World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:7, No:11, 2013

Genetic Algorithm for In-Theatre Military Logistics Search-and-Delivery Path Planning

Jean Berger, Mohamed Barkaoui

Abstract**—**Discrete search path planning in time-constrained uncertain environment relying upon imperfect sensors is known to be hard, and current problem-solving techniques proposed so far to compute near real-time efficient path plans are mainly bounded to provide a few move solutions. A new information-theoretic –based open-loop decision model explicitly incorporating false alarm sensor readings, to solve a single agent military logistics search-and-delivery path planning problem with anticipated feedback is presented. The decision model consists in minimizing expected entropy considering anticipated possible observation outcomes over a given time horizon. The model captures uncertainty associated with observation events for all possible scenarios. Entropy represents a measure of uncertainty about the searched target location. Feedback information resulting from possible sensor observations outcomes along the projected path plan is exploited to update anticipated unit target occupancy beliefs. For the first time, a compact belief update formulation is generalized to explicitly include false positive observation events that may occur during plan execution. A novel genetic algorithm is then proposed to efficiently solve search path planning, providing near-optimal solutions for practical realistic problem instances. Given the run-time performance of the algorithm, natural extension to a closed-loop environment to progressively integrate real visit outcomes on a rolling time horizon can be easily envisioned. Computational results show the value of the approach in comparison to alternate heuristics.

Keywords**—**Search path planning, false alarm, search-and-delivery, entropy, genetic algorithm.

I. Introduction

In combat service support or emergency logistics, the final stage of service/supply delivery on the battlefield/devastated zone known as the "last tactical mile" is undoubtedly very risky. Emerging concepts for military combat service support relies on tactical airborne technology such as versatile and agile unmanned systems to deliver critical supply. The multi-role nature of unmanned aerial vehicles (UAV) to support cargo transportation, reconnaissance and precise tactical delivery tasks in an uncertain hostile and dynamic environment is such an example. Tactical mission delivery typically consists in a supply vehicle (service provider) transporting cargo (tactical supply such as commodities/ammunition) from the forward operating base (FOB) or Battle group to a military unit destination area. Once in the proximity of final destination, the service provider, must locally search and locate the customer unit location to finally deliver the requested supply and then safely return to the FOB.

Military tactical search and delivery mission is roughly pictured in Fig. 1.

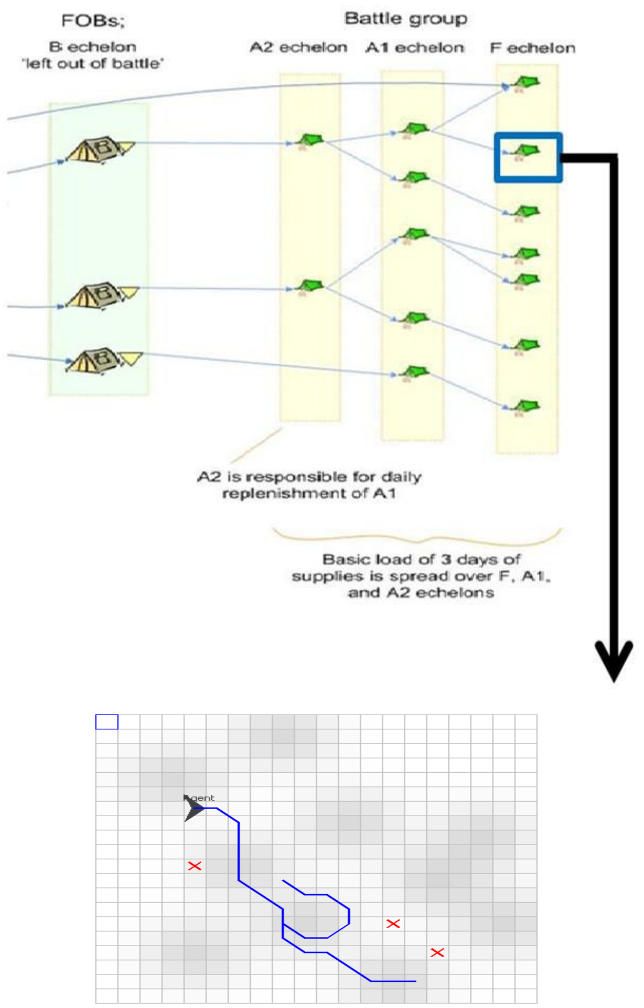

Fig. 1 - In-Theatre Military Logistics Search-and-Delivery Path Planning

As the situation is dynamic and communication bandwidth is limited for security reasons, customer unit location cannot be precisely known in advance by the supply vehicle. As a result, once in the surroundings of the final destination, the challenge is to achieve precise delivery to the targeted unit as precise location may possibly change due a dynamic hostile and uncertain disruptive environment. However, given latest

Jean Berger is with the Defence Research and Development Canada – Valcartier, Quebec City, Qc, Canada (phone: 418-844-4000 4645; fax: 418-844-4538; e-mail:jean.berger@drdc-rddc.gc.ca).

Mohamed Barkaoui is with the Defence Research and Development Canada – Valcartier, Quebec City, Qc, Canada (phone: 418-844-4000 4645; e-mail: mohamed.barkaoui@drdc-rddc.gc.ca).

adversarial actions, intelligence, unit mission plans, expected movement, and gross estimate of required travel time in Theatre, probability distribution of unit presence over the destination area may be derived and exploited by the supply vehicle on duty to carry out the “last tactical mile”. Given the size, speed and high mobility/agility, endurance and autonomy of an unmanned vehicle, the benefits shown by this technology clearly outweigh the cost and risk incurred by traditional manned vehicles demanding continual human expertise and support while operating under stringent temporal constraints. This paper focuses on the search component of the search-and-delivery problem when detecting the unit target at the final destination under imperfect sensing conditions involving false positive and negative detections.

