Rolling-horizon Risk-aware Search in Stochastic Mall Hazards: A Paired Multi-agent Simulation Study

Yiming Wang

Beijing Haidian Foreign Language Tengfei School, Beijing, China
Email: 632469239@qq.com
Manuscript received September 3, 2026; accepted October 11, 2026; published November 9, 2026.

***Abstract*—Dynamic indoor search couples task allocation with navigation because target accessibility and route cost can change during execution. This study compares a joint rolling Policy (P) with three reduced-capability study baselines: static distance planning (B0), open-loop arrival-time risk planning (B1), and fixed-assignment periodic path replanning (B2). P additionally combines completion-event replanning, redistribution of unfinished tasks, and failed-agent work recovery; P−B2 therefore estimates this policy bundle rather than any one component. The confirmatory benchmark comprised 1,080 paired synthetic scenarios and 4320 runs across three layouts, three hazard-spread regimes, two observation-noise levels, and team sizes of three and five. P achieved 60.9% mission success, versus 33.1%, 38.9%, and 50.2% for B0, B1, and B2. Relative to B1, P increased success by 0.220 (95% CI 0.194 to 0.247), reduced cumulative normalized exposure among 411 mutual successes (P−B1 = −0.856; 95% CI −1.231 to −0.500), and met the frozen makespan-ratio limit. Relative to B2, its success advantage was 0.107 (95% CI 0.087 to 0.128). Team-size sensitivity met a pre-specified local diminishing-return criterion but was not globally monotone. An independent rerun reproduced all 27 deterministic output columns exactly. The findings describe a reproducible success-exposure-time tradeoff in a stochastic dynamic-cost benchmark, not real-fire safety: the field is not validated fire physics, and normalized exposure is not physiological dose.**

***Keywords*—multi-agent task allocation, time-dependent path planning, stochastic hazard field, receding horizon, paired simulation, computational reproducibility**

I. Introduction

Indoor search under a changing hazard is more than a shortest-path problem. A team must cover distributed targets while the cost of reaching them evolves. A route selected at mission start can become undesirable, and a target assigned to one agent can later be reached more effectively by another. Navigation and task ownership are therefore coupled. Updating a path while holding its destination fixed may leave part of the available adaptation unused.

Prediction-aware indoor routing evaluates risk at expected traversal time rather than only at planning time [1–3], while MRTA research separates allocation from motion planning [4, 5]. Completion-triggered reallocation and recovery from robot malfunction already appear in repeated-auction research [6]. The bounded gap is therefore not the existence of dynamic reassignment. It is the controlled evaluation of an arrival-time dynamic-cost policy bundle through matched internal ablations, complete failure reporting, and exact-rerun verification. Close firefighter-routing work uses richer prediction and simulation components but different objectives and settings [7].

The research question is: How does a joint rolling-horizon policy change mission success, cumulative normalized exposure, and makespan relative to reduced-capability internal baselines in stochastic single-floor mall search? H1 predicts that P will reduce paired-success exposure relative to B1 while satisfying frozen success and makespan tolerances. H2 predicts that P will exceed B2 in all-run success; because P also adds completion-event updates and failed-agent work recovery, H2 concerns the combined policy increment. H3 specifies a local descriptive pattern of diminishing makespan reductions as team size increases. The contribution is a reproducible policy-level benchmark with matched scenarios, explicit failures, pre-specified estimands, sensitivity analyses, and exact-rerun verification. It is neither a new fire model nor evidence of state-of-the-art algorithmic superiority.

II. Related Work

A* supplies the local route search [8]. D* Lite offers incremental repair after cost changes [9], but this implementation replans from the latest observation to preserve the intended component contrast. Conflict-based Search addresses a different optimal multi-agent pathfinding problem [10]. The present reservation rules only prevent immediate vertex conflicts and edge swaps.

