An Information-Theoretic-based Evolutionary Approach for the Dynamic Search Path Planning Problem

Mohamed Barkaoui, Jean Berger
Canadian Department of National Defence
Defence R&D Canada-Valcartier
Ottawa, Canada

Abdeslem Boukhtouta
Canadian Department of National Defence
Defence R&D Canada-CORA
Quebec, Canada

***Abstract* — A new information-theoretic-based evolutionary approach is proposed to solve the dynamic search path planning problem. Path planning is achieved using an open-loop model with anticipated feedback while dynamically capturing incoming new requests and real action outcomes/observations as exogenous events, to timely adjust search path plans using coevolution. The approach takes advantage of objective function separability and conditional observation probability independence to efficiently minimize expected system entropy, lateness and travel/discovery time respectively. Computational results clearly show the value of the approach in comparison to a myopic heuristics over various problem instances.**

Keywords— Genetic algorithms; unmanned aerial vehicle; information theory; dynamic search path planning.

I. INTRODUCTION

Path planning computational complexity is a major challenge for Unmanned Aerial Vehicles (UAVs) evolving in predominantly uncertain environment. A path solution plan must be computed in real time while a UAV is moving. Should the path planner fail to generate a feasible solution by a predetermined deadline, mission failure would possibly occur.

Despite its intrinsic computational complexity, explicit solutions proposed for Search Path Planning (SPP) are abundant. Many contributions on SPP may for instance be found in the robotics literature, in particular in the robot motion planning ([1], [2]) field, such as terrain acquisition ([3],[4]) and coverage path planning ([5],[6],[7]). Robot motion planning envisioned SPP primarily focusing on coverage problem instances from a constrained shortest path perspective [8],[9],[10]. Typical problem characteristics involve unknown static targets and obstructions in uncertain environment. But, most approaches ultimately resort to a constraint relaxation strategy to reduce computational complexity and mitigate the curse of dimensionality. Alternatively, methods originating from search theory largely propose search techniques such as branch and bound ([11], [12], [13]), A* and related variants. However, in this case, the main challenge in designing efficient heuristics lies on the discovery and computation of tight bounds which remain largely difficult to achieve [14].

The proposed work consists in optimizing path planning to successfully search and detect multiple stationary targets [15]. It relaxes some of the constraints generally assumed or imposed by search approaches developed so far. These relate to partially observable environment state, explicit information feedback exploitation, imperfect sensors (e.g. false-alarm), limited computational resources, or a mix of them.

In this paper, we propose a new information-theoretic-based evolutionary approach to solve the dynamic SPP problem. It defines an anytime/interruptible coevolutionary algorithm coping with dynamic exogenous events, such as UAV action outcomes/observations or new occurring requests. Between two successive events, the coevolutionary algorithm solves an augmented static open-loop SPP problem model, which explicitly captures anticipated information feedback (observation or action outcome). In the expanded model, a fleet of homogeneous UAVs explores a search area to minimize target zone occupancy uncertainty. A prior zone occupancy probability distribution is assumed to be known. Control of the imperfect sensing vehicles is centralized. The separability property of the system entropy objective function, coupled to conditional independence over anticipated observation events, which facilitate efficient objective function pre-computing, permit to derive a new and original decision problem model formulation. The resulting decision model incorporating false alarms and anticipated action outcome feedback naturally lends itself to parallel computing, paving the way to rapidly solve practical size problems. Large horizon problems may on the other hand be tackled dynamically by repeatedly solving a new open-loop problem instance at each time step over a receding horizon, using real information feedback (observation outcome) from the previous episode. The advocated strategy exploits episodic information feedback to opportunistically improve solution quality rather than averaging path solution performance over distant horizons for all possible outcome sequences. We use a Genetic Algorithm (GA) to construct a set of potential paths containing existing zone visits as well as possible future revisits. The algorithm refines current zone visit plans while including potential future revisit request. A comparative computational experiment has been conducted to show potential gain of the proposed approach. Reported results prove its value in comparison to other heuristics.