Search path planning work explicitly accounting for false alarm detections over realistic time horizons remains largely marginal, given the additional computational complexity introduced by this source of uncertainty. Other approaches simply overlook this intrinsic difficulty. For instance, in basic discrete optimal searcher path problems [1], [2], the advocated approach tends to unrealistically oversimplify the problematic, primarily focusing on false negative (overlook) observations only, while ignoring false alarm detections (false positives) to conveniently ensure that the searched target be necessarily found upon positive observation. The underlying discrete optimal searcher path problem involving a stationary target is known to be NP-Hard [3]. Unfortunately, the whole methodology based on probability of success is no longer applicable for false alarm detection problem attributes, as a positive observation can no longer guarantee target detection with certainty. Contributions on decision models and search algorithms focusing more explicitly upon false-positive and false-negative detections are reported in [4]-[8]. However, proposed approaches mostly exploit specific problem characteristics or rely on key constraints relaxation to keep the problem tractable. For instance, in [5], the authors aim at minimizing detection time assuming constant false alarm detection rates for all cells. They do assume a constant false positive/negative detection rates over the grid area to be searched, disregarding terrain features and condition variability. Most importantly, search deadline as an itinerary constraint to successfully find the target is overlooked as well, which turns out to be unrealistic in our setting due to a resource-bounded agent capability (e.g. resource, fuel, flight personnel). The relaxation of search/flight time constraints proves indeed very convenient, to easily and naturally define a simple detection time minimization objective and a stopping criterion to interrupt the search effort when updated target occupancy belief evolve outside predetermined threshold values. However, the approach may require a significant search time in practice. Consequently, when a search time limit is naturally imposed by problem domain contingencies, uncertainty minimization must rather be considered. A similar situation occurs for the sequential eliminating procedure proposed in [6]. Leaning toward the optimization of detection thresholds, the procedure may require a prohibitive run-time, and fails to efficiently exploit a pre-set time limit to optimize search path planning. In other respect, general search-theoretic methods propose procedures mainly inspired from branch and bound, [9], [7], [10], [11] variants. More recent work [8] uses combinations of various greedy search strategies to further reduce complexity. However, despite the development of innovative problem-solving techniques, the problem generally remains computationally hard to solve in near real-time over reasonable or even limited time horizons. Alternatively, Markov decision processes (MDP) modeling represents a common approach to handle false-negative and false-positive observations, resorting to exact and approximate methods to manage uncertainty. But the computational complexity of exact problem-solving techniques used to support sequential decision-making scales exponentially with time horizon. The underlying dynamic programming [12]-[15] and tree–based search techniques [16] may suitably perform under specific constraints but eventually face the curse of dimensionality, presenting poor scalability even for modest size problems. Recent comprehensive surveys on target search problems from search theory and artificial intelligence/ robotic control perspectives may be found in [17], [12], [5] and [13] respectively.

In this paper, a new open-loop information-theoretic –based decision model formulation explicitly accounting for false alarm sensor readings is proposed to solve a single agent military logistics search-and-delivery path planning problem with anticipated feedback. In this context, ‘open-loop with anticipated feedback’ alludes to offline planning while incorporating projected agent observations (visit outcome projection) as opposed to real sensor readings. Anticipated feedback enhances pure open-loop formulations naturally ignoring information feedback, to significantly improve solution quality, while mitigating computational complexity related to costly closed-loop problem formulations. The expected entropy minimization decision model captures uncertainty associated with anticipated false positive (false alarm) observation events over possible scenarios. Projected belief update on unit target occupancy is therefore generalized to explicitly capture possible false positive observation events that could occur during plan execution. In order to efficiently solve the search path planning problem, a novel genetic algorithm is proposed, providing near-optimal solutions for practical realistic problem instances. The algorithm presents new recombination and mutation operators exploiting basic path reconstruction principles. Run-time performance even paves the way to a closed-loop environment extension in which real visit outcome from the previous episode can be dynamically integrated in real-time to update target occupancy belief distribution. In that setting, a complete path solution is gradually expanded, exploiting real observation outcomes, by solving periodically new problem instances over a rolling horizon. Performance comparison over alternate procedures demonstrates the value of the approach.

The remainder of the paper is structured as follows. Section II first introduces problem definition, describing the main characteristics of the open-loop search path planning problem with anticipated feedback. Then a new information-theoretic path planning decision model formulation is presented in Section III. It captures path plan uncertainty (entropy) through projected future agent visits (plan execution) over possible sensor observation outcomes and discusses how path planning could be extended to a dynamic setting using a rolling time horizon. Section IV describes the genetic algorithm proposed to efficiently compute a near-optimal solution. Genetic operators inspired from simple path reconstruction principles are briefly presented. Computational results comparing the value of the proposed method to an alternate myopic problem-solving technique are reported and discussed in Section V. Finally, a conclusion is given in Section VI.

II. Problem Definition

A. General Description

In a military tactical logistics context, the search and delivery path planning problem involves a vehicle supply agent (searcher) searching a targeted customer unit (target) in a bounded environment over a given time horizon in order to deliver supply or services. From a search mission perspective, the agent's goal consists in searching an area (grid) to minimize target cell occupancy uncertainty (entropy), given a prior cell occupancy probability distribution and imperfect sensing capabilities to successfully servicing the unit target. Represented through a grid, the search region characterizes an area defined as a set of cells *N*, describing possible target locations. Presumably occupying a single cell, the precise location of the target is assumed unknown. A prior target location probability density distribution for which cell occupancy probabilities sum up to one can be derived from domain knowledge. It reflects possible individual cell occupancy, defining a grid cognitive map or uncertainty grid. The cognitive map constitutes a knowledge base describing a particular world state, including variables such as target occupancy belief distribution, time, agent position and orientation. An example of a cognitive map is illustrated in Fig. 2 at a specific point in time.