The mTSP and auction-routing literatures clarify route-distribution and team-objective choices [11, 12]. In the Gerkey–Matarić taxonomy, the present benchmark is an ST-SR-TA problem: agents execute tasks sequentially, each task requires one agent, and assignments extend over time [4]. Korsah *et al.* [5] add task-dependency distinctions; this benchmark has no precedence or coalition tasks, although route sequence couples task costs. Market-based coordination and optimization reviews show a much broader MRTA design space than the four tested policies [13, 14]. B0–B2 are therefore reduced-capability study baselines, not representatives of the strongest auction or optimization methods.

Nanjanath and Gini [6] are the closest mechanism precedent: unfinished tasks are re-auctioned after task completion to adapt to execution changes and robot malfunction. The comparison in Table 1 prevents a novelty claim based on mechanisms already demonstrated.

What remains distinctive is the matched component ladder, arrival-time cost integration, failure-retaining estimands, and reproducibility package, not completion-triggered reassignment itself.

Dynamic-routing reviews treat information change, stochasticity, objectives, and update timing as defining dimensions [15, 16]. Receding-horizon studies schedule search-and-rescue agents near spreading hazards, allocate moving tasks, map dynamic irradiance, and handle uncertain

temporal constraints [17–20]. These precedents justify repeated decisions but not universal benefit.

Table 1. Comparison of the closest mechanism precedent and the present P policy

Design	Update Trigger	Allocation and Recovery	Objective or Cost	Validation
Nanjanath and Gini [6]	Task completion	Distributed repeated auctions; failure recovery	Path and completion time	Simulation and physical robots
Present P policy	Every three steps and task completion	Centralized greedy redistribution; failed-agent work recovery	Projected makespan with arrival-time normalized field cost	Paired synthetic simulation and exact rerun

Indoor routing work supports evaluating hazards at expected arrival time [1–3]. Richer frameworks add three-dimensional congestion, digital twins, or coupled fire and evacuation simulation [7, 21, 22].

The synthetic field borrows only neighbor-dependent state evolution from cellular models [23]. It omits heat, smoke transport, toxicity, visibility, ventilation, structural failure, and suppression. FDS documentation illustrates the verification and validation expected of physical fire modeling [24, 25]. This study therefore evaluates a dynamic-cost algorithm, not physical exposure or tenability.

III. Method

A. Environment and Observations

Each scenario is a connected 24 × 24 four-neighbor grid with blocked cells, 16 mandatory targets, and distinct agent entries. At each discrete step an agent moves, waits, or services its current target; service takes two steps. A mission succeeds when all targets are serviced within 120 steps, otherwise makespan is capped at 120. Deterministic reservations reject same-destination moves, edge swaps, and entry into a stationary agent's cell.

The latent hazard has four normalized states: 0, 0.25, 0.80, and 1.0. Code labels denote clear, smoke, fire, and untenable states, but these labels are mnemonic and have no physical units. One task-area cell begins at 0.80 and adjacent traversable cells at 0.25; monotone neighbor-dependent transitions generate slow, medium, or fast spread. One seeded latent trajectory is shared by every policy in a scenario. Table 2 summarizes the operational definitions of the main algorithmic constructs and the interpretations they exclude.

Table 2. Operational definitions and excluded interpretations

Algorithmic Construct	Operational Definition	Not Represented
Hazard state and exposure	Unitless cell state and its accumulated occupancy cost	Temperature, smoke concentration, visibility, dose, or tenability
Agent and failure	Homogeneous grid mover removed after entering state 1.0	Human physiology, equipment reliability, or firefighter behavior
Target and service	Required cell with deterministic two-step service	Victim condition, rescue, suppression, or triage
Mission success	All 16 targets serviced within 120 steps	Fireground safety or operational mission success

Planning uses noisy observations, whereas evaluation uses the shared latent field. Selected observed states shift by one level at noise rate η. An eight-step forecast propagates neighbor effects; forecast-only values stay below 0.95, while observed state-1.0 cells remain impassable. Thus every policy is scored against the same hidden trajectory but plans from imperfect input.

B. Policies

Four policies form an internal reduced-capability ladder, as summarized in Table 3. They are controls for the implemented components rather than state-of-the-art competitors.