DRDC-RDDC-2014-P8

The content of the paper is organized as follows. Problem definition, describing the main characteristics of the dynamic SPP problem, is first introduced. Then, the main solution concept is presented, introducing a new information-theoretic-based evolutionary approach to compute a near-optimal solution. The next section reports and discusses computational results comparing the value of the proposed method with alternative techniques. Finally, a conclusion is given in the last section.

II. Problem Characterization

The targeted closed-loop SPP problem involves a fleet of homogeneous UAVs, searching stationary targets in a bounded environment over a given time horizon. Planning is assumed centralized while considering a hierarchy of objectives. The first objective is to maximize information gain or equivalently to minimize uncertainty or entropy about target occupancy within a given search region, the second consists in minimizing lateness, and the third aims at minimizing target discovery time. The proposed multi-objective hierarchy is aligned to a lexicographic ordering, for which overall quality of computed solutions are ranked against the respective related objectives, in that order. Remaining responsive to dynamic exogenous events, the coevolutionary approach consists in solving an open-loop SPP with anticipated feedback (action outcomes/observations) problem model over a rolling horizon while gradually incorporating information feedback made available. Performed by a single base control station, a path planning episode takes place over a period separating two successive visits. The search region is composed of geographically distributed zones, possibly populated by non-cooperative stationary targets. Targets separately occupy a single zone. Target cardinality and respective positions are assumed unknown. Derived from domain knowledge, a probability density distribution defining individual zone occupancy characterizes prior target location. Occupancy probability distribution is assumed zone-independent. Zones are visited by a homogeneous fleet of UAVs, initially located at a central base station. Paths are assumed to start and end at the central base station. All zones must be visited within a specified time interval and a specific deadline is set to safely complete all surveillance/reconnaissance tasks. Under centralized controlled UAVs act as stand-off imperfect sensing agents collecting sensor readings while periodically communicating state and plan information back and forth with the base control station. UAV's speed and zone visit time are assumed constant. Reflecting vehicle's autonomy, UAV flight time should not exceed a predetermined maximum travel time. The team embraces a simple collision avoidance policy in which members fly at different altitudes. A UAV path solution to the search problem aims at minimizing uncertainty (entropy) over zone target occupancy, lateness and target discovery time respectively

In the current dynamic problem setting, we assume the next UAV destination (intended) to be communicated by the dispatching system at each visit location, or upon zone visit completion. The next destination is determined according to the best computed path plan available. The dispatching system may advise of any new destinations during the problem-solving process as necessary. It should be emphasized that a UAV traveling to its next destination is fully committed to visit its target. Consequently, aspects such as diversions for a moving UAV have not been considered.

A. Observation and sensing model

During episode t, a UAV visits a zone to determine target occupancy. Modeling partial world state agent observability, the observation model governs agent sensor's perception [16]. 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 models uncertainty using conditional probability of detection and false alarm, given zone target vacancy or occupancy state $X \in \{0,1\}$ respectively:

z_t: observation of zone occupancy at the end of period
t : {positive = 1, negative = 0}
$p_c = p(z_t = 1 | X = 1)$ probability of correct observation
$p_f = p(z_t = 1 | X = 0)$ probability of false alarm

These parameters are understandably cell-dependent reflecting specific sensor sensitivity and terrain features and conditions (e.g. landscape, obstacles, clutter, visibility, luminosity). The observation model is assumed to be known by the decision-maker. Agent sensor's range defining visibility or footprint is limited to the zone being searched.

B. Bayesian update

From a real or anticipated UAV sensor observation, local zone target occupancy beliefs ($p(X$=1)) 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 equation (2), p_{t-1} and p_t refer to prior and posterior zone target occupancy probability (belief) respectively.