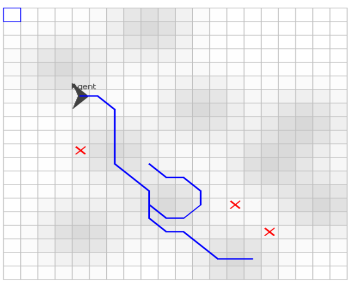

Fig. 2 Uncertainty grid /cognitive map at time step *t*. Belief magnitudes are represented through multi-level shaded cell areas. Projected agent plans are displayed as possible paths

The duration of a cell visit or service time is assumed constant, specifying the period of each episode. An agent can legally move toward its neighboring cells offering eight alternate possible directions at each time step. A search path solution consists in constructing an agent path plan selecting base-level control action moves to minimize uncertainty about target location on the entire grid in order to finally deliver supply or service to the most likely location.

B. Observation Model

Modeling partial world state agent observability, the observation model governs agent sensor's perception. During episode *t*, an agent visits a cell searching for target occupancy. A sensor reading z_t at time *t* may then be either positive (z_t=1) or negative (z_t=0)as determined through a probabilistic observation model. The latter accounts for uncertainty through conditional probability of detection and false alarm, given cell target vacancy or occupancy state $X \in \{0,1\}$ respectively:

z_t: cell occupancy observation at the end of period *t*

$p_c = p(z_t = 1 | X = 1)$ probability of correct observation

$p_f = p(z_t = 1 | X = 0)$ probability of false alarm (false positive)

$p(z = 0 | X = 1) = 1 - p_c$ false negative probability (overlook)

$p(z = 0 | X = 0) = 1 - p_f$ true negative probability

$$\alpha = \frac{p_f}{p_c}, \quad \beta = \frac{1-p_f}{1-p_c}$$

These parameters are understandably cell-dependent reflecting specific sensor sensitivity and terrain features and conditions (e.g. landscape, obstacles, clutter, visibility, luminosity). Agent sensor's range defining visibility or footprint (coverage of observable cells given the current sensor position) is limited to the cell being searched.

C. Bayesian Filtering

From a real or anticipated agent sensor observation, local cell target occupancy belief (*p(X=1)*) (real or anticipated) can be updated using Bayesian filtering:

$$p_t(X \mid z_t) = \frac{p(z_t \mid X)\, p_{t-1}(X)}{p(z_t)} \tag{1}$$

where

$$p(z_t) = \sum_{x \in \{0,1\}} p(z_t \mid X = x)\, p_{t-1}(X = x) \tag{2}$$

In (2), p_{t-1} and p_t(t> 0) refer to prior and posterior cell target occupancy probability (belief) respectively.

D. Entropy

The objective consists in constructing a plan modeled as a sequence of moves to minimize entropy (target occupancy uncertainty) over the entire grid and horizon *T*. From information theory [18], the entropy function *E* is defined as:

$$E = -\sum_{x \in \{0,1\}} p(x) \log_2 p(x) \tag{3}$$

where $p(x)$ specifies the current probability/belief of cell target occupancy, and *x* a binary cell occupancy state. A cell with a zero entropy value means absolute certainty about occupancy or vacancy, whereas a maximum entropy value (1) refers to complete uncertainty as depicted in Fig. 3.

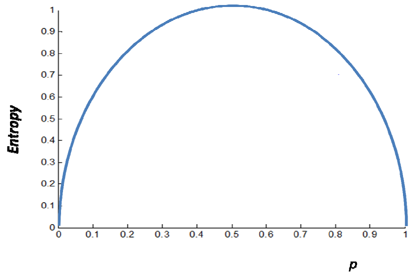

Fig. 3 Entropy defined as a function of probability *p*. It reflects probabilistic uncertainty about a state variable

III. Problem Model Formulation

A. Mathematical Modeling

A mathematical open-loop formulation with anticipated feedback is proposed for the discrete stationary target search and delivery path planning problem. It is based on expected entropy of target cell occupancy over multiple scenarios of possible anticipated observation outcomes, resulting from a *T*-visit agent path across the grid.

The variables and parameters defining the decision model are given as follows:

N: set of cells defining the grid search area

T: mission time horizon

V: set of visited cells, composing the path solution

V_i: maximum number of visits on cell i

$\{s_m, l_m\}_{m\in V}$: assuming l_m visits on cell m, s_m refers to the number of positive sensor observations recorded (number of success) by T.

$\overline{E}_{l_1 l_2 .. l_N}$: objective function defining expected entropy for a path solution involving l_j visits on cell j ({s_j positive and l_j-s_j negative observations}) for all cells such as $\sum_{j\in N} l_j = T$

X_i : occupancy state of a cell i (1: occupancy, 0: vacancy)

$p_{iT}(X_i|\{s_m,l_m\}_{m\in V})$: posterior probability of target unit cell occupancy by the end of period T: Conditional probability at time step T to find cell i in occupancy state X_i given visited cells V (path) and a specific anticipated scenario/history {s_m, l_m} in which s_m positive observations out of l_m visit sensor readings on cell m were recorded.