Table 3. Reduced-capability policy ladder

Policy	Initial Risk Use	Periodic Path Replanning	Completion-event Update	Redistribute Unfinished Tasks	Recover Failed-agent Work
B0: static distance	No	No	No	No	No
B1: open-loop risk-aware	Yes	No	No	No	No
B2: fixed-assignment replanning	Yes	Yes	No	No	No
P: joint rolling policy	Yes	Yes	Yes	Yes	Yes

B0 assigns tasks and routes by distance. B1 uses the initial forecast for risk-aware routes and assignments but never updates them. B2 periodically replans the active route while retaining initial ownership. P combines periodic and completion-event replanning with redistribution of unlocked unfinished tasks. If an agent fails while holding work, B0–B2 terminate because they lack transfer logic, whereas P can recover that work. Thus B0→B1 adds initial risk use, B1→B2 adds periodic route updating, and B2→P adds a three-part policy bundle: completion-event triggering, redistribution, and failure recovery. H2 evaluates that bundle and cannot identify the effect of any one component.

C. Arrival-time Routing and Allocation

Risk-aware A* searches cell-time states (x, τ). Moving to a neighbor or waiting in place has cost

$$c(x,y,\tau) = 1 + \lambda \hat{r}(y,\tau + 1)$$

Here, $\hat{r}$ is forecast risk at arrival and $\lambda = 6$. Values at or above 0.98 are infeasible. Manhattan distance remains admissible because risk costs are nonnegative, but this local

property does not make allocation or the complete mission globally optimal.

The allocator greedily appends unfinished tasks. It estimates candidate completion time and exposure from cached paths, retains assignments within 5% of the minimum projected makespan, selects the lowest estimated exposure, and resolves ties deterministically. The 5% tolerance is computational, not a safety threshold. Exact time-expanded A* is reserved for the next segment because later tasks may be reassigned.

P replans every three steps and after target completion. Completed targets are removed, active service is locked, and other unfinished tasks are redistributed among active agents. B2 uses the three-step path update but neither the completion trigger nor ownership changes. Conflicts are resolved before exposure accrues. Exposure sums latent cell values at agents' final cells each step; state 1.0 removes the agent.

D. Outcomes

The principal outcome profile comprises mission success, cumulative normalized exposure, and makespan. Success is primary when either policy fails. Exposure and makespan are then compared among mutual successes, while unconditional exposure is retained because early termination can artificially reduce accumulated cost. Secondary outcomes include completed tasks, maximum individual exposure, distance, waits, workload variation, replans, reassignments, expanded nodes, failed agents, and wall-clock planning time. Together these variables describe a success-exposure-time tradeoff; they do not estimate a safety frontier.

IV. Experimental Design

The confirmatory design crossed three layouts, three spread regimes, noise rates 0 and 0.15, team sizes 3 and 5, and 30 held-out seeds. This yielded 1080 scenarios and 4320 policy runs. Scenario seeds depended on factors and replicate, not policy, so each comparison shared geometry, latent trajectory, observation errors, demand, and horizon. No seed was excluded; failures remained in analysis.

A retained 360-scenario pilot supported calibration. An earlier discarded pilot revealed that seeds incorrectly depended on risk weight and that long forecasts saturated the hard threshold; both defects were corrected before confirmatory runs, and the discarded file was preserved. Primary settings were then frozen: risk weight 6, replanning interval 3, forecast horizon 8, time limit 120, 16 tasks, and two-step service. The H1 tolerances were also fixed before confirmation. A −0.025 success margin limits acceptable deterioration to 2.5 percentage points (one success per 40 scenarios), and a 1.05 makespan-ratio limit allows at most 5% paired inflation. These are analyst-defined design tolerances, not empirically established minimally important differences or safety thresholds. SHA-256 hashes were stored for simulator, runner, tests, and configurations.