C. Path planning objective

A centralized decision-making process episodically makes a vehicle's SPP decision based on vehicle's position. The objective consists in constructing a plan modeled as a sequence of moves to minimize entropy (target occupancy uncertainty) over the entire region. From information theory, the entropy function E is defined as [17]:

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

where $p(x)$ specifies the current probability/belief of cell target occupancy, and x a binary zone occupancy state. A zone with a zero entropy value means absolute certainty about occupancy or vacancy, whereas a maximum entropy value (1) refers to complete uncertainty. Decision-making is subject to limited computational resources imposed by episode duration.

A threshold of entropy E^* is used to set the number of zone visits. When zone entropy is smaller than E^*, then certainty may be assumed about target occupancy. Otherwise, a new visit is made. The process continues until a maximum number of visits have been conducted. The resort to a maximum number of visits helps prevents excessive zone revisit requests. Visit ordering on a given zone has no impact on zone entropy even if the zone is visited by different UAVs. However, the total lateness and total travel time may be affected.

III. Information-Theoretic Evolutionary approach

The proposed path planning approach is an evolutionary algorithm based on an information-theoretic framework. Path planning ultimately results from the coevolution of two populations of "plan" individuals describing a sequence of zone visits over a given horizon. An individual's score is computed and ranked according to the hierarchical objective structure proposed earlier and based on lexicographic ordering, that is: maximize information gain, minimize lateness and target discovery time.

A. Expected entropy

Path solution quality relies on an information-theoretic approach aimed at minimizing expected system entropy. The approach captures uncertainty related to target occupancy, projecting average entropy resulting from anticipated path plan execution. As target occupancy may be assumed independent with respect to the zones of interest, and that observation outcomes mainly rely on current UAV positions, the objective function is composed of separate contributions involving partitioned subsets of decision variables. Zone entropy strict dependency on local visits conducted only, shows an expected entropy objective function separable. Separability enables expected zone entropy $\overline{E_{z,l}}$ pre-computing as related values mainly depend on sensor characteristics and visit multiplicity (l) on zone z. Expected entropy value alludes to possible future observation outcome scenarios intimately related to imperfect UAV sensors characterizing partial environment observability and ultimately anticipated information feedback resulting from path plan execution. Simulating a given path plan, expected zone entropy is obtained by computing projected average uncertainty (entropy) values over all possible courses of observation outcomes. As zone visit ordering is ultimately invariant with respect to expected zone entropy, symmetry on sequence of observation outcomes may be further exploited. Hence, for a given a number of success (s anticipated positive observations) and failure (l-s anticipated negative observation outcomes) events, scenarios may be segregated in classes for which the probability of occurrence is described by a binomial distribution over possible observation outcomes. Local expected entropy on zone z resulting from a simulated path plan including l visits is defined as follows:

$$\overline{E_{z,l}} = \sum_{s=0}^{l} p_z(s \mid l) E\left(p_z(x=1 \mid s,l)\right) \quad l \geq 1 \tag{4}$$

where

$$p_z(x=1 \mid s,l) = \frac{1}{1+\alpha_z^s \beta_z^{l-s}\left(\dfrac{1}{p_{z,0}(x=1)}-1\right)} \quad ;$$

$$p_z(x=0 \mid s,l) = 1 - p_z(x=1 \mid s,l) \tag{5}$$

$$p_z(s \mid l) = \sum_{x \in X} p_z(s \mid l,x) p_{z,0}(x) \tag{6}$$

$$= p_z(s \mid l, x=1) p_{z,0}(x=1) + p_z(s \mid l, x=0)\left(1 - p_{z,0}(x=1)\right)$$

$$p_z(s \mid l, x=1) = \binom{l}{s}\left(p_z^c\right)^s\left(1-p_z^c\right)^{l-s} \tag{7}$$

$$p_z(s \mid l, x=0) = \binom{l}{s}\left(p_z^f\right)^s\left(1-p_z^f\right)^{l-s} \tag{8}$$