$p_{j0} = p_0(x_j = 1)$: initial probability of target unit cell occupancy

$E_T(\{p_{iT}(X_i|\{s_m,l_m\}_{m\in V})\}_{i\in N})$: system entropy by the end of period T for path cell visits $\{l_m\}_{m\in V}$ for the scenario $\{s_m, l_m\}_{m\in V}$

δ_{ij}: Kronecker delta function (it returns 1 if i=j, 0 otherwise)

The proposed information-theoretic decision model may be formulated as follows:

$$\min_{\{path\}} \overline{E}_{l_1 l_2 \ldots l_N} = \sum_{j\in N} p_0(x_j=1)\left[\begin{array}{l} (\prod_{m\in V}\sum_{s_m=0}^{l_m}) p(s_m|l_m,\{x_i=\delta_{ij}\}_{i\in N})\times \\ \quad E_T(\{p_{iT}(X_i|\{s_m,l_m\}_{m\in V})\}_{i\in N}) \end{array}\right]$$

$$= \sum_{j\in V} p_{j0}\left[(\prod_{m\in V}\sum_{s_m=0}^{l_m}) p(s_m|l_m,x_m=\delta_{mj}) E_T(\{p_{iT} X_i|\{s_m,l_m\}_{m\in V})\}_{i\in N}) \right]$$

$$+ \sum_{j\notin V} p_{j0}\left[(\prod_{m\in V}\sum_{s_m=0}^{l_m}) p(s_m|l_m,x_m=0) E_T(\{p_{iT}(X_i|\{s_m,l_m\}_{m\in V})\}_{i\in N}) \right] \quad (4)$$

$$= \sum_{j\in V} p_{j0}\left[(\prod_{m\in V}\sum_{s_m=0}^{l_m}) p(s_m|l_m,x_m=\delta_{mj}) E_T(\{p_{iT}(X_i|\{s_m,l_m\}_{m\in V})\}_{i\in N}) \right]$$

$$+ (1-\sum_{j\in V} p_{j0})\left[(\prod_{m\in V}\sum_{s_m=0}^{l_m}) p\ (s_m|l_m,x_m=0) E_T(\{p_{iT}(X_i|\{s_m,l_m\}_{m\in V})\}_{i\in N}) \right] \quad (5)$$

in which entropy is averaged over all possible scenarios resulting from anticipated path visit outcomes, that is over all sequence of possible positive/negative observations. The objective is to determine the best path minimizing expected entropy. In (4) the following equivalent notation is implicit:

$$\sum_{s_1=0}^{l_1}\sum_{s_2=0}^{l_2}\cdots\sum_{s_N=0}^{l_N} \sim \left(\prod_{m\in V}\sum_{s_m=0}^{l_m}\right)$$

The alternate form (5) separates the expression in two components considering whether target location is either part of the visited cells or not. As a result, the objective may be reformulated in terms of visited cells only. The probability of a specific scenario for a given path solution (visited cells) is the product of probability of individual cell visit scenarios. Assuming target location in cell j, such probability for a cell k involving s_k positive detections out of $l_k$visits is given by:

$$p(s_k|l_k,\{x_i=\delta_{ij}\}_{i\in N}) = \binom{l_k}{s_k} p^{s_k}(z=1|x_k=\delta_{kj})\, p^{l_k-s_k}(z=0|x_k=\delta_{kj}) \quad (6)$$
$$= p(s_k|l_k,x_k=\delta_{kj})$$

Using the observation model introduced in the previous section and conditional independence, one can easily compute the probability for each visited cell scenario. As the target is located in a unique cell of the search area, we do exploit the following relationships as well:

$$\sum_{j\in N} p_0(x_j=1)=1 \qquad \sum_{j\in N} x_j = 1$$

Bayesian belief update accounting for false alarm and overlook (false negative) observations by the end of period T is therefore explicitly expressed as follows:

$$p_{iT} = p_{iT}(X_i|\{s_i, l_i\}) = \frac{p_{i0}\dfrac{1}{\alpha_i^{s_i}\beta_i^{l_i-s_i}}}{1-\sum_{j\in V} p_{j0}\left(1-\dfrac{1}{\alpha_j^{s_j}\beta_j^{l_j-s_j}}\right)} = \frac{p_{i0}\dfrac{1}{\alpha_i^{s_i}\beta_i^{l_i-s_i}}}{\sum_{j\in V} p_{j0}\dfrac{1}{\alpha_j^{s_j}\beta_j^{l_j-s_j}}} \quad (7)$$

Belief magnitude increases on positive (z=1) and decreases on negative (z=0) observation readings respectively, while being normalized over cell beliefs. One can easily see that dissimilar outcome observations (z_t=0, $z_{t'}$=1) adversely affect belief evolution. α refers to the occurrence of positive detections while β corresponds to negative observation events. Equation (7) generalizes the traditional false alarm–free belief update scheme, and can be formally proved by induction. It is the first time to our knowledge that a general belief update form is formulated explicitly, departing from the recurrent form over consecutive time steps generally proposed and largely used and reported in open literature.

It is worth noticing that the expected entropy objective function is invariant with respect to cell visits order (path ordering) and strictly depends on visited cells and its related posterior probability of target occupancy. This is due to the fact that final entropy only relies on belief updates, and in particular on the number of visits performed on a given cell, and not when those visits took place as shown in (7).