Effects were paired by scenario_id. Success difference was the mean of S_P−S_C over all scenarios. Exposure difference was the mean of E_P−E_C among mutual successes, and the makespan estimand was the median within-scenario ratio T_P/T_C in the same subset. Percentile 95% confidence intervals used 10,000 scenario-level bootstrap resamples with seed 20260801. H1 required the P−B1 success lower bound to be at least −0.025, the exposure upper bound below zero, and the makespan-ratio upper bound no greater than 1.05. H2 required the all-run P−B2 success-difference lower bound to exceed zero and was interpreted as a policy-bundle contrast. H3 required two successive decreases in marginal median makespan improvement and was treated as a descriptive pre-specified criterion. Paired Wilcoxon tests for exposure and makespan used Holm correction across six planned tests; rank-biserial effects were retained, and exact paired-binomial tests summarized discordant successes.

Separate sensitivities tested team sizes 2–8 across four layouts (1680 runs) and crossed replanning intervals 1, 3, and 5 with risk weights 1.5, 3.0, and 6.0 (3240 runs for B1, B2, and P). These finite-grid summaries are descriptive except for the frozen H3 criterion. The full confirmatory configuration was then rerun; after sorting by run_id, 27 deterministic columns had to match exactly, excluding wall-clock time.

V. Results

A. Overall Outcomes

All 4320 confirmatory runs were present, with four policy records per synthetic scenario and no duplicate identifiers or excluded seeds. P succeeded in 658 of 1080 scenarios (60.9%). The reduced-capability study baselines B0, B1, and B2 succeeded in 357 (33.1%), 420 (38.9%), and 542 (50.2%), respectively (Fig. 1). Failure-aware mean makespan fell from 91.68 for B0 to 69.30 for P. Table 4 preserves both unconditional and success-conditioned summaries.

Table 4. Synthetic mission outcomes; exposure is a unitless field cost

Metric	B0	B1	B2	P
Mission success, %	33.1	38.9	50.2	60.9
Failure-aware makespan, mean	91.68	87.41	79.34	69.30
Successful makespan, median	28.0	33.0	34.5	33.0
Exposure, all runs, mean	9.87	7.73	6.73	23.70
Exposure, successful runs, mean	5.49	6.54	6.92	5.84
Planning time, mean, ms	23.60	25.82	50.78	249.66
Tasks completed, mean	12.88	13.16	13.79	15.04

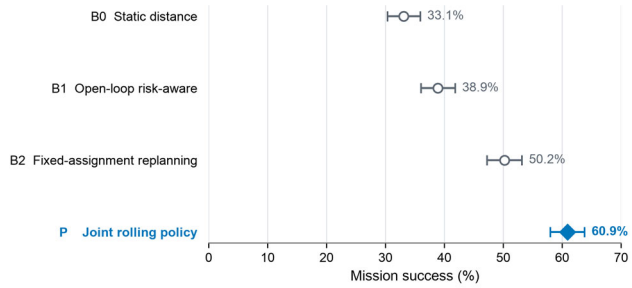


Fig. 1. Mission success by policy across 1080 matched scenarios. Points show single-policy estimates and horizontal lines show Wilson 95% intervals; P is highlighted. These intervals are descriptive, hypothesis decisions use paired differences, and the display has no physical-safety interpretation.

P's unconditional mean exposure was 23.70, much higher than the baselines because it continued difficult scenarios and

completed more targets while some baselines terminated early. This agent-time consequence neither proves benefit nor harm; it motivates reporting success and mutual-success exposure together.

B. Confirmatory Comparisons

The paired results met H1. P−B1 success was 0.220 (95% CI 0.194 to 0.247), above the −0.025 analyst-defined margin. In 411 mutual successes, P−B1 mean exposure was −0.856 normalized units (95% CI −1.231 to −0.500; rank-biserial effect −0.316). The median makespan ratio was 1.000 (95% CI 1.000 to 1.000), within the 1.05 limit; its rank-biserial effect was −0.173. The exact success test gave $p = 2.02 \times 10^{-61}$, and Holm-adjusted Wilcoxon p-values were 4.66×10^{-6} for exposure and 0.0288 for makespan. A median ratio of 1.000 does not imply that every paired difference was zero: the signed-rank test uses the magnitude and direction of nonzero differences across the distribution. H1 rests on frozen effect thresholds and intervals, not the secondary p-values. Fig. 2 shows the paired exposure-difference distributions for P relative to each comparator among mutual successes.