$$\alpha_z = \frac{p_z^f}{p_z^c} \quad ; \quad \beta_z = \frac{1-p_z^f}{1-p_z^c} \tag{9}$$

$$\overline{E_{z,0}} = E_{z,0} \tag{10}$$

Posterior zone z target occupancy belief for an s-success l visit scenario is represented by equation (5). It naturally emerges from expressions (1) and (2) after multiple observations, where $p_{z,0}$ stands for the initial target occupancy belief in zone z. $p_z(s|l,x)$ is a binomial probability distribution of positive observations, giving the probability to obtain s success out of l visits to zone z, conditional on occupancy state x. The probability of correct observation and false alarm rate on zone z are respectively described by p_z^c and p_z^f, as introduced earlier. $E_{z,0}$ reflects actual zone z entropy. The use of homogeneous vehicles and the exploitation of symmetry over equivalent sequence of success/failure events reduce complexity to a linear number of scenarios to be examined, as specific scenario event order does not matter.

B. Problem-solving approach

The main problem objective is to minimize uncertainty or entropy about target occupancy within a given region. Therefore, it seems intuitively appealing to explicitly consider anticipatory information about zones that may require subsequent visits (revisits) in the near future, as part of the heuristic decisions as well. The strategy is to continuously generate plans that are consistent with past decisions while anticipating future requests by considering both existing and potential future visit requests during plan generation. More precisely, the new strategy introduces "dummy" revisits representing revisit request likely projections in UAV paths. The expected number n of such visits for a given zone z is dictated by expressions (11) and (12):

$$\begin{cases} \dfrac{1}{1+\beta_z^n\left(\dfrac{1}{p_{z,0}(x=1)}-1\right)} \le p_1^* & \text{if } p_{z,0}(x=1)\le 0.5 \quad (11) \\ \dfrac{1}{1+\alpha_z^n\left(\dfrac{1}{p_{z,0}(x=1)}-1\right)} \ge p_2^* & \text{otherwise} \quad (12) \end{cases}$$

n represents the minimum number of zone z visits assumed necessary to confirm target occupancy state. User-defined target occupancy probability thresholds p_1^* and p_2^* are used to confirm target vacancy and occupancy respectively.

We use a GA to construct a set of potential paths containing existing zone visits as well as possible future revisits. This ability to capture path construction lookahead without violating temporal constraints confers a significant advantage. A larger number of zone visits can then be considered in building a solution, leading possibly to some quality improvement. The strategy is based on the premise that better solutions may emerge when taking advantage of a possibly larger number of zone visits due to additional requests that may occur in the near future. It is therefore assumed that better opportunities generated by considering possible future requests compensate the cost of myopic scheduling opportunities, ultimately resulting in solution improvement.

Individuals represent expandable solutions capturing currently planned visits, as well as previously serviced zones, while dynamically accommodating incoming zone visit requests. Zones may require several visits before confirming target occupancy state (zone entropy reaches threshold value E^*). Zone entropy is updated after each visit using Eqs (3) and (5). When a zone is visited, outdated planned visits from all solution individuals are then delayed accordingly and updated using a large penalty cost for lateness. Feasible solutions for initial populations are first generated using a sequential insertion heuristic in which zone visits are arbitrarily inserted in random insertion positions within paths. This strategy is fast and simple while ensuring unbiased solution generation.

The processing of new zone visit request is primarily entropy-driven. A new zone service request automatically occurs when zone entropy resulting from the latest visit is still significantly large. If multiple visits are candidates for insertion in planned paths, the first visit to insert will be the one minimizing expected zone entropy (Eq. (4)). The insertion strategy related to zone visit requests is described as follows:

- A new visit request is inserted to minimize solution expected entropy. Visit requests and potential future revisit requests are generated as well through equations (11) and (12), and considered for insertion.
- A new zone visit request along with its potential future revisit requests (using Eqs (11) and (12)), are first considered for insertion in the current plan solution.
- New revisit request on a given zone replaces one of its anticipated revisit, if any; otherwise it will be inserted in one of the planned paths of the current solution, assuming an admissible insertion position. Should such admissible insertion position not exist, a beneficial exchange with planned revisits would be explored. Otherwise the zone revisit request will be stored and reconsidered for future insertion.
- A zone visit can be removed (for exchange purposes) from planned path to insert another visit which could minimize entropy. Removed visits can be subsequently reinserted in the planned paths.
- Anticipated zone revisits that do not occur are ultimately removed from planned paths.
- Zone visit insertion is biased toward entropy minimization first, and then conditionally on at least one feasible insertion position, the resulting solution is further refined to minimize lateness and travel time.

C. *Problem-Solving Process*

The new information-theoretic-based evolutionary approach is performed within previously reported hybrid GA heuristic [18]. Zone entropy and zone revisits are included for consideration in the problem-solving process. The algorithm mainly relies on the basic principles of genetic algorithms, disregarding explicit solution encoding issues for problem representation. In the beginning there is a pre-optimization phase in which solutions are generated as a set of paths visiting known zones. These solutions are expanded to handle incoming visit requests while properly accounting for previously visited zones (history) or committed visits. Since visit ordering on a given zone has no impact on zone entropy, only travel time and lateness are considered in fitness and objective functions to evaluate solution individuals. An exchange procedure exploring alternate path visit insertion swaps from a subset of pending requests to maximize information gain is then executed. The evolutionary process was kept as in [18].

The insertion procedure and the genetic operators (selection, crossover and mutation operators) used in our algorithm are borrowed from [19]. All operators are inspired from the best insertion heuristics for the routing problem. These operators are much more sophisticated than those used in standard GAs. Diversification is implicit to each proposed operator, reducing the need for a specific mutation operator.

The exchange procedure focuses on local path improvement. It attempts to exchange sets of zone visits involving pending visits for insertion. In that scheme, each pending visit is exchanged with alternate planned zone visits in order to generate better path plans. Each current and pending visit pair is explored for swapping using the insertion procedure. Using a "first admissible" improving solution strategy, zone visit exchanges occur as soon as solution quality increases.

IV. COMPUTATIONAL EXPERIMENTS

A computational experiment has been conducted to show the value of the proposed approach. The experiment aims at comparing performance of the proposed evolutionary algorithm to alternate heuristics. Computed solutions from respective methods are reported against differential relative information gain. Differential relative information gain RIG(T) between two heuristics (1 and 2) shown at the end of time horizon T is defined as follows:

$$RIG(T) = \frac{\left|g_T^1 - g_T^2\right|}{\max\left(g_T^1, g_T^2\right)} = \frac{\left|\left(E_0 - \overline{E_T^1}\right) - \left(E_0 - \overline{E_T^2}\right)\right|}{E_0 - \min\left(E_T^1, E_T^2\right)}$$

$$= \frac{\left|\overline{E_T^2} - \overline{E_T^1}\right|}{E_0 - \min\left(E_T^1, E_T^2\right)} \quad (18)$$

where E_0 and $\overline{E_T^i}$ are the initial system entropy and the expected entropy computed by the heuristic i at the end of the time horizon T respectively. The larger the final entropy gap, the better the relative performance.

Computer simulations were conducted under the following conditions. Each problem instance involves 100 zones, randomly distributed over a geographical area. The travel time separating two zones corresponds to their relative Euclidean distance. Zone locations for a problem instance are either generated randomly using a uniform distribution (problem data set R1) or combining randomly distributed and clustered zones (problem data set RC1). The proportion of unknown zone requests is approximately 50% (degree of dynamism). Real-time requests are generated over specific simulation time horizons (15 minutes). Such a scenario involves approximately three requests per minute. The total number of UAVs is fixed in advance for each problem instance. The initial zone entropy value is generated from an exponential distribution.