The expected entropy model can be further approximated by the following model to keep the number of terms in the objective function linear over T (as opposed to an exponential number of contributions, accounting for all possible observation scenarios):

$$\bar{E}_{l_1 l_2 .. l_N} \approx \sum_{j\in V}\left[\begin{array}{l}\sum_{s_j\in\{0,1,l_j-1,l_j\}}\left(p_{j0}pr(s_j|l_j,x_j=1)+\dfrac{1-\sum_{j\in V}p_{j0}}{|V|}pr(s_j|l_j,x_j=0)\right)\times \\ \underline{\prod_{m\in V/\{j\}}}(pr(s_m=0|l_m,x_m=0) \\ \times E_T(\{p_{iT}(X_i|\{s_m=0,l_m\}_{m\in V/\{j\}},s_j,l_j)\}_{i\in N})\end{array}\right] \quad (8)$$

where

$$pr(s_m|l_m, x_m=\delta_{mj}) = \frac{p(s_m|l_m, x_m=\delta_{mj})}{\sum_{s_q\in\{0,1,l_q-1,l_q\}} p(s_q|l_q, x_q=\delta_{qj})} \quad (9)$$

The main idea is to emphasize the most likely conditional events only, restricting the set of values for s_j to its most probable values.

Further simplification consists in focusing on the presumed target location cell j only, while assuming likely negative observations for all other visited cells (i.e. probability of negative observations equals one) leading to the following $|V|$-term expression:

$$\begin{aligned}\bar{E}_{l_1 l_2 ....l_N} &\approx \sum_{j\in N} p_{j0}\left[(\prod_{m\in V}1)E_T(\{p_{iT}(X_i|\{s_m=l_m\delta_{mj},l_m\}_{m\in V})\}_{i\in N})\right] \\ &\approx \sum_{j\in V} p_{j0}\left[E_T(\{p_{iT}(X_i^{\tau}|\{s_m=l_m\delta_{mj},l_m\}_{m\in V})\}_{i\in N})\right] \\ &\quad +(1-\sum_{j\in V}p_{j0})\,E_T(\{p_{iT}(X_i^{\tau}|\{s_m=0,l_m\}_{m\in V})\}_{i\in N})\end{aligned} \quad (10)$$

In this case, the belief update scheme considers imperfect sensors while the objective function does not.

B. Dynamic Planning and Time Horizon

A path can be computed dynamically over receding horizons by using real information feedback as it becomes available and progressively perform a new optimization to periodically improve solution quality. Accordingly, large time horizon problems are solved through repeated fast subproblem optimizations including real information feedback, over receding horizons as pictured in Fig. 4. Time horizonis partitioned in time intervals and corresponding subproblems are sequentially solved over respective episodes of period ΔT. A subproblem solution periodically expands the overall current partial solution by integrating a small local path segment (subperiod δT), while updating the objective function to properly reflect new path move contributions. Limited new move insertions at each time step define overlapping episodes, mitigating the effects of myopic path planning. A new open-loop subproblem is then periodically solved subject to the partial solution constructed so far. The process is then reiterated until the time horizon has been completely covered. The strategy consists in taking advantage of the fast computation of reasonable time horizon subproblems over a limited number of episodes to quickly compute a near optimal solution to the original problem.

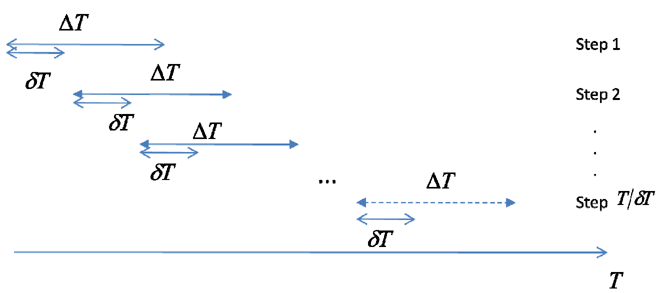

Fig. 4 A large time horizon T is defined over $T/\delta T$ receding horizons of period ΔT. Moves computed in subperiods δT form the final path solution to the original problem

IV. Genetic Algorithm

The searching agent evolves a population of individual solutions through natural selection, recombination and mutation mechanisms over successive generations. Individuals are first represented as chromosomes encoding a feasible path plan corresponding to a sequence of intended actions (physical moves a_{t+1} ... a_{t+T}, each referring to 8 possible directions: *E, NE, N, NW, W,SW, S, SE*) over a time horizon *T* (Fig. 5) for a given episode *t*.

a_{t+1}	a_{t+2}	a_{t+3}	…	a_{t+T-1}	a_{t+T}

Fig. 5 Individual path plan representation at time step *t*

At each generation, the steady-state evolutionary approach consists in using genetic operators to expand the population by λ offspring, and later suppress the worst λ individuals. Recombination and mutation operators are sequentially applied with probability p_{Xover} and p_{Mut} respectively, until all λ new individuals are generated. Those parameters are chosen in order to balance search intensification and diversification. The whole process is repeated until a convergence criterion/condition is met (e.g. maximum number of generations, a maximum run-time or a threshold in solution quality improvement). The algorithm is outlined as follows:

gen=0
Repeat for each new generation
  Evolve Population - build a new generation
    generate λ new offspring using genetic operators (selection, recombination, mutation)
    evaluate fitness of new individuals
    eliminate the λ worst individuals of the expanded population
  gen=*gen*+1
Until(*gen* = *max_gen*)

Return (best computed path plan from Population)

The initial population of path plan individuals is generated randomly, selecting actions proportionally to projected information gain over single-step horizons.

A. Fitness

Fitness characterizes the predisposition of an individual to reproduce. Fitness is defined in terms of expected information gains (differential entropy contributions) while overlooking potential benefits that could result from using intermediate outcomes whenever available. Accordingly, individual reward refers to local information gain or entropy reduction expected by projecting path plan execution over time. It can be expressed as the difference between current entropy (E_0) and final expected entropy ($\overline{E}_T$).An individual fitness is based upon approximate entropy defined by (8).