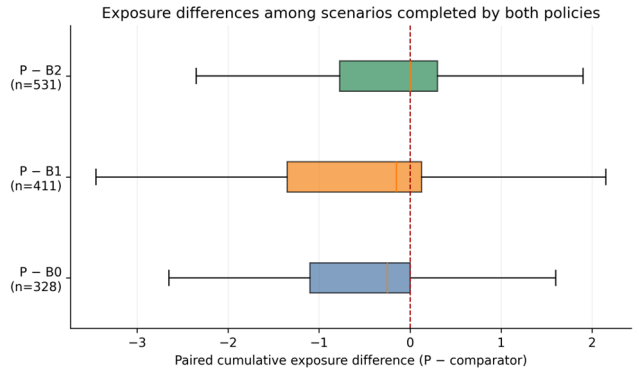


Fig. 2. Conditional P-minus-comparator exposure differences in synthetic scenarios where both policies succeeded. Negative values favor P on unitless field cost; the display does not estimate dose or safety.

H2 was met as a policy-level contrast. Because both P and B2 failed in some scenarios, all-run success was primary. P−B2 success was 0.107 (95% CI 0.087 to 0.128), so its lower bound exceeded zero. In 531 mutual successes, P−B2 exposure was −1.174 units on average (95% CI −1.797 to −0.633; rank-biserial effect −0.151), and the median makespan ratio was 1.000 (rank-biserial effect −0.364). The exact success p-value was 3.61×10^{-26}. This evidence supports the combined addition of completion-event replanning, unfinished-task redistribution, and failed-agent recovery; it does not isolate any one mechanism.

For completeness, P−B0 success was 0.279 (95% CI 0.249 to 0.308). Among 328 mutual successes, exposure was −0.312 units (95% CI −0.596 to −0.031) and the median makespan ratio was 1.000. Because the policies differ in initial risk use, route updates, and allocation, this is only an overall policy contrast.

The P−B2 policy-bundle difference was similar across the three tested spread categories: success differences were 0.111 (95% CI 0.075 to 0.147) in slow spread, 0.103 (0.069 to 0.136) in medium spread, and 0.108 (0.072 to 0.147) in fast spread. These subgroup results describe consistency within the tested grid, not robustness to untested buildings or physical conditions.

C. Sensitivity and Reproducibility

The pre-specified H3 descriptive criterion was met. P's capped median makespan for team sizes 2–8 was 120, 120, 42, 30, 27, 25.5, and 22 steps (Fig. 3). After the transition from capped failure to completion, marginal reductions were 12, 3, and 1.5 steps for sizes 4→5, 5→6, and 6→7. The next improvement rose to 3.5 steps, so returns were not globally monotone. No population-level uncertainty interval was assigned to this finite sensitivity sequence; it cannot justify an operational crew-size recommendation.

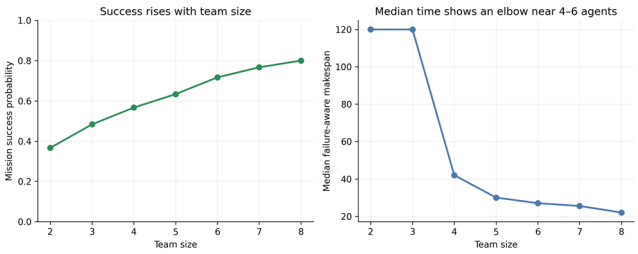


Fig. 3. Descriptive failure-aware median makespan and success across synthetic team sizes; no inferential or operational staffing claim is made.

Across the finite dynamic sensitivity grid, average success for replanning intervals 1, 3, and 5 was 52.5%, 52.5%, and 51.9%. Corresponding exposure means were 30.80, 29.65, and 29.86, while mean planning time declined from 651.40 to 262.45 and 187.34 ms. These rounded aggregates are descriptive and carry no inferential intervals. Updating every step brought little average outcome change but substantial computation; timing remains machine- and implementation-dependent. Fig. 4 shows the planning-time and mission-success trends across the tested replanning intervals and risk weights.