A. Myopic algorithms

The performance of the proposed GA has been compared with a myopic version of our approach (GA_Myopic), in which potential future visit requests associated with unconfirmed target occupancy state are explicitly ignored during plan generation. Also, a greedy one-step limited lookahead method (Greedy) is used for comparison purposes. The greedy procedure consists in myopically planning moves one step ahead of time, progressively visiting the zone with highest expected information gain. After a zone visit, zone entropy is updated accordingly and the next destination is the one maximizing expected information gain. The procedure is then reiterated for each episode over time horizon *T*.

GA and GA_Myopic parameter values are selected according to [19]. GA and GA_myopic both use crossover and mutation rates of 0.8 and 0.6 respectively. The migration parameter refers to the number of (best) solutions mutually exchanged between populations after each generation (Migration parameter = 2). The population size remains fixed (10 individuals) and an elitist strategy is employed for both approaches. The elitist strategy consists to preserve the best two individuals to the next generation (Population overlap per generation =2). Past experience has proved these parameters to be fairly acceptable for a variety of problems, and therefore, turn out to be a natural choice for both approaches. Further information on operator combinations for GA and GA_Myopic can be found in [19]. Parameters in fitness and evaluation functions and in equations (4-12) are given as follows:

$E^*=E(0.96)$, $p_1^* = 0.04$, $p_2^* = 0.96$

$p_z^c = 0.8,\ p_z^f = 0.1\ \ \forall z$

The maximum number of visits per zone is bounded to four. The proposed algorithms have been implemented in C++, using the GAlib GA library [20]. The experiment consisted in performing four simulation runs for each problem instance while limiting run-time (time horizon) to 900 seconds.

B. Results

Solutions can be analyzed based on hierarchical ranking considering information gain first, total lateness second, and then, total travel time. The first analysis of the data focuses on the entropy, since this is the first objective. Since hierarchically the entropy is the first element to minimize, total lateness becomes relevant only if approaches obtain a final solution displaying the same entropy. Travel time is considered to break ties over entropy and total lateness.

Results for 20 problem instances characterizing R1 and RC1 data sets are summarized in Tables I and II. The first column presents problem instance identifiers. Measures of performance are the entropy (Entropy), total lateness (Lateness) and total travel time (Travel Time). We conducted a paired-sample Wilcoxon unsigned-rank test to ascertain if there is any significant difference between GA and the other approaches (with an error-level of 5%).

TABLE I. PERFORMANCE COMPARAISON AND RELATIVE INFORMATION GAIN OF GA OVER GREEDY (ENT=ENTROPY, LAT=LATNESS)

Problem	Greedy			GA			RIG(T)
	Ent.	Lat	Travel Time	Ent.	Lat.	Travel Time	
r101	44.8	925.2	5986.1	29.80	90.13	5817.4	0.65
r102	44.1	954.0	5936.7	29.51	136.7	5807.7	0.63
r103	45.9	519.9	5961.9	28.28	108.0	5898.5	0.72
r104	46.0	595.8	5917.9	26.01	26.64	5927.2	0.75
r105	47.3	1859.	5906.2	30.70	20.97	5930.6	0.75
r106	46.2	1284.	5331.6	29.52	20.44	5381.6	0.72
r107	46.3	843.9	5363.3	29.68	58.68	5275.3	0.72
r108	45.8	95.24	5049.3	31.45	7.73	5065.4	0.67
r109	47.6	2159.	5688.9	28.36	0.00	5631.0	0.79
r110	47.6	544.9	5363.6	29.29	0.00	5384.3	0.78
r111	47.0	392.7	5964.2	26.25	0.00	5947.7	0.79
r112	47.2	247.8	5961.3	25.34	0.00	5938.3	0.80
rc101	47.4	1235.	6200.7	30.29	45.73	6154.2	0.76
rc102	46.9	1093.	6177.7	28.98	59.78	6195.7	0.75
rc103	48.0	480.7	5547.4	30.71	3.62	5621.4	0.78
rc104	45.7	197.0	5918.9	26.69	5.22	5948.1	0.73
rc105	48.0	1032.	6177.8	28.04	37.12	6195.6	0.81
rc106	46.7	826.7	5643.9	31.80	0.00	5639.6	0.71
rc107	47.1	632.4	5563.9	29.32	0.00	5579.0	0.76
rc108	49.0	64.06	4959.1	32.34	0.00	4961.1	0.81
Aver.	**46.7**	**799.3**	**5731.0**	**29.12**	**31.04**	**5715.0**	**0.74**