B. Selection

Fitness values are sorted and normalized using a linear ranking scheme to better control selection pressure [19]. Individual mating is then based upon a fitness-proportional scheme [20].

C. Recombination/Crossover

This genetic operation recombines chromosomes from two selected parents in order to create a child. The proposed recombination operator X_path breeds two parent individuals identifying suitable crossover points and generates an offspring by connecting together head and tail path segments inherited from both parents respectively, truncating control actions when the chromosome length exceeds the planning horizon *T* (see Fig. 6) or appending missing control actions to complete the solution whenever necessary. If parents P1 and P2 have intersecting points, two children are generated by exchanging respective parent sub-segments. Otherwise, closest points x1 (from P1) and x2 (from P2) separating both path parents by the shortest distance are first determined. Should multiple points exist, they would be selected stochastically, biased toward the largest probability values. A child is generated by combining the first segment from P1, to a subroute connecting x1 to x2 by the most direct way and then link to the second segment from P2. In connecting x1 to x2, route construction is biased toward cells with large beliefs, selecting moves with maximum one-step information gain using (10) to minimize run-time computation. Resulting path exceeding *T* moves are removed. Alternatively, if required, missing moves are appended to the end of the route to complete the path using the same information gain -biased procedure to add one cell at a time. A second child can be generated the same way by swapping both parents.

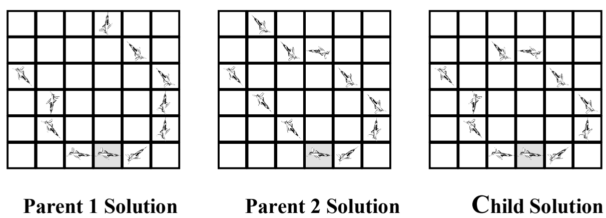

Fig. 6 Crossover operator X_path mating Parent 1 and 2 to generate a new child solution. Parent trajectories are shown to intersect at a cross-over point depicted by the shaded cell. The last control action inherited from Parent 2 is deleted to maintain solution consistency

D. Mutation

Mutation is a natural evolution process in which some individual's genes are randomly modified. Two genetic mutators are proposed, namely, M_path1 and M_path2.

M_path1 first randomly selects a specific visit (*t*> 2) along a path solution, preserving the first segment while reconstructing a new feasible solution from that move on, one cell at a time as shown in Fig. 7. The last path segment reconstruction is information gain -biased, whereby move selection probability is proportional to expected one-step information gain.

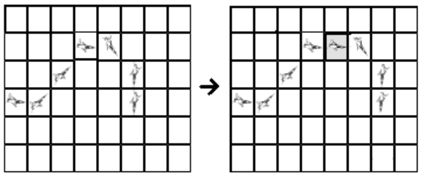

Fig. 7 M_path1 mutation: a move from an individual solution is randomly selected and mutated. A new solution is then reconstructed from that point on (shaded move)

The M_path2 mutation operator randomly removes a short path fragment (some consecutive visits, a few steps from the path endpoint) from a solution and constructs an alternate subpath to locally repair or bridge both disconnected components. The mutator is pictured in Fig. 8. Once again, subpath reconnection is information gain -biased. Path moves are then removed if required.

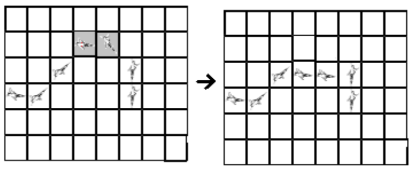

Fig. 8 M_path2 mutation: a subpath segment is removed (shaded area) resulting in two disconnected components. A new solution is then reconstructed linking both orphan components through an alternate subroute

V. Computational Experiment

A computational experiment has been conducted to test the approach over a variety of scenarios. The value of the proposed algorithm is assessed in terms of performance gap and compared to amyopic heuristic. The approach is also measured against solutions using an alternate objective function called covariability reflecting belief dispersion. Comparative solution performance for respective methods/ approaches are reported against relative expected information gain (IG) defined as follows:

$$IG = \frac{\max(\overline{E}_T^{S_1}, \overline{E}_T^{S_2}) - \min(\overline{E}_T^{S_1}, \overline{E}_T^{S_2})}{E_0 - \min(\overline{E}_T^{S_1}, \overline{E}_T^{S_2})} \qquad (11)$$

where S_1 and S_2 correspond to computed solutions from methods 1 and 2 respectively. The larger the gap, the better the relative performance.

A. Myopic Algorithm

The limited look-ahead method consists in myopically planning moves k steps ahead of time, visiting the closest admissible neighbor cell providing the largest gain over a k-step horizon. Accordingly, at each time step, the move with the largest reward is selected first, and the objective function updated accordingly. The procedure is then reiterated for each episode over time horizon $|T|$. Run-time $\in O(|T^k|)$. In our experiment cases for k=1, 2 are explored.

B. Covariability Measure

The algorithm is also measured against solutions using an alternate objective function called covariability (COV) which somewhat reflects belief dispersion and approximate entropy:

$$COV = \sum_{i \neq j \in N} \sum p_{iT} p_{jT} \qquad (12)$$

where p_{It} is the final belief over cell i resulting from path execution for a specific scenario (sequence of observations). We can see from the above function that minimal belief covariability (known target location) leads to minimum entropy.

