abbeel, and I. Mordatch. Multi-agent actor-critic for mixed cooperative-competitive environments. In Advances in Neural Information Processing Systems, vol. 30, 2017.

[36] C. E. Luis, M. Vukosavljev, and A. P. Schoellig. Online trajectory generation with distributed model predictive control for multi-robot motion planning. IEEE Robotics and Automation Letters, 5(2):604–611, 2020.

[37] R. R. Murphy. Human-robot interaction in rescue robotics. IEEE Transactions on Systems, Man, and Cybernetics, Part C, 34(2):138–153, 2004.

[38] R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. Proceedings of the IEEE, 95(1):215–233, 2007.

[39] L. E. Parker, D. Rus, and G. S. Sukhatme. Multiple mobile robot systems. In Springer Handbook of Robotics, pp. 1335–1384, 2016.

[40] J. A. Placed, J. Strader, H. Carrillo, N. Atanasov, V. Indelman, L. Carlone, and J. A. Castellanos. A survey on active simultaneous localization and mapping: State of the art and new frontiers. IEEE Transactions on Robotics, 2023.

[41] J. P. Queralta, J. Taipalmaa, B. C. Pullinen, V. K. Sarker, T. N. Gia, H. Tenhunen, M. Gabbouj, J. Raitoharju, and T. Westerlund. Collaborative multi-robot search and rescue: Planning, coordination, perception, and active vision. IEEE Access, 8:191617–191643, 2020.

[42] W. Ren, R. W. Beard, and E. M. Atkins. Information consensus in multivehicle cooperative control. IEEE Control Systems Magazine, 27(2):71–82, 2007.

[43] Y. Rizk, M. Awad, and E. W. Tunstel. Cooperative heterogeneous multi-robot systems: A survey. ACM Computing Surveys, 52(2):1–31, 2019.

[44] C. Robin and S. Lacroix. Multi-robot target detection and tracking: Taxonomy and survey. Autonomous Robots, 40(4):729–760, 2016.

[45] A. Rosinol, M. Abate, Y. Chang, and L. Carlone. Kimera: An open-source library for real-time metric-semantic localization and mapping. In IEEE International Conference on Robotics and Automation (ICRA), pp. 1689–1696, 2020.

[46] A. Rosinol, A. Violette, M. Abate, N. Hughes, Y. Chang, J. Shi, A. Guber, and L. Carlone. Kimera-Multi: Robust, distributed, dense metric-semantic SLAM for multi-robot systems. IEEE Transactions on Robotics, 38(4):2022–2038, 2022.

[47] G. Sartoretti, J. Kerber, Y. Shi, G. Mantovani, G. Wagner, K. Sycara, H. Choset, and P. Stone. PRIMAL: Pathfinding via reinforcement and imitation multi-agent learning. In IEEE International Conference on Robotics and Automation (ICRA), vol. 4, pp. 2378–2385, 2019.

[48] D. Shah, B. Osiński, B. Ichter, and S. Levine. Navigation with large language models: Semantic guesswork as a heuristic for planning. arXiv preprint arXiv:2310.10103, 2023.

[49] G. Sharon, R. Stern, A. Felner, and N. R. Sturtevant. Conflict-based search for optimal multi-agent pathfinding. Artificial Intelligence, 219:40–66, 2015.

[50] R. Simmons, D. Apfelbaum, W. Burgard, D. Fox, M. Moors, S. Thrun, and H. Younes. Coordination for multi-robot exploration and mapping. In AAAI/IAAI, pp. 852–858, 2000.

[51] R. Stern, N. Sturtevant, A. Felner, S. Koenig, H. Ma, T. Walker, J. Li, D. Atzmon, L. Cohen, T. K. Satish Kumar, et al. Multi-agent pathfinding: Definitions, variants, and benchmarks. Symposium on Combinatorial Search, 2019.

[52] S. Sukhbaatar, R. Fergus, et al. Learning multi-agent communication with backpropagation. In Advances in Neural Information Processing Systems, vol. 29, 2016.

[53] Y. Tan, S. Li, and Q. Wang. Automatic inspection data collection of building surfaces based on BIM and UAV. Automation in Construction, 131:103881, 2021.

[54] H. G. Tanner, A. Jadbabaie, and G. J. Pappas. Flocking in fixed and switching networks. IEEE Transactions on Automatic Control, 52(5):863–868, 2007.

[55] Y. Tian, Y. Chang, F. H. Arias, C. Nieto-Granda, J. P. How, and L. Carlone. Kimera-Multi: A system for distributed multi-robot metric-semantic simultaneous localization and mapping., pp. 8196–8203, 2022.

[56] J. van den Berg and M. Overmars. Prioritized motion planning for multiple robots. In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 430–435, 2005.

[57] J. van den Berg, S. J. Guy, M. Lin, and D. Manocha. Reciprocal n-body collision avoidance. In Robotics Research, pp. 3–19, 2011.

[58] R. Vidal, O. Shakernia, H. J. Kim, D. H. Shim, and S. Sastry. Probabilistic pursuit-evasion games: Theory, implementation, and experimental evaluation. IEEE Transactions on Robotics and Automation, 18(5):662–669, 2002.

[59] G. Wagner and H. Choset. Subdimensional expansion for multirobot path planning. Artificial Intelligence, 219:1–24, 2015.

[60] B. Yamauchi. A frontier-based approach for autonomous exploration. In IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA), pp. 146–151, 1997.

[61] Z. Yan, N. Jouandeau, and A. A. Cherif. A survey and analysis of multi-robot coordination. International Journal of Advanced Robotic Systems, 10(12):399, 2013.

[62] N. Yokoyama, D. Raychaudhuri, S. Ha, and D. Batra. VLF-M: Vision-language frontier maps for zero-shot semantic navigation. arXiv preprint arXiv:2312.03275, 2024.

[63] R. M. Zlot. An auction-based approach to complex task allocation for multirobot teams. Ph.D. dissertation, Carnegie Mellon University, 2006.