TABLE II. PERFORMANCE COMPARAISON AND RELATIVE INFORMATION GAIN OF GA OVER GA_MYOPIC (ENT=ENTROPY, LAT=LATNESS)

Problem	GA_Myopic			GA			RIG(T)
	Ent.	Lat	Travel Time	Ent.	Lat.	Travel Time	
r101	33,2	109,5	6006,0	29,80	90,13	5817,4	0,15
r102	32,5	335,2	6000,9	29,51	136,7	5807,7	0,13
r103	31,0	50,39	5989,4	28,28	108,0	5898,5	0,11
r104	28,5	0.00	5900,4	26,01	26,64	5927,2	0,09
r105	31,0	22,23	5949,6	30,70	20,97	5930,6	0,01
r106	33,7	12,38	5393,7	29,52	20,44	5381,6	0,18
r107	32,9	14,29	5367,2	29,68	58,68	5275,3	0,14
r108	32,8	15,18	5048,1	31,45	7,73	5065,4	0,06
r109	30,7	0.00	5688,7	28,36	0.00	5631,0	0,10
r110	34,0	0.00	5365,5	29,29	0.00	5384,3	0,20
r111	30,0	0.00	5892,4	26,25	0.00	5947,7	0,14
r112	28,2	0.00	5954,5	25,34	0.00	5938,3	0,10
rc101	35,1	1,07	6198,5	30,29	45,73	6154,2	0,22
rc102	31,3	5,36	6228,7	28,98	59,78	6195,7	0,10
rc103	32,3	2,18	5592,8	30,71	3,62	5621,4	0,08
rc104	30,6	10,70	5836,0	26,69	5,22	5948,1	0,15
rc105	33,3	5,09	6227,4	28,04	37,12	6195,6	0,21
rc106	32,3	36,73	5603,4	31,80	0.00	5639,6	0,03
rc107	31,6	30,75	5621,8	29,32	0.00	5579,0	0,10
rc108	34,0	4,71	4950,5	32,34	0.00	4961,1	0,08
Aver.	**31,9**	**32,79**	**5740,8**	**29,12**	**31,04**	**5715,0**	**0,12**

Our approach GA outperforms the greedy procedure as it clearly shows better results for all problem instances. The overall average comparative performance gap for the examined data set is more than 74% as shown in Table 1. Some instances even exhibit an 81% gain. The differential relative entropy clearly demonstrates the value of predictive/advance planning over a limited lookahead myopic attitude. Furthermore, results of the greedy approach may be explained by the ignorance of travel time in selecting next UAV destination. This often leads to long UAV tours to visit remote zones at the expense of other zones due to time constraints. Table II compares GA_Myopic and GA methods showing an alternative average 12% performance gap. In GA_Myopic and GA, visits are inserted into path plan positions minimizing lateness and total travel time. Genetic operators are naturally inclined to generate path plans with the maximum number of visits. GA is shown to significantly improve information gain over GA-Myopic.

V. CONCLUSION

An information-theoretic-based evolutionary approach has been proposed to solve the centralized dynamic SPP problem. Problems having large time horizons may be adapted to a dynamic setting by repeatedly solving new static problem instances over a rolling horizon, incorporating observation outcomes from the last episode. Exploiting system entropy separability, conditional independence of observation events, and considering potential future zone visit requests during plan generation, the proposed approach turns out to be very efficient and competitive when compared to alternate methods.