In this case, the genetic algorithm uses the following objective somewhat mimicking entropy:

$$\min_{\{path\}} < COV > = \sum_{j \in N} p_{j0} \operatorname{cov}_j \qquad (13)$$

$$\operatorname{cov}_j = \sum_{k \in N} p_{kT}(X_k \big| \{s_m = l_m \delta_{mj}, l_m\}_{m \in V})) \sum_{i \in N|\{k\}} p_{iT}(X_i \big| \{s_m = l_m \delta_{mj}, l_m\}_{m \in V})) \qquad (14)$$

A related quantity of interest is the number of contending cells C_T +1 depicting remaining target cell location competitors by the end of time horizon T. Its expected value is given by:

$$< C > = \sum_{j \in N} p_{j0} C_j \qquad (15)$$

$$C_j = \frac{1}{\sum_{k \in N} p_{kT}^2 (X_k \big| \{s_m = l_m \delta_{mj}, l_m\}_{m \in V}))} - 1 \qquad (16)$$

For a given scenario, covariability and contending number of cells relate to one another through the relation: $COV = C_T/(C_T+1)$. This can be illustrated by an example involving exactly k' equally competing cells with probability $1/k'$ for which the number of contending cells clearly shows $C_T + 1 = k'$. C_T is an indicator of the number of remaining cells still to be visited beyond horizon T to successfully find the target.

C. Simulations

Computer simulations were conducted under the following conditions for twelve problem instances:

- Grid size N= 5x5, 10x10, 15x15
- Initial belief probability distributions:
  - belief magnitude: exponential and uniform
  - spatial: clustered and uniform
- $T \approx 0.4\,N$
- Maximum number of visits = 7
- p_c= 0.8 for all cells
- p_f= 0 and 0.1 for all cells
- Genetic algorithm (GA):
  - Population size = 20
  - λ = 10 (Population size/2)

p_{Mut}: (M_path, M_path_local_repair): 50%
p_{Xover}:Crossover rate = 100%
Stopping criterion: time limit of 30 minutes or minimal entropy threshold

The algorithm was implemented in C++ and run on a 0.8 GHz Pentium computer. Given the deficient power of the computer platform at our disposal in comparison to mainstream multiprocessing computer technology (e.g. 8 core, 3 GHz) currently available and used, a 30 minute run-time limit has been set to ensure feasible near-real-time execution in practice.

D. Results

A sample of random simulation results is reported in Table I for two false alarm rate cases and clustered data sets. Each entry corresponds to a separate problem instance. Specific instance, grid size and time horizon scenarios are presented in first, second and third column respectively. Differential performances over false alarm rates p_f=0 and p_f=0.1 scenarios are reported in the fourth column. Genetic algorithm performance is described in terms of expected entropy over initial entropy ratio. Relative information gain gaps differentiating both false alarm rate scenario classes are indicated in the last column. A better information gain performance is naturally shown for a false alarm -free scenario (minimum uncertainty), as final average entropy is expectedly smaller. Imperfect sensors (p_f> 0) always add uncertainty over perfect sensors, degrading relative solution quality (augmenting entropy).

It is surprising to see to what extent performance results may differ in gap magnitudes for slightly different false alarm rate parameters. A false alarm rate of 10% implies minimal information loss of approximately 30% for the data sample examined, translating the importance of uncertainty propagation through possible scenarios and the fact that even modest false alarm rate may ultimately have a significant impact on expected information gain.

TABLE I
RELATIVE GENETIC ALGORITHM PERFORMANCE FOR CLUSTER DATA SETS OVER FALSE ALARM RATES - $p_f \in \{0, 0.1\}$

Problem	|*N*|	Time Horizo n*T*	*Genetic Algorithm* $\overline{E}_T / E_0$ $p_f = 0$	$p_f = 0.1$	*Information gain gap %*
G5x5 clu10	25	10	0.4381	0.7176	49.8
G10x10clu40	100	40	0.2473	0.4684	29.4
G15x15clu80	225	80	0.2519	0.5483	39.6
G15x15clu100	225	100	0.3258	0.5433	32.3

Comparative results with *k*-step myopic procedures (*k*=1,2) and belief covariability minimization is reported in Table II for all problem instances for a false alarm rate p_f= 0.1. As for Table I, columns 1-3 first characterize problem instance, grid size and time horizon. The next table entries include normalized final expected entropy ($\overline{E}_T / E_0$) as well as the final relative expected number of contending cells for the genetic approach. The last two columns report relative performance gaps for 1 and 2 -step limited look-ahead myopic techniques, and the covariability measure -based genetic procedure respectively, in comparison to the proposed genetic algorithm (reference). Reported simulation results show that the entropy-based genetic algorithm generally outperforms both myopic heuristics and, the covariability–based genetic approach. The only exception (-2.3%) simply translates the fact that the genetic algorithm still remains a meta-heuristics and therefore cannot guarantee an optimal solution. In other respect, fluctuations variability in differential performance of the genetic algorithm over myopic procedures are shown to be important ranging from 6.9% to 61% highlighting the value of planning with look-ahead. It is also worth noticing that a top performance gaps (10x10 instances) mostly correlates with the relative final number of contending cells. As expected, the smaller the final number of contending cells, the smaller the uncertainty, and therefore, the larger the performance gap anticipated with alternate methods.

In contrast, separate simulation results show that minimizing covariability is always preferable to cell contention (final number of contending cells at time step *T*) minimization, as the former objective seems better aligned with the entropy function.

Given the state-of-the-art parallel computing technology currently available and the natural inclination of genetic algorithms to be easily/massively parallelized, significant speed-up and further refined approximate expected entropy decision model can be realistically envisioned to efficiently solve this problem.