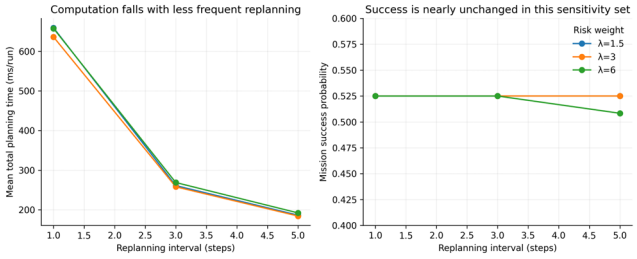


Fig. 4. Descriptive outcome and runtime summaries for synthetic replanning-interval sensitivity; no uncertainty interval or physical-safety meaning is assigned.

The independent rerun contained 4320 records. After stable sorting, all 27 deterministic metadata and outcome columns matched, with zero cell mismatches; only wall-clock time differed. Five tests covering connectivity, seed determinism, avoidance, forecast monotonicity, and cross-policy scenario equality passed. This verifies implementation consistency, not the hazard abstraction.

VI. Discussion

In the sampled stochastic grids, P produced higher success and lower paired-success exposure than open-loop planning while meeting the makespan tolerance. Against fixed-assignment periodic path replanning, its success was 10.7 percentage points higher. The defensible explanation is policy-level: P can update after completion, redistribute unlocked work, and recover tasks from failed agents. The

present design cannot determine how much each feature contributed.

The internal ladder locates changes without establishing state-of-the-art dominance. Adding periodic path updates raised success from B1's 38.9% to B2's 50.2%; adding the P bundle raised it to 60.9%. This is consistent with MRTA's separation of allocation and navigation and with prior completion-triggered failure recovery [6]. It does not establish that the centralized greedy rule is preferable to robust auctions or optimization methods under other costs, objectives, or failure rules.

Unconditional and paired-success exposure answer different questions. The first records all accumulated field cost and rises when a policy keeps agents active; the second compares the same completed cases and favored P. Reporting both prevents early failure from appearing artificially low-cost. Neither represents temperature, toxic dose, injury, or tenability.

The descriptive sensitivities reject a universal "more replanning is better" claim. Intervals 1 and 3 had identical mean success, but interval 1 required roughly 2.5 times the planning time. The team-size curve contained only a local diminishing-return sequence. Without inferential intervals beyond the confirmatory subgroup comparisons, these finite observations cannot prescribe update frequency or staffing.

Several limitations constrain external and comparative validity. The study uses one floor, a regular grid, homogeneous agents, deterministic service time, simple collision handling, a short forecast, and three confirmatory topologies. It omits stairs, communications failure, localization error, congestion physics, equipment, fatigue, suppression, victims, and command constraints. The monotone stochastic field is not calibrated against experiments or FDS, observation noise is not a sensor model, and the greedy allocation can miss better schedules. B0–B2 are internal ablations; no robust-auction or optimization baseline was tested. Wall-clock results depend on the execution environment, and the focused literature review cannot establish exhaustive novelty.

The next experiment should separate the P bundle with a policy that redistributes tasks without completion-event replanning and should add a robust repeated-auction or optimization comparator. Later validation can introduce realistic floor graphs, calibrated fire and smoke fields, heterogeneous capabilities, communication constraints, and congestion-aware motion. The matched-scenario analysis should be retained. Any safety-relevant claim would additionally require physical validation, human-factors review, and operational testing.

VII. Conclusion

This study converted an underdeveloped route-planning proposal into a paired simulation benchmark. Across 1080 matched synthetic scenarios, P exceeded three reduced-capability study baselines in mission success. It met the frozen H1 success, paired-success exposure, and makespan conditions against B1, while the P−B2 contrast supported the combined addition of completion-event updating, unfinished-task redistribution, and failed-agent recovery. Descriptive sensitivity analyses exposed local returns and computational tradeoffs, and an independent rerun exactly reproduced deterministic outputs. The conclusion is narrow: the joint policy bundle improved this implemented dynamic-cost search problem. Its individual mechanisms, superiority to stronger MRTA solvers, and real-fire usefulness remain open questions.