Future work aims at naturally extending the current decision model to capture heterogeneous UAVs, and investigate its practical limitations. Alternate directions consist in adapting the approach to different search objectives such as proportion of target discovery or expected detection time minimization. Other challenges lie in modeling search problem variants incorporating more complex observation models and various target occupancy dependency and domain constraints, possibly infringing separability and symmetry assumptions. Multi-dimensional search problems involving complex domain knowledge modeled as belief networks represent another challenge as well.

REFERENCES

[1] Latombe J.-C. (1991). Robot Motion Planning. The Kluwer International Series in Engineering and Computer Science, vol. 124. Boston, MA, Kluwer.

[2] LaValle S. M. (2006). Planning Algorithms, Cambridge University Press.

[3] Choo C., Smith J., and Nasrabadi N. (1991). An efficient terrain acquisition algorithm for a mobile robot, in Proc. IEEE Int. Conf. Robotics and Automation, Sacramento, CA, pp. 306–311.

[4] Sankaranarayanan A. and Masuda I. (1992). Sensor based terrain acquisition: A new, hierarchical algorithm and a basic theory, in Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems, Raleigh, pp. 1515–1523.

[5] Wong S. and MacDonald B. (2003). A topological coverage algorithm for mobile robots, in Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems, Las Vegas, pp. 1685–1690.

[6] Svennebring J. and Koenig S. (2004). Building terrain-covering ant robots: A feasibility study, Auton. Robots, vol. 16, no. 3, pp. 313–332.

[7] Yang S. and Luo C. (2004). A neural network approach to complete coverage path planning, IEEE Trans. Syst.,Man, Cybern. B, Cybern., vol. 34, no. 1, pp. 718–724.

[8] Wagner W., Lindenbaum M. and Bruckstein A. M. (1999). Distributed covering by ant-robots using evaporating traces, IEEE Trans. Robot. Autom., vol. 15, no. 5.

[9] Rekleitis I., New A. P., Rankin E. S. and Choset H. (2008). Efficient Boustrophedon Multi-Robot Coverage: an algorithmic approach, Ann Math Artif. Intell. 52: pp.109–142.

[10] Agmon N., Hazon N. and Kaminka G. A. (2008). The giving tree: constructing trees for efficient offline and online multi-robot coverage, Ann Math Artif. Intell. 52: pp. 109–142

[11] Martins G. (1993). A new branch-and-bound procedure for computing optimal search paths, Master's Thesis. Naval Postgraduate School.

[12] Washburn A. R. (1998). Branch and bound methods for a search problem, Naval Research Logistics, Vol. 45, pp. 243-257.

[13] Lau H. and Dissanayake G. (2005). Optimal search for multiple targets in a built environment, In Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems

[14] Lau H. (2007). Optimal Search in Structured Environments, PhD Thesis, University of Technology, Sydney

[15] Yang Y., Minai A. and Polycarpou M. (2005). Evidential Map Building Approaches for Multi-UAV Cooperative Search, Proceedings of the American Control Conference.

[16] Berger J., Happe J., Gagne C. and Lau M. (2009). Co-evolutionary information gathering for a cooperative unmanned aerial vehicle team, 12th International Conference on Information Fusion, FUSION '09.

[17] Cover T. and Thomas J. (2006). Elements of Information-Theory, 2nd edition, Wiley.

[18] Berger J., Barkaoui M. and Boukhtouta A. (2007). A hybrid genetic approach for airborne sensor vehicle routing in real-time reconnaissance missions, Aerospace Science and Technology, 11(4), pp. 317-326.

[19] Barkaoui M., Berger J. and Boukhtouta A. (2008). A Hybrid Genetic Approach for the Dynamic Vehicle Routing Problem with Time

Windows, American Journal of Mathematical and Management Sciences, Volume 28, No 1&2, pp. 131-154.

[20] Wall M. (1995). GAlib - A C++ Genetic Algorithms Library, Version 2.4, MIT, Boston.