TABLE II
RELATIVE PERFORMANCE - COVARIABILITY VS. MYOPIC HEURISTIC VS. GENETIC ALGORITHM FOR A SAMPLE DATA SET

Problem (*p_f=0.1*)	|*N*|	**Time Horizon** ***T***	***Genetic Algorithm*** $\overline{E}_T / E_0$	C_T/|*N*| %	**Myopic Heuristic** ***k-move look-ahead*** *k=1*	***IG gap%*** *k=2*	**GA Covariability** ***IG gap %***
G5x5 clu10	25	10	0.7176	10	18	4.2	0.7
G5x5 ran10	25	10	0.8494	47	8.1	8.3	-2.3
G5x5 exp10	25	10	0.6970	22	6.4	6.9	0
G10x10clu40	100	40	0.4684	0.04	81	61	21.3
G10x10exp40	100	40	0.7846	0.31	32	19	11.7
G10x10ran40	100	40	0.8745	0.61	23	24	11.7
G15x15clu80	225	80	0.5483	2	59	53	10.8
G15x15ran80	225	80	0.9030	61	12	10	2.5
G15x15exp80	225	80	0.8520	36	18	8	1.1
G15x15ran100	225	100	0.8769	58	11	10	1.9
G15x15exp100	225	100	0.8080	33	23	15	3.2
G15x15clu100	225	100	0.5433	2	56	29	7.4

VI. CONCLUSION

An innovative information-theoretic –based open-loop decision model with anticipated feedback explicitly accounting for false alarm sensor readings has been proposed to solve a single agent military logistics search-and-delivery path planning problem. Minimizing uncertainty about unit target location, a near optimal path plan is computed considering anticipated possible observation outcomes. For the first time, a general belief update formulation/scheme on target cell occupancy involving false alarms was successfully derived in an explicit form. Then, a novel genetic algorithm was proposed, providing near-optimal solutions to practical size search path planning problems. Computational results prove the algorithm to outperform or favorably compete with alternate approaches. Algorithm performance can be further enhanced using parallel computing to gain additional run-time savings.

Future work envisions search path planning problem extensions including threat risk and moving target in a time-varying hostile environment. Generalized sensor footprint and increasingly complex observation models problem variants are also contemplated.

REFERENCES

[1] S. Benkoski, M. Monticino, and J. Weisinger, “A survey of the search theory literature”, *Naval Research Logistics* 38, 1991, 469–494.

[2] L. Stone, Theory of Optimal Search, 2nd edition, Topics in Operations Research Series, *INFORMS*, 2004.

[3] Trummel, K.E. and Weisinger, J.R., “The complexity of the optimal searcher path problem”, *Operations Research,* 34 (2), 1986, pp.324-327.

[4] Y. Jin, Y. Liao, A. Minai, and M. Polycarpou, “Balancing search and target response in cooperative unmanned aerial vehicle (UAV) Teams”, *IEEE Trans on Sys Man and Cybern. Part B*, 36(3), 2006, pp. 571-587.

[5] T. H. Chung, J. Burdick, “Analysis of search decision making using probabilistic search strategies”, *IEEE Transactions on Robotics*, 28 (10), 2012, pp. 132-144.

[6] K.E. Wilson, R. Szechtman, R., M.P. Atkinson. A sequential perspective on searching for static targets, *European Journal of Operational Research,* 215(1), 2011, pp. 218–226.

[7] A.R Washburn. Search and Detection, 4th edn. Topics in Operations Research Series. *INFORMS*, 2002.

[8] Levner, E. and Kriheli, B. Search and Detection of Failed Components in Repairable Complex Systems under Imperfect Inspections, Advances in Computational Intelligence, *Lecture Notes in Computer Science*, Springer, Vol. 7630, 2013, pp. 399-410

[9] A. R.Washburn, “Branch and Bound Methods for a Search Problem”, *Naval Research Logistics*, 45, 1998, 243-257.

[10] G. Martins, A New Branch-and-Bound Procedure for Computing Optimal Search Paths. Master's Thesis. Naval Postgraduate School, 1993.

[11] H. Lau, and G. Dissanayake, “Optimal search for multiple targets in a built environment”, In *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, Edmonton, Alberta, Canada, 2005, pp. 228-233.

[12] H. Lau, Optimal Search in Structured Environments, *PhD Thesis*, University of Technology, Sydney 2007.

[13] G.A. Hollinger, Search in the Physical World. CMU-RI-TR-10-20, Robotics Institute, *PhD thesis*, Carnegie Mellon University, 2010.

[14] H. Lau, and G. Dissanayake, “Probabilistic search for a moving target in an indoor environment”, In *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2006, pp. 3393-3398.

[15] T. Chung, “On Probabilistic Search Decisions under Searcher Motion Constraints”, *Workshop on Algorithmic Foundation of Robotics VIII*, Guanajuato, Mexico. 2009, pp. 501-516.

[16] G.A. Hollinger, and S. Singh, “GSST: Anytime Guaranteed Search with Spanning Trees”, *Autonomous Robots*, 29(1), 2010, pp. 99-118.

[17] J.O. Royset and H. Sato, “Route Optimization for Multiple Searchers”, *Naval Research Logistics*, 57 (8), 2010, pp. 701-717.

[18] T. Cover, and J. Thomas, *Elements of Information-Theory,* 2nd edition, Wiley, 2006.

[19] Potvin, J.-Y. and Bengio, S., 1996. The Vehicle Routing Problem with Time Windows Part II: Genetic Search, *INFORMS Journal on Computing* 8, pp. 165–172.

[20] Goldberg, D., 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning.* New York: Addison-Wesley.