Data and Code Availability

Run-level confirmatory, sensitivity, and rerun CSV files; configurations; hashes; simulation and analysis code; tables; and figures are included in the project directory. results/confirmatory/runs.csv is the confirmatory source of truth, and results/reproducibility/runs.csv is the independent rerun.

Ethics Statement

The study used synthetic maps, agents, and stochastic fields and involved no people, animals, personal data, or operational deployment.

AI Assistance Disclosure

An OpenAI language model assisted with question refinement, literature organization, methodological critique, simulation and analysis code drafting, statistical interpretation, figures, and manuscript composition. Bibliographic identities were checked against DOI, publisher, proceedings, institutional, or government records. Reported numerical results came from executed code and an independent rerun. The author must inspect the code, read centrally cited full texts, verify interpretations, comply with institutional and venue disclosure rules, and assume final scholarly responsibility.

Responsible-use Statement

The software is a research prototype for simulated dynamic-cost search. It is not a certified incident-command tool, does not predict real tenability, and must not direct emergency operations.

Conflict of Interest

The author declares no conflict of interest.

Author Contributions

Yiming Wang: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Validation, Visualization, Writing—original draft, and Writing—review & editing. The authors approved the final manuscript and accepts responsibility for all aspects of the work.

Reference

[1] J. Wang, H. Zhao, and S. Winter, "Integrating sensing, routing and timing for indoor evacuation," *Fire Safety Journal*, vol. 78, pp. 111–121, 2015. doi: 10.1016/j.firesaf.2015.08.009

[2] M. Choi and S. Chi, "Optimal route selection model for fire evacuations based on hazard prediction data," *Simulation Modelling Practice and Theory*, vol. 94, pp. 321–333, 2019. doi: 10.1016/j.simpat.2019.04.002

[3] G.-R. Shih and P.-H. Tsai, "Safest-path planning approach for indoor fire evacuation," *International Journal of Disaster Risk Reduction*, vol. 93, 103760, 2023. doi: 10.1016/j.ijdrr.2023.103760

[4] 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*, vol. 23, no. 9, pp. 939–954, 2004.

doi: 10.1177/0278364904045564
[5] G. A. Korsah, A. Stentz, and M. B. Dias, "A comprehensive taxonomy for multi-robot task allocation," *The International Journal of Robotics Research*, vol. 32, no. 12, pp. 1495–1512, 2013. doi: 10.1177/0278364913496484
[6] M. Nanjanath and M. Gini, "Repeated auctions for robust task execution by a robot team," *Robotics and Autonomous Systems*, vol. 58, no. 7, pp. 900–909, 2010. doi: 10.1016/j.robot.2010.03.011
[7] J. Wei, C. Cui, and X. Yang, "Dynamic firefighting route planning for efficient evacuation in complex subway stations: A deep learning-enhanced robust optimization approach," *Simulation Modelling Practice and Theory*, vol. 146, 103223, 2026. doi: 10.1016/j.simpat.2025.103223
[8] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968. doi: 10.1109/TSSC.1968.300136
[9] S. Koenig and M. Likhachev, "D* Lite," in *Proc. 18th Nat. Conf. Artificial Intelligence (AAAI)*, 2002, pp. 476–483. https://publications.ri.cmu.edu/d-lite
[10] G. Sharon, R. Stern, A. Felner, and N. R. Sturtevant, "Conflict-based search for optimal multi-agent pathfinding," *Artificial Intelligence*, vol. 219, pp. 40–66, 2015. doi: 10.1016/j.artint.2014.11.006
[11] T. Bektaş, "The multiple traveling salesman problem: An overview of formulations and solution procedures," *Omega*, vol. 34, no. 3, pp. 209–219, 2006. doi: 10.1016/j.omega.2004.10.004
[12] M. G. Lagoudakis *et al.*, "Auction-based multi-robot routing," in *Proc. Robotics: Science and Systems I*, 2005. doi: 10.15607/RSS.2005.I.045
[13] M. B. Dias, R. Zlot, N. Kalra, and A. Stentz, "Market-based multirobot coordination: A survey and analysis," in *Proc. of the IEEE*, vol. 94, no. 7, 2026, pp. 1257–1270. doi: 10.1109/JPROC.2006.876939
[14] H. Chakraa, F. Guérin, E. Leclercq, and D. Lefebvre, "Optimization techniques for multi-robot task allocation problems: Review on the state-of-the-art," *Robotics and Autonomous Systems*, vol. 168, 104492, 2023. doi: 10.1016/j.robot.2023.104492
[15] V. Pillac, M. Gendreau, C. Guéret, and A. L. Medaglia, "A review of dynamic vehicle routing problems," *European Journal of Operational Research*, vol. 225, no. 1, pp. 1–11, 2013. doi: 10.1016/j.ejor.2012.08.015
[16] H. N. Psaraftis, M. Wen, and C. A. Kontovas, "Dynamic vehicle routing problems: Three decades and counting," *Networks*, vol. 67, no. 1, pp. 3–31, 2016. doi: 10.1002/net.21628
[17] Y. Emam, S. Wilson, M. Hakenberg, U. Münz, and M. Egerstedt, "A receding horizon scheduling approach for search & rescue scenarios," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 3451–3456, 2020. doi: 10.1016/j.ifacol.2020.12.1680
[18] J. G. Martin, M. Hanif, T. Hatanaka, J. M. Maestre, and E. F. Camacho, "Predictive receding-horizon multi-robot task allocation with moving tasks," in *Proc. 2022 European Control Conference (ECC)*, 2022, pp. 2030–2035. doi: 10.23919/ECC55457.2022.9838127
[19] J. G. Martin, M. Hanif, T. Hatanaka, J. M. Maestre, and E. F. Camacho, "Predictive receding-horizon multi-robot task allocation applied to the mapping of direct normal irradiance in a thermosolar power plant," *Solar Energy*, vol. 263, 111911, 2023. doi: 10.1016/j.solener.2023.111911
[20] S. Choudhury, J. K. Gupta, M. J. Kochenderfer, D. Sadigh, and J. Bohg, "Dynamic multi-robot task allocation under uncertainty and temporal constraints," *Autonomous Robots*, vol. 46, no. 1, pp. 231–247, 2022. doi: 10.1007/s10514-021-10022-9
[21] R. Bhardwaj, A. Bhargava, and V. Kumar, "INFED: Enhancing fire evacuation dynamics through 3D congestion-aware indoor navigation framework," *Simulation Modelling Practice and Theory*, vol. 136, 103010, 2024. doi: 10.1016/j.simpat.2024.103010
[22] J. K. N. Tan, S. Zhang, A. W.-K. Law, and S. H. Cheung, "Digital-twin enabled evacuation to improve individual and community resilience of building occupants against indoor fires: Framework and route-finding algorithm," *Developments in the Built Environment*, vol. 22, 100672, 2025. doi: 10.1016/j.dibe.2025.100672
[23] W. Jiang *et al.*, "Modelling of wildland-urban interface fire spread with the heterogeneous cellular automata model," *Environmental Modelling & Software*, vol. 135, 104895, 2021. doi: 10.1016/j.envsoft.2020.104895
[24] K. B. McGrattan, R. J. McDermott, C. G. Weinschenk, and G. P. Forney, "Fire Dynamics Simulator user's guide," *NIST Special Publication*, vol. 1019, no. 6, pp. 1–339, 2013. https://www.nist.gov/publications/fire-dynamics-simulator-users-guide-sixth-edition
[25] National Institute of Standards and Technology. (2026). FDS-SMV manuals (FDS 6.11.1; SMV 6.11.2). [Online]. Available: https://pages.nist.gov/fds/manuals.html

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).